

PhD in Clinical & Experimental Medicine

**UNIVERSITÀ DEGLI STUDI DEL PIEMONTE ORIENTALE
“AMEDEO AVOGADRO”**

Dipartimento di Medicina Traslazionale

Corso di Dottorato di Ricerca in Medicina Clinica e Sperimentale

Ciclo XXVI

**Molecular Determinants of Follicular Lymphoma Transformation
to Diffuse Large B cell Lymphoma**

SSD (Settore Scientifico Disciplinare) della tesi MED/15

Coordinatore

Chiar.ma Prof.ssa Marisa Gariglio

Tutor

Chiar.mo Prof. Gianluca Gaidano

Dottorando

Dr. Marco Fangazio

INDEX

INTRODUCTION	5
MATERIALS AND METHODS.....	10
Study panel.....	10
DNA extraction.....	12
Analysis of IGHV-IGHD-IGHJ rearrangements	12
Whole-Exome capture and next-generation sequencing.....	13
High-density SNP array analysis	18
Dominant clone analysis	20
Analysis of recurrence	22
Algorithm for detection of Mutations with Common Focal genomic alterations (MutComFocal)	23
MutSigCV analysis of whole exome sequencing data	24
Functional annotation clustering of tFL target genes.....	24
Unselected FL cases.....	25
Sequencing analysis of ASHM target genes	26
<i>De novo</i> DLBCL cases.....	28
Fluorescence in situ hybridization (FISH)	28
Immunohistochemistry	29
Statistical analysis.....	30

RESULTS	31
Divergent evolution of FL and tFL from a common mutated precursor.....	31
Recurrent genetic lesions	34
Genetic lesions shared between FL and tFL	35
Genetic lesions specifically associated with tFL	38
Non-phase specific lesions.....	40
The genomic landscape of tFL is unique but more related to GCB-DLBCL	42
DISCUSSION AND CONCLUSIONS	44
FIGURES.....	48
Figure 1. Whole exome sequencing and mutation discovery approach.....	48
Figure 2. FL and tFL display shared and unique genomic aberrations.....	49
Figure 3. Distribution of shared and phase-specific SNVs in the discovery panel.....	50
Figure 4. FL and tFL arise through divergent evolution in most patients.....	52
Figure 5. tFL displays a complex genomic profile relative to FL.....	53
Figure 6. Recurrently mutated genes in tFL.	54
Figure 7. Distribution of recurrent genetic lesions in tFL.....	55
Figure 8. Significant regions of copy number alterations in the tFL genome.....	57
Figure 9. Mutations of histone modification and chromatin remodeling genes.....	58
Figure 10. Biallelic loss of CDKN2A/B is specifically acquired during transformation.....	59
Figure 11. Recurrent genetic lesions of <i>MYC</i> in tFL.	61
Figure 12. ASHM is associated with transformation.	62

Figure 13. Mutation features of ASHM target genes.	63
Figure 14. tFL displays a unique genomic profile that partially overlaps with GCB-DLBCL.	64
TABLES	65
Table I. Biological and phenotypic features of the discovery cases.	66
Table II. Results of Illumina Hi-Seq2000 sequencing and mapping after whole exome capture.	67
Table III. SNVs identified by whole exome sequencing in the 12 discovery FL-tFL pairs.	68
Table IV. Segments of tumor-acquired copy number alterations identified in the discovery panel.	80
Table V. Complete list of genes found mutated in >10% of tFL, with additional selected genes of relevance within the same pathway.	91
Table VI. Significant regions of deletion and amplification in tFL according to the GISTIC algorithm.	92
Table VII. SNVs identified by whole exome sequencing in genes mutated in >5% of cases.	93
REFERENCES	102
ACKNOWLEDGMENTS	110
PUBLICATIONS RELATED TO THE PhD PROJECT/COLLABORATIONS	111

INTRODUCTION

Follicular lymphoma (FL) is the second most common nodal B cell lymphoma, comprising ~20% of all new cases with a median age of 60 years at the time of diagnosis. (Swerdlow et al., 2008). While initially characterized by slow progression and high response rates to a variety of chemotherapeutic agents, FL shows a continuous pattern of relapses with decreasing sensitivity to anti-tumor therapy, ultimately leading to death (Fisher et al., 2005).

Morphologically, FL is defined as a clonal proliferation of malignant germinal center (GC) B cells with a partially follicular pattern. Malignant GC cells are usually mingled with nonmalignant cells such as T cells, follicular dendritic cells (FDCs), and macrophages (Swerdlow et al., 2008). Based on the absolute number of centroblasts (large or small) that are counted in at least ten high-power microscopic fields (hpf, 40x) in ten different neoplastic follicles, FL is classified into three different grades. Grade 1 and grade 2 cases have a marked predominance of centrocytes with only a few centroblasts (grade 1 = 0-5 centroblasts/hpf; grade 2 = 6-15 centroblasts/hpf). Grade 3 cases (>15 centroblasts/hpf) are further subdivided according to the proportion of centrocytes: in grade 3A, centrocytes are still present, while in grade 3B follicles are composed entirely of large blastic cells (centroblasts or immunoblasts) (Swerdlow et al., 2008). The grading of FL is clinically relevant as it guides different treatment plans and thus represents a keystone at the time of diagnosis. In particular, grade 3 FL (especially grade 3B with sheets of blasts) is considered an aggressive lymphoma and is typically treated like a DLBCL (Dreyling et al., 2011). conversely, the distinction between grade 1 and grade 2 is discouraged, given the virtual absence of differences in clinical behavior (Swerdlow et al., 2008).

The notion that FL arises from GC B cells is supported by the observations that neoplastic cells express GC surface markers like BCL6 and CD10, are subject to ongoing somatic hypermutation (SHM), are organized in structures that architecturally resemble normal GC follicles, and have a characteristic gene expression profile of centrocytes and/or centroblasts (Alizadeh et al., 2000; Kuppers, 2005; Swerdlow et al., 2008).

The genetic hallmark of 80 to 90% of FL cases is the characteristic t(14;18)(q32;q21) translocation. It occurs following a double-strand break that juxtaposes the gene encoding the heavy chain of the immunoglobulin gene (IGH) at chromosomal band 14q32 to the *BCL2* gene on the long arm of chromosome 18, cytoband 21 (Tsujimoto et al., 1985). Other less common variants have also been described, such as t(2;18)(p12;q21) and t(18;22)(q21;q11), both of which involve breakpoints in the loci of genes encoding the immunoglobulin light chains (kappa and lambda respectively) (Bentley et al., 2005). In all cases, the translocation results in the deregulated expression of the BCL2 protein, which is normally not expressed in GC B cells, leading to the activation of downstream anti-apoptotic programs. *BCL2* translocations have been attributed to defective RAG-mediated recombination events that may occur early during B cell development, when VDJ recombination takes place in the bone marrow, or later in the GC environment due to reactivation of the RAG genes. In the first model, the t(14;18) is believed to occur as the first genetic hit but does not prevent the further differentiation of the B cell. Upon exiting the bone marrow, naïve B cells carrying the t(14;18) colonize secondary lymphoid organs where they undergo the GC reaction. These translocated B cells, however, have a survival advantage due to their constitutive expression of the BCL2 protein and thus resistance to apoptosis (Roulland et al., 2011). The translocated B cells then are postulated to acquire secondary genetic lesions in the GC, under the influence of the activation-induced cytidine

deaminase (AID) enzyme, which normally mediates SHM and class switch recombination. The involvement of this enzyme in lymphomagenesis has been elegantly demonstrated in a mouse model where the genetic deletion of *AID* prevented BCL6-driven GC-derived lymphomas (Pasqualucci et al., 2008). This hypothesis of FL pathogenesis recognizes naive t(14;18)-positive circulating B cells as potential tumor-initiating cells, but is challenged by the observation that similar cells can be detected in the blood of 50%–70% of healthy individuals who are not prone to developing FL (Limpens et al., 1995).

Besides translocations, more recent genome wide high-throughput sequencing technologies have allowed a comprehensive characterization of the FL genome (Li et al., 2014; Morin et al., 2011). In particular, it has been shown that FL carries recurrent mutations of histone-modifying genes, including those encoding for the histone methyltransferases *MLL2* and *EZH2* (Morin et al., 2010; Morin et al., 2011), and the histone acetyltransferases *CREBBP* and *EP300* (Morin et al., 2011; Pasqualucci et al., 2011a). FL is also characterized by recurrent copy number abnormalities, in particular losses affecting the short arm of chromosome 1 and the long arm of chromosome 6 (Cheung et al., 2009). The 1p36 region is frequently affected by heterozygous deletions and copy-neutral loss of heterozygosity (cn-LOH); one candidate target of these lesions is the *TNFRSF14* gene, a member of the TNF-receptor superfamily, which was also found recurrently targeted by truncating mutations (Cheung et al., 2009). Furthermore, the systematic screening of genes within commonly deleted regions on the long arm of chromosome 6 revealed both *TNFAIP3/A20*, a negative regulator of NF- κ B signaling, and *EPHA7* a tyrosine kinase receptor, as potential tumor suppressors in FL (Oricchio et al., 2011).

Clinically, the classical presentation of patients with FL includes an asymptomatic lymphadenopathy, such that most cases are already at an advanced stage at the time of diagnosis

and 50% of them have bone marrow involvement leading to cytopenia. One of the currently most used models for the clinical stratification of FL patients is the Follicular Lymphoma International Prognostic Index (FLIPI), which was established in 2004 from a multicenter collaboration (Solal-Celigny et al., 2004). This prognostic system includes five parameters (age, stage, anemia, number of involved nodal areas, and serum lactate dehydrogenase) and stratifies patients into three prognostic groups with different 10-year overall survival (OS) rates, i.e. 71%, 51%, and 36% (Solal-Celigny et al., 2004). Most patients, however, develop over time increasing resistance to therapy with progressively shorter time to progression and lower response rate. Disease transformation to a more aggressive histologic type (most commonly DLBCL) is crucial in the history of these patients and remains a challenge for the physician, as it often signifies a terminal event (Al-Tourah et al., 2008; Bastion et al., 1997; Montoto et al., 2007). Histologic documentation is mandatory for diagnosing FL transformation, which otherwise can only be clinically suspected, but not proven; furthermore, an unequivocal definition of transformation requires the demonstration of a clonal relationship between the original FL and the subsequent DLBCL, by documenting the use of the same immunoglobulin gene rearrangement (Lossos and Gascoyne, 2011). FL transformation has been reported to occur in 16% to 70% of patients over time, depending on the study, with a rate of transformation calculated to be 3% per year, and is associated with a poor survival post-transformation (less than 2 years) (Al-Tourah et al., 2008; Bastion et al., 1997; Montoto et al., 2007).

Even though the notion of the FL transformation was described more than twenty years ago, few studies have been conducted to date in order to elucidate the genetic mechanisms driving the transformation process by using longitudinal series of paired FL and transformed FL (tFL) with documented clonal relationship between the two phases (De Jong et al., 1988; Lossos

and Gascoyne, 2011). These studies identified a number of genetic, epigenetic and tumor micro-environmental alterations that are associated to FL transformation, and include: i) mutations affecting p53 (Lo Coco et al., 1993; Sander et al., 1993); ii) inactivation of the CDKN2A/p16 gene by point mutations, deletions of locus on the short arm of chromosome 9, and epigenetic modification of the promoter leading to diminished p16 expression in the transformed biopsies (Elenitoba-Johnson et al., 1998; Pinyol et al., 1998). iii) MYC deregulation. While the identification of MYC translocations is rare in FL biopsies at the time of diagnosis, one study suggested an association with an increased risk of transformation (Christie et al., 2008). Moreover, gene expression profiles of longitudinal samples indicate that FL transformation is associated with increase in expression of MYC and its downstream targets (Lossos et al., 2002), which in the majority of the cases were unrelated to genetic aberration, (Lossos et al., 2002). Finally, analysis of selected genes in a subset of cases revealed an association between progression to DLBCL and aberrant somatic hypermutation (ASHM) (Rossi et al., 2006), a mechanism of genetic instability resulting from the abnormal functioning of the physiologic SHM process that operates in GC B cells (Pasqualucci et al., 2001).

However, these studies were based on small number of cases and used a candidate gene approach, as opposed to an unbiased, genome-wide analysis.

Our study was aimed at examining the history of clonal evolution during FL transformation to DLBCL and comprehensively identifying molecular determinants that underlie this process, and may in turn be translated into more effective therapies.

MATERIALS AND METHODS

Study panel

A discovery panel of 24 sequential fresh-frozen biopsies (12 pairs) obtained at FL diagnosis and at transformation to DLBCL (Table I) and a screening panel of 27 tFL samples were selected on the basis of the following criteria: i) histologic diagnosis of FL grade 1, 2 or 3a according to the World Health Organization Classification of Tumors of the Hematopoietic and Lymphoid Tissues (Swerdlow et al., 2008), in the absence of previous treatment; ii) for the tFL phase, histologic diagnosis of DLBCL, with $\leq 10\%$ residual areas of FL; iii) evidence of clonal relationship between the FL and the tFL sample, as assessed based on analysis of the rearranged immunoglobulin loci and, in two cases, by sequencing of the *BCL2-IGH* junction; iv) at least 6 months interval from initial FL diagnosis to transformation. Patients with FL grade 3b, composite lymphoma and/or the presence of diffuse areas in the FL biopsy were excluded. All diagnoses were independently reviewed by two pathologists. DNA extracted from both discovery and screening biopsies were used for whole-exome sequencing (WES) and SNP6.0 array analysis. Matched normal DNA was available for 4 of the 12 discovery pairs and was documented to lack contaminating tumor cells by PCR amplification of the clonally rearranged tumor-associated Ig genes as well as by SNP array analysis of the corresponding loci. Using these 12 pairs, we estimate a 99% probability of detecting mutations that affect genes at 30% prevalence and 93% probability for genes at 20% prevalence.

Discovery panel. The discovery panel comprised 24 sequential fresh-frozen biopsies (12 pairs) obtained at FL diagnosis and at transformation to DLBCL from the archives of the Division of Hematology, Department of Translational Medicine at the Amedeo Avogadro

University of Eastern Piedmont, the Department of Pathology & Cell Biology at Columbia University, the Department of Pathology at Northwestern University, the Department of Molecular Biotechnology and Health Science at the University of Torino, and the Pathology & Lymphoid Malignancies Unit at the San Raffaele Scientific Institute in Milan, after approval by the respective Institutional Review Board (Exempt Human Subject Research of anonymized/deidentified existing pathological specimens, under regulatory guideline 45 CFR 46.101(b)(4)). All FL biopsies were taken at initial diagnosis, i.e. from previously untreated patients. First line treatment in the paired FL/tFL cases was according to standard protocols, and included R-CHOP based regimens in most cases (Table I). The fraction of tumor cells in the tissue sample was assessed by SNP array genomic profiling performed on the same material utilized for WES, and was estimated to be >65% in 14 cases and $\geq 40\%$ in all cases, with one exception (6FL, <20% tumor cells). Matched normal DNA was available for 4 of the 12 pairs, and was documented to lack contaminating tumor cells by PCR amplification of the clonally rearranged tumor-associated immunoglobulin genes as well as by SNP array analysis of the *IG* loci. Using these 12 pairs, we estimate a 99% probability of detecting mutations that affect genes at 30% prevalence and 93% probability of detecting mutations that affect genes at 20% prevalence.

Screening panel. For the screening phase, 27 tFL samples were selected based on the criteria described above. A matched FL biopsy was available for 12 of the cases, either as formalin-fixed paraffin-embedded (FFPE) tissue (n=5) or as frozen material (n=7), and was used to verify whether tFL-associated mutations were already present at FL diagnosis by targeted Sanger re-sequencing. In 3 additional patients, matched normal tissue was available to document the somatic origin of the mutations. The 12 patients with available matched FL material and the

3 patients with available matched normal DNA were not included in the discovery panel because the DNA quality of the FL phase was inadequate for WES and/or SNP array analysis.

DNA extraction

High-molecular-weight genomic DNA was extracted from frozen tumor and matched normal samples according to standard procedures, quantitated by the Quant-iT™ PicoGreen® reagent (Invitrogen), and verified for integrity by gel electrophoresis. Clonality was established by amplification of the rearranged immunoglobulin genes and by SNP array analysis of chromosomal regions 14q32.33, 2p11.2 and 22q11.2. The QIAamp DNA FFPE Tissue Kit (Qiagen, Germantown, MD) was used to extract high quality DNA from FFPE tissue.

Analysis of IGHV-IGHD-IGHJ rearrangements

PCR amplification of *IGHV-IGHDIGHJ* rearrangements was performed on high-molecular-weight genomic DNA using IGHV leader primers or consensus primers for the IGHV FR1 along with appropriate IGHJ primers, and sequenced directly as previously described (Rossi et al., 2011). Sequences were analyzed using the IMGT databases and the IMGT/V-QUEST tool (version 3.2.17, Université Montpellier 2, CNRS, LIGM, Montpellier, France) as well as the NCBI/IGBLAST tool (<http://www.ncbi.nlm.nih.gov/igblast/>) (Ye et al., 2013), in order to define *IGHV* gene and allele usage, percentage of identity to the closest germline *IGHV* allele, VH CDR3 length and composition (including *IGHD* and *IGHJ* gene usage) and *IGHD* gene reading frame.

Whole-Exome capture and next-generation sequencing

Purified high-molecular-weight genomic DNA (~3 µg) from the 12 FL and 39 tFL samples (n = 12 from the discovery panel and 27 from the screening panel) was enriched in protein-coding sequences with the Agilent Sure Select Human 51Mb All Exon v4 Kit (Agilent Technologies) according to the manufacturer's protocol. The resulting target-enriched pool was normalized and combined (four-plex) before high-throughput paired-end (2 × 100 bp) sequencing was performed on the Illumina HiSeq2000 System at Centrillion Biosciences. The analysis produced on average 67.5 million passed-filter paired-end reads per sample (range, 51.7 to 111.6; Table II). After filtering for duplicate reads (defined as reads with identical start and orientation), sequences were aligned to the reference human genome hg19 assembly (GRCh37) with the Burrows-Wheeler Aligner tool (version 0.5.9). The mean coverage depth (i.e., the mean number of reads covering the target exome of a haploid reference) was 80.8× (range, 45.8 to 119.4) with an average of 89.3% of the captured region covered at >10× (range, 83.8 to 93.4) and 73.9% covered at >30× (range, 53.9 to 84.5; Table II). Sequence variants, including nucleotide substitutions and small insertions and/or deletions, were obtained independently for each tumor and normal sample with the Statistical Algorithm for Variant Identification (Trifonov et al., 2013b) and were independently validated by conventional Sanger sequencing.

Identification of sequence variants. The SAVI (Statistical Algorithm for Variant Identification) algorithm was used to identify sequence variants from WES data, independently for each tumor and normal sample (Rossi et al., 2012; Trifonov et al., 2013b). Briefly, we first constructed empirical priors for the distribution of variant allele frequencies in each sample; from that prior, we obtained a corresponding high-credibility interval (posterior probability $\geq 1 - 10^{-5}$) for the frequency of each variant and a high-credibility interval for the corresponding

change in frequency between the tumor and normal sample or the FL and tFL phase. Variants were considered absent if detected in $\leq 2\%$ of the sequence reads; the 3% lower detection threshold was defined based on the estimated one percent error rate of the sequencing technology and a binomial distribution for the errors, given a mean coverage depth of $>80X$ across samples (Table II). Candidate variants were then filtered for i) systematic errors known to be associated with the Illumina sequencing technology, ii) variants observed in only one strand, and iii) variants mapping to multiple loci in the genome, which may reflect captured pseudogenes and regions of low complexity. To this end, each variant with a flanking 35-base context sequence around its genomic position was mapped to the hg19 reference using the BLAST algorithm, and only variant reads with “unique mappability” were retained, that is, we required the 71-base sequence to uniquely map to the reference genome with only one mismatch. A comparison between the heterozygous dbSNP call rate obtained by sequencing (high confidence only) and the reference set of heterozygous dbSNP calls detected by Affymetrix SNP6.0 array analysis in the same cases showed that, at this level of resolution, we accurately detected both alleles in $>94\%$ of the heterozygous SNPs (Table II).

Identification of somatic SNVs. Candidate somatic SNVs (non-silent and silent) were obtained by subtracting i) variants present in the corresponding paired normal DNA, available in 4 discovery samples and 3 screening samples; ii) known population polymorphisms present in the NCBI dbSNP (Build 137) and Ensembl databases (<http://www.ensembl.org/info/genome/variation/index.html>), unless known to represent clinical SNPs; iii) variants present in any one of 105 exomes from unaffected individuals that were analyzed at Columbia University (Columbia 105). In pairs lacking matched normal DNA, sequence variants were also filtered out if present in both the FL and tFL biopsy (as these may

represent constitutional variants), unless they affected 52 target genes that were previously documented to have pathogenic significance in lymphoid malignancies and/or were reported as mutated in >10% of DLBCL cases by multiple independent DLBCL genome studies (Lohr et al., 2012; Morin et al., 2011; Pasqualucci et al., 2011b; Zhang et al., 2013):

<i>ARID1A</i>	<i>CD58</i>	<i>FAS</i>	<i>MYC</i>	<i>PRDM1</i>	<i>TNFRSF11A</i>
<i>B2M</i>	<i>CD79A</i>	<i>FBXO11</i>	<i>MYD88</i>	<i>RHOH</i>	<i>TNFRSF14</i>
<i>BAFFR</i>	<i>CD79B</i>	<i>FOXO1</i>	<i>NFKBIA</i>	<i>S1P2</i>	<i>TP53</i>
<i>BCL10</i>	<i>CIITA</i>	<i>GNA13</i>	<i>NFKBIE</i>	<i>SGK1</i>	<i>TRAF2</i>
<i>BCL2</i>	<i>CREBBP</i>	<i>GNAI2</i>	<i>NOTCH1</i>	<i>SOCS1</i>	<i>TRAF3</i>
<i>BCL6</i>	<i>EP300</i>	<i>KLHL6</i>	<i>PAX5</i>	<i>STAT6</i>	<i>TRAF5</i>
<i>BCL7A</i>	<i>EPHA7</i>	<i>MEF2B</i>	<i>PCDH7</i>	<i>TET2</i>	<i>ZNF521</i>
<i>CARD11</i>	<i>ETS1</i>	<i>MLL2</i>	<i>PCL0</i>	<i>TMEM30A</i>	
<i>CCND3</i>	<i>EZH2</i>	<i>MLL3</i>	<i>PIM1</i>	<i>TNFAIP3</i>	

Other genes reported in recent DLBCL sequencing efforts (n~3700) were not indiscriminately included, as the significance of these events is unclear. This conservative approach was taken because of the large number of shared FL/tFL variants identified in the absence of matched germline DNA, which most likely represent private polymorphisms or sequencing/mapping artifacts. This notion was validated experimentally by comparing the number of shared SNVs that were called when interrogating the entire set of ~3700 genes in 4 FL/tFL pairs with available matched normal DNA (#2, #5, #11, #23), before and after subtraction of their paired germline SNVs. By this approach, 1901 shared non-silent SNVs were identified in the four cases; while 1410 were subsequently subtracted because annotated in public databases and/or in our database of 105 Normal DNAs, only 58 of the remaining 490 SNVs were truly absent in the matched normal DNA (FDR = 88.2%). Conversely, only 4 events were

inappropriately called as somatic when restricting the geneset to the 52 selected genes (FDR = 17.4%).

	N of identified FL/tFL shared SNVs	N after subtraction of dbSNP137 and Normal 105	N after subtraction of matched normal (true somatic)	FDR
All published*	1901	490	58	88.2%
52 Selected Genes	56	23	19	17.4%

* Morin et al., 2011; Pasqualucci et al., 2011b; Lohr et al., 2012; Zhang et al., 2013.

Because constitutional variants are present by definition in each cell/sample of an individual patient, phase-specific variants that were not explained by CN loss or cnLOH of the affected region in the paired sample were considered as somatically acquired also in the absence of paired normal DNA.

In the screening panel, the somatic origin of the mutations identified was proven directly in 5 cases, for which adequate material was available from paired normal DNA or from the paired FL sample (for tFL-specific mutations). In the remaining 22 cases, the somatic origin of the mutations is suggested, although not proven, by their absence in several databases of germline variants, including the NCBI dbSNP database (Build 137), the Ensembl database, and the Columbia 105 normal exomes.

Validation of candidate somatic mutations by Sanger sequencing. Candidate somatic mutations were independently confirmed by conventional Sanger-based re-sequencing of PCR products obtained from normal, FL and tFL DNA using primers specific for the exon encompassing the variant. PCR primers were either derived from previously published studies (Parsons et al., 2011) or designed in the Primer 3 program (<http://frodo.wi.mit.edu/primer3/>) to

anneal ≥ 50 bp upstream or downstream of target exon boundaries, and were filtered using UCSC *In Silico* PCR to exclude pairs yielding more than a single product. All PCR primers and conditions are available upon request. Purified amplicons were sequenced directly (Genewiz, Inc, South Plainfield, NJ), and compared to the corresponding germline sequences using the Mutation Surveyor Version 4.0.8 software package (SoftGenetics, State College, PA; <http://www.softgenetics.com>) after automated and/or manual curation. Whole-genome-amplified DNA obtained using the Repli-g Mini kit (Qiagen, Germantown, MD) was used as template in the first round of validation. Candidate somatic mutations were then confirmed from both strands on independent PCR products obtained from high-molecular-weight genomic DNA. In selected samples showing more than one sequence variant within the same gene (e.g., *MLL2*, *TNFRSF14*), the allelic distribution of the mutations was determined by sequencing cloned PCR products (≥ 10 individual clones/sample) encompassing both variants and amplified from genomic DNA or cDNA. The validation screen revealed a false positive rate of 2.9% ($n = 10/344$) for variants called in $>20\%$ of the reads and 18.3% ($n = 15/82$) for variants called in 3-20% of the reads (before normalization for tumor content), consistent with the sensitivity of the Sanger sequencing approach (Figure 1). The false negative rate was 2.6% ($n = 3/117$), confirming the reported high specificity and sensitivity of the approach (Figure 1) (Rossi et al., 2012).

High-confidence somatic nonsynonymous mutations were tested for their functional consequences *in silico* by using two amino acid substitution prediction algorithms, including SIFT (Sorting Intolerant From Tolerant) (<http://sift.jcvi.org/>) and PolyPhen-2 (Polymorphism Phenotyping) (genetics.bwh.harvard.edu/pph2/) (Tchernitchko et al., 2004). Genes found mutated were verified for their presence in the COSMIC database

(<http://www.sanger.ac.uk/genetics/CGP/cosmic>) and in the Cancer Gene Census database (<http://www.sanger.ac.uk/genetics/CGP/Census/>). The Oncotator web application tool was used for comprehensive annotation of variant calls with data relevant to cancer (<http://www.broadinstitute.org/oncotator/>).

High-density SNP array analysis

Genome-wide DNA profiles from 12 FL, 39 tFL and 6 matched normal samples were obtained from high-molecular-weight genomic DNA using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) and standard protocols from Affymetrix, as described (Pasqualucci et al., 2011a). Data analysis and identification of regions of abnormal copy numbers was performed using the dChipSNP software (www.dchip.org) (Lin et al., 2004) and a normalization procedure, according to a published workflow (Mullighan et al., 2007; Pounds et al., 2009). The circular binary segmentation (CBS) algorithm was then applied to the SNP array data to identify candidate genomic regions of amplification and deletion, using the following criteria: i) mean log₂ ratios of ≥ 0.2 or ≤ -0.2 (≥ 0.17 or ≤ -0.17 in cases where the % of tumor cells in the biopsy was estimated to be $<30\%$ or for subclonal lesions); ii) ≥ 10 SNP markers (SNP and/or CNV) within a segment (Mullighan et al., 2007; Venkatraman and Olshen, 2007). To exclude calls of genomic gains or loss arising from inherited genomic copy number variants (CNVs) in the unpaired samples, 149 normal DNAs from four independent studies (Fabbri et al., 2011; Mullighan et al., 2007; Pasqualucci et al., 2011b; Rossi et al., 2012) as well as 230 normal DNAs from the HapMap Project were also analyzed in dChip and visualized simultaneously with the tumor samples. Alterations present in the pool of reference samples were presumed to be inherited and therefore eliminated. In addition, CNVs were excluded if the same

regions were present in the Database of Genomic Variants (<http://projects.tcag.ca/cgi-bin/variation/gbrowse/hg18/>). Adjacent regions separated by <10 markers and/or by germline CNVs were considered to be part of the same aberration if showing comparable probe intensity values. Genes (within the RefSeq database) were considered to be affected by deletions if one or more of their exons were encompassed by the aberration. The following linear thresholds were used to annotate copy number aberrations, which take into account the presence of contaminating normal cells in the biopsies: heterozygous deletion: ≤ 1.70 ; homozygous deletion: ≤ 0.9 ; copy number gain: ≥ 2.3 ; gene amplification: ≥ 4 . Additionally, the presence of both copy number gains and copy number losses affecting distinct subsets of exons within a given gene in the same tumor sample was annotated as a breakpoint, as this pattern could reflect an unbalanced rearrangement or a rearrangement with loss of the reciprocal translocation partner.

Identification of regions of cnLOH. Loss-of-heterozygosity (LOH) analysis was computed in dChipSNP (www.dchip.org) (Lin et al., 2004) using a Hidden Markov Model and an inferred LOH call threshold of 0.5, as previously reported (Mullighan et al., 2007). In cases lacking matched normal DNA, 149 non tumoral samples from our historical cohorts were used as normal reference (Fabbri et al., 2011; Mullighan et al., 2007; Pasqualucci et al., 2011b; Rossi et al., 2012). Segments of LOH were then exported from the dChipSNP software, visually inspected and manually curated in order to retain only segments of copy number neutral LOH; in samples lacking paired normal DNA, non-telomeric lesions that were ≤ 24 Mb in size were disregarded because of the impossibility to confirm the somatic origin of the alteration, as per previous studies (O'Keefe et al., 2010).

Identification of focal minimal common regions (MCR) of aberration. MCRs of aberration were identified as described (Pasqualucci et al., 2011b), using the curated list of

aberrant segments obtained from the SNP6.0 array analysis and the following criteria: i) the consensus interval across all affected tumors is smaller than 1Mb and/or encompasses a maximum of 3 genes; ii) for MCRs of gain, the entire coding region of the gene is included. These criteria exclude MCRs that contain ≤ 3 genes but derive from the overlap between large segments, as well as regions devoid of coding genes, which were not considered in the present study.

Genomic Identification of Significant Targets in Cancer (GISTIC) analysis. In order to identify relevant regions of CNA, the GISTIC method described by Beroukhi et al. (Beroukhi et al., 2007) was applied to the curated segmentation results as described. The method accounts for multiple-hypothesis testing using the false-discovery rate (FDR) framework to assign a q-value to each aberration. In this study, we used the default q-value cutoff of 0.25. For each significant region, a peak region was identified, which is the part of the aberrant region with greatest amplitude and frequency of alteration and with minimal q-value. In addition, a wide peak was determined using a leave-one-out algorithm to allow for errors in the boundaries in a single sample (the wide peak boundaries are more robust for identifying the most likely gene targets in the region).

Dominant clone analysis

For the purpose of reconstructing the history of clonal evolution during FL transformation to DLBCL, we first estimated the percentage of tumor cells in the biopsy on the basis of the inferred CN value at the clonally rearranged Ig loci (i.e., the region of intrachromosomal deletional recombination at 14q32, 2p11, and 11q22). Then, the allelic frequency of each SNV was corrected for the fraction of tumor cells in the biopsy by calculating

its expected frequency, the 95% CIs from its observed frequency, the total depth at the variant position, and the tumor content of the sample, assuming a binomial distribution.

Mutations were classified as clonal if the fraction of variant reads (upon correction for the percentage of tumor cells in the specimen) was > 20 and were classified as subclonal otherwise. Because this analysis focuses on the history of the dominant tumor clone, only SNVs that were clonally represented in at least one disease phase were considered. Then, SNVs were assigned to one of the following three categories: (1) shared mutations (i.e., mutations that are detected in the major clone of both diseases phases), (2) FL-specific mutations (i.e., mutations present in the major clone of the FL phase and either completely absent or present at subclonal levels in the paired tFL specimen, provided the difference in the corrected frequencies between the two phases was statistically significant at $p < 0.05$ [as explained below] and after excluding that its absence in the paired tFL specimen was not due to CN loss or cnLOH involving the same genomic region), and (3) tFL-specific mutations (i.e., mutations present in the major clone of the tFL phase and absent or subclonal in the FL biopsy provided the difference in the corrected frequencies between the two phases was statistically significant at $p < 0.05$ and could not be explained by CN loss or cnLOH in the latter). To assess whether the difference in frequencies between the two phases was significant, we considered a binomial distribution and calculated the probability of observing the variant in the pre-transformation (or post-transformation) phase, given the coverage depth at that position, the observed frequency of variant reads in the paired sample, and the estimated tumor content of the samples. WES and SNP6.0 analysis were used to exclude that the absence of FL-specific SNVs in the longitudinal tFL biopsy was due to recombination events, such as CN losses or cnLOH affecting the same region; the few mutations belonging to this category were conservatively classified as “shared” (in cases with paired

normal DNA) or were excluded (in cases without paired normal DNA). Mutations that were clonally represented in the tFL phase but could be detected in a small fraction of reads in the FL phase (with $p < 0.05$) were defined as DLBCL-specific, given that they most likely reflect a pre-existing smaller clone within the FL major clone, whereas mutations that were clonally represented in the FL phase but were also detected in a small fraction of the tFL population (with $p < 0.05$) were considered as FL-specific, given that they may reflect residual FL cells amidst the DLBCL clone, which were not selected during progression and expansion.

Analysis of recurrence

Genes identified by the integrated WES and SNP6.0 approach in the 39 tFL cases were ranked based on the overall frequency of affected cases (combining SNVs and CNAs, the latter only if the gene was comprised within a focal MCR of aberration encompassing a maximum of 3 genes). The same dataset was interrogated by three independent algorithms, including MutComFocal (Frattini et al., 2013; Trifonov et al., 2013a), MutSigCV (for SNV only) (Lawrence et al., 2013), and GISTIC (for CNA only) (Beroukhi et al., 2007). Genes affected in >10% of samples and scored as significant by at least one independent algorithm are included in the recurrence analysis, along with a few additional genes that, although mutated at lower frequencies, are of functional relevance within the same pathways. With four exceptions (*CARD11*, *CD79B*, *IRF4* and *EBF1*, recognized by MutComFocal), none of the genes observed at <10% frequency was scored as significant.

Algorithm for detection of Mutations with Common Focal genomic alterations (MutComFocal)

As an additional tool to rank genes found affected by genetic lesions, we used MutComFocal, a Bayesian approach that assigns a driver score to each gene by integrating point mutations and CNA data (Frattini et al., 2013; Trifonov et al., 2013a). In general, MutComFocal uses three different strategies. First, the focality component of the score is inversely proportional to the size of the genomic lesion to which a gene belongs, and thus prioritizes more focal genomic lesions. Second, the recurrence component of the MutComFocal score is inversely proportional to the total number of genes altered in a sample, which prioritizes samples with a smaller number of altered genes. Third, the mutation component of the score is inversely proportional to the total number of genes mutated in a sample, which achieves the twofold goal of prioritizing mutated genes on one hand and prioritizing samples with a smaller number of mutations on the other. The recurrence and focality scores can be interpreted as the posterior probabilities that a gene is driving the selection of the disease under two different priors, one global and one local in nature. The recurrence score is higher if a gene participates in many samples that do not have too many altered genes, whereas the focality score is higher if the gene participates in many focal lesions. Besides lending strong support to the inference of a gene as a potential driver, the directionality of the CNA (amplification or deletion) informs us of the probable behavior of the candidate gene as an oncogene or tumor suppressor, respectively. The validity of MutComFocal in identifying functional genomic alterations in the absence of functional data has been proven in previous studies (Frattini et al., 2013; Pasqualucci et al., 2011a).

MutSigCV analysis of whole exome sequencing data

As an independent approach to isolate relevant targets of genetic lesions, the full set of putative somatic SNVs (synonymous and nonsynonymous) and indels identified by SAVI in the 39 tFL cases was analyzed employing the MutSigCV algorithm, which identifies genes that are mutated more often than expected by chance (<http://www.broadinstitute.org/cancer/cga/mutsig>) (Lawrence et al., 2013). Since all mutated genes in our list of variants were fully captured and covered at sequencing depth >10, we used the MutSigCV full coverage table, which tabulates the number of nucleotides in the human reference genome by gene, category (CpG transitions, CpG transversions, C:G transitions, C:G transversions, A:T transitions, A:T transversions, or null+indel mutations), and effect (silent, non-silent, or non-coding). We also employed the default covariate table provided by MutSigCV, which contains for each gene its global expression, DNA replication time, and the HiC statistic. A false-discovery rate $q < 0.15$ was used to identify the genes mutated at a significant frequency (Lawrence et al., 2013).

Functional annotation clustering of tFL target genes

In order to determine whether genes affected by genetic lesions in tFL were enriched in specific functional categories, the genes resulting from the WES analysis and those included in focal MCRs of aberration encompassing 1-3 genes were merged in a unique gene list (n=773), and functional annotation clustering analysis was performed using the DAVID software (<http://david.abcc.ncifcrf.gov>) (Huang et al., 2009a; Huang et al., 2009b) based on Gene Ontology terms. This tool groups terms with similar annotations in clusters, and subsequently ranks them by Enrichment Score; clusters were then selected to retain only those including at least one significant ($p < 0.05$) term. The DAVID software was also used to run a Functional

Annotation chart analysis based on KEGG, BIOCARTA, PANTHER, REACTOME and BBID pathways. Moreover, the overlap between the gene list and the canonical pathways of the Molecular Signatures Database (MSigDB) was calculated by the compute overlap tool of the GSEA software (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>) (Subramanian et al., 2005).

Unselected FL cases

Whole exome sequencing data from 23 unselected FL were used to interrogate 48 recurrent genes as an indirect support for the identification of tFL-specific lesions. This study was based on patients enrolled between 2007 and 2012 into two separate lymphoma repositories at the University of Michigan Comprehensive Cancer Center (IRBMED # HUM00007985 and IRBMED # HUM00017055). The study IRBMED # HUM00007985 procured material from patients with lymphoma through fine needle aspiration of externally palpable lymph nodes, and written informed consent was obtained from all patients before enrollment. The study IRBMED # HUM00017055 procured patient material during times of clinically indicated lymphoma surgeries. Genomic research on all specimens was approved through IRBMED # HUM00005467 (Genomic analysis of B-cell Non-Hodgkin's lymphoma). Briefly, paired normal (CD3+) and tumor DNAs from purified flow-sorted immunoglobulin light chain restricted cells were enriched in coding sequences using the Illumina TruSeq Exome Enrichment Kit, which covers 20,794 genes (88.3% of the RefSeq database, version March 2011) compared to 20,935 genes covered by the Illumina SureSelect V4 platform. Libraries were prepared according to the manufacturer's recommendations and sequenced on the Illumina HiSeq 2000 platform, with paired-end 96bp reads chemistry.

Data were analyzed using a bioinformatics pipeline developed by the DNA sequencing and bioinformatics core of the University of Michigan (unpublished data). The analysis documented comparable quality control parameters relative to those obtained from the FL/tFL cases, including an average sequencing depth of 41x (range, 23x-81x) and 90% of bases in the target region covered by at least 10 reads. Moreover, comparison of the heterozygous dbSNP call rates obtained by WES using this pipeline and the set of heterozygous dbSNP calls detected by Affymetrix SNP6.0 array analysis of the same cases (n=12 normal CD3+ cell samples or, in parallel, 12 normal CD3+ samples and 12 B-cell tumor samples combined) disclosed a specificity for SNV calling of 99.4% and a sensitivity of 90%. Thus, the two pipelines provided comparable sensitivities in identifying SNVs. Consistently, the proportion of cases harboring non-silent mutations at representative “shared” FL/tFL genes was not significantly different between the 12 discovery FL cases analyzed by SAVI and the 23 unselected FL cases:

	N of mutated FL (discovery)	%	N of mutated FL (unselected)	%	<i>p value</i>
<i>CREBBP</i>	7/12*	58.3	14/23**	60.9	1.000
<i>MLL2</i>	9/12*	75.0	11/23**	47.8	0.163
<i>BCL2</i>	9/12*	75.0	11/23	47.8	0.163
<i>TNFRSF14</i>	4/12*	33.0	5/23*	21.7	0.686

* all independently validated by Sanger sequencing; **validated by Sanger sequencing in all cases tested

Sequencing analysis of ASHM target genes

Mutational analysis of representative, previously identified bona fide targets of ASHM (*PIM1*, *PAX5*, *RhoH/TTF*, *MYC*, *BCL7A*, *CIITA*, and *SOCS1*) was performed in all 39 tFL and 15 paired FL samples on selected regions previously shown to contain > 90% of the mutations found in DLBCL, as described (Mottok et al., 2009; Pasqualucci et al., 2001; Vakiani et al.,

2008) (oligonucleotide sequences and PCR conditions are also available upon request). These seven genes were selected based on the following well-established criteria for SHM targets: i) mutation distribution within 2Kb from the transcription start site; ii) absence in normal GC B cells; iii) presence of multiple mutations in the same case, at frequencies significantly higher than the mammalian background mutation rate; iv) preference for transitions over transversions and G/C vs A/T nucleotide targeting. The *BCL6* 5' sequences, which represent a physiologic target of SHM in normal GC B cells (Pasqualucci et al., 1998), and the *BCL2* ~2Kb region encompassing exons 1-2, which has been shown to accumulate mutations in *BCL2*-translocated cases under the influence of the juxtaposed IG enhancer sequences (Saito et al., 2009; Schuetz et al., 2012), were also included as controls. Purified PCR products were sequenced directly from both strands at Genewiz, Inc., and chromatograms were analyzed using the Mutation Surveyor Version 4.0.8 software package. All mutations identified were confirmed on independent PCR products obtained from high-molecular-weight genomic DNA. Nucleotide changes due to previously reported polymorphisms or occurring more than once in separate cases were considered as polymorphic variants, unless their somatic origin could be proven.

Analysis of ongoing ASHM. To examine whether ASHM has been active during clonal expansion of the tFL dominant clone, we investigated four tFL samples for the presence of intraclonal variation in their mutated *PAX5* sequences. To this end, the *PAX5* hypermutable domain was amplified from genomic DNA using the Pfu Turbo Cx Hotstart DNA polymerase (Agilent Technologies, Inc; Santa Clara, CA) according to published protocols (Pasqualucci et al., 2001). Purified pre-GC naive B cells and the fibroblast cell line IMR91 were processed in parallel as control for the Pfu error rate. PCR products were gel purified and cloned into the Zero Blunt TOPO vector (Life Technologies, Grand Island, NY) following the manufacturer's

instructions. For each PCR product, 61 to 128 clones were sequenced and analyzed. Mutation frequencies were calculated by dividing the total number of unique mutations (i.e., changes observed in individual subclones, but not in the dominant clone) by the total number of base pairs sequenced.

***De novo* DLBCL cases**

A panel of 102 at diagnosis DLBCLs (50 GCB-DLBCL and 52 ABC-DLBCL) with no previous history of FL was used to compare the mutation profile of tFL to that of *de novo* DLBCL. This panel partially overlaps with previously published series from our group, and its molecular characterization, including gene expression profile-based classification (available for all cases) and SNP6.0 array profiling (available for 79 of the 102 cases), was obtained as previously described (Compagno et al., 2009; Pasqualucci et al., 2011a; Pasqualucci et al., 2011b). Mutational analysis of selected genes was performed by PCR amplification and Sanger sequencing as reported (Compagno et al., 2009; Pasqualucci et al., 2011a; Pasqualucci et al., 2011b); for genes that had not been previously published, oligonucleotide sequences and PCR conditions are available upon request. Mutation data for the *SGK1* and *GNA13* genes are derived from Morin et al., Nature 2011.

Fluorescence in situ hybridization (FISH)

The presence of chromosomal translocations affecting *BCL2*, *MYC* and *BCL6* was assessed by FISH analysis, using commercially available break-apart probes obtained from Abbott Molecular (Des Plaines, IL). All probes were verified for the specificity of hybridization on normal metaphase preparations. Hybridizations were performed on archival methanol-acetic

acid fixed metaphase/interphase preparations and FFPE sections previously subjected to clinical testing for lymphoma work up. Methanol-acetic acid fixed cells were hybridized using standard FISH protocols, and the signal patterns were scored in at least 200 interphase cells or at least 10 metaphases counterstained with 4',6'-diamidino-2-phenylindole (DAPI). In case of paraffin embedded tissues, 4-5 μ m thick tissue sections were cut onto adhesive-coated slides. Tissue sections were baked overnight at 60°C before hybridization. The slides were subjected to protease treatment using a paraffin pretreatment kit (Abbott Molecular, Des Plaines, IL). FISH was performed using standard methods. Hybridization signals were scored on at least 100 distinct nuclei from each tissue section on DAPI-stained slides. All analyses were performed on a Nikon Eclipse 600 microscope attached to a Cytovision imaging analysis workstation. Separation of fused signal or other deviations from the normal signal pattern were considered as indicative of rearrangement or CNA within the gene.

Immunohistochemistry

Expression of the CD10, BCL6, BCL2, IRF4, MYC and p16 proteins was analyzed by immunohistochemistry on FFPE material from primary biopsies, according to standard methods. The following antibodies were used: anti-CD10 (SP67, Ventana, Tucson, AZ); anti-MYC (EP121, Epitomics, Burlingame, CA); anti-p16 (Ventana, Tucson, AZ); anti-BCL2 (sc-783), anti-BCL6 (sc-858) and anti-IRF4 (sc-6059) (all from Santa Cruz Biotechnology, Santa Cruz, CA). Samples were scored based on percentage labeling (1-100%) of the tumor cells in the expected subcellular compartment; a 20% cut-off was used to define samples as positive for expression of the p16 protein. Analysis of MYC expression was performed according to the criteria published in (Green et al., 2012; Hu et al., 2013).

Statistical analysis

The chi-square test and, when appropriate, the Fisher's exact test were used to assess whether differences in categorical variables (e.g., the frequency of mutated cases in separate disease groups) were significant. Continuous variables, not assuming normal distribution, were compared by the Mann-Whitney U test. Statistical significance was defined as p-value <.05. The analysis was performed with the Statistical Package for the Social Sciences (SPSS) software v.19.0 (Chicago, IL). To assess the degree of similarity between the genomic profile of tFL and the two described subtypes of *de novo* DLBCL, we first computed the distance between the frequency of thirty-four alterations –selected because of functional relevance in either GCB, ABC or tFL– in the three tumor types (Euclidean distance, Ward). The statistical significance of these observations was then measured by resampling 100,000 times from the frequencies of each genetic aberration and the number of samples used in each group (binomial sampling). At each sampling, we computed the distances between the three groups, and we then evaluated how many times the original observations were confirmed. The results showed that tFL appears distinct from both subtypes of *de novo* DLBCL; however, 99.8% of the times, the distance between tFL and *de novo* GCB-DLBCL was closer than that between tFL and ABC-DLBCL.

RESULTS

Divergent evolution of FL and tFL from a common mutated precursor

To investigate whether transformation of FL evolves as a linear process (i.e., through the emergence of an aggressive subclone from the initial dominant FL population) or derives from the divergent evolution of an ancestral common precursor cell (CPC) that acquired distinct mutations to become a FL or a tFL, we integrated massively parallel WES and genome-wide high-resolution SNP array analysis in a “discovery panel” of sequential FL and tFL biopsies obtained from 12 patients, including four with available matched normal DNA (Tables I and II and Figure 1). In all cases, investigation of the rearranged immunoglobulin (Ig) genes by Sanger sequencing and/or SNP array analysis confirmed the clonal relationship between the two phases, whereas the inferred copy-number value at the segment of deletional recombination within the Ig loci was used to quantify the percentage of tumor cells in the biopsy (Bergsagel and Kuehl, 2013), allowing to normalize the data for clonal representation (Table I). Fluorescence in situ hybridization (FISH) analysis was used to assess the presence of chromosomal translocations affecting *BCL2*, *MYC*, and *BCL6*.

We extrapolated the evolutionary history of transformation by defining genomic alterations that are present in the dominant clone of both pre- and post-transformation specimens (“shared lesions”) and contrasting them to those that are present exclusively in the FL or tFL biopsy (“phase-specific lesions”). This analysis allows to discriminate between a linear, sequential model, wherein the tFL dominant clone will maintain all lesions present in the FL dominant clone, along with additional tFL-acquired alterations, and a divergent evolution model,

which postulates the existence of lesions that are unique to the dominant clone of the FL or the tFL in addition to the set of shared alterations (Materials and Methods and Figure 1).

Overall, we found 52 clonally represented, shared copy-number aberrations (CNAs; average, 4.3 per sample; range, 0 to 19 per sample) and 234 shared single-nucleotide variants (SNVs), including silent and non-silent mutations (average, 38.5 per sample in the four patients with matched normal DNA; in the remaining eight pairs, shared SNVs were only considered if they affected 52 genes that have been previously validated as functional targets of somatic mutations in lymphoid malignancies, because of the exceedingly high number of variants that are predicted in the absence of matched normal DNA, most likely reflecting private SNPs not reported in public databases; see the Materials and Methods). The presence of shared genetic alterations was documented in all sample pairs analyzed, confirming the original clonal relationship between the FL and tFL sample (Figure 2, left).

In addition to shared lesions, all tFL cases harbored unique mutations and CNAs that were not present in the major FL clone at diagnosis, indicating acquisition during the transformation process or selection of a minor subclone, the size of which was below the detection threshold of the methodologies used. The number of tFL-specific lesions ($n = 709$ SNVs and 291 CNAs, including 119 losses and 172 gains) was widely heterogeneous across different patients, ranging from 24 to 161 per case (average, 83 per sample) (Figure 2, right; see also Figure 3A and Table III). Importantly, unique, clonally represented events were also detected in 10 of 12 baseline FL biopsies ($n = 327$, including 229 SNVs and 98 CNAs; Figure 2, middle, and Figure 3A). The presence of FL-specific lesions was not due to CN loss or copy-neutral loss of heterozygosity affecting the same region in the sequential tFL biopsy, as

documented by both SNP array and WES analysis. Thus, these events had been acquired independently by the dominant FL clone, consistent with divergent evolution.

Evidence of nonlinear evolution was also observed at the individual gene level. As an example, both FL and tFL of patient #23 harbored biallelic *MLL2* mutations in the dominant clone, but only one of the two events (S286fs) was shared between the pre- and post-transformation biopsy, consistent with its presence in the common ancestor clone, whereas distinct mutations were detected in the second allele of the FL (R2687*) and tFL (R280_splice) specimen, indicating that they had been acquired independently by the ancestor clone during evolution to these two diseases (Figures 3B-3D).

Overall, 10 of 12 patients analyzed (83.3%; 95% confidence interval [CI], 55% to 95%) showed a mutation pattern suggestive of divergent evolution, indicating that this is the predominant mode in the history of FL transformation (Figure 4). The remaining two patients (#8 and #12) did not harbor FL-specific events; furthermore, a significant proportion of tFL-specific lesions (13 of 63 in patient #8 and 34 of 61 in patient #12) could be detected at low frequencies (4% to 15%) in the FL specimen, suggesting that the tFL arose from a minor subclone within the dominant FL population, which subsequently acquired additional mutations in a linear evolution pattern (Figure 4 and Table III).

With one exception (patient #17), the number of events acquired by the tFL dominant clone (including CNAs and SNVs) was significantly higher than that acquired by the FL dominant clone, ranging from >50-fold in case #6 to 2-fold in case #11 ($p < 0.005$) and underscoring the genomic complexity of the tFL genome (Figures 2 and 5A). Notably, at least one of these ten patients did not receive any treatment between the original FL diagnosis and transformation, indicating that the higher mutation load of tFL does not simply reflect the

consequence of the mutagenic effect or the selective pressure of chemotherapy. In at least two patients, several chromosomes displayed convoluted intrachromosomal rearrangements due to alternating gains and losses of genomic material, frequently accounting for over ten switches per chromosome (Figure 5B and Table IV). Although the sequencing approach adopted in our study (WES, as opposed to whole-genome sequencing) prevents from distinguishing true chromothripsis from localized lesions that occurred progressively (Korbel and Campbell, 2013), these data highlight a remarkable genomic instability in tFL cases in comparison to both FL and other lymphoid malignancies (Fabbri et al., 2011; Mullighan et al., 2007; Rossi et al., 2012).

Collectively, these findings support a divergent evolution model in a significant proportion of patients undergoing transformation, whereby FL and tFL arise from a common mutated ancestor through the independent acquisition of distinct genetic lesions.

Recurrent genetic lesions

In order to identify lesions potentially relevant for transformation among the large number of candidates that emerged from the analysis of the discovery panel (710 unique genes, including those targeted by non-silent SNVs, small indels, and/or CNAs if they were within minimal common regions involving a maximum of three loci), we extended the WES and SNP array analysis to 27 additional tFL cases (screening panel; combined, 39 tFL cases). Then, these data were interrogated for the presence of recurrent alterations, a common readout for functionally relevant genes, and by three independent analytical methods for the identification of key targets of functional genomic alterations: (1) MutComFocal, a recently developed computational algorithm that isolates candidate cancer genes from high throughput CN and SNV data (Trifonov et al., 2013a), (2) MutSigCV, a tool that analyzes SNV data in order to identify

genes mutated more often than expected by chance (Lawrence et al., 2013), and (3) GISTIC, a computational approach identifying significant targets of somatic CNAs.

Figure 6 illustrates the overall proportion of tFL cases harboring genetic lesions in genes altered at $\geq 10\%$ frequency and recognized as functionally relevant targets by at least one of the three independent approaches, with few additional genes of functional annotation within the same pathway (see Figures 7-9 and Table V for the full list of genes found mutated in $\geq 10\%$ of cases; see also Table VI). Genes are grouped into biological categories and annotated in order to indicate whether alterations were also found in the diagnostic FL biopsy (available in 24 of the 39 cases) or predominantly acquired or selected at transformation. Collectively, these aberrations point to a number of biological programs and signaling pathways that are either dysregulated early in the putative FL and tFL precursor cell (i.e., “shared” lesions) or selected during transformation (“tFL-acquired” lesions).

Genetic lesions shared between FL and tFL

The most commonly affected genes in both FL and tFL were those encoding for histone modification enzymes, including methyltransferases and acetyltransferases (36 of 39 cases [92.3%]). In line with previous findings in unselected FL (Morin et al., 2010; Morin et al., 2011; Pasqualucci et al., 2011a; Pasqualucci et al., 2011b), the H3K4 trimethyltransferase *MLL2* was mutated in 26 of 39 tFL cases (66.7%) with 36 truncating events and nine missense mutations (Figures 6-7 and 9 and Table VII). These lesions were already present at FL diagnosis in all but one patient and were never lost at transformation, consistent with an early acquisition by the CPC. The activity of the *MLL2*-containing complex was also impaired by mutually exclusive alterations of *KDM6B*, encoding for an H3K4 histone demethylase interacting with *MLL2* (n = 3

of 39 cases, including two SNVs and one homozygous deletion) and *MLL3* (n = 3 of 39 cases; Figures 7 and 9 and Table VII). Additionally, one-fourth of tFL cases (n = 10 of 39 [25.6%]) harbored *EZH2* gain-of-function mutations that almost invariably replace the hotspot tyrosine residue Y641 within the protein SET domain (n = 9 of 10 [90.0%], Table VII). *EZH2* mutations have been reported in 7% unselected FL cases and 22% *de novo* GCB-DLBCL (Morin et al., 2010; Morin et al., 2011), where they increase H3K27 levels through altered substrate specificity.

Another class of chromatin modifiers was represented by the acetyltransferases *CREBBP* (n = 21 of 39 patients [53.8%]; 19 point mutations and two focal deletions) and *EP300* (6 of 39 cases [15.4%], of which five mutated and one deleted; Figures 6-7 and 9). In both genes, the mutation pattern was highly reminiscent of what has been reported in unselected FL and *de novo* DLBCL with respect to the inactivating nature of the lesions, the evidence of mutational hotspots (R1446 in five patients, F1484 in two patients, Y1503 in two patients, and Δ S1680 in three patients, Table VII), and the predominantly monoallelic distribution (19 of 21 affected cases), indicating a haploinsufficient tumor suppressor role (Pasqualucci et al., 2011a).

Programmed cell death was the second largest program dysregulated in both FL and tFL and, thus, presumably in the common ancestor clone. In addition to *BCL2* translocations, detected in 27 of 33 tFL (81.8%) and invariably shared between the two phases (n = 18 informative pairs), all t(14;18)-positive cases harbored multiple somatic point mutations within the ~2 kb region downstream of the *BCL2* transcription initiation sites (Figures 6 and 7 and Table VII), reflecting the activity of the AID-dependent SHM process driven by the juxtaposed Ig enhancer (Lohr et al., 2012; Saito et al., 2009).

The *FAS* gene was disrupted in 13 of 39 tFL cases (33.3%) because of inactivating mutations (n = 4 of 39 [10.2%]) and genomic deletions (n = 9 of 39 [23.1%], including two focal homozygous events; Figures 6 and 7). In the three affected patients with available pre- and post-transformation biopsy, these lesions were always detectable at FL diagnosis, suggesting their presence in the putative CPC (Table VII). Interestingly, none of the 23 unselected FL exomes harbored *FAS* mutations, giving rise to the possibility that these lesions represent an early marker for transformation. *FAS* encodes for a receptor protein that acts as a major mediator of apoptosis in GCB cells carrying low-affinity and self-reactive antigen receptors (van Eijk et al., 2001). With the exception of one amino acid substitution removing the initiating methionine (M1T), all *FAS* mutations (Y232*, P217_splice, and D317V) cluster in exons 7 to 9, which encode for the protein intracytoplasmic tail (Table VII). This domain is required for the assembly of the death-inducing signaling complex, and its truncation will result in the functional loss of normal *FAS* signaling by a dominant-negative effect (Siegel et al., 2000), as documented in patients with autoimmune lymphoproliferative syndrome (Holzelova et al., 2004). A deleterious effect was also predicted for the D317V amino acid change on the basis of the PolyPhen 2 algorithm (Table VII). *FAS* was identified as a relevant target of genomic deletions by two independent algorithms, including GISTIC (Figure 8 and Table VI) and MutComFocal.

Altogether, these findings identify the disruption of pathways affecting chromatin modifier functions and resistance to apoptosis as recurrent lesions common to FL and tFL, and, thus, presumably occurring early during the initial clonal expansion of the putative precursor clone.

Genetic lesions specifically associated with tFL

The most common genomic aberration specifically acquired during progression to tFL was the loss of *CDKN2A/B*, two tumor suppressor genes whose protein products (p14-ARF, p16-INK4A, and p15-INK4B) play major roles as negative regulators of cell-cycle G1 progression and as stabilizers of the tumor suppressor p53 (Sherr, 2004). Overall, 46.1% of tFL cases (n = 18 of 39) carry genomic aberrations affecting these loci, including 17 CN losses (n = 6 heterozygous and 11 homozygous; Figure 10A) and a nonsense C72* mutation combined with cnLOH (Figure 7 and Table VII). In most deleted cases (n = 10 of 17), the loss of genetic material encompassed ≤ 3 genes, identifying a minimal common region smaller than 10 kb and exquisitely restricted to the *CDKN2A/B* locus. Biallelic *CDKN2A/B* alterations were never present at FL diagnosis, indicating a specific role during transformation (Figure 10B).

The loss of *CDKN2A/B* may impinge on different biological programs, including DNA damage responses (via the p14-ARF/p53 pathway) and cell-cycle regulation (via the RB/p16 tumor-suppressive pathway). As expected, immunohistochemical staining for p16 expression confirmed its complete loss in the neoplastic lymphocytes of all biallelically deleted tFL cases (n = 8 of 8); however, two of three monoallelically deleted cases and a significant proportion of wild-type (WT) tFL cases (n = 5 of 15 [33.3%]) were also p16-negative (data not shown), suggesting the involvement of epigenetic mechanisms of inactivation. Although the limited number of cases prevents statistical analysis, *CDKN2A/B* biallelic lesions tend to be mutually exclusive with biallelic deletions and/or mutations of *TP53*, observed in 7 of 39 tFL cases (17.9%) but absent at FL diagnosis (n = 3 informative cases; Figure 10C). These observations suggest that *CDKN2A/B* loss may contribute to FL transformation by affecting both cell-cycle regulation and p53-dependent DNA damage responses, thus promoting genomic instability.

Consistent with this hypothesis, patients exhibiting dysregulation of the ARF/p53 axis via biallelic alterations of *CDKN2A/B* and/or *TP53* were characterized by a significantly higher number of CNAs in comparison to patients that harbor WT alleles (average $n = 45.0$ versus 20.5 , Mann Whitney U test, $p = 0.03$; Figure 10D).

Genetic lesions deregulating *MYC*, namely chromosomal translocations ($n = 6$ of 24 tFL cases with available FISH data [25.0%]), copy-number gains and/or amplifications ($n = 13$ of 39 [33.3%]), and point mutations reflecting the activity of ASHM (Figures 7 and 11 and Table VII) were the second most common tFL-specific lesions. Although low copy-number gains could be occasionally observed in the original FL biopsy, high CN amplifications, translocations, and point mutations were either completely absent ($n = 13$ cases) or only detected in a minor subclone within the dominant FL population (patient #8). Deregulated *MYC* oncogenic activity may provide multiple advantages to the cancer cell through its pleiotropic function in cell growth, metabolism, and genetic instability.

Also enriched at transformation were biallelic mutations and/or deletions encompassing *B2M* ($n = 5$ of 39) and *CD58* ($n = 2$ of 39), two genes involved in the control of immune recognition by cytotoxic T lymphocytes and natural killer cells, respectively, and previously shown to be recurrently inactivated in *de novo* DLBCL (Challa-Malladi et al., 2011) (Figures 6 and 7 and Tables VI and VII). *B2M* genomic aberrations were specifically acquired or selected at transformation ($n = 3$ informative cases) and, accordingly, *B2M* as well as *CD58* were not mutated in 23 unselected FL exomes analyzed, implicating escape from immune surveillance as a contributor to the transformation process.

Finally, multiple point mutations, small deletions, and duplications were identified in the $5'$ sequences of several recognized ASHM target genes, including *PIM1*, *PAX5*, *RhoH/TTF*,

MYC, *BCL7A*, *CIITA*, and *SOCS1* (overall, 34 of 39 cases [87.1%]; Figure 12). These lesions display typical features of AID-mediated activity (Figure 13) and, depending on the genomic configuration of the involved locus, were variably distributed in coding and/or noncoding sequences. These changes were not detected in the pre-transformation biopsy, indicating that they had been specifically acquired or selected at transformation (Figure 12). These data point to a malfunction of SHM occurring late in the disease, although the prolonged exposure of the precursor cell to the potentially deleterious environment of the GC reaction may also favor the accumulation of lesions.

Non-phase specific lesions

A number of genetic alterations were variably observed as shared or phase-specific events, suggesting heterogeneous contributions to disease pathogenesis. Consistent with previous reports (Lossos and Gascoyne, 2011), *TNFRSF14*, encoding for a member of the TNF receptor superfamily that signals to T cells with stimulatory or inhibitory effects depending on the ligand, was disrupted in 22 of 39 tFL (56.4%) and 9 of 17 FL cases (52.9%) due to a combination of truncating mutations (n = 14, distributed in 11 tFL and 3 FL cases), genomic deletions (n = 15 tFL and 6 FL cases), and cnLOH (n = 4 tFL cases and 1 FL case; Figure 7 and Table VII). Importantly, all mutated tFL cases had lost the residual WT allele because of deletion or cnLOH. Although multiple other candidate genes are encompassed by the large heterozygous deletions affecting chromosomal region 1p36.31, this typical pattern of biallelic inactivation documented the specific involvement of *TNFRSF14* in these lesions (Figure 7).

STAT6, a DNA binding transcription factor implicated in IL4 and IL13-mediated responses, was the target of heterozygous somatic point mutations in 9 of 39 (23.1%) tFL cases.

Analysis of the diagnostic FL biopsy revealed the presence of the same mutation in three of five available cases, whereas the remaining two had acquired this lesion at transformation; *STAT6* mutations were observed in unselected FL cases (1 of 23 [4.3%]), consistent with a role in both disease phases. Notably, all *STAT6* mutations cluster within the protein DNA binding domain, identifying a mutational hotspot of possible functional relevance at residue D419, which was substituted to G (four cases), H (one case), and Y (one case; Figure 7 and Table VII), as previously reported in primary mediastinal B cell lymphoma (Ritz et al., 2009).

Several genes encoding for core histones were also frequently mutated in both tFL (n = 11 of 39 [28.2%]) and FL (n = 6 of 12 [50%]) cases, *HIST1H1E* being the most common target (n = 7 of 39 tFL [17.9%], 2 of 12 FL, [16.6%], and 2 of 23 unselected FL cases [8.7%]; Tables V and VII). In addition, > 10% of tFL cases harbored loss-of-function (nonsense and frameshift) mutations disrupting the *ARID1A* (n = 7 of 39 [17.9%]) and *ARID1B* (n = 1 of 39 [2.6%]) genes, two components of the SWI-SNF chromatin remodeling complex, which was recently shown to take part in maintaining the pluripotency of stem cells as well as in reprogramming somatic cells (Ho et al., 2009).

Finally, a number of recurrent lesions are expected to interfere with signaling pathways that are triggered in response to engagement of the BCR or CXCR4 receptors; these include oncogenic mutations of *CARD11* and *CD79B* (4 of 39 tFL [10.3%] and 3 of 39 tFL cases [7.7%], respectively), alterations in negative and positive regulators of NF- κ B (*TNFAIP3*, biallelically lost in 6 of 39 tFL cases, and *TRAF2*- and Nck-interacting kinase [*TNIK*], mutated in 4 of 39 cases), truncating mutations of *GNA13* (n = 8 of 39 tFL [20.5%] versus 1 of 23 unselected FL cases [4.3%]), and point mutations of *FOXO1* (n = 6 of 39 tFL [15.4%] versus 0 of 23 unselected FL cases; Tables V and VII); deregulation of the latter two genes may impinge on the PI3K

pathway as well as RhoGTPase responses, and has been observed in *de novo* DLBCL (Morin et al., 2013; Trinh et al., 2013). Altogether, these results point to several signaling pathways that are recurrently dysregulated in FL and/or tFL, although their individual contribution is not specifically restricted to a discrete phase of disease pathogenesis (see the Discussion and Conclusion).

The genomic landscape of tFL is unique but more related to GCB-DLBCL

In order to determine whether tFL and *de novo* DLBCL represent pathogenically different diseases, we compared the genomic profile of the 39 tFL cases to that of 102 *de novo* DLBCL cases representative of the two major molecular subtypes; i.e., GCB and activated B cell (ABC) DLBCL (Alizadeh et al., 2000). When analyzed by unsupervised hierarchical clustering with the frequency of aberrations at 34 informative targets, tFL appears to be closer to GCB- than ABC-DLBCL (see the Materials and Methods). Common features of the two diseases include the presence of *BCL2* rearrangements, *REL* amplifications, *EZH2*, *GNA13*, and *TNFRSF14* mutations (Morin et al., 2010; Morin et al., 2011; Pasqualucci et al., 2011b) along with the absence of typical ABC-DLBCL-specific aberrations (*MYD88* mutations and *PRDM1* inactivation) (Mandelbaum et al., 2010; Morin et al., 2013; Ngo et al., 2011; Pasqualucci et al., 2006) (Figure 14). However, unique combinations of genetic lesions were found in tFL, which are otherwise never observed in GCB-DLBCL (e.g., biallelic deletions of *CDKN2A/B*; 28.2% of tFL cases versus 0% of GCB-DLBCL, $p < 0.01$). Moreover, tFL tends to be enriched in alterations that are generally less frequent in *de novo* DLBCL, such as *STAT6* mutations (23.1% versus 2.4%), *ARID1A* mutations (17.9% versus 7.2%), and *FAS* mutations and/or deletions (33.3% versus 16.7%). Particularly, although observed in both tFL and *de novo* DLBCL,

aberrations of *MLL2*, *CREBBP*, and *BCL2* were significantly enriched in cases derived from FL transformation in comparison to GCB-DLBCL ($p = 0.02$, 0.0006 , and 0.0001), suggesting that at least a subset of GCB-DLBCL arises from the precursor cell postulated by the results described above. In conclusion, the genome of tFL appears more similar to GCB-DLBCL but, overall, is unique in comparison to both subtypes of *de novo* DLBCL.

DISCUSSION AND CONCLUSIONS

This study reports a comprehensive characterization of the coding genome of tFL, an aggressive disease with dismal prognosis, the pathogenesis of which has been incompletely understood so far. The goal of the study was to take advantage of systematic, genome-wide approaches in order to address key questions related to this aggressive condition that have remained unanswered or had been previously only studied by examining individual genes. These include: *i)* tracing the evolutionary history of the dominant clone during transformation from indolent FL, *ii)* providing an assessment of the range and frequency of genetic aberrations that are associated with this event, *iii)* identifying genomic changes as potential genetic drivers of transformation, and *iv)* elucidating the relationship between DLBCL deriving from FL transformation and DLBCL arising *de novo*.

The first finding of the present study is that, although all FL-tFL sample pairs have a clear clonal relationship, the dominant tFL clone arises in most patients from a mutated CPC through the acquisition of independent genetic events, consistent with divergent evolution. The existence of this CPC cannot be physically demonstrated, but it can be postulated on the basis of the presence of a set of lesions that are shared between FL and tFL, and is consistent with previous studies based on the analysis of the rearranged Ig genes (Carlotti et al., 2009) as well as with recent work on FL progression (Green et al., 2013). Indeed, the pattern of branching clonal evolution from FL to tFL and the early involvement of genes encoding for epigenetic modifiers were confirmed in a recently published, independent study (Okosun et al., 2014).

This study does not address the topography of the tFL clone and/or the intraclonal architecture of these tumors, which require the backtracking of tFL-specific lesions in the

diagnostic FL sample by ultrahigh deep sequencing analysis (Ding et al., 2012; van Eijk et al., 2001; Walter et al., 2012; Welch et al., 2012). Thus, additional studies will be needed in order to clarify whether the tFL clone temporally developed after FL diagnosis or whether it can be already detected as a minor subclone within the FL diagnostic sample.

Notably, the two most prominent programs deregulated in the precursor clone –i.e., epigenetic modifications and resistance to apoptosis– represent actionable targets, and a number of drugs are already being tested in the clinic (e.g., BCL2 inhibitors, histone deacetylase inhibitors, and EZH2 inhibitors) (Sawas et al., 2011). If these lesions are essential to the survival of the fully transformed tumor cells, and if the tFL clone is not already present at FL diagnosis, then the development of combination regimens incorporating drugs that specifically target this group of alterations in early stages of FL may lead to the elimination of the precursor clone, possibly preventing transformation.

This study reveals that, although the genome of tFL is significantly more complex in comparison to FL, no unifying genetic lesion is selected during transformation to DLBCL. Nonetheless, the recurrent alteration of genes involved in the control of cell-cycle progression (*CDKN2A/B* and *MYC*) and DNA damage responses (alternative biallelic loss of *TP53* and *CDKN2A*) suggest that loss of genetic stability and deregulated proliferation are critical steps in tFL development. Lesions affecting *CDKN2A/B* have been observed in several terminally aggressive lymphoid malignancies, including relapsed acute lymphoblastic leukemia (Mullighan et al., 2008), Richter syndrome (Fabbri et al., 2013), and ABC-DLBCL (Lenz et al., 2008), consistent with a central role in the acquisition of a more aggressive phenotype. The existence of tFL cases that lack p16 expression despite being devoid of *CDKN2A* genetic lesions suggests an even broader involvement of this pathway through alternative epigenetic mechanisms (e.g.,

promoter hypermethylation or posttranscriptional modifications). Furthermore, the inactivation of epigenetic modifiers such as *CREBBP/EP300* and *MLL2* may also contribute to the deregulation of cell growth and DNA damage responses by interfering with p53 acetylation and activation (Pasqualucci et al., 2011a).

The genomic complexity of tFL appears to be remarkably high with respect to other hematologic malignancies, as exemplified by the presence of numerous CNAs and the evidence of ASHM. The latter may represent a major mechanism for transformation, as previously suggested for *de novo* DLBCL (Pasqualucci et al., 2001). Interestingly, although the physiologic SHM process is known to operate in FL (Bahler and Levy, 1992), ASHM was only observed in the dominant clone of tFL, pointing to a disruption occurring late during evolution to or selection of a more aggressive disease. Although the analysis presented here did not uncover any apparent lesions in genes that are directly involved in SHM, the alteration of histone marks owing to genetic lesions in histone modification and chromatin remodeling enzymes may induce chromatin conformation changes that favor the accessibility of non-physiologic genomic target regions to the AID mutator. Thus, additional studies interrogating the entire genome as well as the epigenome of this cancer will be necessary in order to conclusively address this question.

Finally, and consistent with observations obtained by gene expression profiling, our data highlight tFL as a distinct disease that, although more similar to GCB-DLBCL, harbors unique combinations of oncogenic and tumor suppressor lesions in comparison to *de novo* DLBCL. Previous studies have suggested that transformation proceeds through two distinct pathways: one characterized by a high-proliferation signature, and a second where T cell and follicular dendritic-associated genes predominate (Davies et al., 2007; Lossos et al., 2002). We did not observe statistically significant mutual exclusion between genetic lesions affecting these two

classes of mutated genes (e.g., *MYC* and *CDKN2A/B* versus *TNFRSF14* and *STAT6*). However, integrated genomic and transcriptional profiling of larger cohorts of patients will be needed in order to address this question. The unique tFL genomic landscape, combining alterations that are specific to GCB- and ABC-DLBCL as well as lesions that are uncommon in *de novo* DLBCL, could be responsible, at least in part, for its poor response to standard anti-DLBCL regimens. The results herein suggest the potential usefulness of combining current immunochemotherapeutic regimens that harness the proliferative and genetically unstable phenotype of tFL with more specific approaches targeting some of the pathways recurrently altered in tFL, as it could be the case for CDK4/6 activation deriving from the loss of *CDKN2A* or NF- κ B activation deriving from *TNFAIP3* deletions and/or mutations.

FIGURES

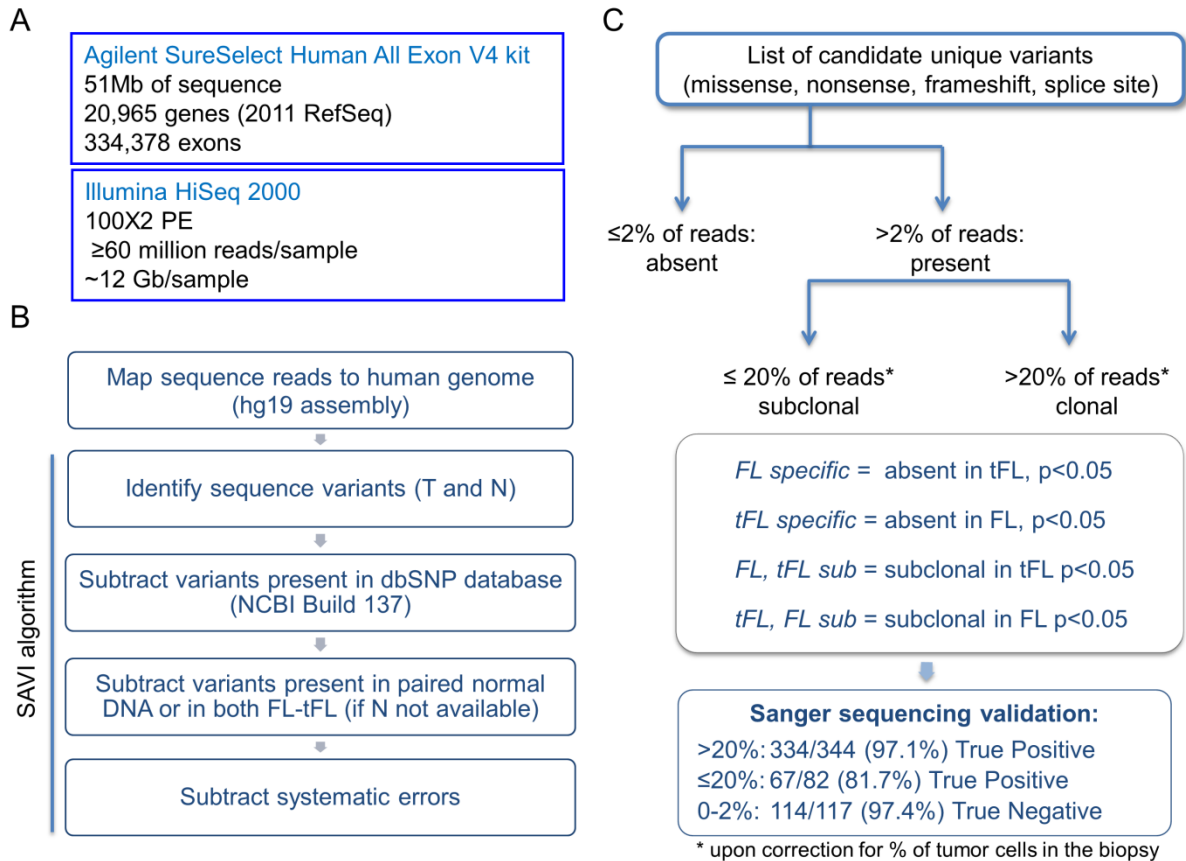


Figure 1. Whole exome sequencing and mutation discovery approach. (A) Schematics of the sequence capture protocol used for enrichment of non-repetitive coding regions and massively parallel sequencing of FL, tFL and matched normal DNA samples. The technical features of the Agilent 51Mb capture array and the expected performance of the Illumina HiSeq2000 instrument are indicated. (B) Computational pipeline adopted for the identification of tumor-specific, non-silent sequence variants after mapping to the human genome assembly hg19 (NCBI Build 37). (C) Final validation procedure and class assignment of candidate non-silent SNVs.

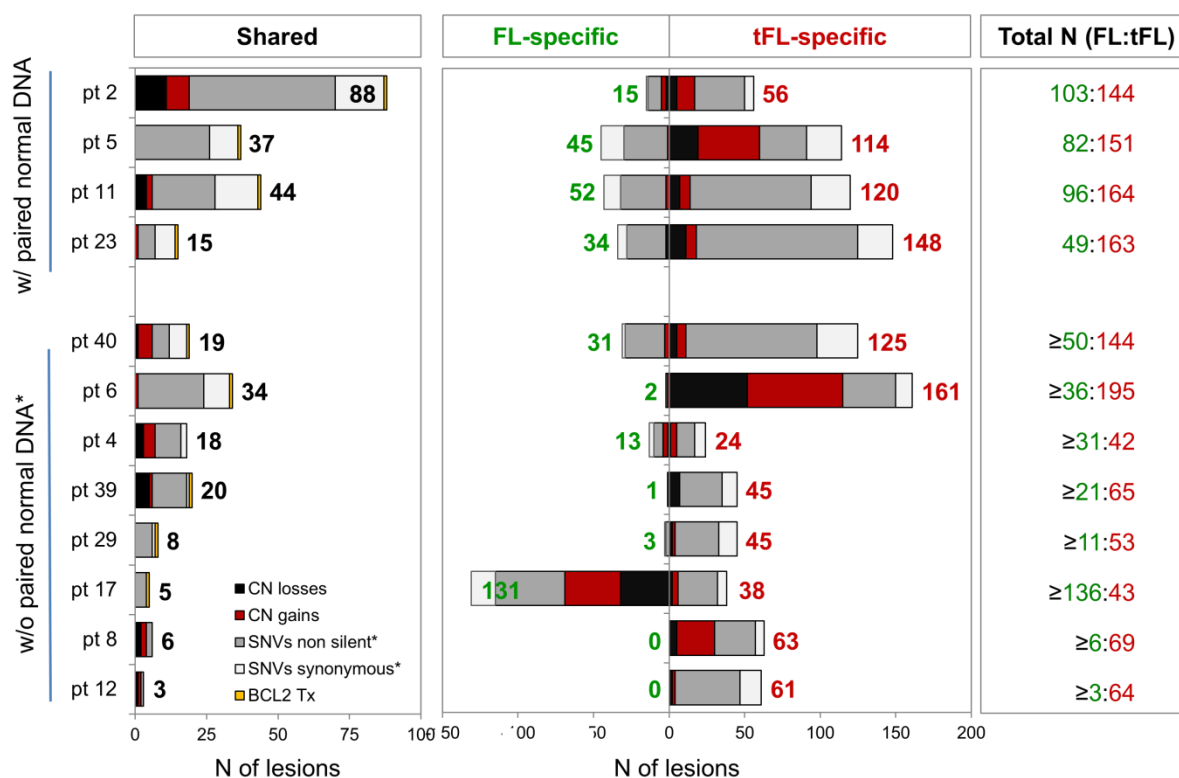


Figure 2. FL and tFL display shared and unique genomic aberrations. Overall load of genetic lesions identified by WES and CN analysis in the dominant clone of the 12 discovery cases. Color codes denote distinct types of aberrations (Tx, translocation). *In cases lacking matched normal DNA, shared SNVs are limited to those affecting 52 selected genes with well-established roles in lymphomagenesis (see Materials and Methods); thus, the total number of genetic lesions in these patients (right column) most likely represents an underestimate. FL-specific SNVs that could be due to genomic loss or cnLOH of the same region in the tFL phase were excluded.

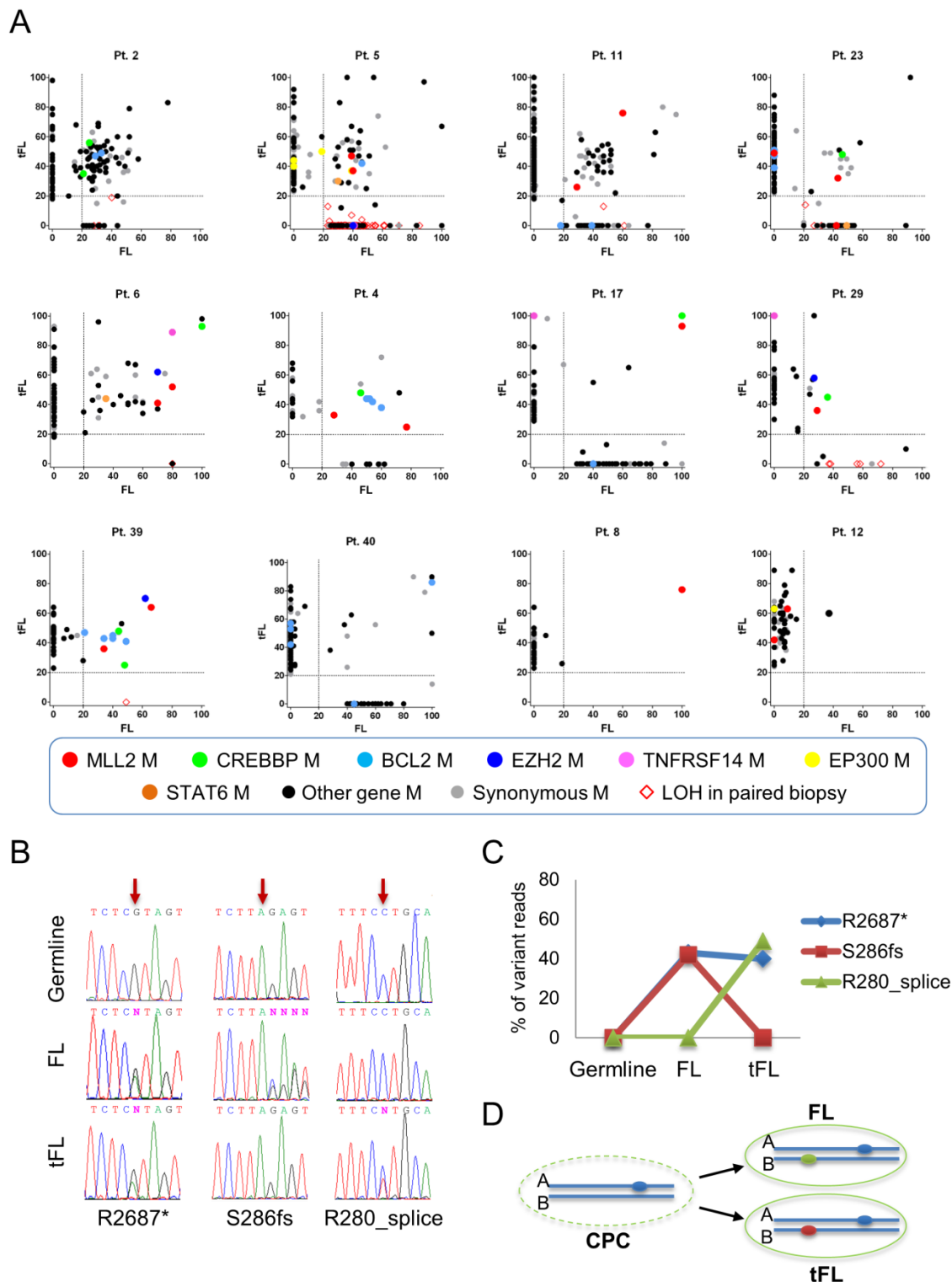


Figure 3. Distribution of shared and phase-specific SNVs in the discovery panel.

(A) Normalized allele frequencies of somatic SNVs (synonymous, missense, nonsense,

frameshift, splice-site) identified in the 12 paired FL/tFL discovery cases. Frequencies are corrected for the percentage of tumor cells in the biopsy, as estimated by SNP6.0 array analysis of the same cases (see also Materials and Methods). Note that the dispersion plots show all somatic SNVs identified by the SAVI algorithm, including i) clonal and subclonal events, ii) “phase-specific” SNVs that were located in regions targeted by genomic loss or cnLOH in the paired specimen (diamonds), and thus excluded from the dominant clonal analysis; and iii) “phase-specific” SNVs that were deemed to be shared because the difference in allele frequencies between the two phases was not significant. Thus, the total number of events per case is different from those shown in Figure 2, where only clonal events that could be unequivocally assigned to one of the three categories (shared, FL-specific, tFL-specific) are illustrated. (B) DNA sequencing traces obtained by direct sequencing of PCR products amplified from patient #23, carrying three mutations in the *MLL2* gene: a point mutation (R2687*) present in both FL and tFL, a FL-specific S286fs mutation, and a tFL-specific R280_splice mutation, consistent with divergent evolution. (C) Percentage of reads carrying the three *MLL2* variants in FL, tFL and matched constitutional DNA. (D) Model for the history of clonal evolution in patient #23, as reconstructed based on the *MLL2* mutation pattern. In the figure, the dotted circle indicates the postulated common precursor clone (CPC); A and B denote the two *MLL2* alleles, and individual SNVs are indicated by filled dots, color-coded as in (C).

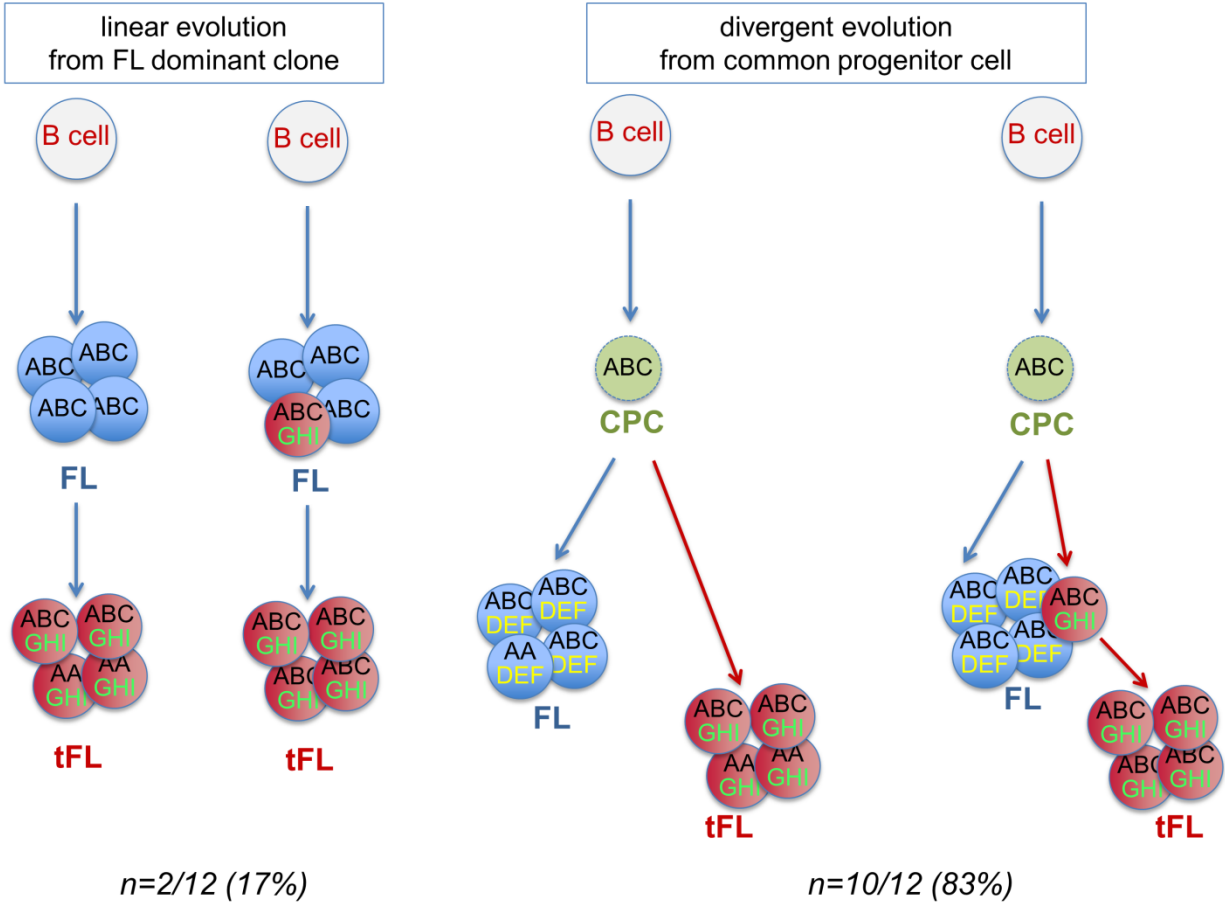


Figure 4. FL and tFL arise through divergent evolution in most patients. Inferred models of clonal evolution during FL transformation. The original B cell clone is on top; blue and red circles depict the FL and tFL clones, respectively, whereas green, dotted circles represent the postulated common mutated precursor cell (CPC). In the linear model (2 of 12 patients analyzed), the tFL dominant clone originates directly from the FL dominant clone after the acquisition of additional mutations. In the divergent model (10 of 12 patients analyzed), the FL and tFL dominant clones derive from a CPC through the independent acquisition of distinct mutations. In both scenarios, the tFL dominant clone may be present as a minor subclone already at the time of the FL biopsy. ABC, mutations shared between FL and tFL; DEF, mutations unique to the FL dominant clone; GHI, mutations unique to the tFL dominant clone.

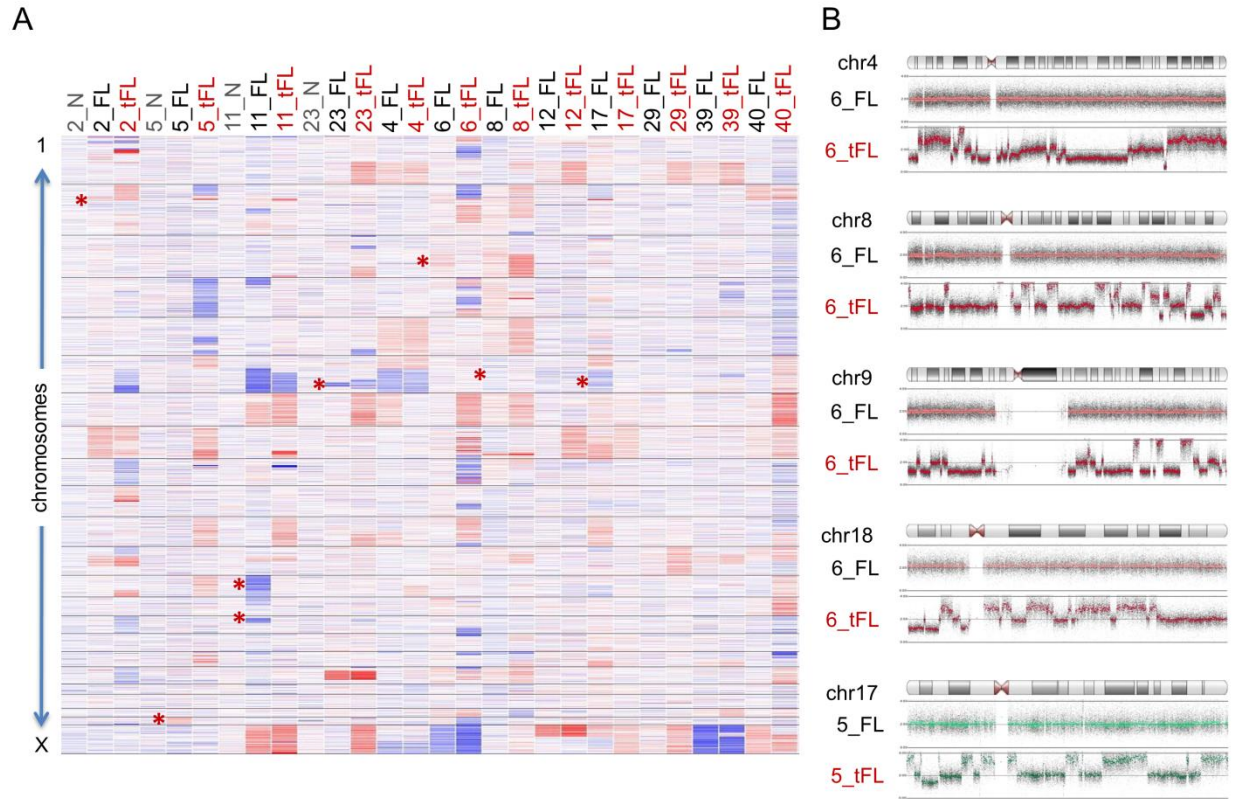


Figure 5. tFL displays a complex genomic profile relative to FL. (A) dChipSNP heatmap showing median-smoothed log2 CN ratio in the 12 discovery FL/tFL pairs. In the red-blue scale, white corresponds to a normal (diploid) CN log-ratio, blue is deletion and red is gain. Asterisks indicate representative FL-specific CNAs. (B) Total copy number plots of representative chromosomes characterized by high numbers of genomic switches in two tFL samples, as compared to their paired FL biopsy. The ideogram of the corresponding chromosome is aligned above the plots.

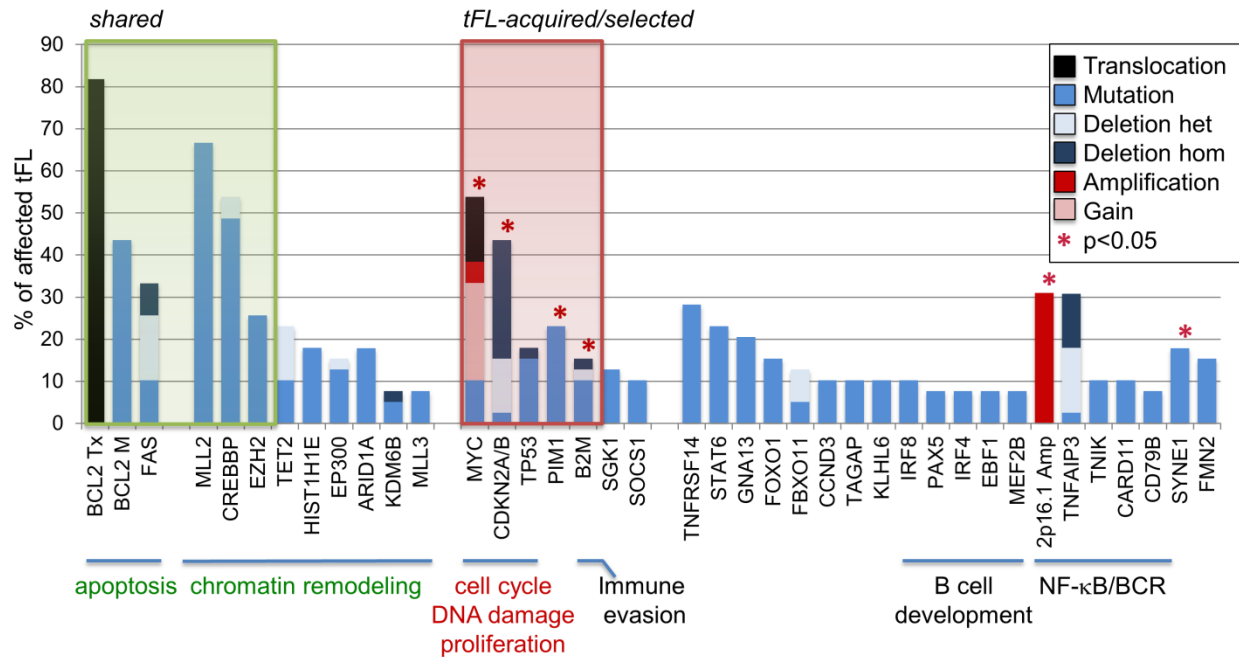


Figure 6. Recurrently mutated genes in tFL. Proportion of tFL cases (n = 39) carrying genetic lesions (SNVs and CNAs) in genes altered at $\geq 10\%$ frequency and deemed significant by one of three independent algorithms (see the Supplemental Information and Table S5); additional genes of functional relevance in the same pathways are also shown. Deletions and gains were only computed if defining MCRs of aberration encompassing a maximum of three genes. Green shade highlights alterations consistently shared between the FL and tFL sample, and red shade identifies alterations enriched in tFL (* $p < 0.05$ when comparing tFL frequency versus unselected FL frequency; note that TP53 was considered as a tFL-specific target despite its borderline p value because it was invariably unmutated in the diagnostic FL sample of all three informative tFL cases and because of the broad literature data. The remaining genes did not appear to be phase-specific or could not be unequivocally assigned to a given category because of the relatively small numbers.

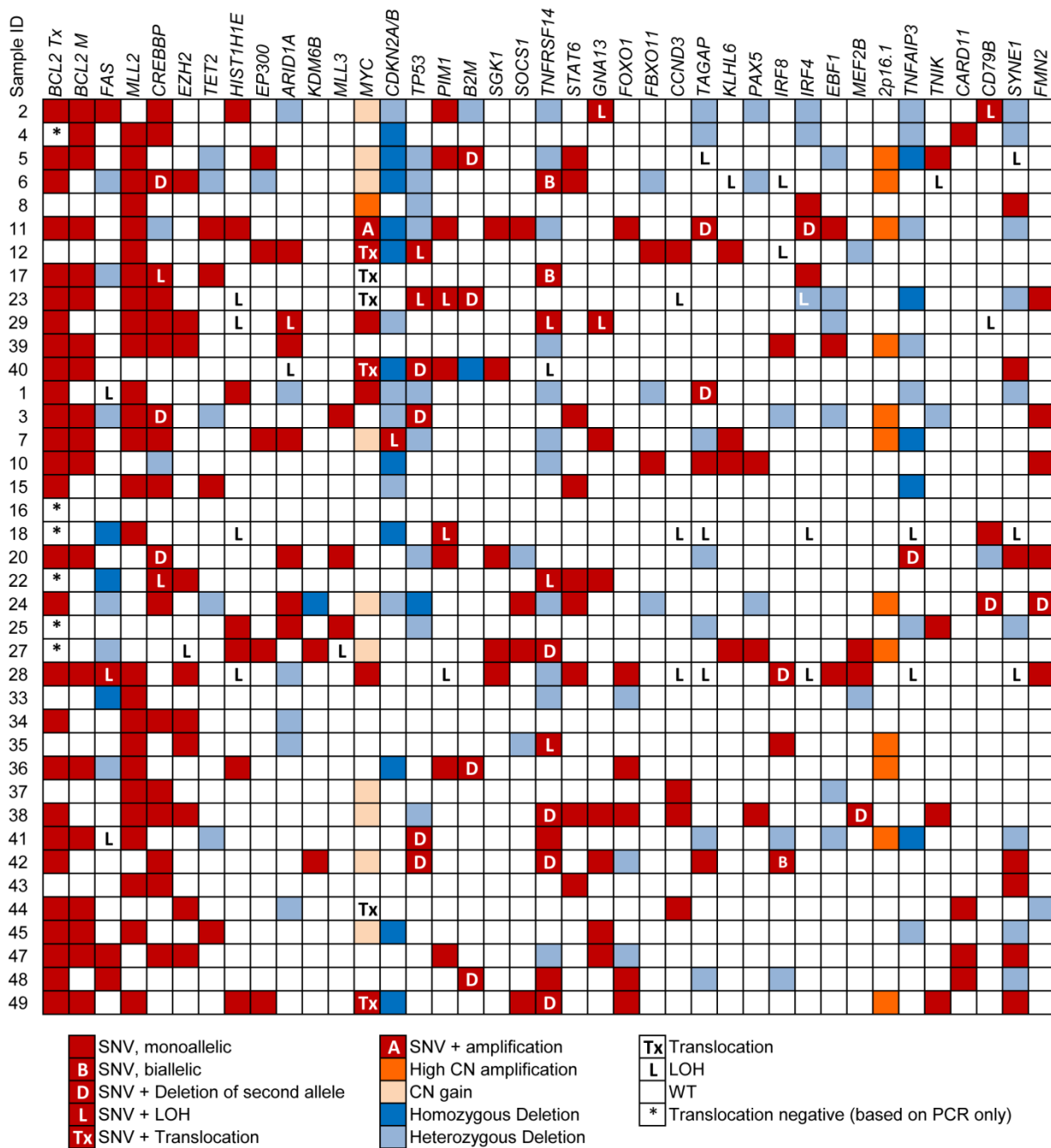


Figure 7. Distribution of recurrent genetic lesions in tFL. The heatmap shows the genomic status of recurrently affected genes in the 39 tFL cases analyzed (non-coding SNVs not included). BCL2 mutations in patients #3, 7 and 10 were detected by Sanger sequencing but not by WES, possibly due to the high mutation load of these alleles, which are targeted by SHM.

Note that, at variance with Figure 6, heterozygous deletions are shown even if very large and not defining a minimal common region of aberration.

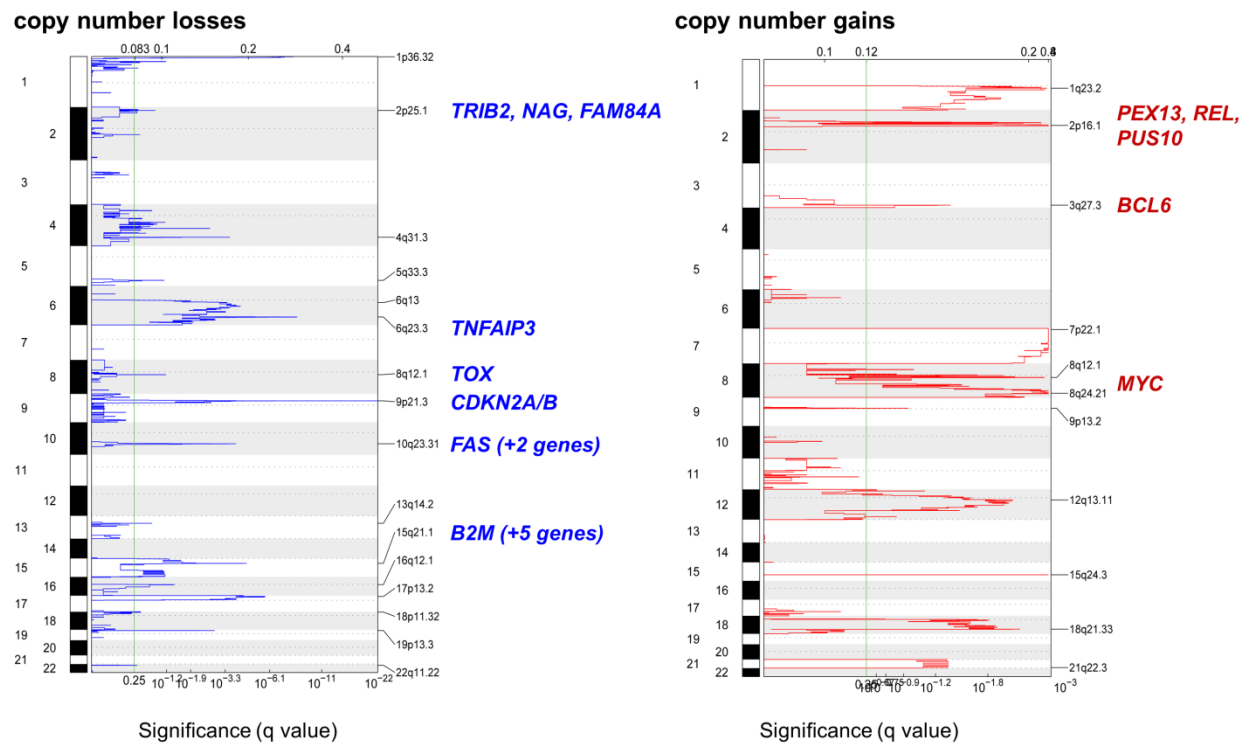


Figure 8. Significant regions of copy number alterations in the tFL genome.

Statistical analysis of copy number gains (red) and copy number losses (blue) in 39 tFL cases according to the GISTIC analysis tool. False discovery rates for each aberration (q values, X axis) are plotted at each genomic position across the 22 chromosomes (Y axis). The green line represents the q value cutoff for significance (0.25). Centromere positions in each chromosome are indicated by dotted lines, and the location of the significant peak regions (16 deletions and 11 gains) is indicated to the right of each panel; gene names are provided only for focal regions encompassing ≤ 3 genes, unless indicated.

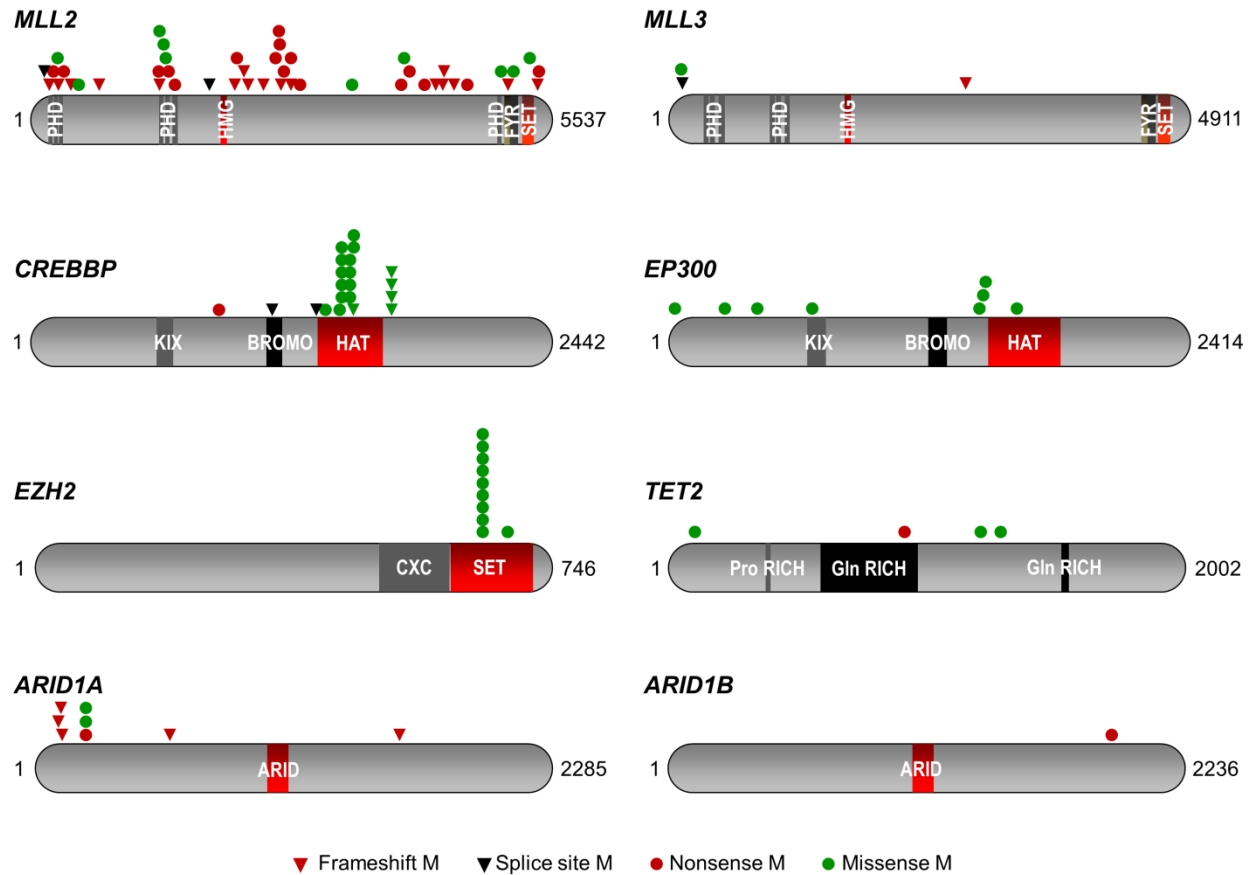


Figure 9. Mutations of histone modification and chromatin remodeling genes.

Schematic representation of 8 commonly mutated histone/chromatin modification enzymes in tFL, with their protein functional domains. Color-coded symbols indicate different types of mutations.

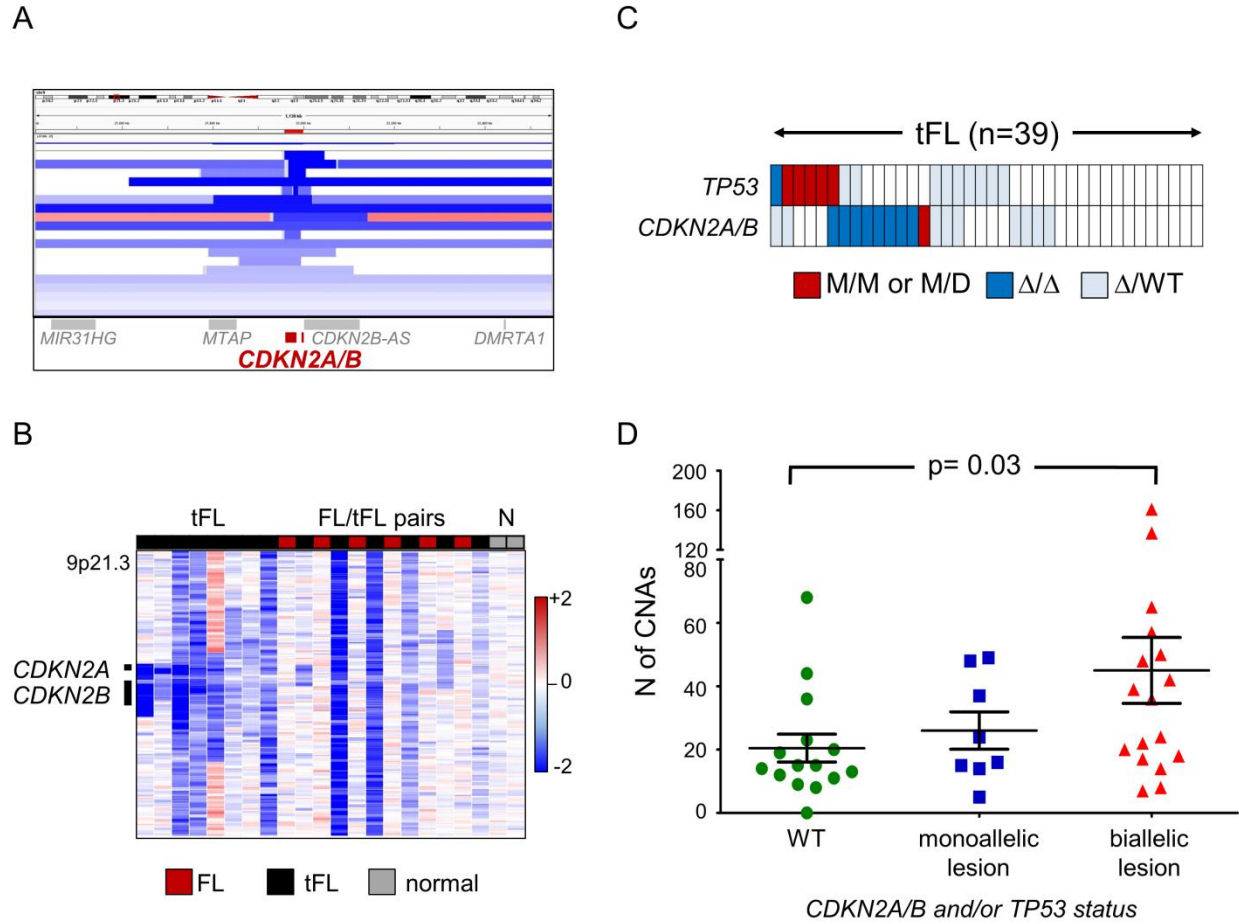


Figure 10. Biallelic loss of CDKN2A/B is specifically acquired during transformation. (A) Segmentation data from 18 DLBCL samples harboring CDKN2A/B deletions, visualized with the Integrative Genomics Viewer (<http://www.broadinstitute.org/igv>). Each track represents one sample, wherein white denotes a normal (diploid) copy number, blue indicates a region of copy-number loss, and red corresponds to a region of copy-number gain. The inferred copy number and the corresponding color intensity may vary across samples because of the presence of nontumor cells infiltrating the biopsy. Individual genes in the region are aligned in the bottom panel, and the area defined by the red bar at the top corresponds to the CDKN2A/B locus. (B) dChipSNP heatmap showing median-smoothed log2 CN ratio in 14 tFL biopsies harboring CDKN2A/B deletions, including six with matched FL biopsies and two

control DNAs (N). A vertical black bar indicates the location of the CDKN2A/B loci. (C) Relative distribution of biallelic lesions affecting CDKN2A/B and TP53. Columns correspond to individual patients, and the genomic status of the two genes is color-coded as indicated. (D) Samples carrying biallelic alterations of CDKN2A/B and/or TP53 are characterized by a significantly higher number of CN aberrations (Mann-Whitney U test, $p = 0.03$).

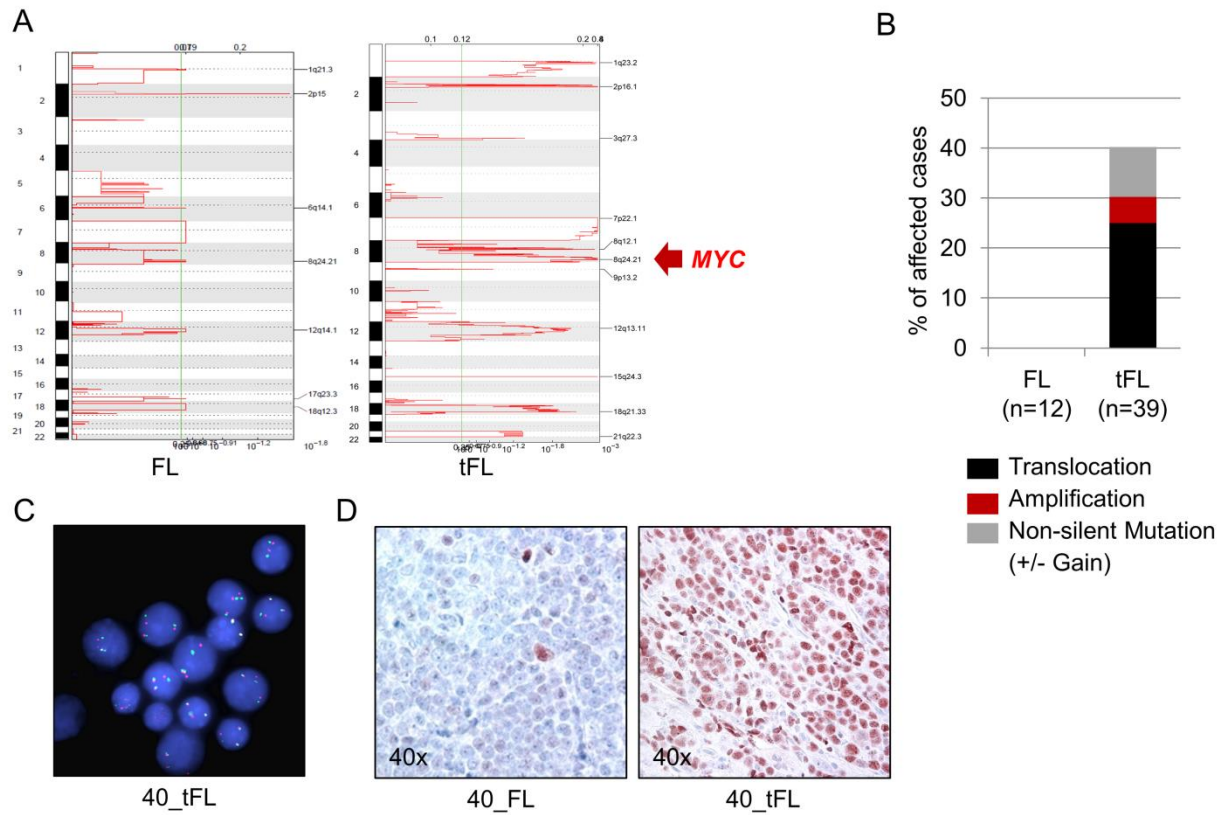


Figure 11. Recurrent genetic lesions of MYC in tFL. (A) GISTIC analysis of CN gains in FL and tFL cases (see also Figure 8). (B) Percentage of cases carrying genetic lesions of MYC in FL and tFL. (C) FISH analysis with a MYC break-apart probe in tFL #40. (D) IHC analysis of MYC expression in the pre- and post-transformation biopsy of the same patient.

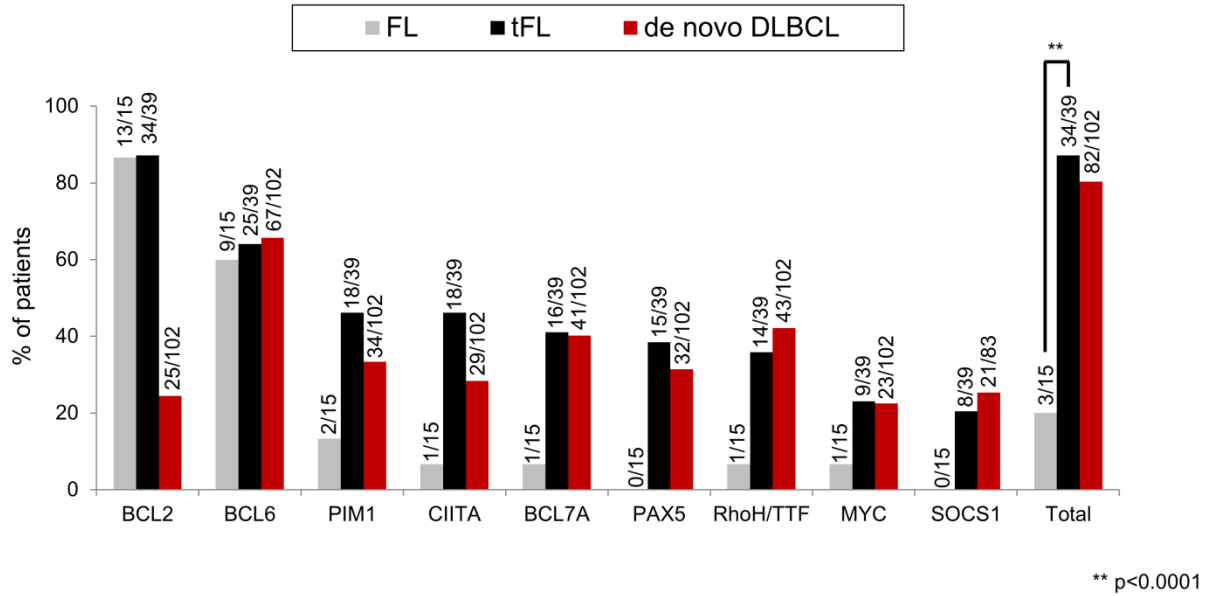


Figure 12. ASHM is associated with transformation. Proportion of cases harboring mutations in known ASHM targets; the *BCL6* intron 1 region, a physiologic target of SHM in GC B cells, and the *BCL2* sequences, which accumulate mutations in translocated alleles under the influence of the juxtaposed *IGH* promoter, are shown as controls. The three mutated FL cases harbored one single event each.

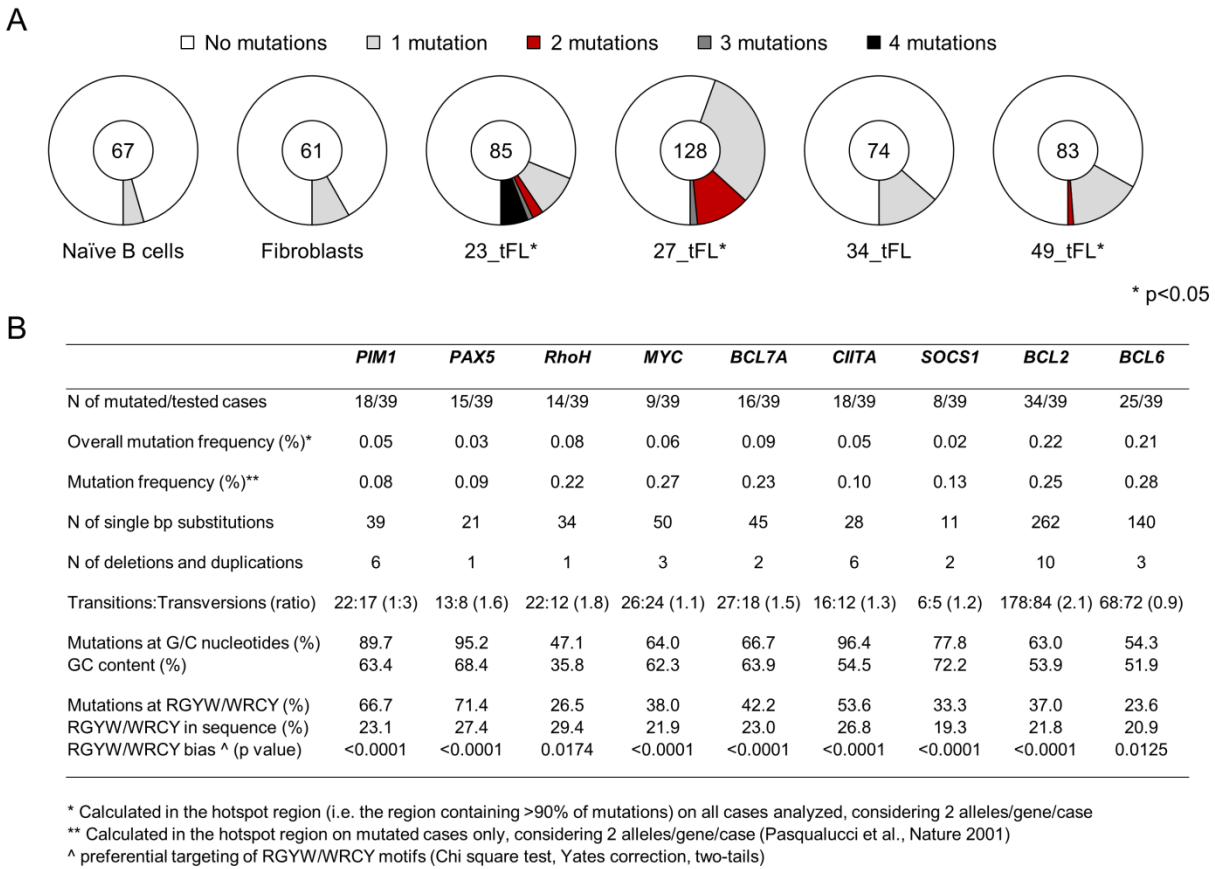


Figure 13. Mutation features of ASHM target genes. (A) Level of intraclonal variation in the *PAX5* sequences of four representative tFLs. Data from each patient are presented as pie charts and indicate the proportion of clones (out of a total of 68-128) that contain 0, 1, 2, etc., unique mutations (mutations present in the major clone are not included). Naïve B cells and the fibroblast cell line IMR91 were analyzed as control for the Pfu polymerase error. (B) Features of mutations affecting the 5' sequences of known ASHM target genes in tFL.

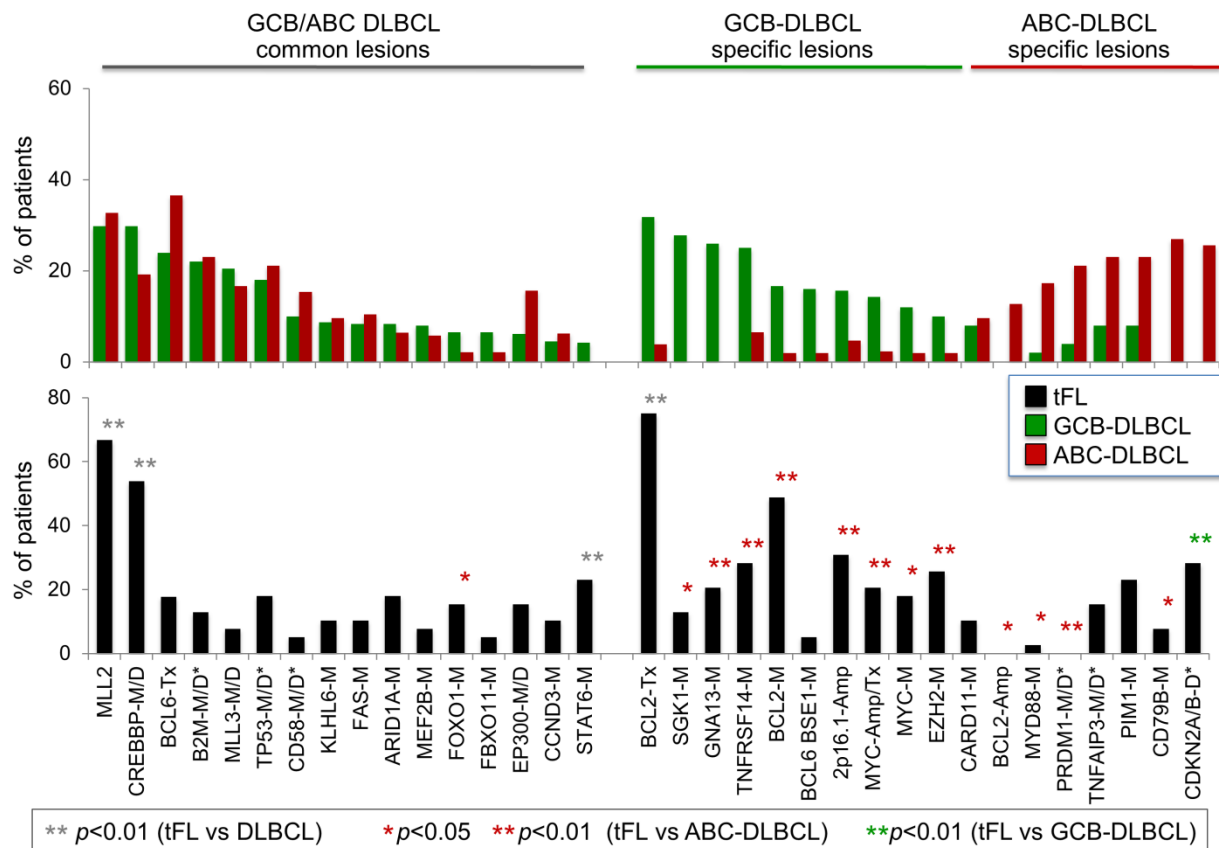


Figure 14. tFL displays a unique genomic profile that partially overlaps with GCB-DLBCL. Percentage of cases carrying the indicated genetic lesions in *de novo* DLBCL (top: green bars, GCB-DLBCL; n = 50; red bars, ABC-DLBCL; n = 52) and tFL (bottom: black bars; n = 39). Asterisks indicate statistically significant differences (one-tailed Fisher's exact test; *p < 0.05, **p < 0.01). Mutation frequencies for *SGK1* and *GNA13* are derived from Morin et al., 2011. M, mutation (missense, nonsense, frameshift, and splice-site); D, deletion; G, gain; AMP, amplification; M/D*, biallelic inactivation; Tx, translocation; BSE1, binding site in exon 1.

TABLES

Table I. Biological and phenotypic features of the discovery cases

Sample ID	Diagnosis	Year at diagnosis	Tissue site	% Tumor cells*	<i>IGHV</i> rearrangement (<i>IGHV/IGHD/IGHJ</i>)	<i>BCL2</i> R	Treatment regimen following FL and tFL diagnosis^^	<i>BCL2</i> IHC	<i>BCL6</i> IHC	<i>CD10</i> IHC	<i>MUM1</i> IHC	% <i>IGHV</i> homology
2_FL	FL grade I-II	2006	LN	65%	<i>IGHV3-9*01/D1-14*01/J4*02</i>	POS	R-CHOP based	>70%	30-70%	>70%	<10%	67%
2_tFL	DLBCL	2010	LN	90%	<i>IGHV3-9*01/D1-14*01/J4*02</i>	POS	ifosfamide-carboplatin-etoposide	>70%	>70%	>70%	<10%	67%
2_N	germline	na	PBMC	<1%	NEG	-	na	-	-	-	-	-
4_FL	FL grade I-II	1995	LN	50%	<i>IGHV3-11*03/D1-1*01/J4*02</i>	NEG**	CHOP	>70%	>70%	>70%	<10%	85%
4_tFL	DLBCL	1999	LN	50%	<i>IGHV3-11*03/D1-1*01/J4*02</i>	NEG**	dexamethasone-cytarabine-cisplatin	>70%	>70%	>70%	10-30%	85%
5_FL	FL grade II (80%)/IIIA (20%)	2008	LN	70%	<i>IGHV3-23*01/D3-22*01/J5*02</i>	POS	watch & wait (W&W)	70%	10%	70%	10%	87%
5_tFL	DLBCL	2010	spleen	70%	<i>IGHV3-23*01/D3-22*01/J5*02</i>	POS	R-COMP14 + Rituximab	70%	70%	70%	20%	88%
5_N	germline	na	saliva	<1%	NEG	-	na	-	-	-	-	-
6_FL	FL grade II	2002	LN	<20%	<i>IGHV3-23/D1-1/J4*02</i>	POS	not available	50%	90%	50%	20%	91%
6_tFL	DLBCL	2003	LN	90%	<i>IGHV3-23/D1-1/J4*02</i>	POS	not available	90%	80%	90%	0%	91%
8_FL	FL grade I-II	1999	LN	40%	<i>IGHV3-30*03/D1-7*01/J1*01</i>	NEG	R-CHOP + Zevalin	90%	5%	20%	5%	97%
8_tFL	DLBCL	2007	LN	70-80%	<i>IGHV3-30*03/D1-7*01/J1*01</i>	NEG	R-DHAP + R-MAD + ABMT + FEAM + ABMT	100%	20%	10%	40%	97%
11_FL	FL grade IIIA	2003	abdominal mass	80%	<i>IGHV3-74*01/D3-3*02/J5*02</i>	POS	R-CHOP based	90%	ND	90%	ND	79%
11_tFL	DLBCL (80%)/FL (20%)	2012	LN	>90%	<i>IGHV3-74*01/D3-3*02/J5*02</i>	POS	R-ODHA + EDX + FEAM + ABMT	90%	90%	90%	90%	79%
11_N	germline	na	BM	<1%	NEG	-	na	-	-	-	-	-
12_FL	FL grade I	2004	LN	75%	<i>IGHV4-30-4*01/D2-21*01/J6*03</i>	NEG	R-FND	80%	0%	30%	0%	94%
12_tFL	DLBCL	2004	abdominal mass	80%	<i>IGHV4-30-4*01/D2-21*01/J6*03</i>	NEG	R-CHOP + ABMT	0%	80%	90%	50%	85%
17_FL	FL grade I	2005	LN	50%	<i>IGHV3-21*01/D2-21*01/J1*02</i>	POS	R-CHOP	100%	100%	100%	90%	95%
17_tFL	DLBCL	2009	LN	50%	<i>IGHV3-21*01/D2-21*01/J1*02</i>	POS	not available	100%	90%	100%	80%	94%
23_FL	FL grade II	1997	skin	90%	<i>IGHV3-48*03/D2-2*02/J5*02</i>	POS	W&W - R-FND + C + anti-CD22 + ABMT + R	90%	90%	90%	10%	92%
23_tFL	DLBCL	2004	abdominal mass	>80%	<i>IGHV3-48*03/D2-2*02/J5*02</i>	POS	allogeneic BMT	90%	80%	90%	90%	90%
23_N	germline	na	BM	<1%	NEG	-	na	-	-	-	-	-
29_FL	FL grade II	2007	LN	55%	<i>IGHV3-23*01/D1-26*01/J4*02</i>	POS	W&W - R-COMP + R	90%	90%	100%	0%	84%
29_tFL	DLBCL	2010^	PE	80%	<i>IGHV3-23*01/D1-26*01/J4*02</i>	POS	R-OxDHA, Panabinostat	100%	0%	0%	0%	84%
39_FL	FL grade I	2006	LN	45%	<i>IGHV3-23*01/D1-14*01/J1*01</i>	POS	R-CVP	>80%	>80%	>80%	<5%	81%
39_tFL	DLBCL	2009	LN	75%	<i>IGHV3-23*01/D1-14*01/J1*01</i>	POS	na	>80%	>80%	>80%	<5%	81%
40_FL	FL grade I	2010	LN	40%	ND***	POS	R-CHOP based	>80%	>80%	>80%	<5%	ND
40_tFL	DLBCL	2011	soft tissue	90%	ND***	POS	patient died one month after diagnosis	>80%	30-40%	>70%	<6%	ND

Abbreviations: na, not applicable; ND, not determined; LN, lymph node; PBMC, peripheral blood mononuclear cells; BM, bone marrow; PE, pleural effusion; R, rearrangement; IHC, immunohistochemistry (% positive cells)

^^ R=Rituximab; CHOP=Cyclophosphamide, Doxorubicine, Vincristine, Prednisone; DHAP=Dexamethasone, Cytarabin, Cisplatinum; FND=Fludarabine, Mitoxantrone, Dexamethasone; COMP=Cyclophosphamide, Vincristine, Myocet (nonpegylated liposomal doxorubicin), Prednisone; OxDHA=Oxidized Docosahexaenoic Acid; FEAM=Fotemustine, Etoposide, Cytarabine, Melphalan; ABMT=Autologous Bone Marrow Transplantation; MAD=Mitoxantrone-Cytarabine-Dexamethasone; CVP=Cyclophosphamide, Vincristine, Prednisone

* assessed by analysis of the rearranged IG loci on high-density SNP arrays. The absence of tumor cells in the normal samples was also excluded by PCR amplification of the *IGHV* rearrangement

** based on PCR amplification of the *BCL2-IGH* MBR (FISH analysis failed in this case)

*** the clonal relationship between FL and tFL in this case was confirmed by analysis of the rearranged IG loci on high-density SNP arrays

^ Study sample obtained in 2012

Table II. Results of Illumina Hi-Seq2000 sequencing and mapping after whole exome capture

Sample ID	Target Region Coverage (%)			Mean Depth [^]	N of Reads*	% of heterozygous dbSNP detected**
	≥ 10X	≥ 20X	≥ 30X			
2_FL	89.4	84.6	78.0	93.8	71.9	95.0
2_tFL	89.3	84.3	77.7	95.5	73.6	94.0
2_N	89.6	85.0	78.8	95.9	73.5	95.0
4_FL	89.5	85.5	80.0	106.3	85.0	94.7
4_tFL	89.4	85.4	80.0	106.2	82.8	94.4
5_FL	89.2	83.2	75.6	90.1	71.2	92.5
5_tFL	88.7	83.3	76.3	91.6	71.8	94.8
5_N	87.0	75.8	63.5	59.5	111.6	90.9
6_FL	89.6	85.7	80.2	106.3	84.7	95.7
6_tFL	88.8	83.7	77.1	99.0	76.6	93.8
8_FL	89.6	85.0	78.9	101.9	80.2	95.7
8_tFL	90.1	86.7	82.0	119.4	91.7	95.3
11_FL	90.2	85.5	79.1	102.7	78.1	95.1
11_tFL	90.4	86.0	80.2	109.6	86.6	94.2
11_N	90.5	86.2	80.5	109.5	83.5	95.5
12_FL	88.4	80.5	71.1	73.2	61.4	94.8
12_tFL	83.8	68.8	53.9	45.8	60.8	93.2
17_FL	87.0	76.2	63.9	56.9	59.3	94.9
17_tFL	88.5	80.1	69.8	66.8	58.1	95.4
23_FL	88.1	79.2	68.5	64.7	60.1	95.0
23_tFL	88.4	80.2	70.3	69.4	62.4	94.8
23_N	90.3	85.1	78.3	90.4	64.0	94.4
29_FL	87.9	78.7	68.1	66.6	55.3	94.6
29_tFL	88.6	80.9	71.4	74.5	59.3	95.0
39_FL	85.5	72.3	58.5	50.6	59.6	93.8
39_tFL	88.0	78.8	68.2	68.1	55.4	94.1
40_FL	88.9	81.5	72.4	75.2	60.8	95.9
40_tFL	89.0	82.58	74.45	83.33	58.3	92.7
1_tFL	89.1	83.3	76.0	97.3	77.5	94.9
1_N	90.0	86.2	81.0	88.4	86.0	95.5
3_tFL	89.6	85.3	79.6	111.3	86.5	92.8
3_N	89.8	85.6	79.9	104.7	82.4	94.4
7_tFL	90.2	86.4	81.3	115.1	88.0	94.5
10_tFL	87.2	77.0	65.3	60.2	53.4	93.8
15_tFL	89.5	83.3	75.3	80.6	66.6	96.4
18_tFL	87.6	77.7	65.9	60.4	53.5	87.9
20_tFL	93.0	87.2	76.4	64.2	51.7	94.2
22_tFL	93.4	89.2	82.2	73.3	66.1	95.3
24_tFL	92.9	87.5	78.2	67.4	53.1	93.4
25_tFL	93.2	89.8	84.5	77.5	66.6	95.1
26_tFL	93.4	89.4	83.2	75.2	67.2	95.9
27_tFL	93.4	88.6	80.4	70.7	60.5	94.7
28_tFL	93.2	87.3	77.4	66.4	57.5	92.4
33_tFL	88.2	79.4	69.2	70.2	63.8	95.2
37_tFL	88.7	80.5	70.7	72.1	62.2	95.0
34_tFL	88.7	80.8	71.3	74.3	59.7	95.1
35_tFL	89.2	82.7	74.5	83.1	61.4	95.7
36_tFL	90.1	85.0	78.0	88.2	60.2	97.0
38_tFL	88.0	78.8	68.0	65.7	57.8	95.4
41_tFL	88.0	79.1	68.7	68.9	61.4	93.0
42_tFL	89.1	82.0	73.3	79.1	61.6	96.1
43_tFL	89.6	83.7	76.1	81.5	56.5	96.2
44_tFL	88.1	79.0	68.3	68.8	62.1	95.3
45_tFL	87.7	78.2	67.0	63.9	53.8	94.8
47_tFL	87.2	76.5	64.5	59.8	58.7	95.4
48_tFL	88.8	81.1	71.8	75.7	59.1	95.7
49_tFL	87.7	78.0	67.2	66.8	55.7	93.8

[^] Mean number of sequence reads covering the target exome

* Passed filter paired-end reads, in millions

** As assessed by comparison with Affymetrix SNP6.0 array data obtained from the same sample

Table III. SNVs identified by whole exome sequencing in the 12 discovery FL-tFL pairs

	Gene ID	Exon^	Strand	Genome Change^^	cDNAChange	ProteinChange	Variant Classification	WES FL				WES tFL				Class	Refseq prot Id	PPH2 Class	Sanger validation
								Var Depth	Total Depth	Freq	Adj Freq	Var Depth	Total Depth	Freq	Adj Freq				
2	ADAM9	13	+	g.chr8:38912084G>T	c.1387G>T	p.D463Y	Missense_Mutation	0	145	0	0	18	113	16	18	tFL specific (subclonal)	NP_003807	neutral	Yes
2	ADAMTS9	11	-	g.chr3:64633665C>T	c.1661G>A	p.R554Q	Missense_Mutation	0	127	0	0	34	105	32	35	tFL specific	NP_891550	neutral	Yes
2	ADK	9	+	g.chr10:76360204G>A	c.830G>A	p.R277Q	Missense_Mutation	14	80	18	24	27	68	40	45	Shared	NP_006712	deleterious	Yes
2	AKNAD1	13	-	g.chr1:109363230C>T	c.2186G>A	p.R729Q	Missense_Mutation	25	105	24	32	45	93	48	53	Shared	NP_689976	deleterious	Yes
2	ANKMY1	4	-	g.chr2:241465828C>T	c.721G>A	p.A241T	Missense_Mutation	26	148	18	24	61	169	36	40	Shared	NP_057636	neutral	Yes
2	APC2	13	+	g.chr19:1461961G>A	c.1639_splice	p.E547_splice	Splice_Site	0	39	0	0	20	38	53	58	tFL specific	NP_005874	nd	nd
2	B3GAT2	2	-	g.chr6:71603923C>T	c.644G>A	p.R215Q	Missense_Mutation	0	67	0	0	20	76	26	30	tFL specific	NP_542780	deleterious	Yes
2	BAP1	11	-	g.chr3:52439126C>A	c.1116G>T	p.Q372H	Missense_Mutation	41	171	24	32	84	185	45	50	Shared	NP_004647	deleterious	Yes
2	BCL2	1	-	g.chr18:60985896C>A	c.4G>T	p.A2S	Missense_Mutation	40	184	22	29	59	141	42	47	Shared	NP_000624	deleterious	nd
2	BCL2	1	-	g.chr18:60985881T>G	c.19A>C	p.T7P	Missense_Mutation	59	234	25	33	76	171	44	49	Shared	NP_000624	deleterious	nd
2	BIRC6	50	+	g.chr2:32730200C>G	c.9628C>G	p.L3210V	Missense_Mutation	64	242	26	35	1	397	0	0	FL specific	NP_057336	nd	nd
2	BMP7	1	-	g.chr20:55840986G>A	c.193C>T	p.P6S5	Missense_Mutation	33	166	20	27	66	193	34	38	Shared	NP_001710	neutral	Yes
2	C6orf204	5	-	g.chr6:11883247T>C	c.1241A>G	p.Y414C	Missense_Mutation	0	427	0	0	108	338	32	36	tFL specific	NP_001035940	deleterious	nd
2	C9orf170	2	+	g.chr9:89771500C>T	c.181C>T	p.R61W	Missense_Mutation	1	68	1	0	22	37	59	65	tFL specific	NP_001001709	none	nd
2	C9orf84	4,5	-	g.chr9:114518616A>C	c.659T>G	p.L220R	Missense_Mutation	0	102	0	0	16	45	36	40	tFL specific	NP_775792	neutral	Yes
2	C9orf86	14	+	g.chr9:139734668C>G			Splice_Site	29	144	20	27	57	132	43	48	Shared	NP_078994	deleterious	Yes
2	CD79B	5,6	-	g.chr17:62006656_62006657insA	c.619_620insT	p.Y207fs	Frame_Shift_Ins	106	181	59	78	198	263	75	83	Shared	NP_000617	deleterious	Yes
2	CHCHD6	7	+	g.chr3:126676332G>A	c.640G>A	p.E214K	Missense_Mutation	50	186	27	36	63	181	35	39	Shared	NP_115719	deleterious	Yes
2	CREBBP	29,30	-	g.chr16:3781323_3781325delAAG	c.5040_5042delCTT	p.L1681del	In_Frame_Del	7	46	15	21	11	36	31	35	Shared	NP_004371	deleterious	Yes
2	CREBBP	26,27	-	g.chr16:3786704A>G	c.4507T>C	p.Y1503H	Missense_Mutation	31	166	19	25	84	164	51	56	Shared	NP_004371	none	Yes
2	CSDE1	7,8	-	g.chr1:115275306G>T	c.969C>A	p.D323E	Missense_Mutation	25	135	19	26	90	187	48	53	Shared	NP_001007554	neutral	Yes
2	CXCR4	1,2	-	g.chr2:136872621_136872622delAG	c.876_877delCT	p.F292fs	Frame_Shift_Del	23	106	22	32	42	93	46	51	Shared	NP_003458	deleterious	Yes
2	DDX3Y	7	+	g.chrY:15025670T>A	c.578T>A	p.I193N	Missense_Mutation	0	126	0	0	114	130	88	98	tFL specific	NP_001116137	deleterious	nd
2	DOCK5	10	+	g.chr8:25159875T>A	c.881T>A	p.V294D	Missense_Mutation	35	155	23	31	0	183	0	0	FL specific	NP_079216	neutral	nd
2	DOCK5	46	+	g.chr8:25257463G>T	c.4741G>T	p.E1581*	Nonsense_Mutation	33	162	20	27	44	162	27	30	Shared	NP_079216	deleterious	Yes
2	DPPA2	5	-	g.chr3:109026987C>T	c.550G>A	p.A184T	Missense_Mutation	70	369	19	25	132	319	41	46	Shared	NP_620170	neutral	Yes
2	DSC3	8	-	g.chr18:28598745T>G	c.964A>C	p.I322L	Missense_Mutation	30	149	20	27	25	120	21	24	Shared	NP_001932	neutral	Yes
2	ETV7	3	-	g.chr6:36343719C>T	c.236G>A	p.G79E	Missense_Mutation	20	141	14	19	65	146	45	50	tFL specific	NP_057219	neutral	Yes
2	FADS1	1	-	g.chr11:61584160T>C	c.242A>G	p.E81G	Missense_Mutation	17	76	22	30	38	106	36	40	Shared	NP_037534	deleterious	Yes
2	FAIM	4,5	+	g.chr3:138351871A>C	c.490A>C	p.I164L	Missense_Mutation	26	78	33	44	0	78	0	0	FL specific	NP_001028204	neutral	nd
2	FANCD2	29	+	g.chr3:10119819C>A	c.2914C>A	p.L972I	Missense_Mutation	36	194	19	25	0	224	0	0	FL specific	NP_149075	deleterious	nd
2	FAS	8,9	+	g.chr10:90773895T>A	c.696T>A	p.Y232*	Nonsense_Mutation	36	181	20	27	65	174	37	41	Shared	NP_000034	deleterious	Yes
2	FCGBP	6	-	g.chr19:40419599C>G	c.3395G>C	p.C1132S	Missense_Mutation	0	174	0	0	44	121	36	40	tFL specific	NP_003881	deleterious	Yes
2	G3BP2	3	-	g.chr4:76582773T>C	c.319A>G	p.K107E	Missense_Mutation	30	106	28	37	54	106	51	57	Shared	NP_036429	deleterious	Yes
2	GCKR	5	+	g.chr2:27721592T>G	c.356T>G	p.V119G	Missense_Mutation	0	66	0	0	30	104	29	32	tFL specific	NP_001477	neutral	Yes
2	GNA13	3	-	g.chr17:63014369A>T	c.561_splice	p.P187_splice	Splice_Site	0	54	0	0	52	76	68	75	tFL specific	NP_006563	deleterious	Yes
2	GREB1L	30	+	g.chr18:19100755C>A	c.5579C>A	p.A1860D	Missense_Mutation	8	105	8	11	14	80	18	20	Shared (subclonal)	NP_001136438	deleterious	nd
2	GRID2	11	+	g.chr4:94376882C>T	c.1615C>T	p.R539C	Missense_Mutation	16	70	23	31	36	60	60	67	Shared	NP_001501	deleterious	Yes
2	GRID2IP	19	-	g.chr7:6541685C>T	c.3215G>A	p.S1072N	Missense_Mutation	0	120	0	0	46	112	41	46	tFL specific	NP_001138590	deleterious	Yes
2	HIST1H1E	1	+	g.chr6:26157109C>T	c.491C>T	p.A164V	Missense_Mutation	19	43	44	58	17	41	41	45	Shared	NP_005312	none	Yes
2	HIST1H2AC	1	+	g.chr6:26124578T>A	c.118T>A	p.Y40N	Missense_Mutation	20	91	22	30	43	89	48	53	Shared	NP_003503	deleterious	Yes
2	HIST1H2BF	1	+	g.chr6:26199883A>G	c.97A>G	p.S33G	Missense_Mutation	18	129	14	19	27	76	36	39	tFL specific	NP_003513	none	Yes
2	HIST1H3B	1	-	g.chr6:26031910G>A	c.379C>T	p.L127F	Missense_Mutation	2	87	2	0	19	82	23	26	tFL specific	NP_003528	deleterious	Yes
2	HNRNPU	1	-	g.chr1:245027102G>A	c.508C>T	p.Q170*	Nonsense_Mutation	101	359	28	37	162	404	40	44	Shared	NP_114032	deleterious	Yes
2	HS3ST2	1	+	g.chr16:22826360G>C	c.429G>C	p.L143F	Missense_Mutation	0	195	0	0	34	161	21	24	tFL specific	NP_006034	neutral	Yes
2	ITGAM	16	+	g.chr16:31332814T>G	c.1868T>G	p.I623S	Missense_Mutation	19	174	11	15	66	181	36	40	tFL specific	NP_000623	neutral	nd
2	JHDM1D	12	-	g.chr7:139801764C>T	c.1625G>A	p.C542Y	Missense_Mutation	12	51	24	32	20	59	34	37	Shared	NP_085150	deleterious	Yes
2	KIF6	13	-	g.chr6:39507923G>A	c.1501C>T	p.R501C	Missense_Mutation	89	258	34	45	113	297	38	42	Shared	NP_659464	neutral	Yes
2	MEGF9	1	-	g.chr9:123476565_123476567delGGC	c.46_48delGCC	p.A23del	In_Frame_Del	1	85	1	0	13	76	17	19	tFL specific	NP_001073966	deleterious	nd
2	MGAM	37	+	g.chr7:141765161G>A	c.4511G>A	p.R1504Q	Missense_Mutation	0	33	0	0	8	31	26	29	tFL specific	NP_004659	deleterious	nd
2	MIOX	4	+	g.chr22:50926402C>A	c.265C>A	p.P89T	Missense_Mutation	9	75	12	16	22	36	61	68	Shared	NP_060054	deleterious	Yes
2	MS4A1	5	+	g.chr11:60234430A>T	c.574_splice	p.G192_splice	Splice_Site	1	103	1	0	33	46	72	79	tFL specific	NP_690605	deleterious	Yes
2	MTSSL1	15	-	g.chr16:70697674T>G	c.2150A>C	p.D717A	Missense_Mutation	0	12	0	0	8	24	33	38	tFL specific	NP_612392	neutral	nd
2	MX1	1	+	g.chr21:															

2	SLC1A7	7	-	g.chr1:53558439A>G	c.818T>C	p.V273A	Missense_Mutation	16	66	24	32	45	111	41	45	Shared	NP_006662	neutral	Yes
2	SLC4A7	5	-	g.chr3:27477893A>G	c.548T>C	p.L183P	Missense_Mutation	33	175	19	25	0	174	0	0	FL specific	NP_003606	deleterious	nd
2	SMC2	3	+	g.chr9:106860766G>A	c.358G>T	p.V120I	Missense_Mutation	0	123	0	0	47	78	60	67	tFL specific	NP_001036016	neutral	Yes
2	SPTBN4	30	+	g.chr19:41073888C>A	c.6656C>A	p.A2219E	Missense_Mutation	9	33	27	35	19	39	49	54	Shared	NP_066022	none	Yes
2	STK3	6	-	g.chr8:99718847G>A	c.532C>T	p.R178C	Missense_Mutation	47	121	39	52	51	95	54	60	Shared	NP_006272	deleterious	nd
2	STXBP2	11	+	g.chr19:7707699C>T	c.950C>T	p.T317I	Missense_Mutation	0	53	0	0	18	45	40	44	tFL specific	NP_008880	neutral	Yes
2	SVIL	14,18	-	g.chr10:29782207T>A	c.3955A>T	p.M1319L	Missense_Mutation	0	46	0	0	18	64	28	31	tFL specific	NP_068506	neutral	Yes
2	TBC1D4	1	p.	g.chr13:76055882G>A	c.22C>T	p.Q8*	Nonsense_Mutation	45	228	20	27	74	213	35	39	Shared	NP_055647	neutral	Yes
2	TIE1	12	+	g.chr1:43778152C>T	c.1807C>T	p.R603C	Missense_Mutation	16	67	24	32	0	47	0	0	FL specific	NP_005415	deleterious	nd
2	TNFRSF11A	9	+	g.chr18:60036496G>A	c.1346G>A	p.C449Y	Missense_Mutation	33	177	19	25	38	129	29	32	Shared	NP_003830	deleterious	Yes
2	TRIM8	1	+	g.chr10:104404615_104404636del22bp	c.241_262del22bp	p.L81fs	Frame_Shift_Del	0	422	0	0	75	406	19	21	tFL specific	NP_112174	Yes	Yes
2	TSPAN5	1	-	g.chr4:99579329A>G	c.49T>C	p.Y117H	Missense_Mutation	10	42	24	32	12	28	43	48	Shared	NP_005714	deleterious	Yes
2	TTBK1	13	+	g.chr6:43250722T>G	c.2244T>G	p.D748E	Missense_Mutation	16	103	16	21	1	72	1	0	FL specific	NP_115927	neutral	nd
2	UHRF1BP1L	5	-	g.chr12:100492218A>C	c.440T>G	p.L147R	Missense_Mutation	0	250	0	0	33	198	17	19	tFL specific (subclonal)	NP_055869	deleterious	Yes
2	XBP1	1	-	g.chr22:29196505A>T	c.8T>A	p.V3E	Missense_Mutation	4	29	14	19	8	27	30	33	Shared	NP_001073007	deleterious	Yes
2	XPO1	14	-	g.chr2:61719472C>T	c.1711G>A	p.E571K	Missense_Mutation	0	303	0	0	192	327	59	65	tFL specific	NP_003391	deleterious	Yes
2	ZFP36L1	2	-	g.chr14:69257097A>T	c.170T>A	p.L57Q	Missense_Mutation	26	70	37	49	20	55	36	40	Shared	NP_004917	neutral	Yes
2	ZNF267	4	+	g.chr16:31927203G>C	c.1633G>C	p.E545Q	Missense_Mutation	48	161	30	40	73	154	47	52	Shared	NP_003405	deleterious	Yes
4	AUTS2	1	+	g.chr7:69064912T>G	c.273T>G	p.F91L	Missense_Mutation	0	25	0	0	8	25	32	64	tFL specific	NP_056385	neutral	nd
4	BCL2	1	-	g.chr18:60985679G>T	c.221C>A	p.T74N	Missense_Mutation	55	186	30	60	34	179	19	38	Shared	NP_000624	neutral	nd
4	BCL2	1	-	g.chr18:60985644G>A	c.256C>T	p.L86F	Missense_Mutation	51	187	27	54	32	150	21	42	Shared	NP_000624	neutral	nd
4	BCL2	1	-	g.chr18:60985305G>A	c.595C>T	p.L199F	Missense_Mutation	49	192	26	52	42	195	22	44	Shared	NP_000624	neutral	nd
4	BCL2	1	-	g.chr18:60985508C>G	c.392C>G	p.A131G	Missense_Mutation	73	294	25	50	74	341	22	44	Shared	NP_000624	neutral	Yes
4	C6orf174	1	-	g.chr6:127837683C>G	c.77G>C	p.R26P	Missense_Mutation	1	60	2	0	9	57	16	32	tFL specific	NP_001012279	none	nd
4	CARD11	5	-	g.chr7:2979559C>T	c.688G>A	p.D230N	Missense_Mutation	105	291	36	72	71	294	24	48	Shared	NP_115791	deleterious	Yes
4	CCAR1	10	+	g.chr10:70513622_70513624delATG	c.1132_1134delATG	p.M379del	In_Frame_Del	0	176	0	0	26	156	16	32	tFL specific	NP_060707	Yes	Yes
4	CREBBP	26,27	-	g.chr16:3786760A>G	c.4451T>C	p.F1484S	Missense_Mutation	24	103	23	46	21	89	24	48	Shared	NP_004371	none	Yes
4	CSMD2	23	-	g.chr1:34166260T>C	c.3470A>G	p.N1157S	Missense_Mutation	0	69	0	0	14	82	17	34	tFL specific	NP_443128	neutral	Yes
4	DCC	18	+	g.chr18:50923807T>C	c.2818T>C	p.Y940H	Missense_Mutation	0	88	0	0	13	65	20	41	tFL specific	NP_005206	neutral	Yes
4	DCHS2	25	-	g.chr4:155157060A>G	c.7379T>C	p.I2460T	Missense_Mutation	2	308	1	0	55	317	17	34	tFL specific	NP_060109	neutral	Yes
4	HIST1H2BI	1	+	g.chr6:26273535C>T	c.332C>T	p.A1111V	Missense_Mutation	44	172	26	52	0	174	0	0	FL specific	NP_003516	deleterious	nd
4	LPL	3	+	g.chr8:19809308C>A	c.278C>A	p.P93Q	Missense_Mutation	35	120	29	58	1	122	1	0	FL specific	NP_000228	neutral	nd
4	MLL2	31	-	g.chr12:49434889G>A	c.6664C>T	p.Q2222*	Nonsense_Mutation	18	46	39	77	6	48	12	25	Shared	NP_003473	Yes	Yes
4	MLL2	34	-	g.chr12:49432396G>A	c.8743C>T	p.R2915*	Nonsense_Mutation	10	73	14	28	10	61	16	33	Shared	NP_003473	Yes	Yes
4	PCDH15	20,21	-	g.chr10:55721653C>T	c.2869_splice	p.G957_splice	Splice_Site	0	92	0	0	23	135	17	34	tFL specific	NP_149045	Yes	nd
4	PLVAP	1	-	g.chr19:17487901C>T	c.197G>A	p.R66Q	Missense_Mutation	27	133	20	40	0	122	0	0	FL specific	NP_112600	deleterious	nd
4	RAB7L1	1	-	g.chr1:205744066G>C	c.19C>G	p.L7V	Missense_Mutation	0	239	0	0	49	283	17	34	tFL specific	NP_003920	neutral	Yes
4	SFI1	14,15	+	g.chr22:31998256G>A	c.1603G>A	p.E535K	Missense_Mutation	1	73	1	0	16	75	21	43	tFL specific	NP_001007468	deleterious	Yes
4	SHISA7	4	-	g.chr19:55944810G>A	c.1330C>T	p.R444W	Missense_Mutation	0	56	0	0	15	44	34	68	tFL specific	NP_001138648	deleterious	nd
4	SYK	10,11	+	g.chr9:93650064T>A	c.1615T>A	p.Y539N	Missense_Mutation	24	95	25	50	0	79	0	0	FL specific	NP_003168	deleterious	nd
4	TAF1	21	+	g.chrX:70613247C>T	c.3145C>T	p.R1049C	Missense_Mutation	0	153	0	0	27	125	22	44	tFL specific	NP_002178	deleterious	nd
4	TMEM30A	4,5	-	g.chr6:75969072G>A	c.676C>T	p.R226*	Nonsense_Mutation	44	141	31	62	0	170	0	0	FL specific	NP_060717	Yes	nd
4	TNFAIP3	8	+	g.chr6:138202192_138202193delITG	c.2109_2110delITG	p.D703fs	Frame_Shift_Del	15	58	26	52	0	59	0	0	FL specific	NP_062681	nd	Yes
4	WFS1	7	+	g.chr4:6303891C>T	c.2369C>T	p.S790L	Missense_Mutation	0	301	0	0	54	192	28	56	tFL specific	NP_001139325	neutral	Yes
5	ABCA1	9	-	g.chr9:107599817C>A	c.1086G>T	p.L362F	Missense_Mutation	8	36	22	32	0	27	0	0	FL specific (tFL LOH)	NP_005493	neutral	nd
5	ADAM23	7	+	g.chr2:207412210C>A	c.778C>A	p.Q260K	Missense_Mutation	6	28	21	30	0	35	0	0	FL specific	NP_003803	neutral	nd
5	ADAMTS15	8	-	g.chr11:130343699G>A	c.2836G>A	p.V946I	Missense_Mutation	15	49	31	44	0	96	0	0	FL specific	NP_620686	neutral	nd
5	ADCY4	19	-	g.chr14:24791830C>A	c.2427G>T	p.Q809H	Missense_Mutation	8	36	22	32	0	45	0	0	FL specific	NP_640340	deleterious	nd
5	ADCY4	18	-	g.chr14:24792213G>T	c.2239C>A	p.L747I	Missense_Mutation	5	15	33	47	0	53	0	0	FL specific	NP_640340	deleterious	nd
5	ARHGAP4	1	-	g.chrX:153191649_153191660del12bp			Translation_Start_Site	24	97	25	36	29	72	41	58	Shared	NP_001657	nd	nd
5	ARID1A	3	+	g.chr1:27057727C>T	c.1435C>T	p.Q479*	Nonsense_Mutation	65	279	23	33	0	278	0	0	FL specific	NP_006006	Yes	Yes
5	ARID1B	20	+	g.chr6:157528087C>T	c.5758C>T	p.Q1920*	Nonsense_Mutation	26	120	22	31	79	136	58	83	Shared	NP_059989	Yes	Yes
5	B2M	1	+	g.chr15:45003781_45003782delCT	c.37_38delCT	p.L13fs	Frame_Shift_Del	0	65	0	0	9	22	41	59	tFL specific	NP_004039	Yes	Yes
5	BCL2	1	-	g.chr18:60985363C>G	c.537G>C	p.E179D	Missense_Mutation	88	277	32	46	76	260	29	42	Shared	NP_000624	neutral	Yes
5	C12orf50	10	-	g.chr12:88379699G>A	c.1054C>T	p.R352C	Missense_Mutation	38	139	27	38	53	148	36	51	Shared	NP_689802	neutral	nd
5	C16orf7	10	-	g.chr16:89776997C>T	c.1255G>A	p.V419I	Missense_Mutation	60	191	31	44	76	196	39	56	Shared	NP_004904	nd	nd
5	CCDC6	1	-	g.chr10:61666050T>G	c.133A>G	p.K45E	Missense_Mutation	1	71	1	0	12	61	20	29	tFL specific	NP_005427	neutral	nd
5	CD36	4	+	g.chr7:80292358T>G	c.482T>G	p.L161R	Missense_Mutation	23	82	28	40	19	53	36	51	Shared	NP_001120916	neutral	nd
5	CDH9	2	-	g.chr5:26915882G>A	c.379C>T	p.R127C	Missense_Mutation	61	195	31	44	44	132	33	47	Shared	NP_057363	deleterious	nd
5	CDKAL1	7	+	g.chr6:20846305G>A	c.639_splice	p.G139_splice	Splice_Site	130	404	32	46	0	427	0	0	FL specific (tFL LOH)	NP_060244	nd	nd
5	CDR2	5	-	g.chr16:22359001A>C	c.650T>G	p.V217G	Missense_Mutation	1	159	1	0	39	157	25	36	tFL specific	NP_001793	neutral	nd
5	CELSR3	1	-	g.chr3:48697682C>T	c.2386G>A	p.G796S	Missense_Mutation	0	102	0	0	22	77	29	42	tFL specific	NP_001398	deleterious	nd
5	CENPB	1	-	g.chr20:3766275C>T	c.856G>A	p.A286T	Missense_Mutation	26	76	34	49	21	126	17	25	Shared	NP_001801	deleterious	nd
5	CNTN2	1	+	g.chr1:205022366C>A	c.53C>A	p.A18D	Missense_Mutation	0	34	0	0	6	32	19	28	tFL specific	NP_005067	neutral	nd
5	DDRGN1	2	-	g.chr20:3183906A>C	c.248T>G	p.V83G	Missense_Mutation	11	259	4	6	32	182	18	26	Shared (subclonal)	NP_076424	deleterious	nd
5	DMBT1	9	+	g.chr10:124338236A>T	c.640A>T	p.S214C	Missense_Mutation	56	173	32	46	0	199	0	0	FL specific	NP_015568	deleterious	nd
5	E2F2	2	-	g.chr1:23850874C>A	c.358_splice	p.T120_splice	Splice_Site	5	16	31	44	0	25	0	0	FL specific	NP_004082	nd	nd
5	EP300	5	+	g.chr22:41525969T>C	c.1244T>C	p.L415P	Missense_Mutation	59	216	27	39	52	200	26	37	Shared	NP_001420	deleterious	Yes
5	EP300	27	+	g.chr22:41566470T>G	c.4347T>G	p.H1449Q	Missense_Mutation	0	350	0	0	84	299	28	40	tFL specific	NP_001420	none	Yes
5	EP300	27	+	g.chr22:41566571A>G	c.4448A>G	p.Y1483C	Missense_Mutation	0	221	0	0	51	166	31	44	tFL specific	NP_001420	none	Yes
5	EP300	27	+	g.chr22:41566522T>A	c.4399T>A	p.Y1467N	Missense_Mutation	39	310	13	19	93	262	35	50	Shared	NP_001420		

5	HUWE1	34	-	g.chrX:53615357T>A	c.4599A>T	p.L1533F	Missense_Mutation	10	34	29	41	0	23	0	0	FL specific	NP_113584	deleterious	nd
5	IL11	5	-	g.chr19:55877505G>A	c.470C>T	p.P157L	Missense_Mutation	16	23	70	100	15	32	47	67	Shared	NP_000632	deleterious	nd
5	ITCH	8	+	g.chr20:33030093A>G	c.1073A>G	p.E358G	Missense_Mutation	0	79	0	0	20	81	25	36	tFL specific	NP_113671	deleterious	nd
5	KIAA0889	4	-	g.chr20:35444652A>G	c.479T>C	p.M160T	Missense_Mutation	0	161	0	0	74	192	39	56	tFL specific	NP_954650	deleterious	nd
5	KIAA1211	5	+	g.chr4:57180703G>C	c.1035G>C	p.R345S	Missense_Mutation	8	25	32	46	0	18	0	0	FL specific (tFL LOH)	NP_065773	none	nd
5	KIAA1211	5	+	g.chr4:57180711delA	c.1043delA	p.E348fs	Frame_Shift_Del	8	19	42	61	0	19	0	0	FL specific (tFL LOH)	NP_065773	none	nd
5	KIAA1211	5	+	g.chr4:57180712G>T	c.1044G>T	p.E348D	Missense_Mutation	8	19	42	61	0	19	0	0	FL specific (tFL LOH)	NP_065773	none	nd
5	KIF20B	19	+	g.chr10:91497909A>T	c.3311A>T	p.E1104V	Missense_Mutation	24	110	22	32	8	95	8	12	FL specific (tFL subclonal)	NP_057279	deleterious	nd
5	LNPEP	2,3	+	g.chr5:96320847A>T	c.924A>T	p.E308D	Missense_Mutation	11	32	34	48	0	22	0	0	FL specific (tFL LOH)	NP_005566	deleterious	nd
5	MAGEC1	2	+	g.chrX:140995690T>G	c.2500T>G	p.F834V	Missense_Mutation	207	354	58	83	0	266	0	0	FL specific	NP_005453	none	nd
5	MAST4	28	+	g.chr5:66460745T>C	c.5171T>C	p.V1724A	Missense_Mutation	38	186	20	28	43	138	31	45	Shared	NP_055998	neutral	nd
5	MKI67	12	-	g.chr10:129901103C>T	c.9001G>A	p.G3001S	Missense_Mutation	32	151	21	30	53	143	37	53	Shared	NP_002408	none	nd
5	MLL2	53	-	g.chr12:49415877_49415878delTT	c.16469_16470delAA	p.K5490fs	Frame_Shift_Del	49	178	28	40	66	258	26	37	Shared	NP_003473	none	Yes
5	MLL2	5	-	g.chr12:49433328_49433329insA	c.8118_8119insT	p.T2706fs	Frame_Shift_Ins	13	47	28	39	22	66	33	47	Shared	NP_003473	none	Yes
5	MLLT6	11	+	g.chr17:36873796T>C	c.1763T>C	p.L588P	Missense_Mutation	0	109	0	0	9	43	21	30	tFL specific	NP_005928	deleterious	nd
5	NAA11	1	-	g.chr4:80246910A>C	c.122T>G	p.L41R	Missense_Mutation	0	144	0	0	60	118	51	73	tFL specific	NP_116082	deleterious	nd
5	NDST2	6	-	g.chr10:75565345C>A	c.1746G>T	p.Q582H	Missense_Mutation	7	23	30	43	0	42	0	0	FL specific	NP_003626	deleterious	nd
5	NDST2	6	-	g.chr10:75565380C>A	c.1711G>T	p.E571*	Nonsense_Mutation	7	34	21	31	0	75	0	0	FL specific	NP_003626	deleterious	nd
5	NOTCH2	5	-	g.chr1:120529605G>T	c.852C>A	p.C284*	Nonsense_Mutation	5	21	24	36	0	21	0	0	FL specific	NP_077719	deleterious	nd
5	OR10K2	1	-	g.chr1:158390578G>A	c.79C>T	p.L27F	Missense_Mutation	0	162	0	0	58	146	40	57	tFL specific	NP_001004476	deleterious	nd
5	PARP14	12	+	g.chr3:122433232delA	c.3956delA	p.E1319fs	Frame_Shift_Del	30	115	26	37	21	107	20	29	Shared	NP_060024	deleterious	nd
5	PGR	1	-	g.chr11:100998942G>A	c.860C>T	p.A287V	Missense_Mutation	0	153	0	0	58	203	29	42	tFL specific	NP_000917	neutral	nd
5	PHF20L1	14	+	g.chr8:133844626G>A	c.1891G>A	p.V631I	Missense_Mutation	33	164	20	28	0	300	0	0	FL specific	NP_057102	deleterious	nd
5	PHF23	5	-	g.chr17:7139100T>C	c.1057A>G	p.M353V	Missense_Mutation	143	489	29	42	0	342	0	0	FL specific (tFL LOH)	NP_077273	deleterious	nd
5	PIK3R1	9,14	+	g.chr5:67592015G>A	c.1831G>A	p.E611K	Missense_Mutation	9	30	30	42	0	42	0	0	FL specific	NP_852664	neutral	nd
5	PIM1	1	+	g.chr6:37138423G>C	c.72G>C	p.K24N	Missense_Mutation	0	217	0	0	82	347	24	34	tFL specific	NP_002639	neutral	Yes
5	PLEC	25	-	g.chr8:145003317C>A	c.3625G>A	p.G1209C	Missense_Mutation	9	43	21	30	0	77	0	0	FL specific	NP_958782	none	nd
5	PLEC	1	-	g.chr8:145024684G>C	c.191C>G	p.T64S	Missense_Mutation	8	45	18	27	0	53	0	0	FL specific	NP_958782	none	nd
5	PSMD14	10	+	g.chr2:162267814A>T	c.836A>T	p.D279V	Missense_Mutation	38	182	21	30	0	218	0	0	FL specific	NP_005796	deleterious	nd
5	PYGL	18	-	g.chr14:51375672A>G	c.2179T>C	p.Y727H	Missense_Mutation	9	69	13	19	35	83	42	60	tFL specific	NP_002854	deleterious	nd
5	RAB33A	1	+	g.chrX:129306245G>A	c.209G>A	p.G70D	Missense_Mutation	40	65	62	88	49	72	68	97	Shared	NP_004785	deleterious	nd
5	RLTPR	22	+	g.chr16:67683987C>A	c.2118C>A	p.D706E	Missense_Mutation	0	510	0	0	117	449	26	37	tFL specific	NP_001013860	deleterious	nd
5	ROSI	43	-	g.chr6:117609660C>T	c.7039G>A	p.D2347N	Missense_Mutation	183	625	29	41	0	462	0	0	FL specific	NP_002935	deleterious	nd
5	RXFP1	16	+	g.chr4:159567991A>G	c.1394A>G	p.D465G	Missense_Mutation	42	157	27	39	65	139	47	67	Shared	NP_067647	deleterious	nd
5	SEC24B	10,11	+	g.chr4:110437783G>A	c.2113G>A	p.E705K	Missense_Mutation	0	51	0	0	9	37	24	35	tFL specific	NP_006314	neutral	nd
5	SEC24B	10,11	+	g.chr4:110437797G>A	c.2126_splice	p.K709_splice	Splice_Site	0	43	0	0	8	36	22	32	tFL specific	NP_006314	neutral	nd
5	SEC24B	10,11	+	g.chr4:110437798T>G	c.2126_splice	p.K709_splice	Splice_Site	0	43	0	0	8	35	23	33	tFL specific	NP_006314	neutral	nd
5	SIK3	18	-	g.chr11:116732038G>T	c.2059C>A	p.P687T	Missense_Mutation	7	34	21	31	0	28	0	0	FL specific	NP_079440	deleterious	nd
5	SIK3	9	-	g.chr11:116746598G>T	c.1049C>A	p.A350E	Missense_Mutation	8	37	22	31	0	37	0	0	FL specific	NP_079440	neutral	nd
5	SLAIN2	7	+	g.chr4:48422366A>G	c.1585A>G	p.T529A	Missense_Mutation	89	265	34	48	0	167	0	0	FL specific (tFL LOH)	NP_065897	neutral	nd
5	SLC22A1	6	+	g.chr6:160557637G>A	c.1016G>A	p.R339H	Missense_Mutation	39	101	39	55	0	94	0	0	FL specific (tFL LOH)	NP_003048	deleterious	nd
5	SLC6A1	10	+	g.chr3:11070971G>A	c.1256G>A	p.R419H	Missense_Mutation	0	127	0	0	45	111	41	58	tFL specific	NP_003033	deleterious	nd
5	SLC06A1	13	-	g.chr5:101709112G>T	c.2104C>A	p.P702T	Missense_Mutation	28	136	21	30	0	89	0	0	FL specific (tFL LOH)	NP_775759	neutral	nd
5	STAT6	11	-	g.chr12:57496661T>C	c.1256A>G	p.D419G	Missense_Mutation	52	242	21	30	65	304	21	30	Shared	NP_003144	neutral	nd
5	SYNJ2	17	+	g.chr6:158497754A>C	c.2389A>G	p.T797P	Missense_Mutation	1	60	2	0	9	39	23	33	tFL specific	NP_003889	neutral	nd
5	TFDP3	1	-	g.chrX:132351690A>C	c.598T>G	p.L200V	Missense_Mutation	0	54	0	0	52	89	58	83	tFL specific	NP_057605	deleterious	nd
5	THAP10	1	-	g.chr15:71184332C>T	c.280G>A	p.A94T	Missense_Mutation	0	56	0	0	20	79	25	35	tFL specific	NP_064532	neutral	nd
5	TM4SF5	3	+	g.chr17:4685831G>A	c.292G>A	p.V98M	Missense_Mutation	16	67	24	34	0	54	0	0	FL specific (tFL LOH)	NP_003954	neutral	nd
5	TNFRSF14	4	+	g.chr1:2491260A>G	c.305_splice	p.A102_splice	Splice_Site	42	147	29	41	0	87	0	0	FL specific (tFL LOH)	NP_003811	neutral	Yes
5	TNFRSF14	4	+	g.chr1:2491315delG	c.358delG	p.G120fs	Frame_Shift_Del	27	113	24	35	0	75	0	0	FL specific	NP_003811	neutral	Yes
5	TNIK	23	-	g.chr3:170811704G>A	c.2645C>T	p.T882M	Missense_Mutation	56	157	36	51	46	140	33	47	Shared	NP_055843	neutral	nd
5	TP53	4	-	g.chr17:7578461C>A	c.469G>T	p.V157F	Missense_Mutation	16	56	29	41	0	74	0	0	FL specific (tFL LOH)	NP_001119584	deleterious	Yes
5	TRIM35	1	-	g.chr8:27168745C>T	c.8G>A	p.R3Q	Missense_Mutation	11	24	46	65	0	24	0	0	FL specific	NP_741983	neutral	nd
5	TTL13	4	+	g.chr3:9859358G>T	c.345G>T	p.L115F	Missense_Mutation	0	64	0	0	6	38	16	24	tFL specific	NP_001021100	deleterious	nd
5	TUSC3	1	+	g.chr8:15397991C>T	c.52C>T	p.R18W	Missense_Mutation	0	27	0	0	22	73	30	42	tFL specific	NP_006756	none	nd
5	UBAC1	5	-	g.chr9:138838114C>A	c.544_splice	p.A182_splice	Splice_Site	13	72	18	26	0	69	0	0	FL specific (tFL LOH)	NP_057256	deleterious	nd
5	UTRN	65	+	g.chr6:145142097A>C	c.9437A>C	p.H3146P	Missense_Mutation	2	93	2	0	25	68	37	53	tFL specific	NP_009055	deleterious	nd
5	VP541	18,19	-	g.chr7:38796559C>T	c.1574G>A	p.G525D	Missense_Mutation	0	151	0	0	27	96	28	40	tFL specific	NP_055211	neutral	nd
5	ZCCHC11	7	-	g.chr1:52956441C>T	c.1351G>A	p.A451T	Missense_Mutation	58	148	39	55	7	69	10	14	FL specific (tFL subclonal)	NP_056084	neutral	nd
5	ZCCHC7	4	+	g.chr9:37305667G>C	c.907G>C	p.D303H	Missense_Mutation	65	203	32	46	73	247	30	43	Shared	NP_115602	neutral	nd
5	ZMYM3	12	-	g.chrX:70467305C>T	c.2204G>A	p.W735*	Nonsense_Mutation	14	15	93	100	0	21	0	0	FL specific	NP_963893	deleterious	nd
6	ABCA13	10	+	g.chr7:48280528T>A	c.1127T>A	p.L376Q	Missense_Mutation	0	53	0	0	18	68	26	29	tFL specific	NP_689914	deleterious	Yes
6	ABCC6	30	-	g.chr16:16244597C>T	c.4241G>A	p.R1414H	Missense_Mutation	0	63	0	0	28	80	35	39	tFL specific	NP_001162	deleterious	nd
6	AIM2	3	-	g.chr1:159035995T>G	c.521A>C	p.Q174P	Missense_Mutation	32	348	9	45	111	269	41	46	Shared	NP_004824	deleterious	Yes
6	ALDH1L1	5	-	g.chr3:125873443C>T	c.674G>A	p.R225H	Missense_Mutation	0	126	0	0	30	82	37	41	tFL specific	NP_036322	deleterious	nd
6	AMACR	1	-	g.chr5:34008097C>G	c.28G>C	p.E10Q	Missense_Mutation	0	111	0	0	29	105	28	31	tFL specific	NP_055139	deleterious	nd
6	ATG2A	37	-	g.chr11:64664968T>C	c.5068_splice	p.G1690_splice	Splice_Site	0	132	0	0	43	108	40	45	tFL specific	NP_055919	deleterious	nd
6	B3GAT1	2	-	g.chr11:134253908G>A	c.287C>T	p.P96L	Missense_Mutation	0	70	0</									

6	DSEL	1	-	g.chr18:65178982T>A	c.2894A>T	p.K965M	Missense_Mutation	0	316	0	0	59	209	28	31	tFL specific	NP_115536	neutral	nd
6	ELL2		-	g.chr5:95297410T>C	c.16A>G	p.T6A	Missense_Mutation	1	28	4	20	10	33	30	35	Shared	NP_036213	neutral	nd
6	EZH2	14,15	-	g.chr7:14850872T>G	c.1922A>G	p.Y641S	Missense_Mutation	51	355	14	70	243	436	56	62	Shared	NP_004447	deleterious	Yes
6	FAM129A	14	-	g.chr1:184764994G>A	c.1904C>T	p.A635V	Missense_Mutation	8	105	8	40	36	99	36	40	Shared	NP_443198	neutral	Yes
6	FAM189A2	5	+	g.chr9:71992569C>T	c.403C>T	p.R135C	Missense_Mutation	19	197	10	50	67	110	61	68	Shared	NP_001121080	deleterious	nd
6	GCC2	19	+	g.chr2:109109241C>G	c.4442C>G	p.T1481R	Missense_Mutation	0	207	0	0	61	263	23	26	tFL specific	NP_852118	deleterious	nd
6	IGSF1	12	-	g.chrX:130411881G>C	c.2269C>G	p.R757G	Missense_Mutation	0	82	0	0	40	49	82	91	tFL specific	NP_001546	neutral	nd
6	KATNAL1	2	-	g.chr13:30854364delA			Splice_Site	0	58	0	0	35	85	41	45	tFL specific	NP_001014402	neutral	nd
6	LMTK3	12	-	g.chr19:49001165C>G	c.3248G>C	p.R1083T	Missense_Mutation	25	237	11	55	66	174	38	42	Shared	NP_001073903		Yes
6	LY75	2	-	g.chr2:160755249A>G	c.416T>C	p.V139A	Missense_Mutation	14	237	6	30	120	250	48	53	Shared	NP_002340	deleterious	nd
6	MBOAT7	5	-	g.chr19:54684833G>A	c.511C>T	p.R171C	Missense_Mutation	1	21	5	26	15	40	38	43	Shared	NP_077274	deleterious	Yes
6	MCOLN1	9	+	g.chr19:7593809G>A	c.1087G>A	p.V363M	Missense_Mutation	0	129	0	0	36	96	38	43	tFL specific	NP_065394	neutral	nd
6	MLKL	6	-	g.chr16:74712849T>C	c.986A>G	p.H329R	Missense_Mutation	0	109	0	0	24	71	34	38	tFL specific	NP_689862	deleterious	nd
6	MLL2	44	-	g.chr12:49422905C>T	c.14190G>A	p.W4730*	Nonsense_Mutation	20	147	14	70	61	165	37	41	Shared	NP_003473		Yes
6	MLL2	32	-	g.chr12:49433247G>A	c.8200C>T	p.R2734*	Nonsense_Mutation	18	114	16	80	42	89	47	52	Shared	NP_003473		Yes
6	NME7	7,8	-	g.chr1:169206928G>A			Splice_Site	0	243	0	0	72	220	33	37	tFL specific	NP_037462		nd
6	NR1P1	1	-	g.chr21:16337866T>C	c.2648A>G	p.N883S	Missense_Mutation	0	416	0	0	90	393	23	26	tFL specific	NP_003480	neutral	nd
6	P4HB	1	-	g.chr17:79818257T>C	c.91A>G	p.K31E	Missense_Mutation	7	64	11	55	38	63	60	67	Shared	NP_000909	neutral	nd
6	PCLO	5	-	g.chr7:82583626C>G	c.6643G>C	p.D2215H	Missense_Mutation	35	259	14	70	81	243	33	37	Shared	NP_149015	none	nd
6	PKP4	15	+	g.chr2:159523021G>A	c.2674G>A	p.V892M	Missense_Mutation	0	73	0	0	27	77	35	39	tFL specific	NP_003619		nd
6	PLCZ1	11	-	g.chr12:18847846T>C	c.1459A>G	p.R487G	Missense_Mutation	0	56	0	0	12	28	43	48	tFL specific	NP_149114	deleterious	nd
6	PNP	2	+	g.chr14:20940517C>G	c.62C>G	p.T21S	Missense_Mutation	0	204	0	0	74	224	33	37	tFL specific	NP_000261	neutral	nd
6	POM121L12	1	+	g.chr7:53104171G>T	c.807G>T	p.E269D	Missense_Mutation	0	236	0	0	74	283	26	29	tFL specific	NP_872401		nd
6	PTF1A	1	+	g.chr10:23481695_23481696delTA	c.236_237delTA	p.L79fs	Frame_Shift_Del	2	57	4	21	13	68	19	21	Shared	NP_835455		nd
6	PTPRB	10,12	-	g.chr12:70965679C>A	c.2377G>T	p.G793*	Nonsense_Mutation	0	417	0	0	136	348	39	43	tFL specific	NP_002828		nd
6	RIMS2	14	+	g.chr8:105001612C>T	c.2341C>T	p.R781C	Missense_Mutation	4	67	6	30	59	68	87	96	Shared	NP_055492	deleterious	nd
6	SCAI	3,4	-	g.chr9:127818245G>A	c.140C>T	p.A47V	Missense_Mutation	0	196	0	0	64	191	34	38	tFL specific	NP_001138349	neutral	nd
6	SLCO3A1	6	-	g.chr15:92669413G>A	c.1297G>A	p.A433T	Missense_Mutation	0	260	0	0	75	120	62	69	tFL specific	NP_037404	deleterious	nd
6	SORCS1	23	-	g.chr10:108366932G>T	c.3157C>A	p.L1053M	Missense_Mutation	0	91	0	0	51	115	44	49	tFL specific	NP_443150	neutral	nd
6	SOX6	3	-	g.chr11:16256170delT	c.532delA	p.M178fs	Frame_Shift_Del	0	153	0	0	29	178	16	18	tFL specific (subclonal)	NP_001139291		nd
6	STAT6	11	-	g.chr12:57496661T>C	c.1256A>G	p.D419G	Missense_Mutation	19	271	7	35	121	306	40	44	Shared	NP_003144	neutral	nd
6	TCTN1	13	+	g.chr12:111085103A>C	c.1582A>C	p.N528H	Missense_Mutation	0	149	0	0	55	127	43	48	tFL specific	NP_001076006	deleterious	nd
6	TDRD7	16	+	g.chr9:100258071T>A	c.3203T>A	p.V1068E	Missense_Mutation	0	165	0	0	41	248	17	19	tFL specific (subclonal)	NP_055105	deleterious	nd
6	TIRAP	2	+	g.chr11:126162978A>T	c.674T>A	p.E225V	Missense_Mutation	0	196	0	0	46	81	57	63	tFL specific	NP_001034750		nd
6	TNFRSF14	4	+	g.chr1:2491310delT	c.353delT	p.V118fs	Frame_Shift_Del	20	123	16	80	119	149	80	89	Shared	NP_003811		Yes
6	TUBD1	8	-	g.chr17:57937757C>A	c.1288G>T	p.G430*	Nonsense_Mutation	0	32	0	0	16	34	47	53	tFL specific	NP_057345		nd
6	ZBTB11	11	-	g.chr3:101370487C>T	c.2685G>A	p.W895*	Nonsense_Mutation	0	54	0	0	9	31	29	32	tFL specific	NP_055230		nd
6	ZNF521	3	-	g.chr18:22806240C>T	c.1642G>A	p.V548M	Missense_Mutation	39	165	24	100	107	121	88	98	Shared	NP_056276	neutral	nd
6	ZNF544	4	+	g.chr19:58772937G>A	c.965G>A	p.G322E	Missense_Mutation	24	235	10	50	79	216	37	41	Shared	NP_055295	neutral	nd
6	ZNF563	4	-	g.chr19:12429698G>A	c.1141C>T	p.R381*	Nonsense_Mutation	0	95	0	0	10	56	18	20	tFL specific (subclonal)	NP_660319		nd
8	ACOT4	1	+	g.chr14:74059001A>C	c.338A>C	p.D113A	Missense_Mutation	2	29	7	19	10	56	18	26	Shared (subclonal)	NP_689544	neutral	nd
8	ANKS1A	15	+	g.chr6:35047347G>C	c.2338G>C	p.D780H	Missense_Mutation	0	84	0	0	21	112	19	27	tFL specific	NP_056060	neutral	Yes
8	ARPP21	16	+	g.chr3:35785389G>A	c.1964G>A	p.R655Q	Missense_Mutation	0	128	0	0	33	143	23	33	tFL specific	NP_057384	deleterious	nd
8	CYP2S1	4	+	g.chr19:41704523G>T	c.650G>T	p.G217V	Missense_Mutation	0	311	0	0	82	463	18	26	tFL specific	NP_085125	neutral	Yes
8	DMD	13,14,21,24,25,75	-	g.chrX:31165405T>A	c.10784A>T	p.Q3595L	Missense_Mutation	0	170	0	0	66	239	28	40	tFL specific	NP_003997	deleterious	Yes
8	FHOD3	24	+	g.chr18:34349311G>A	c.4108G>A	p.V1370I	Missense_Mutation	0	149	0	0	41	196	21	30	tFL specific	NP_079411	neutral	Yes
8	FOXN1	2	+	g.chr17:26851587delC	c.190delC	p.P64fs	Frame_Shift_Del	0	268	0	0	100	389	26	37	tFL specific	NP_003584		Yes
8	HEATR3	4	+	g.chr16:50104185G>A	c.496G>A	p.V166M	Missense_Mutation	2	58	3	8	28	87	32	45	tFL specific	NP_891552	neutral	nd
8	IL10RA	3	+	g.chr11:117860335G>A	c.367G>A	p.V123M	Missense_Mutation	0	40	0	0	7	41	17	26	tFL specific	NP_001549	deleterious	Yes
8	IRF4	2	+	g.chr6:394888C>A	c.284C>A	p.T95K	Missense_Mutation	2	166	1	0	79	224	35	50	tFL specific	NP_002451	deleterious	Yes
8	KBTBD5	1	+	g.chr3:42727568G>A	c.458G>A	p.R153H	Missense_Mutation	0	458	0	0	194	593	33	47	tFL specific	NP_689606	deleterious	Yes
8	LINGO4	1	-	g.chr1:151774608G>T	c.573C>A	p.N191K	Missense_Mutation	0	104	0	0	31	166	19	27	tFL specific	NP_001004432	deleterious	Yes
8	LMO7	27	+	g.chr13:76427376G>T	c.3814G>T	p.A1272S	Missense_Mutation	0	248	0	0	81	299	27	39	tFL specific	NP_056667	deleterious	nd
8	MLL2	10	-	g.chr12:49445934G>C	c.1532C>G	p.P511R	Missense_Mutation	61	107	57	100	69	129	53	76	Shared	NP_003473	none	Yes
8	NEBL	12	-	g.chr10:21134220T>G	c.1194A>C	p.L398F	Missense_Mutation	0	112	0	0	16	100	16	23	tFL specific	NP_006384	deleterious	nd
8	NGF	1	-	g.chr1:115829136G>A	c.281C>T	p.P94L	Missense_Mutation	0	119	0	0	36	153	24	34	tFL specific	NP_002497	deleterious	Yes
8	NRAP	34,35	-	g.chr10:115364403C>T	c.4192G>A	p.A1398T	Missense_Mutation	0	52	0	0	17	81	21	30	tFL specific	NP_932326	neutral	Yes
8	NTRK1	15,16	+	g.chr1:156849951T>C	c.2205_splice	p.E735_splice	Splice_Site	0	65	0	0	19	105	18	26	tFL specific	NP_002520		Yes
8	PTPRK	8	-	g.chr6:128411074T>G	c.1226A>C	p.D409A	Missense_Mutation	0	128	0	0	29	130	22	32	tFL specific	NP_002835	deleterious	Yes
8	PTPRK	8	-	g.chr6:128411075C>T	c.1225G>A	p.D409N	Missense_Mutation	0	128	0	0	30	128	23	33	tFL specific	NP_002835	deleterious	nd
8	RC3H1	4	-	g.chr1:173951999C>T	c.634G>A	p.E212K	Missense_Mutation	0	321	0	0	74	401	18	26	tFL specific	NP_742068	deleterious	Yes
8	SLITRK3	1	-	g.chr3:164906338C>A	c.2281G>T	p.E761*	Nonsense_Mutation	0	153	0	0	37	236	16	23	tFL specific	NP_055741		nd
8	SPIB	6	+	g.chr19:50931489C>T	c.685C>T	p.R229C	Missense_Mutation	0	66	0	0	63	139	45	64	tFL specific	NP_003112	deleterious	Yes
8	SULF1	13	+	g.chr8:70540007A>T	c.1948A>T	p.I650F	Missense_Mutation	0	120	0	0	25	92	27	39	tFL specific	NP_055985	deleterious	Yes
8	SYCP2	32	-	g.chr20:58450428A>C	c.3247T>G	p.W1083G	Missense_Mutation	0	73	0	0	21	80	26	38	tFL specific	NP_055073	deleterious	Yes
8	SYNE1	138,139	-	g.chr6:152461107T>G	c.25436A>C	p.E8479A	Missense_Mutation	0	180	0	0	33	162	20	29	tFL specific	NP_892006	neutral	Yes
8	TBCD	35	+	g.chr17:80895232G>A	c.3277G>A	p.A1093T	Missense_Mutation	0	129	0	0	36	177	20	29	tFL specific	NP_005984	neutral	Yes
8	TDRD5	11	+	g.chr1:179620154G>C	c.1953G>C	p.L651F	Missense_Mutation	0	69	0	0	21	75	28	40	tFL specific	NP_775804	deleterious	Yes
8	ZRANB2	2	-	g.chr1:71544337A>C	c.109_splice	p.E37_splice	Splice_Site	1	213	0	0	90	281	32	46	tFL specific	NP_976225		Yes
11	ACTB	1	-	g.chr7:5569256G>C	c.33C>G	p.D11E	Missense_Mutation	44	155	28	31	0	141	0	0	FL specific	NP_001092	deleterious	nd
11	ADAMTS5	1	-	g.chr21:28338172C>T	c.539G>A	p.R180H	Missense_Mutation	84	165	51	57	0	186	0	0	FL specific	NP_008969	deleterious	nd
11	ADAMTSL4	3	+	g.chr1:150525															

11	C4orf41	10	+	g.chr4:184603956A>T	c.1185A>T	p.R395S	Missense_Mutation	88	252	35	39	0	239	0	0	FL specific	NP_068761	deleterious	nd
11	C7orf72	4	+	g.chr7:50173778C>A	c.803C>A	p.A268D	Missense_Mutation	0	78	0	0	25	107	23	26	tFL specific	NP_001155306		nd
11	CADM3	5,6	+	g.chr1:159163783C>T	c.644C>T	p.S215F	Missense_Mutation	73	221	33	37	67	177	38	42	Shared	NP_001120645	deleterious	nd
11	CBL	6	+	g.chr11:119146718G>A	c.881G>A	p.R294Q	Missense_Mutation	38	120	32	35	0	242	0	0	FL specific	NP_005179	deleterious	nd
11	CCDC102A	4	-	g.chr16:57552623C>T	c.952G>A	p.E318K	Missense_Mutation	0	147	0	0	77	158	49	54	tFL specific	NP_149989	deleterious	nd
11	CD22	12	+	g.chr19:35837071A>C	c.2345A>C	p.E782A	Missense_Mutation	0	118	0	0	49	95	52	58	tFL specific	NP_001762	neutral	Yes
11	CDKN1C	1	-	g.chr11:2906224C>G	c.496G>C	p.A166P	Missense_Mutation	0	12	0	0	6	25	24	26	tFL specific	NP_000067	none	nd
11	CFTR	14	+	g.chr7:117232464T>C	c.2243T>C	p.I748T	Missense_Mutation	105	378	28	31	0	459	0	0	FL specific	NP_000483	neutral	nd
11	CFTR	23	+	g.chr7:117282630T>C	c.3856T>C	p.F1286L	Missense_Mutation	36	116	31	34	0	113	0	0	FL specific	NP_000483	deleterious	nd
11	CLTC	20	+	g.chr17:57758663C>A	c.3073C>A	p.Q1025K	Missense_Mutation	151	369	41	46	0	375	0	0	FL specific	NP_004850	neutral	nd
11	CNOT6L	2	-	g.chr4:78697521A>C	c.31T>G	p.Y11D	Missense_Mutation	0	84	0	0	15	36	42	47	tFL specific	NP_653172	deleterious	nd
11	CTXN1	1	-	g.chr19:7990185G>C	c.239C>G	p.A80G	Missense_Mutation	0	57	0	0	20	65	31	35	tFL specific	NP_996664	deleterious	nd
11	DEM1	1	+	g.chr1:40981132G>A	c.916G>A	p.E306K	Missense_Mutation	0	150	0	0	76	172	44	49	tFL specific	NP_073611	neutral	Yes
11	DENND4C	25	+	g.chr9:19361848C>T	c.4556C>T	p.P1519L	Missense_Mutation	1	381	0	0	239	282	85	94	tFL specific	NP_060395	deleterious	nd
11	DPP10	19	+	g.chr2:116548875A>C	c.1643A>C	p.Q548P	Missense_Mutation	126	172	73	81	71	166	43	48	Shared	NP_065919	deleterious	nd
11	DUSP4	1	-	g.chr8:29207666T>C	c.130A>G	p.K44E	Missense_Mutation	182	391	47	52	194	454	43	48	Shared	NP_001385	neutral	nd
11	DYNC1H1	67	+	g.chr14:102508602A>G	c.12252A>G	p.I4084M	Missense_Mutation	38	134	28	31	0	158	0	0	FL specific	NP_001367	deleterious	nd
11	EBF1	3	-	g.chr5:158523356C>A	c.350G>T	p.S117I	Missense_Mutation	0	39	0	0	27	59	46	51	tFL specific	NP_076870	deleterious	nd
11	EBF1	2	-	g.chr5:158524410T>C	c.135_splice	p.S45_splice	Splice_Site	20	57	35	39	0	55	0	0	FL specific	NP_076870		nd
11	ELMO1	13	-	g.chr7:37172780G>C	c.1146C>G	p.D382E	Missense_Mutation	0	27	0	0	12	21	57	64	tFL specific	NP_055615	neutral	nd
11	ENOSF1	11,13,12	-	g.chr18:677782C>T	c.1009G>A	p.E337K	Missense_Mutation	29	75	39	43	20	59	34	37	Shared	NP_059982	deleterious	nd
11	EPHA7	12	-	g.chr6:93967203C>T	c.2149G>A	p.G717R	Missense_Mutation	0	22	0	0	6	21	29	31	tFL specific	NP_004431	deleterious	Yes
11	ETAA1	5	+	g.chr2:67630450_67630455del6bp	c.636_641del6bp	p.212_214NK6>K	In_Frame_Del	0	172	0	0	48	126	38	42	tFL specific	NP_061875		Yes
11	ETAA1	5	+	g.chr2:67630455G>A	c.641G>A	p.R214K	Missense_Mutation	0	174	0	0	22	136	16	18	tFL specific (subclonal)	NP_061875	deleterious	nd
11	EVC	11	+	g.chr4:5758029G>C	c.1503G>C	p.Q501H	Missense_Mutation	0	100	0	0	53	130	41	46	tFL specific	NP_714928	neutral	Yes
11	EXOC2	7	-	g.chr6:599127T>G	c.841A>C	p.K281Q	Missense_Mutation	1	208	0	0	89	271	33	37	tFL specific	NP_060773	deleterious	nd
11	EYS	23	-	g.chr6:65301177G>C	c.4583C>G	p.A1528G	Missense_Mutation	0	112	0	0	54	142	38	42	tFL specific	NP_001136272	none	nd
11	FAM133A	1	+	g.chrX:92964683delA	c.265delA	p.K89fs	Frame_Shift_Del	1	241	0	0	204	252	81	90	tFL specific	NP_775969		nd
11	FAM160A1	5	+	g.chr4:152551014C>T	c.1139C>T	p.P380L	Missense_Mutation	173	439	39	43	196	429	46	51	Shared	NP_001103447	deleterious	nd
11	FAM194B	13	-	g.chr13:46115619A>C	c.2069T>G	p.M690R	Missense_Mutation	0	167	0	0	98	257	38	42	tFL specific	NP_872348	deleterious	nd
11	FAT4	1	+	g.chr4:126238027T>C	c.461T>C	p.V154A	Missense_Mutation	0	204	0	0	85	214	40	44	tFL specific	NP_078858	neutral	nd
11	FBN2	4	-	g.chr5:127863625T>C	c.472A>G	p.T158A	Missense_Mutation	0	237	0	0	94	209	45	50	tFL specific	NP_001990	none	Yes
11	FBXO11	22	-	g.chr2:48035373A>C	c.2416T>G	p.C806G	Missense_Mutation	0	197	0	0	78	250	31	35	tFL specific	NP_079409	deleterious	Yes
11	FGB	8	+	g.chr4:155491660G>A	c.1334G>A	p.R445K	Missense_Mutation	38	209	18	20	0	179	0	0	FL specific (subclonal)	NP_005132	deleterious	nd
11	FOXF2	1	+	g.chr6:1391102A>G	c.920A>G	p.D307G	Missense_Mutation	8	50	16	18	1	46	2	0	FL specific (subclonal)	NP_001443	deleterious	nd
11	FOXO1	1	-	g.chr13:41240126C>G	c.224G>C	p.S75T	Missense_Mutation	0	50	0	0	12	54	22	25	tFL specific	NP_002006	neutral	Yes
11	GATA3	2	+	g.chr10:8100283G>T	c.257G>T	p.R86L	Missense_Mutation	0	355	0	0	186	554	34	38	tFL specific	NP_002042	deleterious	nd
11	GPR137C	1	+	g.chr14:53019881C>T	c.16C>T	p.P6S	Missense_Mutation	9	25	36	40	12	29	41	47	Shared	NP_001093122	none	Yes
11	GPRIN3	1	-	g.chr4:90168981G>A	c.2281C>T	p.R761*	Nonsense_Mutation	42	108	39	43	47	129	36	40	Shared	NP_938022		nd
11	GTSE1	6	+	g.chr22:46712141C>T	c.1264C>T	p.Q422*	Nonsense_Mutation	0	97	0	0	58	117	50	56	tFL specific	NP_057510		nd
11	HAPLN1	3	-	g.chr5:82940420C>A	c.537G>T	p.Q179H	Missense_Mutation	32	68	47	52	11	35	31	36	Shared	NP_001875	deleterious	Yes
11	HIST1H1B	1	-	g.chr6:27834917C>T	c.391G>A	p.A131T	Missense_Mutation	0	346	0	0	130	351	37	41	tFL specific	NP_005313	none	Yes
11	HIST1H1E	1	+	g.chr6:26156744G>C	c.126G>C	p.E42D	Missense_Mutation	1	189	1	0	69	221	31	35	tFL specific	NP_005312	neutral	Yes
11	HIST1H1E	1	+	g.chr6:26157247A>G	c.629A>G	p.K210R	Missense_Mutation	33	89	37	41	0	100	0	0	FL specific	NP_005312	none	Yes
11	HIST1H2BO	1	+	g.chr6:27861568C>T	c.328C>T	p.H110Y	Missense_Mutation	39	93	42	47	37	114	32	36	Shared	NP_003518	deleterious	nd
11	HIST1H31	1	-	g.chr6:27839715G>C	c.379C>G	p.L127V	Missense_Mutation	51	141	36	40	0	157	0	0	FL specific	NP_003524	deleterious	nd
11	HOXD1	1	+	g.chr2:177054044A>G	c.515A>G	p.D172G	Missense_Mutation	0	329	0	0	125	358	35	39	tFL specific	NP_078777	deleterious	nd
11	HRC	1	-	g.chr19:49657498A>T	c.997T>A	p.S333T	Missense_Mutation	140	350	40	44	0	399	0	0	FL specific	NP_002143	neutral	nd
11	IER5	1	+	g.chr1:181058502C>T	c.464C>T	p.A155V	Missense_Mutation	0	32	0	0	22	31	71	79	tFL specific	NP_057629	none	nd
11	IFRD1	11	+	g.chr7:112112860C>T	c.1210C>T	p.P404S	Missense_Mutation	81	262	31	35	0	270	0	0	FL specific	NP_001007246	deleterious	nd
11	IL4I1	2	-	g.chr19:50399269C>T	c.55G>A	p.V19M	Missense_Mutation	42	123	34	38	0	122	0	0	FL specific	NP_690863	none	nd
11	IL4R	6	+	g.chr16:27367183T>A	c.725T>A	p.I242N	Missense_Mutation	0	160	0	0	61	151	40	44	tFL specific	NP_000409	neutral	Yes
11	IQSEC3	1	+	g.chr12:247814C>T	c.376C>T	p.R126W	Missense_Mutation	0	142	0	0	63	147	43	48	tFL specific	NP_056047	none	Yes
11	IRF4	2	+	g.chr6:394899T>C	c.295T>C	p.C99R	Missense_Mutation	0	185	0	0	92	215	43	48	tFL specific	NP_002451	deleterious	nd
11	ITPKB	1	-	g.chr1:226923671C>G	c.1489G>C	p.D497H	Missense_Mutation	0	146	0	0	39	135	29	33	tFL specific	NP_002212	deleterious	nd
11	ITPKB	1	-	g.chr1:226924292C>T	c.868G>A	p.A290T	Missense_Mutation	1	150	1	0	47	159	30	33	tFL specific	NP_002212	neutral	Yes
11	ITPR2	21	-	g.chr12:26807062T>A	c.16C>T		Splice_Site	6	36	17	19	5	33	15	17	Shared (subclonal)	NP_002214		nd
11	IVNS1ABP	1	-	g.chr1:185278558T>C	c.68A>G	p.N23S	Missense_Mutation	163	489	33	37	0	416	0	0	FL specific	NP_006460	deleterious	nd
11	KANK3	4	+	g.chr19:8398962T>G	c.1466A>C	p.D489A	Missense_Mutation	6	34	18	21	0	53	0	0	FL specific	NP_940873	deleterious	nd
11	KCNJ16	1	+	g.chr17:68128488T>C	c.260T>C	p.I87T	Missense_Mutation	0	432	0	0	131	383	34	38	tFL specific	NP_061128	neutral	nd
11	KIAA1429	13	-	g.chr8:95523712T>C	c.3091A>G	p.M1031V	Missense_Mutation	45	101	45	50	38	77	49	55	Shared	NP_056311	neutral	nd
11	KIF11	21	+	g.chr10:94410193T>A	c.2958T>A	p.Y986*	Nonsense_Mutation	0	69	0	0	19	37	51	57	tFL specific	NP_004514		nd
11	KIF12	4	-	g.chr9:116858391C>T	c.421G>A	p.G141R	Missense_Mutation	31	67	46	51	25	62	40	44	Shared	NP_612433	deleterious	Yes
11	KIF4A	27	+	g.chrX:69626912A>G	c.3242A>G	p.K1081R	Missense_Mutation	34	85	40	44	0	87	0	0	FL specific	NP_036442	deleterious	nd
11	KMO	10	+	g.chr1:241731850C>G	c.860C>G	p.S287C	Missense_Mutation	24	53	45	50	0	35	0	0	FL specific	NP_003670	neutral	nd
11	KRT77	1	-	g.chr12:53097041T>A	c.178A>T	p.S60C	Missense_Mutation	0	107	0	0	24	63	38	42	tFL specific	NP_778253	neutral	nd
11	LIPG	3	+	g.chr18:47093866A>T	c.334A>T	p.R112*	Nonsense_Mutation	18	108	17	19	0	80	0	0	FL specific (subclonal)	NP_006024		nd
11	LRRC7	19	+	g.chr1:70504890A>T	c.3269A>T	p.Q1090L	Missense_Mutation	0	135	0	0	49	149	33	<				

11	PIM1	4	+	g.chr6:37138960C>G	c.300C>G	p.F100L	Missense_Mutation	0	199	0	0	70	194	36	40	tFL specific	NP_002639	neutral	nd
11	PIM1	4	+	g.chr6:37139210C>T	c.550C>T	p.L184F	Missense_Mutation	0	90	0	0	24	87	28	31	tFL specific	NP_002639	deleterious	nd
11	PQLC1	1	-	g.chr18:77710731G>A	c.196C>T	p.L66F	Missense_Mutation	0	194	0	0	67	245	27	30	tFL specific	NP_079354	neutral	nd
11	PSMB4	1	+	g.chr1:151372065T>G	c.2T>G	p.M1R	Missense_Mutation	0	48	0	0	21	50	42	47	tFL specific	NP_002787	deleterious	nd
11	RANBP3L	8	-	g.chr5:36260951T>G	c.600A>C	p.E200D	Missense_Mutation	40	96	42	47	38	91	42	47	Shared	NP_659437	deleterious	nd
11	RCC1	8,9	+	g.chr1:28863328G>A	c.1007G>A	p.S336N	Missense_Mutation	0	23	0	0	14	42	33	37	tFL specific	NP_001260	neutral	Yes
11	SGK1	5,7	-	g.chr6:134494491A>G	c.338T>C	p.L113P	Missense_Mutation	1	283	0	0	156	253	62	69	tFL specific	NP_005618	deleterious	Yes
11	SGK1	4,6	-	g.chr6:134494640A>G	c.293T>C	p.P98S	Missense_Mutation	0	103	0	0	16	58	28	31	tFL specific	NP_005618	deleterious	Yes
11	SGK1	4,6	-	g.chr6:134494655G>C	c.278C>G	p.A93G	Missense_Mutation	0	91	0	0	22	63	35	39	tFL specific	NP_005618	deleterious	nd
11	SGK1	4,6	-	g.chr6:134494671A>C	c.262T>G	p.S88A	Missense_Mutation	1	91	1	0	21	62	34	38	tFL specific	NP_005618	neutral	nd
11	SGK1	2,4	-	g.chr6:134495654G>C	c.147C>G	p.C49W	Missense_Mutation	0	76	0	0	13	51	25	28	tFL specific	NP_005618	deleterious	nd
11	SGK1	1	-	g.chr6:134495897T>A	c.49A>T	p.M17L	Missense_Mutation	1	119	1	0	106	162	65	72	tFL specific	NP_005618	deleterious	nd
11	SLC15A1	1	-	g.chr13:99404845G>A			Splice_Site	0	116	0	0	50	160	31	34	tFL specific	NP_005064		nd
11	SLC30A4	1	-	g.chr15:45814456C>T	c.97G>A	p.D33N	Missense_Mutation	124	242	51	57	0	363	0	0	FL specific	NP_037441	deleterious	nd
11	SLC5A5	1	+	g.chr19:17983136C>T	c.8C>T	p.A3V	Missense_Mutation	0	23	0	0	6	37	16	19	tFL specific (subclonal)	NP_000444	neutral	nd
11	SLC6A17	9	+	g.chr1:110738225G>T	c.1510G>T	p.A504S	Missense_Mutation	13	50	26	29	22	68	32	36	Shared	NP_001010898	deleterious	Yes
11	SMAD9	4,5	-	g.chr13:37427768C>T	c.1048G>A	p.V350M	Missense_Mutation	0	29	0	0	20	59	34	37	tFL specific	NP_001120689	neutral	nd
11	SOC31	1	-	g.chr16:11349022T>G	c.314A>C	p.D105A	Missense_Mutation	0	32	0	0	14	30	47	52	tFL specific	NP_003736	deleterious	Yes
11	SORBS2	7,9	-	g.chr4:186545415C>T	c.1156G>A	p.D386N	Missense_Mutation	79	193	41	46	0	222	0	0	FL specific	NP_066547	deleterious	nd
11	SOX1	1	+	g.chr13:112722519C>G	c.547C>G	p.H183D	Missense_Mutation	7	22	32	36	0	12	0	0	FL specific	NP_005977	deleterious	nd
11	SOX5	6,13,14	-	g.chr12:23689561_23689563delTAT	c.1812_1814delATA	p.Y606del	In_Frame_Del	0	115	0	0	26	149	17	19	tFL specific (subclonal)	NP_008871		Yes
11	SSFA2	4	+	g.chr2:182761670G>T	c.343G>T	p.D115Y	Missense_Mutation	21	28	75	82	11	20	55	63	Shared	NP_001123917	deleterious	nd
11	TAGAP	3	-	g.chr6:159464691A>G			Splice_Site	1	109	1	0	51	81	63	70	tFL specific	NP_473455		nd
11	TCTEX1D1	2	+	g.chr1:67236111G>A	c.161G>A	p.R54H	Missense_Mutation	0	232	0	0	110	270	41	46	tFL specific	NP_689878	neutral	nd
11	THBS1	2	+	g.chr15:39874489T>A	c.163T>A	p.S55T	Missense_Mutation	1	209	0	0	92	244	38	42	tFL specific	NP_003237	neutral	nd
11	TIE1	15	+	g.chr1:43782921C>T	c.2461C>T	p.R821W	Missense_Mutation	35	84	42	47	38	97	39	43	Shared	NP_005415	deleterious	Yes
11	TLR2	1	+	g.chr4:154625042T>G	c.983T>G	p.L328R	Missense_Mutation	0	263	0	0	99	242	41	46	tFL specific	NP_003255	deleterious	Yes
11	TM4SF18	1	-	g.chr3:149051090delA	c.80delT	p.L27fs	Frame_Shift_Del	0	313	0	0	138	354	39	43	tFL specific	NP_620141		Yes
11	TNFAIP3	1	+	g.chr6:138192608A>T	c.244A>T	p.K82*	Nonsense_Mutation	31	45	69	77	1	74	1	0	FL specific	NP_006281		Yes
11	TRIM42	2	+	g.chr3:140401494G>A	c.532G>A	p.E178K	Missense_Mutation	1	307	0	0	116	277	42	47	tFL specific	NP_689829	deleterious	Yes
11	TRIML2	7	-	g.chr4:189012827T>A	c.864A>T	p.K288N	Missense_Mutation	58	333	17	19	0	287	0	0	FL specific (subclonal)	NP_775824	deleterious	nd
11	UHRF2	7	+	g.chr9:6481712A>C	c.1230A>C	p.K410N	Missense_Mutation	0	21	0	0	22	38	58	64	tFL specific	NP_690856	deleterious	nd
11	VMP1	10	+	g.chr17:57915726C>T	c.1045C>T	p.L349F	Missense_Mutation	0	136	0	0	34	106	32	36	tFL specific	NP_112200	deleterious	nd
11	VMP1	11	+	g.chr17:57917181G>T	c.1130G>T	p.C377F	Missense_Mutation	0	49	0	0	26	48	54	59	tFL specific	NP_112200	neutral	nd
11	XIRP2	8	+	g.chr2:168107995C>A	c.10093C>A	p.Q3365K	Missense_Mutation	120	253	47	52	94	206	46	51	Shared	NP_689594	neutral	Yes
11	YTHDF2	2	+	g.chr1:29064197A>G			Splice_Site	0	52	0	0	21	72	29	33	tFL specific	NP_057342		Yes
11	ZC3H12D	1	-	g.chr6:149795562G>A	c.118C>T	p.L40F	Missense_Mutation	0	50	0	0	54	71	76	85	tFL specific	NP_997243	deleterious	nd
11	ZC3H18	1	+	g.chr16:88644006G>T	c.475G>T	p.G159C	Missense_Mutation	0	21	0	0	5	21	24	26	tFL specific	NP_653205	none	Yes
11	ZFP64	9	-	g.chr20:50701809G>C			Splice_Site	0	58	0	0	22	58	38	42	tFL specific	NP_955459		nd
11	ZNF226	3	+	g.chr19:44677011A>C	c.143_splice	p.G48_splice	Splice_Site	0	252	0	0	102	202	50	56	tFL specific	NP_001027545		nd
11	ZNF467	4	-	g.chr7:149462893C>T	c.698G>A	p.R233H	Missense_Mutation	0	134	0	0	30	128	23	26	tFL specific	NP_997219	deleterious	nd
11	ZNF676	3	-	g.chr19:22363391C>G	c.1128G>C	p.K376N	Missense_Mutation	0	38	0	0	7	32	22	25	tFL specific	NP_001001411	neutral	nd
12	ARID1A	1	+	g.chr1:27023216_27023222del7bp	c.322_328del7bp	p.K108fs	Frame_Shift_Del	0	47	0	0	15	33	45	56	tFL specific	NP_006006		Yes
12	EP300	10	+	g.chr22:41537176C>T	c.2003C>T	p.S668F	Missense_Mutation	2	84	2	0	21	41	51	63	tFL specific	NP_001420	deleterious	Yes
12	EPB41L4A	14	-	g.chr5:111541128C>T	c.1252G>A	p.G418S	Missense_Mutation	4	166	2	0	55	110	50	62	tFL specific	NP_071423	neutral	nd
12	GFRAL	4	+	g.chr6:55214931C>T	c.358C>T	p.R120C	Missense_Mutation	0	68	0	0	13	30	43	55	tFL specific	NP_997293	neutral	nd
12	MLL2	39	-	g.chr12:49425091_49425113del23bp	c.13375_13397del23bp	p.G4459fs	Frame_Shift_Del	3	257	1	0	52	153	34	42	tFL specific	NP_003473		Yes
12	MYC	2	+	g.chr8:128750512C>T	c.4C>T	p.P2S	Missense_Mutation	0	128	0	0	16	82	20	25	tFL specific	NP_002458	deleterious	Yes
12	MYC	2	+	g.chr8:128750526C>G	c.18C>G	p.S6R	Missense_Mutation	0	130	0	0	17	81	21	27	tFL specific	NP_002458	neutral	Yes
12	MYC	2	+	g.chr8:128750543A>C	c.35A>C	p.Y12S	Missense_Mutation	0	111	0	0	16	81	20	25	tFL specific	NP_002458	deleterious	Yes
12	MYC	2	+	g.chr8:128750577C>G	c.69C>G	p.F23L	Missense_Mutation	1	105	1	0	14	67	21	27	tFL specific	NP_002458	deleterious	Yes
12	MYC	2	+	g.chr8:128750810T>A	c.302T>A	p.M101K	Missense_Mutation	2	105	2	0	30	99	30	37	tFL specific	NP_002458	deleterious	Yes
12	NOX4	11,12	-	g.chr11:89106626T>G	c.1109A>C	p.H370P	Missense_Mutation	1	49	2	0	13	18	72	89	tFL specific	NP_058627	deleterious	nd
12	OTOF	14	-	g.chr2:26705333T>C	c.1520A>G	p.D507G	Missense_Mutation	2	127	2	0	33	65	51	64	tFL specific	NP_919224	deleterious	nd
12	RAPGEF1	11	-	g.chr9:134497269A>T	c.1768T>A	p.Y590N	Missense_Mutation	0	30	0	0	7	17	41	50	tFL specific	NP_005303	deleterious	nd
12	TBC1D8B	6	+	g.chrX:106069318G>T	c.886G>T	p.G296*	Nonsense_Mutation	0	53	0	0	15	35	43	54	tFL specific	NP_060222		nd
12	TRIM36	10	-	g.chr5:114462531T>C	c.1856A>G	p.D619G	Missense_Mutation	3	207	1	0	49	108	45	56	tFL specific	NP_061170	deleterious	nd
12	MYC	2	+	g.chr8:128751015C>A	c.507C>A	p.S169R	Missense_Mutation	8	312	3	4	104	272	38	48	tFL specific	NP_002458	neutral	Yes
12	KLHL6	3	-	g.chr3:183226202_183226203insA	c.553_554insT	p.S185fs	Frame_Shift_Ins	15	444	3	4	154	295	52	65	tFL specific	NP_569713		Yes
12	AFAP1L2	14	-	g.chr10:116060439G>A	c.1553C>T	p.T518I	Missense_Mutation	2	61	3	4	22	46	48	60	tFL specific	NP_001001936	neutral	nd
12	SLC26A2	2	+	g.chr5:149361256C>A	c.2100C>A	p.C700*	Nonsense_Mutation	6	159	4	5	43	89	48	60	tFL specific	NP_000103		nd
12	LRP6	6	-	g.chr12:12334248C>T	c.1102G>A	p.D368N	Missense_Mutation	16	390	4	5	118	260	45	56	tFL specific	NP_002327	deleterious	nd
12	MYC	2	+	g.chr8:128750886A>C	c.378A>C	p.K126N	Missense_Mutation	9	220	4	5	70	205	34	43	tFL specific	NP_002458	deleterious	Yes
12	PLEKHG5	15,17	-	g.chr1:6530328T>C	c.1936A>G	p.S646G	Missense_Mutation	7	183	4	5	67	125	54	68	tFL specific	NP_001036128	deleterious	nd
12	SH3PXD2B	13	-	g.chr5:171766536G>A	c.1573C>T	p.R525W	Missense_Mutation	8	208	4	6	63	132	48	60	tFL specific	NP_001017995	deleterious	nd
12	HSP90AB1	8	+	g.chr6:44219478G>A	c.1319G>A	p.G440E	Missense_Mutation	9	252	4	6	72	153	47	59	tFL specific	NP_031381	deleterious	nd
12	MYC	2	+	g.chr8:128750653A>G	c.145A>G	p.I49V	Missense_Mutation	2	54	4	6	10	45	22	28	tFL specific	NP_002458	deleterious	nd
12	CCND3	2,4,5	-	g.chr6:41903731G>A	c.826C>T	p.Q276*	Nonsense_Mutation	3	72	4	6	32	66	48	60	tFL specific	NP_001751		Yes
12	MYC	2	+	g.chr8:128750713C>T	c.205C>T	p.L69F	Missense_Mutation	2	52	4	6	18	49	37	46	tFL specific	NP_002458	neutral	Yes
12	TP53	5	-	g.chr17:7578274T>C	c.575A>G	p.Q192R	Missense_Mutation	4	93	4	6	25	67	37	46	tFL specific	NP_001119584	deleterious	Yes
12	MYC	3	+	g.chr8:128752698G>C	c.859G>C	p.A287P	Missense_Mutation	2	39	5	6	20	54	37	46	tFL specific	NP_002458	neutral	nd
12	STAT5B																		

12	BRD8	5	-	g.chr5:137506818C>T	c.268G>A	p.E90K	Missense_Mutation	29	269	11	15	75	166	45	56	tFL specific	NP_631938	deleterious	nd
12	STAT6	11	-	g.chr12:57496661T>C	c.1256A>G	p.D419G	Missense_Mutation	59	208	28	37	66	138	48	60	Shared	NP_003144	neutral	nd
17	AMELX	3,4,5	+	g.chrX:11316732C>A	c.209C>A	p.S70Y	Missense_Mutation	94	243	39	71	0	277	0	0	FL specific	NP_001133	deleterious	nd
17	ANKIB1	19	+	g.chr7:92028136C>T	c.3143C>T	p.A1048V	Missense_Mutation	13	60	22	40	0	56	0	0	FL specific	NP_061877	neutral	nd
17	ARHGEF3	6,8	-	g.chr3:56785161C>A	c.536_splice	p.E179_splice	Splice_Site	14	68	21	38	0	60	0	0	FL specific	NP_062455	deleterious	nd
17	ARRDC1	8	+	g.chr9:140509562T>C	c.1274T>C	p.V425A	Missense_Mutation	12	70	17	31	0	83	0	0	FL specific	NP_689498	neutral	nd
17	ATP9A	15	-	g.chr20:50256007T>G	c.1543A>C	p.T515P	Missense_Mutation	8	33	24	44	0	46	0	0	FL specific	NP_006036	neutral	nd
17	BCL2	1	-	g.chr18:60985849C>G	c.51G>C	p.K17N	Missense_Mutation	48	215	22	40	0	258	0	0	FL specific	NP_000624	neutral	Yes
17	C14orf49	8	-	g.chr14:95912314C>A	c.1564G>T	p.V522L	Missense_Mutation	0	36	0	0	10	43	23	42	tFL specific	NP_689805	neutral	nd
17	C1orf85	1	-	g.chr1:156265436T>G	c.1A>C	p.M1L	Missense_Mutation	0	43	0	0	17	58	29	53	tFL specific	NP_653181	deleterious	nd
17	C5orf38	1	+	g.chr5:2752463G>C	c.85G>C	p.A29P	Missense_Mutation	0	119	0	0	24	110	22	40	tFL specific	NP_848664	none	nd
17	C5orf38	1	+	g.chr5:2752467G>T	c.89G>T	p.R30L	Missense_Mutation	0	120	0	0	27	118	23	42	tFL specific	NP_848664	none	nd
17	CACNA1H	33,34	+	g.chr16:1270496delC	c.6564delC	p.S2188fs	Frame_Shift_Del	0	51	0	0	14	52	27	49	tFL specific	NP_066921	nd	nd
17	CDC42BPA	18,19	-	g.chr1:227261695G>A	c.2605C>T	p.R869C	Missense_Mutation	0	34	0	0	8	30	27	48	tFL specific	NP_003598	deleterious	nd
17	CHRM3	1	+	g.chr1:240072501C>T	c.1750C>T	p.R584C	Missense_Mutation	0	87	0	0	19	111	17	31	tFL specific	NP_000731	deleterious	nd
17	CREBBP	26,27	-	g.chr16:3786763A>T	c.4448T>A	p.I1483N	Missense_Mutation	31	51	61	100	40	70	57	100	Shared	NP_004371	none	Yes
17	CXorf22	9	+	g.chrX:35984693T>G	c.1422T>G	p.C474W	Missense_Mutation	20	71	28	51	0	116	0	0	FL specific	NP_689845	deleterious	nd
17	CXorf64	2	+	g.chrX:125955252G>A	c.631G>A	p.V211I	Missense_Mutation	18	88	20	36	0	154	0	0	FL specific	NP_001116188	neutral	nd
17	DSG4	12	+	g.chr18:28986058C>T	c.1655C>T	p.T552M	Missense_Mutation	17	51	33	60	0	77	0	0	FL specific	NP_817123	neutral	Yes
17	E1F4A3	7	-	g.chr17:78112839T>C	c.709A>G	p.I237V	Missense_Mutation	22	50	44	79	0	54	0	0	FL specific	NP_055555	neutral	Yes
17	EML6	14	+	g.chr2:55096338delG	c.2072delG	p.R691fs	Frame_Shift_Del	0	42	0	0	15	72	21	38	tFL specific	NP_001034842	nd	nd
17	FO XK2	1	+	g.chr17:80477815_80477826del12bp	c.51_62del12bp	p.17_21GGGAG>G	In_Frame_Del	0	17	0	0	8	19	43	79	tFL specific	NP_004505	nd	nd
17	GALNTL6	1	+	g.chr4:172735864C>A	c.133C>A	p.Q45K	Missense_Mutation	0	72	0	0	16	79	20	37	tFL specific	NP_001030017	neutral	nd
17	GLB1L	3	-	g.chr2:220107600A>G	c.280T>C	p.Y94H	Missense_Mutation	32	132	24	44	0	131	0	0	FL specific	NP_078782	deleterious	nd
17	GNAI2	5	+	g.chr3:50293622A>G	c.465_splice	p.Y155_splice	Splice_Site	9	32	28	33	3	50	6	8	FL specific (tFL subclonal)	NP_002061	nd	nd
17	GNAI2	5	+	g.chr3:50293701A>T	c.542A>T	p.K181M	Missense_Mutation	34	83	41	49	9	89	10	13	FL specific (tFL subclonal)	NP_002061	deleterious	nd
17	GPR1	1	-	g.chr2:207041089G>A	c.883C>T	p.L295F	Missense_Mutation	21	79	27	49	0	98	0	0	FL specific	NP_005270	neutral	nd
17	GPS2	2	-	g.chr17:7217803_7217805delTTA	c.204_splice	p.Q68_splice	Splice_Site	38	94	40	72	0	157	0	0	FL specific	NP_004480	nd	nd
17	GRK5	13	+	g.chr10:121207741C>A	c.1373C>A	p.A458D	Missense_Mutation	0	77	0	0	13	76	17	31	tFL specific	NP_005299	neutral	nd
17	HNRNPUI	1	-	g.chr1:245027580_245027592del13bp	c.18_30del13bp	p.V6fs	Frame_Shift_Del	14	67	21	38	0	105	0	0	FL specific	NP_114032	nd	nd
17	HTR4	6	-	g.chr5:147863841G>A	c.1176C>T	p.A393V	Missense_Mutation	14	70	20	36	0	68	0	0	FL specific	NP_000861	nd	nd
17	IRF4	1	+	g.chr6:393222C>T	c.70C>T	p.L24F	Missense_Mutation	0	167	0	0	65	189	34	62	tFL specific	NP_002451	deleterious	nd
17	IRX4	5	-	g.chr5:1878451G>T	c.1192C>A	p.Q398K	Missense_Mutation	18	87	21	39	0	93	0	0	FL specific	NP_057442	deleterious	nd
17	ISM1	6	+	g.chr20:13279768G>A	c.1057G>A	p.G353R	Missense_Mutation	55	177	31	56	0	190	0	0	FL specific	NP_543016	deleterious	nd
17	JAKMIP2	1	-	g.chr5:147051337C>G	c.33G>C	p.K11N	Missense_Mutation	15	86	17	31	0	90	0	0	FL specific	NP_055605	deleterious	nd
17	KCNE4	1	+	g.chr2:223918006C>T	c.458C>T	p.T153M	Missense_Mutation	20	68	29	53	1	57	2	0	FL specific	NP_542402	deleterious	nd
17	KIAA0586	15	+	g.chr14:58934513A>C	c.2087A>C	p.H696P	Missense_Mutation	13	47	28	51	0	49	0	0	FL specific	NP_055564	nd	nd
17	KIAA1370	5	-	g.chr15:52902156C>G	c.955G>C	p.G319R	Missense_Mutation	39	139	28	51	0	162	0	0	FL specific	NP_062546	deleterious	nd
17	KIF26A	12	+	g.chr14:104644133A>G	c.5008A>G	p.T1670A	Missense_Mutation	6	30	20	37	0	43	0	0	FL specific	NP_056471	deleterious	nd
17	KRII	4	-	g.chr19:10673464C>A	c.342G>T	p.K114N	Missense_Mutation	54	221	24	44	0	272	0	0	FL specific	NP_075384	neutral	nd
17	KRT10	7	-	g.chr17:38975051G>A	c.1736C>T	p.S579F	Missense_Mutation	15	46	33	60	0	55	0	0	FL specific	NP_000412	none	nd
17	MAP2K1	3	+	g.chr15:66729153T>A	c.361T>A	p.C121S	Missense_Mutation	0	61	0	0	10	63	16	29	tFL specific	NP_002746	deleterious	nd
17	MLL2	39	-	g.chr12:49426351C>T	c.12137G>A	p.G4046E	Missense_Mutation	32	56	57	100	42	82	51	93	Shared	NP_003473	nd	Yes
17	MYH2	1	-	g.chr17:10451033C>T	c.204_splice	p.A68_splice	Splice_Site	48	293	16	29	0	353	0	0	FL specific	NP_001093582	nd	nd
17	MYH7	25	-	g.chr14:23889433T>C	c.3347A>G	p.E1116G	Missense_Mutation	0	160	0	0	30	159	19	35	tFL specific	NP_000248	deleterious	nd
17	MYO5C	18	-	g.chr15:52537567A>G	c.2162T>C	p.L721S	Missense_Mutation	13	44	30	55	0	56	0	0	FL specific	NP_061198	deleterious	nd
17	NEUROD4	1	+	g.chr12:55421024A>C	c.801A>C	p.Q267H	Missense_Mutation	57	219	26	47	0	297	0	0	FL specific	NP_067014	deleterious	nd
17	NFKBIA	2	-	g.chr14:35872966A>G	c.266T>C	p.L89P	Missense_Mutation	16	29	55	64	16	33	48	65	Shared	NP_065390	deleterious	nd
17	NOTCH1	34	-	g.chr9:139390617G>A	c.7574C>T	p.P2525L	Missense_Mutation	40	119	34	40	58	143	41	55	Shared	NP_060087	deleterious	nd
17	NPAT	10	-	g.chr11:10805599G>A	c.866C>T	p.T289M	Missense_Mutation	116	237	49	89	0	249	0	0	FL specific	NP_002510	neutral	Yes
17	NUMA1	13	-	g.chr11:71727183G>C	c.1366C>G	p.H456D	Missense_Mutation	0	49	0	0	8	46	17	30	tFL specific	NP_006176	deleterious	Yes
17	ODF3L1	4	+	g.chr15:76019546C>T	c.490C>T	p.P164S	Missense_Mutation	35	102	34	62	0	134	0	0	FL specific	NP_787077	deleterious	nd
17	PCDH8	1	-	g.chr13:53420819C>T	c.1753G>A	p.G585S	Missense_Mutation	0	64	0	0	21	77	27	49	tFL specific	NP_002581	deleterious	nd
17	PER2	19	-	g.chr2:239160261C>T	c.3253G>A	p.D1085N	Missense_Mutation	0	32	0	0	9	46	20	36	tFL specific	NP_073728	deleterious	nd
17	PFAS	16	+	g.chr17:8167890G>A	c.2020G>A	p.V674M	Missense_Mutation	0	67	0	0	12	77	16	29	tFL specific	NP_036525	deleterious	nd
17	PKLR	6	-	g.chr1:155264430G>A	c.808C>T	p.R270*	Nonsense_Mutation	15	62	24	44	0	81	0	0	FL specific	NP_000289	nd	nd
17	RAPGEF1	23	-	g.chr9:134455710T>C	c.3023A>G	p.N1008S	Missense_Mutation	14	65	22	41	0	75	0	0	FL specific	NP_005303	deleterious	nd
17	RG57	2	-	g.chr1:241261969A>G	c.172T>C	p.S58P	Missense_Mutation	27	103	26	47	0	117	0	0	FL specific	NP_002915	deleterious	nd
17	SLC14A2	18	+	g.chr18:43258938A>T	c.2513_splice	p.A838_splice	Splice_Site	0	47	0	0	9	57	16	29	tFL specific	NP_009094	nd	nd
17	SLC20A1	7	+	g.chr2:113417241G>A	c.1509G>A	p.W503*	Nonsense_Mutation	0	123	0	0	29	161	18	33	tFL specific	NP_005406	nd	Yes
17	SMOC1	1	+	g.chr14:70346459T>C	c.64T>C	p.S22P	Missense_Mutation	5	17	29	53	0	24	0	0	FL specific	NP_071420	none	nd
17	SNTG2	9	+	g.chr2:1204829G>A	c.632G>A	p.R211K	Missense_Mutation	0	30	0	0	14	68	21	38	tFL specific	NP_061841	neutral	nd
17	SNX10	5	+	g.chr7:26411458T>G	c.329T>G	p.L110R	Missense_Mutation	6	33	18	32	0	38	0	0	FL specific	NP_037454	deleterious	nd
17	TECTA	18	+	g.chr11:121038769G>T	c.5593G>T	p.G1865C	Missense_Mutation	19	89	21	39	0	86	0	0	FL specific	NP_005413	deleterious	nd
17	THUMP2D	3	-	g.chr2:39997134C>T	c.388G>A	p.D130N	Missense_Mutation	0	90	0	0	21	129	16	29	tFL specific	NP_079540	neutral	nd
17	TLR7	2	+	g.chrX:12906053C>A	c.2426C>A	p.A809D	Missense_Mutation	32	155	21	38	0	161	0	0	FL specific	NP_057646	deleterious	nd
17	TMEM134	6,7	-	g.chr11:67232169T>C	c.12137G>A	p.G4046E	Missense_Mutation	22	131	17	31	0	107	0	0	FL specific	NP_079400	nd	nd
17	TNFRSF14	3	+	g.chr1:2489823_2489824delGT	c.220_221delGT	p.V74fs	Frame_Shift_Del	0	14	0	0	15	22	68	100	tFL specific	NP_003811	nd	Yes
17	TP53	5	-	g.chr17:7578211C>A	c.638G>T	p.R213L	Missense_Mutation	13	31	42	76	0	53	0	0	FL specific	NP_001119584	deleterious	Yes
17	TRPV4	1	-	g.chr12:110252313G>C	c.289C>G	p.P97A	Missense_Mutation	22	59	37	68	0	51	0	0	FL specific	NP_067638	deleterious	nd
17	U2SURP	23	+	g.chr3:142757747T>A	c.2329T>A	p.S777T	Missense_Mutation	0	27	0	0	6	27	22	41				

23	ARAP3	9	-	g.chr5:141051821A>C	c.1433T>G	p.L478R	Missense_Mutation	0	110	0	0	33	96	34	46	tFL specific	NP_071926	deleterious	nd
23	ARHGEF10	25	+	g.chr8:1882059C>T	c.3173C>T	p.A1058V	Missense_Mutation	0	50	0	0	17	41	41	55	tFL specific	NP_055444	neutral	nd
23	ASAP2	14	+	g.chr2:9496359T>G	c.1212T>G	p.F404L	Missense_Mutation	0	90	0	0	35	93	38	51	tFL specific	NP_003878	deleterious	nd
23	ASH1L	4	+	g.chr1:155408508T>C	c.5438A>G	p.Y1813C	Missense_Mutation	1	111	1	0	47	172	27	36	tFL specific	NP_060959	deleterious	nd
23	ATP6AP1	9	+	g.chrX:153663735G>A	c.1087G>A	p.G363R	Missense_Mutation	86	113	76	89	0	242	0	0	FL specific	NP_001174	deleterious	nd
23	B2M	1	+	g.chr15:45003808C>T	c.64C>T	p.Q22*	Nonsense_Mutation	0	19	0	0	9	13	69	88	tFL specific	NP_004039		Yes
23	BAG5	1,2	-	g.chr14:104027396C>T	c.106G>A	p.D36N	Missense_Mutation	0	223	0	0	90	246	37	49	tFL specific	NP_004864	neutral	nd
23	BCL2	1	-	g.chr18:60985279G>A			Splice_Site	0	116	0	0	48	125	38	51	tFL specific	NP_000624		nd
23	BCL2	1	-	g.chr18:60985292C>T	c.608G>A	p.S203N	Missense_Mutation	0	119	0	0	47	127	37	49	tFL specific	NP_000624		nd
23	BCL2	1	-	g.chr18:60985566A>G	c.334T>C	p.F112L	Missense_Mutation	0	180	0	0	79	214	37	50	tFL specific	NP_000624	deleterious	Yes
23	BCL2	1	-	g.chr18:60985670G>A	c.230C>T	p.A777V	Missense_Mutation	0	125	0	0	41	118	35	47	tFL specific	NP_000624	neutral	nd
23	BCL2	1	-	g.chr18:60985880G>A	c.20C>T	p.T71	Missense_Mutation	0	241	0	0	72	250	29	39	tFL specific	NP_000624	neutral	Yes
23	C12orf34	2	+	g.chr12:110206541G>T	c.807G>T	p.L269F	Missense_Mutation	0	48	0	0	22	59	37	49	tFL specific	NP_116218	deleterious	nd
23	C18orf34	12	-	g.chr18:30847010delT	c.1279delA	p.T427fs	Frame_Shift_Del	0	175	0	0	37	191	19	25	tFL specific	NP_001098998		nd
23	CARD8	7	-	g.chr19:48718624C>T	c.1003G>A	p.A335T	Missense_Mutation	0	83	0	0	32	73	44	59	tFL specific	NP_055774	neutral	nd
23	CDC16	18	+	g.chr13:115037900G>T	c.1845G>T	p.M615I	Missense_Mutation	0	90	0	0	50	131	38	51	tFL specific	NP_003894	deleterious	nd
23	CFTR	7	+	g.chr7:117176731A>C			Splice_Site	0	33	0	0	10	47	21	28	tFL specific	NP_000483		nd
23	CHRD1L2	2	-	g.chr11:74429842A>C	c.118T>G	p.Y40D	Missense_Mutation	20	64	31	37	0	66	0	0	FL specific		deleterious	nd
23	CLDN5	1	-	g.chr22:19511372A>T	c.662T>A	p.L221H	Missense_Mutation	0	120	0	0	41	113	36	48	tFL specific	NP_003268	deleterious	nd
23	CNKSR3	13	-	g.chr6:154727579T>C	c.1577A>G	p.K526R	Missense_Mutation	0	61	0	0	18	57	32	42	tFL specific	NP_775786	deleterious	nd
23	COL6A1	21	+	g.chr21:47417392T>C	c.1456T>C	p.S486P	Missense_Mutation	0	74	0	0	27	79	34	46	tFL specific	NP_001839	none	nd
23	CREBBP	26,27	-	g.chr16:3786708_3786710delCTC	c.4501_4503delGAG	p.E1501del	In_Frame_Del	41	110	37	46	48	133	36	48	Shared	NP_004371		Yes
23	CRLF1	7	-	g.chr19:18705206C>T	c.1063G>A	p.G355S	Missense_Mutation	0	32	0	0	11	35	31	41	tFL specific	NP_004741	neutral	nd
23	CRYBA2	4	-	g.chr2:219855024T>A	c.544A>T	p.T182S	Missense_Mutation	0	57	0	0	19	59	32	42	tFL specific	NP_476435	neutral	nd
23	CTBP2	1	-	g.chr10:126715493G>T	c.836C>A	p.T279N	Missense_Mutation	0	152	0	0	47	138	34	45	tFL specific	NP_001320		Yes
23	CTNND2	3	-	g.chr5:11565099G>A	c.244C>T	p.R82*	Nonsense_Mutation	0	69	0	0	23	69	33	44	tFL specific	NP_001323		nd
23	DHX8	10	+	g.chr17:41576264G>T	c.1335G>T	p.E445D	Missense_Mutation	24	75	32	37	1	79	1	0	FL specific	NP_004932	deleterious	nd
23	DST	38	-	g.chr6:56426110T>C	c.7949A>G	p.N2650S	Missense_Mutation	0	45	0	0	30	65	46	61	tFL specific	NP_001138241	deleterious	nd
23	DST	13	-	g.chr6:56498957A>C	c.3495T>G	p.D1165E	Missense_Mutation	0	199	0	0	69	194	36	48	tFL specific	NP_001138241	deleterious	nd
23	DTNA	3	+	g.chr18:32374090C>T	c.238C>T	p.R80C	Missense_Mutation	17	64	27	32	0	99	0	0	FL specific (tFL LOH)	NP_001381	deleterious	nd
23	DTX1	1	+	g.chr12:113496076G>A	c.79G>A	p.V27M	Missense_Mutation	0	22	0	0	9	25	36	48	tFL specific	NP_004407	deleterious	nd
23	DTX1	2	+	g.chr12:113515546G>A	c.577G>A	p.G193S	Missense_Mutation	0	52	0	0	19	56	34	46	tFL specific	NP_004407	deleterious	nd
23	EPAS1	3	+	g.chr2:46583316G>A	c.244G>A	p.E82K	Missense_Mutation	0	65	0	0	29	65	45	59	tFL specific	NP_001421	neutral	nd
23	ERBB2	1,2	+	g.chr17:37863270G>A	c.101G>A	p.R34Q	Missense_Mutation	18	36	50	58	20	47	43	56	Shared	NP_004439	neutral	nd
23	EXOC6B	19	-	g.chr2:72606914C>T	c.2066G>A	p.R689Q	Missense_Mutation	0	20	0	0	8	14	57	75	tFL specific	NP_056004		nd
23	FAM129B	4	-	g.chr9:130287433G>A	c.325C>T	p.R109W	Missense_Mutation	0	59	0	0	25	56	45	61	tFL specific	NP_073744		nd
23	FAM5C	1	-	g.chr1:190423926A>T	c.95T>A	p.V32D	Missense_Mutation	0	115	0	0	41	143	29	39	tFL specific	NP_950252	deleterious	nd
23	FBXO11	7	-	g.chr2:48060189_48060190delTG	c.702_703delCA	p.D234fs	Frame_Shift_Del	44	106	42	50	0	103	0	0	FL specific	NP_079409		Yes
23	FCRL2	7	-	g.chr1:157736749T>G	c.1175A>C	p.Y392S	Missense_Mutation	0	178	0	0	102	192	53	71	tFL specific	NP_110391	neutral	nd
23	FLNB	16	+	g.chr3:58095850delC	c.2437delC	p.P813fs	Frame_Shift_Del	0	49	0	0	8	21	38	50	tFL specific	NP_001448		nd
23	FMN2	1	+	g.chr1:240255566G>C	c.157G>C	p.G53R	Missense_Mutation	0	45	0	0	24	77	31	42	tFL specific	NP_064450	none	nd
23	FN1	36,37,38	-	g.chr2:216240050T>C	c.6044A>G	p.Q2015R	Missense_Mutation	0	40	0	0	20	48	42	55	tFL specific	NP_997647	deleterious	nd
23	FREM2	5	+	g.chr13:39357329C>T	c.5764C>T	p.Q1922*	Nonsense_Mutation	0	18	0	0	11	38	29	39	tFL specific	NP_997244	NP_specific	nd
23	FRMD3	14	-	g.chr9:85863366C>A	c.1261G>T	p.A421S	Missense_Mutation	0	61	0	0	27	66	41	55	tFL specific	NP_777598	neutral	nd
23	GABRA6	9	+	g.chr5:161128513T>G	c.1096T>G	p.S366A	Missense_Mutation	57	244	23	27	0	129	0	0	FL specific (tFL LOH)	NP_000802	neutral	nd
23	GALNTL5	4	+	g.chr7:151684319G>T	c.611G>T	p.R204I	Missense_Mutation	0	39	0	0	10	53	19	25	tFL specific	NP_660335	deleterious	nd
23	GPSM1	4	+	g.chr9:139231500G>C	c.549G>C	p.E183D	Missense_Mutation	12	42	29	33	0	50	0	0	FL specific	NP_001139110	neutral	nd
23	GRIN2D	12	+	g.chr19:48946661G>C	c.3478G>C	p.G1160R	Missense_Mutation	0	18	0	0	6	18	33	44	tFL specific	NP_000827	none	nd
23	GRIN3B	2	+	g.chr19:1003277C>G	c.575C>G	p.T192R	Missense_Mutation	28	78	36	42	0	67	0	0	FL specific	NP_619635	neutral	Yes
23	GTF2IRD1	4	+	g.chr7:73932631G>A	c.584G>A	p.R195H	Missense_Mutation	2	156	1	0	164	275	60	80	tFL specific	NP_057412	deleterious	nd
23	GTF3C5	5	+	g.chr9:135927516C>T	c.838C>T	p.L280F	Missense_Mutation	0	208	0	0	61	199	31	41	tFL specific	NP_036219	deleterious	nd
23	HBP1	2	+	g.chr7:106823014A>C	c.366A>C	p.Q122H	Missense_Mutation	0	29	0	0	21	48	44	59	tFL specific	NP_036389	deleterious	nd
23	HIST1H2BN	1	+	g.chr6:27806813C>T	c.374C>T	p.S125F	Missense_Mutation	0	68	0	0	25	89	28	38	tFL specific	NP_003511	deleterious	nd
23	HIST1H3D	1	-	g.chr6:26197090C>T	c.389G>A	p.R130H	Missense_Mutation	0	98	0	0	26	88	30	40	tFL specific	NP_003521	neutral	nd
23	ID2	1	+	g.chr2:8822446C>T	c.151C>T	p.P51S	Missense_Mutation	0	131	0	0	44	127	35	47	tFL specific	NP_002157	deleterious	nd
23	INPP5B	17	-	g.chr1:38339779T>A	c.1837A>T	p.T613S	Missense_Mutation	14	39	36	44	0	44	0	0	FL specific	NP_005531	neutral	nd
23	ITFG3	11	+	g.chr16:314830C>T	c.1468C>T	p.R490C	Missense_Mutation	0	149	0	0	43	120	36	48	tFL specific	NP_114428	deleterious	nd
23	KCNMA1	25	-	g.chr10:78651335C>T	c.3290G>A	p.R1097H	Missense_Mutation	0	23	0	0	8	34	24	32	tFL specific	NP_001154824	deleterious	nd
23	KDM3B	22	+	g.chr5:137766109C>T	c.5065C>T	p.H1689Y	Missense_Mutation	0	80	0	0	35	86	41	55	tFL specific	NP_057688	deleterious	nd
23	KIAA0195	14	+	g.chr17:73488828G>A	c.1870G>A	p.V624I	Missense_Mutation	0	369	0	0	142	336	42	56	tFL specific	NP_055553	neutral	nd
23	KIAA0391	6	+	g.chr14:35739741delG	c.1559delG	p.W520fs	Frame_Shift_Del	0	80	0	0	33	84	39	52	tFL specific	NP_055487		nd
23	KIAA1045	2	+	g.chr9:34972445C>T	c.481C>T	p.R161C	Missense_Mutation	0	60	0	0	32	84	38	51	tFL specific	NP_056112	deleterious	nd
23	KIAA1407	9	-	g.chr3:113729769C>G	c.1263G>C	p.E421D	Missense_Mutation	0	149	0	0	53	163	33	44	tFL specific	NP_065868	deleterious	nd
23	KIAA1797	10	+	g.chr9:20789369G>A	c.1217G>A	p.C406Y	Missense_Mutation	1	63	2	0	22	52	42	56	tFL specific	NP_060264	neutral	nd
23	KIF1A	26	-	g.chr2:241686727C>A	c.2686G>T	p.A896S	Missense_Mutation	0	41	0	0	14	48	29	39	tFL specific	NP_004312	neutral	nd
23	KLHL21	1	-	g.chr1:6662679C>T	c.199G>A	p.G67R	Missense_Mutation	0	45	0	0	17	61	28	38	tFL specific	NP_055666	deleterious	nd
23	KRT6B	4	-	g.chr12:52843573T>A	c.881A>T	p.E294V	Missense_Mutation	0	10	0	0	5	11	45	59	tFL specific	NP_005546	deleterious	nd
23	LPCAT2	10	+	g.chr16:55583260C>T	c.1007C>T	p.T336I	Missense_Mutation	0	159	0	0	65	152	43	57	tFL specific	NP_060309	neutral	nd
23	LZTS2	2	+	g.chr10:102763576A>G	c.721A>G	p.S241G	Missense_Mutation	0	285	0	0	104	298	35	47	tFL specific	NP_115805	neutral	nd
23	MCM7	3,7	-	g.chr7:99695823C>A	c.811G>T	p.V271F	Missense_Mutation	34	89	38	45	0	133	0	0	FL specific	NP_005907	deleterious	nd
23	ME1	4	-	g.chr6:84061780T>G			Splice_Site	26	82	32	38	0	86	0	0	FL specific	NP_002386		nd
23	MLL2	32	-	g.chr12:49433388G>A	c.8059C>T	p.R2687*	Nonsense_Mutation												

23	PER3	6	+	g.chr1:7858732A>T	c.789A>T	p.E263D	Missense_Mutation	18	46	39	47	0	57	0	0	FL specific	NP_058515	none	nd
23	PHF21A	3	-	g.chr11:46098303A>C	c.153_splice	p.Q51_splice	Splice_Site	0	81	0	0	22	75	29	39	tFL specific	NP_001095272		nd
23	PIK3C2A	3	-	g.chr11:17172104C>A	c.1268G>T	p.S423I	Missense_Mutation	0	101	0	0	25	94	27	36	tFL specific	NP_002636	neutral	nd
23	PIKFYVE	4	+	g.chr2:209142398T>A	c.579T>C	p.N193K	Missense_Mutation	0	94	0	0	33	89	37	49	tFL specific	NP_055855	deleterious	nd
23	PIM1	4	+	g.chr6:37139072G>C	c.412G>C	p.A138P	Missense_Mutation	0	108	0	0	55	128	43	57	tFL specific	NP_002639	neutral	Yes
23	PKDREJ	1	-	g.chr22:46653186A>T	c.6034T>A	p.L2012M	Missense_Mutation	0	161	0	0	37	126	29	39	tFL specific	NP_006062	neutral	nd
23	PLCB4	15	+	g.chr20:9374309A>T	c.1398A>T	p.K466N	Missense_Mutation	0	43	0	0	13	44	30	40	tFL specific	NP_877949	deleterious	nd
23	PLD1	25,26	+	g.chr3:171321022A>T	c.3071T>A	p.V1024E	Missense_Mutation	0	145	0	0	44	178	25	33	tFL specific	NP_002653	neutral	nd
23	PRDM16	17	+	g.chr1:3350370_3350375del6bp	c.3826_3831del6bp	p.L*1276del	In_Frame_Del	0	42	0	0	14	30	47	62	tFL specific	NP_071397		nd
23	PSME1	5	+	g.chr14:24606747_24606749delAAG	c.265_267delAAG	p.K91del	In_Frame_Del	0	142	0	0	41	140	29	39	tFL specific	NP_006254		nd
23	PTCHD1	3	+	g.chrX:23410999C>A	c.1364C>A	p.A455E	Missense_Mutation	81	104	78	92	106	126	84	100	Shared	NP_775766	deleterious	nd
23	PTPRS	7,10	-	g.chr19:5244358G>A	c.1124C>T	p.P375L	Missense_Mutation	32	73	44	52	0	61	0	0	FL specific	NP_002841	deleterious	nd
23	PTPRT	3	-	g.chr20:41419852A>G	c.469T>C	p.W157R	Missense_Mutation	0	88	0	0	35	93	38	51	tFL specific	NP_008981		nd
23	RAB3GAP2	11	-	g.chr1:220368644C>G	c.1036G>C	p.A346P	Missense_Mutation	0	68	0	0	16	91	18	24	tFL specific	NP_036546	deleterious	nd
23	RHBDF2	3	-	g.chr17:74475216T>C	c.503A>G	p.E168G	Missense_Mutation	18	40	45	53	0	48	0	0	FL specific	NP_078875	deleterious	nd
23	RIMS1	2	+	g.chr6:72678694T>G	c.173T>G	p.V58G	Missense_Mutation	0	71	0	0	10	60	17	23	tFL specific	NP_055804	deleterious	nd
23	ROCK2	3	-	g.chr2:11426700C>T	c.289G>C	p.V97I	Missense_Mutation	0	251	0	0	102	278	37	49	tFL specific	NP_004841	neutral	nd
23	RRAGC	2	-	g.chr1:39322649A>G	c.343T>C	p.W115R	Missense_Mutation	34	93	37	44	0	107	0	0	FL specific	NP_071440	deleterious	nd
23	SEC24B	8,9	+	g.chr4:110433120C>T	c.1784C>T	p.P595L	Missense_Mutation	28	67	42	49	0	52	0	0	FL specific	NP_006314	deleterious	nd
23	SEMA6A	13	-	g.chr5:115813732C>T	c.1546G>A	p.E516K	Missense_Mutation	0	208	0	0	83	217	38	51	tFL specific	NP_065847	neutral	nd
23	SP100	17	+	g.chr2:231339161_231339164del4bp	c.1600_splice	p.R534_splice	Splice_Site	0	74	0	0	30	79	38	51	tFL specific	NP_003104		nd
23	SRCIN1	7	-	g.chr17:36718132C>T	c.1267G>A	p.A423T	Missense_Mutation	0	51	0	0	26	69	38	51	tFL specific	NP_079524	neutral	nd
23	ST3GAL6	8	+	g.chr3:98510813T>A	c.885T>A	p.N295K	Missense_Mutation	0	162	0	0	59	167	35	47	tFL specific	NP_006091	deleterious	nd
23	ST8SIA3	1	+	g.chr18:55020100G>A	c.23G>A	p.R8Q	Missense_Mutation	0	139	0	0	35	162	22	29	tFL specific	NP_056963	neutral	nd
23	STAT6	11	-	g.chr12:57496671C>G	c.1246G>C	p.G416R	Missense_Mutation	99	234	42	49	1	231	0	0	FL specific	NP_003144	deleterious	nd
23	SYNGAP1	8	+	g.chr6:33406062G>T	c.1380G>T	p.K460N	Missense_Mutation	26	77	34	40	0	90	0	0	FL specific	NP_006763	deleterious	nd
23	SYT10	3	-	g.chr12:33559745T>G	c.1056A>C	p.K352N	Missense_Mutation	0	35	0	0	15	31	48	63	tFL specific	NP_945343	deleterious	nd
23	TIMELESS	17	-	g.chr12:568817153C>A	c.2197G>T	p.A733S	Missense_Mutation	0	99	0	0	28	83	34	45	tFL specific	NP_003911	neutral	nd
23	TMCO3	3	+	g.chr13:114154286A>G	c.638A>G	p.N213S	Missense_Mutation	0	23	0	0	13	24	54	73	tFL specific	NP_060375	neutral	nd
23	TMED8	6	-	g.chr14:77808242_77808244delTGTT	c.848_850delACA	p.283_284DS>G	In_Frame_Del	20	82	24	29	0	77	0	0	FL specific	NP_998766		nd
23	TMEM27	6	-	g.chrX:15646170G>T	c.593C>A	p.S198Y	Missense_Mutation	0	127	0	0	78	166	47	63	tFL specific	NP_065716	neutral	nd
23	TMEM86B	1	-	g.chr19:55740064C>A	c.46G>T	p.A16S	Missense_Mutation	8	20	40	48	0	30	0	0	FL specific	NP_776165	neutral	nd
23	TNIP1	10	-	g.chr5:150422217G>A	c.1018C>T	p.Q340*	Nonsense_Mutation	0	34	0	0	11	37	30	41	tFL specific	NP_006049		nd
23	TOR1AIP2	4	-	g.chr1:179815448C>A	c.1171G>T	p.D391Y	Missense_Mutation	0	205	0	0	56	293	19	25	tFL specific	NP_659471	deleterious	nd
23	TP53	8	-	g.chr17:7576851A>G	c.993_splice	p.Q331_splice	Splice_Site	0	128	0	0	41	117	35	47	tFL specific	NP_001119584		Yes
23	TRPM1	26	-	g.chr15:31294516C>T	c.4321G>A	p.V1441I	Missense_Mutation	0	186	0	0	85	227	37	49	tFL specific	NP_002411	neutral	nd
23	TTL6	6,13	-	g.chr17:46862342G>T	c.1983C>A	p.N661K	Missense_Mutation	0	24	0	0	13	41	32	43	tFL specific	NP_001124390	deleterious	nd
23	UBE2CBP	1	-	g.chr6:83775407A>T	c.31T>A	p.F11I	Missense_Mutation	0	168	0	0	60	197	30	40	tFL specific	NP_944602	deleterious	nd
23	UCK1	3	-	g.chr9:134404963C>T	c.277G>A	p.D93N	Missense_Mutation	0	280	0	0	85	275	31	41	tFL specific	NP_113620	deleterious	nd
23	UPF1	1	+	g.chr19:18943078delG	c.60delG	p.E20fS	Frame_Shift_Del	0	70	0	0	28	83	34	45	tFL specific	NP_002902		nd
23	USP25	2	+	g.chr21:17135247C>T	c.83C>T	p.T28M	Missense_Mutation	0	206	0	0	69	197	35	46	tFL specific	NP_037528	deleterious	nd
23	USPL1	8	+	g.chr13:31232694A>C	c.2480A>C	p.K827T	Missense_Mutation	65	201	32	38	0	200	0	0	FL specific	NP_005791	neutral	nd
23	WDR47	10	-	g.chr1:109525909G>C	c.2090C>G	p.A697G	Missense_Mutation	27	64	42	49	0	79	0	0	FL specific	NP_001136023	deleterious	nd
23	WDR59	24	-	g.chr16:74920225C>A	c.2489G>T	p.S830I	Missense_Mutation	0	28	0	0	9	27	33	44	tFL specific	NP_085058	neutral	nd
23	ZBED4	1	+	g.chr22:50280489C>G	c.3179C>G	p.A1060G	Missense_Mutation	0	95	0	0	27	68	40	53	tFL specific	NP_055653	neutral	nd
23	ZBTB40	11	+	g.chr1:22839511G>C	c.2556G>C	p.K852N	Missense_Mutation	37	99	37	43	0	95	0	0	FL specific	NP_001077090	deleterious	nd
23	ZCCHC7	1	+	g.chr9:37126480G>C	c.151G>C	p.E51Q	Missense_Mutation	0	130	0	0	46	122	38	51	tFL specific	NP_115602	deleterious	nd
23	ZDHHC18	1	+	g.chr1:27153443A>G	c.148A>G	p.S50G	Missense_Mutation	12	57	21	25	14	84	17	23	Shared	NP_115659	none	nd
23	ZFP90	4	+	g.chr16:68597622_68597626del5bp	c.932_936del5bp	p.Q311fs	Frame_Shift_Del	31	89	35	41	0	110	0	0	FL specific	NP_597715		nd
23	ZNF425	4	-	g.chr7:148802440G>A	c.523C>T	p.R175fs	Missense_Mutation	0	154	0	0	50	201	25	33	tFL specific	NP_001001661	neutral	nd
23	ZNF687	6	+	g.chr1:151262602_151262615del14bp	c.2969_2982del14bp	p.V990fs	Frame_Shift_Del	0	99	0	0	25	132	19	25	tFL specific	NP_065883		nd
23	ZNF74	5	+	g.chr22:20761144T>G	c.1821T>G	p.D607E	Missense_Mutation	0	76	0	0	25	84	30	40	tFL specific	NP_003417	neutral	nd
29	ABCC4	23	-	g.chr13:95726486A>G	c.2899T>C	p.S967P	Missense_Mutation	0	44	0	0	16	44	36	44	tFL specific	NP_005836	neutral	nd
29	ADAM30	1	-	g.chr1:120437102T>A	c.1858A>T	p.K620*	Nonsense_Mutation	9	117	8	15	46	98	47	59	tFL specific	NP_068566		nd
29	ANKLE2	9	-	g.chr12:13331202T>C>A	c.1665G>T	p.K555N	Missense_Mutation	0	139	0	0	85	256	33	41	tFL specific	NP_055929	deleterious	nd
29	APC	8	+	g.chr5:112151265G>A	c.908G>A	p.R303K	Missense_Mutation	0	14	0	0	12	27	44	56	tFL specific	NP_001120983	deleterious	nd
29	ARID1A	5	+	g.chr1:27087512_27087513insC	c.2086_2087insC	p.S696fs	Frame_Shift_Ins	0	116	0	0	95	105	90	100	tFL specific	NP_006006		Yes
29	C4BP4	11	+	g.chr1:207317893C>T	c.1625C>T	p.T542I	Missense_Mutation	0	24	0	0	14	22	64	77	tFL specific	NP_000706	neutral	nd
29	CHD8	17	-	g.chr14:21871370A>T	c.2683T>A	p.Y895N	Missense_Mutation	0	68	0	0	29	69	42	52	tFL specific	NP_065971	deleterious	nd
29	COBL	3	-	g.chr7:51261156T>C	c.376A>G	p.T126A	Missense_Mutation	5	72	7	13	41	80	51	64	tFL specific	NP_056013	neutral	nd
29	CPE	1	+	g.chr4:166300396C>G	c.23C>G	p.A8G	Missense_Mutation	0	29	0	0	9	20	45	58	tFL specific	NP_001864	none	nd
29	CREBBP	29,30	-	g.chr16:3781324_3781326delAGG	c.5039_5041delCCT	p.S1680del	In_Frame_Del	4	21	19	36	12	32	37	45	Shared	NP_004371		Yes
29	DAG1	2	+	g.chr3:49570156G>C	c.2212G>C	p.D738H	Missense_Mutation	0	65	0	0	22	59	37	47	tFL specific	NP_004384	deleterious	nd
29	DCTN4	8,9,10	-	g.chr5:150102491A>T	c.886T>A	p.F296I	Missense_Mutation	0	143	0	0	47	113	42	53	tFL specific	NP_057305	deleterious	nd
29	DIRAS3	1	-	g.chr1:68512353C>T	c.628G>A	p.E210K	Missense_Mutation	52	107	49	89	7	91	8	10	FL specific (tFL subclonal)	NP_004666	neutral	nd
29	DYRK1A	6	+	g.chr21:38862603A>G	c.791A>G	p.K264R	Missense_Mutation	0	123	0	0	46	106	43	54	tFL specific	NP_001387	deleterious	nd
29	EZH2	14,15	-	g.chr7:148508727T>A	c.1922A>T	p.Y641F	Missense_Mutation	34	231	15	27	116	250	46	58	Shared	NP_004447	deleterious	Yes
29	GNAI3	3	-	g.chr17:63014369A>G	c.561_splice	p.P187_splice	Splice_Site	0	47	0	0	55	60	92	100	tFL specific	NP_006563		Yes
29	GSTZ1	6,8,9	+	g.chr14:77797515A>T	c.628A>T	p.T210S	Missense_Mutation	0	68	0	0	39	77	51	64	tFL specific	NP_665877	deleterious	nd
29	HIST1H2BK	1	-	g.chr6:27114204G>A	c.374C>T	p.A125V	Missense_Mutation	13	63	21	38	0	61	0	0	FL specific (tFL LOH)	NP_542160	neutral	nd
29	HTR2A	3	-	g.chr13:47409763G>T	c.625C>A	p.P209Y	Missense_Mutation	0	45	0									

29	PRMT6	1	+	g.chr1:107599741T>A	c.404T>A	p.V135D	Missense_Mutation	12	39	31	56	0	52	0	0	FL specific (tFL LOH)	NP_060607	deleterious	nd
29	SAMD9L	1	-	g.chr7:92763590C>T	c.1695G>A	p.W565*	Nonsense_Mutation	0	131	0	0	62	150	41	51	tFL specific	NP_688916		nd
29	SKIV2L	18	+	g.chr6:31933633G>A	c.2045G>A	p.R682H	Missense_Mutation	10	66	15	27	58	65	89	100	Shared*	NP_008860	deleterious	nd
29	TBC1D20	8	-	g.chr20:419481C>T	c.961G>A	p.A321T	Missense_Mutation	0	56	0	0	25	63	40	50	tFL specific	NP_653229	neutral	nd
29	TMEM119	1	-	g.chr12:108985565T>A	c.595A>T	p.M199L	Missense_Mutation	0	143	0	0	108	171	63	79	tFL specific	NP_859075	neutral	nd
29	TNFRSF14	2	+	g.chr1:2489197C>A	c.102C>A	p.C34*	Nonsense_Mutation	17	86	20	37	0	125	0	0	FL specific (tFL LOH)	NP_003811		Yes
29	TNFRSF14	5	+	g.chr1:2492097C>A	c.495C>A	p.C165*	Nonsense_Mutation	0	126	0	0	156	167	93	100	tFL specific	NP_003811		Yes
29	TYRP1	4	+	g.chr9:12702400A>T	c.1043A>T	p.Y348F	Missense_Mutation	12	83	14	26	38	82	46	57	Shared*	NP_000541	neutral	nd
29	ZNF276	3,4	+	g.chr16:89789807G>A	c.696G>A	p.W232*	Nonsense_Mutation	0	52	0	0	20	44	45	56	tFL specific	NP_001106997		nd
29	ZNF462	2	+	g.chr9:109686573C>A	c.380C>A	p.T127N	Missense_Mutation	18	140	13	24	37	97	38	47	Shared*	NP_067047	none	nd
39	ARID1A	18	+	g.chr1:27101552_27101553delAA	c.4834_4835delAA	p.K1612fs	Frame_Shift_Del	17	81	21	46	41	103	40	53	Shared	NP_006006		nd
39	BARD1	6	-	g.chr2:215632225C>T	c.1549G>A	p.G517R	Missense_Mutation	0	101	0	0	39	109	36	48	tFL specific	NP_000456	deleterious	nd
39	BCL2	1	-	g.chr18:60985412T>G	c.488A>C	p.N163T	Missense_Mutation	39	179	22	49	80	255	31	41	Shared	NP_000624	neutral	Yes
39	BCL2	1	-	g.chr18:60985547T>C	c.353A>G	p.Q118R	Missense_Mutation	18	120	15	34	62	191	32	43	Shared	NP_000624	neutral	Yes
39	BCL2	1	-	g.chr18:60985508G>A	c.392C>T	p.A131V	Missense_Mutation	28	157	18	40	82	257	32	43	Shared	NP_000624	neutral	Yes
39	BCL2	1	-	g.chr18:60985499A>G	c.401T>C	p.V134A	Missense_Mutation	29	159	18	40	86	251	34	45	Shared	NP_000624	neutral	Yes
39	BCL2	1	-	g.chr18:60985311C>G	c.589G>C	p.G197R	Missense_Mutation	7	82	9	21	49	140	35	47	Shared	NP_000624		nd
39	BNC2	7	-	g.chr9:16419316C>T	c.2971G>A	p.D991N	Missense_Mutation	0	86	0	0	19	113	17	23	tFL specific	NP_060107	deleterious	nd
39	C9orf93	23	+	g.chr9:15874651T>A	c.3590T>A	p.V1197E	Missense_Mutation	1	35	3	7	20	62	32	43	tFL specific	NP_775821	deleterious	nd
39	CREBBP	25,26	+	g.chr16:3788617C>T	c.4337G>A	p.R1446H	Missense_Mutation	7	32	22	48	5	26	19	25	Shared	NP_004371	none	Yes
39	CREBBP	22,23	-	g.chr16:3794893A>T	c.3982_splice	p.R1328_splice	Splice_Site	14	69	20	44	28	78	36	48	Shared	NP_004371		Yes
39	CYP4F2	6	-	g.chr19:16000349C>T	c.802G>A	p.D268N	Missense_Mutation	0	35	0	0	10	41	24	32	tFL specific	NP_001073	neutral	nd
39	EBF1	2	-	g.chr5:158524075G>C	c.198C>G	p.N66K	Missense_Mutation	0	44	0	0	14	58	24	32	tFL specific	NP_076870	deleterious	nd
39	EZH2	16,17	-	g.chr7:148506467G>C	c.2030C>G	p.A677G	Missense_Mutation	7	25	28	62	20	38	53	70	Shared	NP_004447	deleterious	Yes
39	FEZF1	4	-	g.chr7:121942390C>A	c.1089G>T	p.K363N	Missense_Mutation	3	81	4	9	33	89	37	49	tFL specific	NP_001019784	deleterious	nd
39	FOXP4	7	+	g.chr6:41555548C>A	c.890C>A	p.T297N	Missense_Mutation	0	23	0	0	14	29	48	64	tFL specific	NP_001012426	neutral	nd
39	IRF8	8	+	g.chr16:85954866_85954872del7bp	c.1259_1265del7bp	p.E420fs	Frame_Shift_Del	0	38	0	0	14	56	25	34	tFL specific	NP_002154		Yes
39	ITGAV	5,7	+	g.chr2:187500869G>A	c.692G>A	p.S231N	Missense_Mutation	0	139	0	0	58	155	37	49	tFL specific	NP_002201	deleterious	nd
39	ITPR1	41	+	g.chr3:4810375C>T	c.5762C>T	p.A1921V	Missense_Mutation	0	20	0	0	6	21	29	40	tFL specific	NP_001093422	neutral	nd
39	MEGF6	19	-	g.chr1:3417807G>A	c.2380C>T	p.R794C	Missense_Mutation	0	17	0	0	13	29	45	60	tFL specific	NP_001400	neutral	nd
39	MLL2	3	-	g.chr12:49448310C>T	c.400_splice	p.G134_splice	Splice_Site	13	88	15	34	46	172	27	36	Shared	NP_003473		Yes
39	MLL2	48	-	g.chr12:49420474C>A	c.15275G>T	p.C5092F	Missense_Mutation	28	94	30	66	64	133	48	64	Shared	NP_003473		Yes
39	MS4A1	2	+	g.chr11:60230493delG	c.178delG	p.G60fs	Frame_Shift_Del	0	31	0	0	11	25	44	60	tFL specific	NP_690605		nd
39	MS4A1	4	+	g.chr11:60233497T>A	c.440T>A	p.L147*	Nonsense_Mutation	0	102	0	0	45	141	32	43	tFL specific	NP_690605		nd
39	NTRK2	17,18	+	g.chr9:87636203C>T	c.2320C>T	p.R774*	Nonsense_Mutation	0	28	0	0	13	55	24	32	tFL specific	NP_001018074		nd
39	OR5R1	1	-	g.chr11:56184966A>T	c.743T>A	p.V248D	Missense_Mutation	1	58	2	0	18	56	32	43	tFL specific	NP_001004744	deleterious	nd
39	PARP1	20	-	g.chr1:226551735C>A	c.2695G>T	p.D899Y	Missense_Mutation	0	117	0	0	51	185	28	37	tFL specific	NP_001609	deleterious	nd
39	PDPN	1	+	g.chr1:13910569C>T	c.269C>T	p.A90V	Missense_Mutation	0	66	0	0	29	92	32	43	tFL specific	NP_006465	neutral	nd
39	PIK3CG	2	+	g.chr7:106513046G>C	c.2060G>C	p.R687T	Missense_Mutation	0	147	0	0	74	197	38	51	tFL specific	NP_002640	neutral	nd
39	PRRT3	3	-	g.chr3:9988311C>T	c.2546G>A	p.R849Q	Missense_Mutation	0	26	0	0	12	34	35	47	tFL specific	NP_997234	neutral	nd
39	PTPN13	20	+	g.chr4:87684231G>T	c.3905G>T	p.S1302I	Missense_Mutation	29	129	22	49	0	124	0	0	FL specific (tFL LOH)	NP_542414	deleterious	Yes
39	QRICH2	15	-	g.chr17:74274239C>A	c.4450G>T	p.D1484Y	Missense_Mutation	0	21	0	0	11	36	31	42	tFL specific	NP_115510	deleterious	nd
39	RDH14	1	-	g.chr2:18741648A>C	c.191T>G	p.L64R	Missense_Mutation	0	15	0	0	9	26	35	46	tFL specific	NP_065956	deleterious	nd
39	RP1L1	3	-	g.chr8:10469653C>T	c.1955G>A	p.S652N	Missense_Mutation	1	40	2	0	10	39	26	35	tFL specific	NP_849188	deleterious	Yes
39	SCARF2	11	-	g.chr22:20780222G>A	c.2056C>T	p.P686S	Missense_Mutation	0	62	0	0	20	54	37	50	tFL specific	NP_699165	deleterious	nd
39	SPTAN1	36,37	+	g.chr9:131374059G>A	c.4825G>A	p.V1609I	Missense_Mutation	0	39	0	0	17	39	44	59	tFL specific	NP_003118	neutral	nd
39	TBR1	6	+	g.chr2:162280011C>T	c.1322C>T	p.A441V	Missense_Mutation	2	42	5	12	23	69	33	44	tFL specific	NP_006584	deleterious	nd
39	TMEM57	6	+	g.chr1:25785202G>C	c.973G>C	p.A325P	Missense_Mutation	0	43	0	0	22	69	32	43	tFL specific	NP_060672	neutral	nd
39	UTP14C	1	+	g.chr13:52603717G>T	c.777G>T	p.K259N	Missense_Mutation	0	81	0	0	43	118	36	48	tFL specific	NP_067677	deleterious	nd
39	ZNF845	2	+	g.chr19:53848778A>T	c.35A>T	p.D12V	Missense_Mutation	0	43	0	0	19	61	31	42	tFL specific	NP_612383	deleterious	nd
40	ADAM19	12	-	g.chr5:156929927A>G	c.1192T>C	p.S398P	Missense_Mutation	0	388	0	0	203	280	72	80	tFL specific	NP_150377		nd
40	AOB	26	-	g.chr2:21234882T>A	c.4858A>T	p.N1620Y	Missense_Mutation	0	261	0	0	54	218	25	28	tFL specific	NP_000375	deleterious	nd
40	BCL2	1	-	g.chr18:60985725G>C	c.175C>G	p.P59A	Missense_Mutation	32	74	43	100	150	192	78	86	Shared	NP_000624	neutral	Yes
40	BCL2	1	-	g.chr18:60985411G>C	c.489C>G	p.N163K	Missense_Mutation	51	276	18	45	0	463	0	0	FL specific	NP_000624	neutral	Yes
40	BCL2	1	-	g.chr18:60985644G>A	c.256C>T	p.L86F	Missense_Mutation	0	107	0	0	79	156	51	57	tFL specific	NP_000624	neutral	nd
40	BCL2	1	-	g.chr18:60985646G>A	c.254C>T	p.A85V	Missense_Mutation	0	111	0	0	76	158	48	53	tFL specific	NP_000624	neutral	nd
40	BCL2	1	-	g.chr18:60985883C>G	c.17G>C	p.R6T	Missense_Mutation	1	219	0	0	146	384	38	42	tFL specific	NP_000624	deleterious	Yes
40	BCR	1	+	g.chr22:23523305G>A	c.158G>A	p.R53H	Missense_Mutation	10	63	16	40	0	68	0	0	FL specific	NP_004318	deleterious	nd
40	BCR	1	+	g.chr22:23523233G>A	c.86G>A	p.G29D	Missense_Mutation	0	50	0	0	13	32	41	45	tFL specific	NP_004318	deleterious	nd
40	BCR	1	+	g.chr22:23523690G>C	c.543G>C	p.E181D	Missense_Mutation	25	147	17	43	0	206	0	0	FL specific	NP_004318	neutral	nd
40	BCR	1	+	g.chr22:23523730C>G	c.583G>C	p.E195Q	Missense_Mutation	30	148	20	50	0	232	0	0	FL specific	NP_004318	neutral	nd
40	BCR	1	+	g.chr22:23523318C>G	c.171C>G	p.I57M	Missense_Mutation	16	66	24	60	0	77	0	0	FL specific	NP_004318	neutral	nd
40	BCR	1	+	g.chr22:23523568G>A	c.421G>A	p.E141K	Missense_Mutation	0	70	0	0	37	92	40	45	tFL specific	NP_004318	neutral	nd
40	BCR	1	+	g.chr22:23523249G>T	c.102G>T	p.E34D	Missense_Mutation	13	48	27	68	0	38	0	0	FL specific	NP_004318	neutral	nd
40	BCR	1	+	g.chr22:23523224G>A	c.77G>A	p.R26H	Missense_Mutation	15	47	32	80	0	26	0	0	FL specific	NP_004318	neutral	nd
40	BCR	1	+	g.chr22:23523997G>C	c.850G>C	p.V284L	Missense_Mutation	0	84	0	0	44	133	33	37	tFL specific	NP_004318	deleterious	nd
40	BMP7	1	-	g.chr20:55840918C>G	c.261C>G	p.Y87*	Nonsense_Mutation	0	153	0	0	99	201	49	54	tFL specific	NP_001710		nd
40	BMP7	1	-	g.chr20:55840929C>T	c.250C>A	p.L84M	Missense_Mutation	0	164	0	0	77	194	40	45	tFL specific	NP_001710	neutral	nd
40	BMP7	1	-	g.chr20:55840958C>T	c.221G>A	p.G74D	Missense_Mutation	0	143	0	0	82	181	45	50	tFL specific	NP_001710	neutral	nd
40	BTG2	1	+	g.chr1:203274808G>A	c.74G>A	p.R25K	Missense_Mutation	23	99	23	58	0	133	0	0	FL specific	NP_006754	neutral	nd
40	C1orf129	14	+	g.chr1:170967384A>G	c.1565A>G	p.D522G	Missense_Mutation	29	174	17	43	0	181	0	0	FL specific	NP_079339	deleterious	nd
40	CDH19	11	-	g.chr18:64															

40	DTX1	1	+	g.chr12:113496037G>A	c.40G>A	p.G14S	Missense_Mutation	0	25	0	0	17	41	41	45	tFL specific	NP_004407	deleterious	nd	
40	DTX1		1	+	g.chr12:113496145A>G	c.148A>G	p.N50D	Missense_Mutation	0	14	0	0	5	18	28	33	tFL specific	NP_004407	deleterious	nd
40	EFTUD2	22,23	-	g.chr17:42930934C>T	c.2417G>A	p.G806D	Missense_Mutation	0	60	0	0	30	87	34	38	tFL specific	NP_004238	deleterious	nd	
40	ENTPD2	5	-	g.chr9:139945377T>A	c.751A>T	p.R251W	Missense_Mutation	0	18	0	0	7	19	37	39	tFL specific	NP_982293	deleterious	nd	
40	EPS15L1	14	-	g.chr19:16515512C>T	c.1315G>A	p.E439K	Missense_Mutation	21	122	17	43	0	179	0	0	FL specific	NP_067058	neutral	nd	
40	FADS2	1	+	g.chr11:61595887G>T	c.25G>T	p.E9*	Nonsense_Mutation	0	33	0	0	15	43	35	40	tFL specific	NP_004256		nd	
40	FAM102A	1	-	g.chr9:130742267C>G			Splice_Site	0	40	0	0	32	71	45	50	tFL specific	NP_001030331		nd	
40	FAM102A	1	-	g.chr9:130742271C>G	c.146G>C	p.R49T	Missense_Mutation	9	43	21	52	0	72	0	0	FL specific	NP_001030331	deleterious	nd	
40	FAM102A	1	-	g.chr9:130742352G>A	c.65C>T	p.T22I	Missense_Mutation	0	54	0	0	39	116	34	38	tFL specific	NP_001030331	deleterious	nd	
40	FAM190A	20,22,23	+	g.chr4:91321254T>A	p.L526*		Nonsense_Mutation	0	96	0	0	29	75	39	43	tFL specific	NP_001138537		nd	
40	FAM9A	6	-	g.chrX:8763451C>T	c.499G>G	p.D167N	Missense_Mutation	0	184	0	0	48	118	41	46	tFL specific	NP_777611	deleterious	nd	
40	FOXP2	8,10	+	g.chr7:114298209C>G	c.1355C>G	p.A452G	Missense_Mutation	0	244	0	0	93	422	22	25	tFL specific	NP_055306	none	nd	
40	GNAI3	4	-	g.chr17:63010387delA	c.1122delT	p.L374fs	Frame_Shift_Del	19	80	24	60	0	67	0	0	FL specific	NP_006563		nd	
40	GNAI3	4	-	g.chr17:63010388A>T	c.1121T>A	p.L374H	Missense_Mutation	20	81	25	62	0	67	0	0	FL specific	NP_006563	deleterious	nd	
40	GNAI2	5	+	g.chr3:50293703A>G	c.544A>G	p.T182A	Missense_Mutation	0	104	0	0	53	115	46	51	tFL specific	NP_002061	deleterious	nd	
40	GNAI2	5	+	g.chr3:50293712A>G	c.553A>G	p.I185V	Missense_Mutation	0	103	0	0	53	117	45	50	tFL specific	NP_002061	neutral	nd	
40	GNA13	7	-	g.chr7:80091620C>T	c.729G>A	p.M243I	Missense_Mutation	0	110	0	0	25	111	23	26	tFL specific	NP_001095856	deleterious	nd	
40	GPR112	22	+	g.chrX:135496371delG	c.9090delG	p.K3030fs	Frame_Shift_Del	0	112	0	0	42	145	29	32	tFL specific	NP_722576		nd	
40	HOXC9	1	+	g.chr12:54394378C>T	c.406C>T	p.R136C	Missense_Mutation	7	39	18	45	0	47	0	0	FL specific	NP_008828	deleterious	nd	
40	HTATSF1	9	+	g.chrX:135593004G>T	c.1100G>T	p.R367I	Missense_Mutation	0	53	0	0	16	51	31	34	tFL specific	NP_001156752	deleterious	nd	
40	HYAL4	1	+	g.chr7:12350889T>C	c.572T>C	p.F191S	Missense_Mutation	1	147	1	3	75	176	43	48	tFL specific	NP_036401	deleterious	nd	
40	ID3	1	-	g.chr1:23885745A>T	c.173T>A	p.V58D	Missense_Mutation	32	77	42	100	88	108	81	90	Shared	NP_002158	deleterious	nd	
40	IL22RA2	4,5	-	g.chr6:137468891G>A	c.610C>T	p.R204*	Nonsense_Mutation	26	57	46	100	27	60	45	50	Shared	NP_443194		nd	
40	ITPKB	1	-	g.chr1:226925103C>G	c.57G>C	p.E19D	Missense_Mutation	0	180	0	0	93	245	38	42	tFL specific	NP_002212	neutral	nd	
40	KCNA1	1	+	g.chr12:5020943C>A	c.399C>A	p.D133E	Missense_Mutation	1	192	1	3	47	196	24	27	tFL specific	NP_000208	neutral	nd	
40	KCNJ3	1	+	g.chr2:155555927C>T	c.640C>T	p.R214W	Missense_Mutation	21	106	20	50	0	145	0	0	FL specific	NP_002230	deleterious	nd	
40	KIAA0232	5	+	g.chr4:6863628A>T	c.1519A>T	p.T507S	Missense_Mutation	0	109	0	0	43	107	40	44	tFL specific	NP_055558	deleterious	nd	
40	KLHL14	1	-	g.chr18:30349993G>C	c.562C>G	p.Q188E	Missense_Mutation	0	109	0	0	88	170	52	58	tFL specific	NP_065856	deleterious	nd	
40	KNG1	10	+	g.chr3:186460008A>T	c.1823A>T	p.D608V	Missense_Mutation	0	323	0	0	111	237	47	52	tFL specific	NP_001095886	deleterious	nd	
40	LRP1B	80	-	g.chr2:141083445A>T	c.12226T>A	p.L4076M	Missense_Mutation	0	119	0	0	17	53	32	36	tFL specific	NP_061027	deleterious	nd	
40	LRN3	1	+	g.chr7:110764563A>G	c.1735A>G	p.N579D	Missense_Mutation	1	174	1	3	50	199	25	28	tFL specific	NP_001093130	deleterious	nd	
40	MAGEB1	1	+	g.chrX:30268653C>T	c.43C>T	p.R15C	Missense_Mutation	0	103	0	0	37	96	39	44	tFL specific	NP_002354	deleterious	nd	
40	MFHAS1	1	-	g.chr8:8750182C>G	c.387G>C	p.E129D	Missense_Mutation	0	50	0	0	27	58	47	53	tFL specific	NP_004216	deleterious	nd	
40	MFHAS1	1	-	g.chr8:8750244G>A	c.325C>T	p.L109F	Missense_Mutation	11	49	22	55	0	53	0	0	FL specific	NP_004216	deleterious	nd	
40	MICALCL	2	+	g.chr11:12316129C>T	c.1151C>T	p.P384L	Missense_Mutation	0	329	0	0	124	340	36	40	tFL specific	NP_116256	neutral	nd	
40	MID2	6	+	g.chrX:107159293T>A	c.1135T>A	p.F379I	Missense_Mutation	0	131	0	0	37	145	26	29	tFL specific	NP_036348	deleterious	nd	
40	MYC	2	+	g.chr8:128750677_128750679delCCC	c.169_171delCCC	p.P57del	In_Frame_Del	0	73	0	0	22	51	43	48	tFL specific	NP_002458		Yes	
40	MYC	2	+	g.chr8:128750681_128750689del9bp	c.173_181del9bp	p.58_61TPPL>M	In_Frame_Del	0	80	0	0	22	52	42	46	tFL specific	NP_002458		Yes	
40	MYC	2	+	g.chr8:128750692delT	c.184delT	p.S62fs	Frame_Shift_Del	0	87	0	0	22	51	43	48	tFL specific	NP_002458		Yes	
40	MYC	2	+	g.chr8:128750696C>A	c.188C>A	p.P63H	Missense_Mutation	0	90	0	0	22	50	44	49	tFL specific	NP_002458	deleterious	nd	
40	MYC	2	+	g.chr8:128750698_128750702del5bp	c.190_194del5bp	p.S64fs	Frame_Shift_Del	0	88	0	0	22	52	42	46	tFL specific	NP_002458		Yes	
40	MYC	2	+	g.chr8:128750852T>C	c.344T>C	p.F115S	Missense_Mutation	0	222	0	0	198	326	61	68	tFL specific	NP_002458	deleterious	Yes	
40	MYC	3	+	g.chr8:128752656G>A	c.817G>A	p.D273N	Missense_Mutation	0	49	0	0	39	58	67	74	tFL specific	NP_002458	neutral	nd	
40	NLGN4X	4	-	g.chrX:5821895delT	c.824delA	p.K275fs	Frame_Shift_Del	0	97	0	0	40	82	49	54	tFL specific	NP_851849		nd	
40	NOX3	12	-	g.chr6:155728292A>T	c.1552T>A	p.F518I	Missense_Mutation	0	44	0	0	11	35	31	36	tFL specific	NP_056533	deleterious	nd	
40	OXTR	2	-	g.chr3:8794766C>T	c.1067G>A	p.R356H	Missense_Mutation	0	52	0	0	31	74	42	47	tFL specific	NP_000907	deleterious	nd	
40	PAK1	13	-	g.chr11:77043800C>T	c.1526G>A	p.R509K	Missense_Mutation	0	121	0	0	51	206	25	28	tFL specific	NP_002567	neutral	nd	
40	PAK1	13	-	g.chr11:77043879C>T	c.1447G>A	p.E483K	Missense_Mutation	0	69	0	0	31	138	22	25	tFL specific	NP_002567	neutral	nd	
40	PAPD7	12	+	g.chr5:6754975C>T	c.1546C>T	p.H516Y	Missense_Mutation	0	24	0	0	14	32	44	48	tFL specific	NP_008930	deleterious	nd	
40	PCDH9	3,4	-	g.chr13:66878911T>A	c.3590A>T	p.Q1197L	Missense_Mutation	35	231	15	38	155	309	50	56	Shared*	NP_982354	neutral	nd	
40	PEAR1	7	+	g.chr1:156877825C>T	c.884C>T	p.P295L	Missense_Mutation	0	70	0	0	25	78	32	36	tFL specific	NP_001073940	deleterious	nd	
40	PECR	2	-	g.chr2:216931116C>T	c.250G>A	p.E84K	Missense_Mutation	1	257	0	0	72	194	37	41	tFL specific	NP_060911	deleterious	nd	
40	PGM2L1	9	-	g.chr11:74056511T>G	c.74056511T>G		Splice_Site	0	189	0	0	53	114	46	51	tFL specific	NP_775853		nd	
40	PIM1	1	+	g.chr6:37138419C>G	c.68C>G	p.T23S	Missense_Mutation	36	210	17	43	200	349	57	63	Shared	NP_002639	neutral	Yes	
40	PIM3	4	+	g.chr22:50355429C>G	c.586C>G	p.L196V	Missense_Mutation	0	25	0	0	6	21	29	31	tFL specific	NP_001001852	deleterious	nd	
40	POU2AF1	1	-	g.chr11:111249884T>G			Splice_Site	11	67	16	40	0	139	0	0	FL specific	NP_006226		nd	
40	PSMB2	1	-	g.chr1:36107027A>T	c.7T>A	p.Y3N	Missense_Mutation	0	49	0	0	16	50	32	35	tFL specific	NP_002785	deleterious	nd	
40	RPA4	1	+	g.chrX:96139489T>A	c.180T>A	p.F60L	Missense_Mutation	0	260	0	0	113	379	30	33	tFL specific	NP_037479	neutral	nd	
40	RPA4	1	+	g.chrX:96139782T>C	c.473T>C	p.I158T	Missense_Mutation	0	187	0	0	83	271	31	35	tFL specific	NP_037479	neutral	nd	
40	RRAGB	3	+	g.chrX:55748629T>G	c.137T>G	p.M46R	Missense_Mutation	0	62	0	0	19	68	28	32	tFL specific	NP_057740	deleterious	nd	
40	SCAF1	6	+	g.chr19:50155358G>C	c.1712G>C	p.R571P	Missense_Mutation	5	17	29	71	0	36	0	0	FL specific	NP_067051	none	nd	
40	SCCPDH	10	+	g.chr1:246927579T>A	c.1022T>A	p.F341Y	Missense_Mutation	1	181	1	2	47	128	37	41	tFL specific	NP_057086	neutral	nd	
40	SCG2	1	-	g.chr2:224463460G>A	c.541C>T	p.R181C	Missense_Mutation	3	211	1	3	47	110	43	48	tFL specific	NP_003460	deleterious	nd	
40	SEMA3C	14	-	g.chr7:80387798A>T	c.1492T>A	p.L498M	Missense_Mutation	0	83	0	0	47	145	32	36	tFL specific	NP_006370	neutral	nd	
40	SEMA4B	1	+	g.chr15:90744833T>C	c.23T>C	p.L8P	Missense_Mutation	9	41	22	55	0	50	0	0	FL specific	NP_064595		nd	
40	SGK1	2,4	-	g.chr6:134495653_134495676del24bp	c.125_148del24bp	p.IANN5YAC42del	In_Frame_Del	0	123	0	0	20	89	23	26	tFL specific	NP_005618		nd	
40	SLC30A4	1	-	g.chr15:45814505_45814531del27bp	c.22_48del27bp	p.KRLKSMLRK8del	In_Frame_Del	7	194	4	10	45	73	62	69	tFL specific	NP_037441		nd	
40	SLC6A2	4	+	g.chr16:55719076C>A	c.666C>A	p.H222Q	Missense_Mutation	0	118	0	0	86	181	48	53	tFL specific	NP_001034	deleterious	nd	
40	SOX9	1	+	g.chr17:70117806C>T	c.274C>T	p.P92S	Missense_Mutation	0	133	0	0	56	131	43	48	tFL specific	NP_000337	deleterious	nd	
40	SREK1	3,5	+	g.chr5:65458423A>T			Splice_Site	0	99	0	0	20	81	25	28	tFL specific	NP_631907		nd	
40	SRPX2	7	+	g.chrX:99921825G>A	c.856G>A	p.A286T	Missense_Mutation	0	153	0	0	48	189	25	28	tFL specific	NP_055282	deleterious	nd	
40	SULF1	1	+	g.chr8:70476307C>T	c.97C															

40	ZNF775	2	+	g.chr7:150094402C>T	c.833C>T	p.P278L	Missense_Mutation	0	62	0	0	73	146	50	55	tFL specific	NP_775951	neutral	nd
----	--------	---	---	---------------------	----------	---------	-------------------	---	----	---	----------	----	-----	----	-----------	--------------	-----------	---------	----

^ If multiple isoforms have been described for the gene, the exon number is provided for all isoforms

^^ Numbering according to the Human Genome hg19 assembly

* These variants may represent germline polymorphisms.

PPH2, PolyPhen-2 algorithm prediction; nd, not done

SNVs that were clonal in one disease phase and subclonal in the paired biopsy were classified as shared when the difference in variant allele frequency between the two phases was not significant (see Materials and Methods)

Table IV. Segments of tumor-acquired copy number alterations identified in the discovery panel

Sample ID	Chr	Cytoband	Start^	End^	Copy number	Segment size (Mb)	Aberration	N of genes	Genes in segment (if 10 or less)	N of miRs
2_FL	1	p36.33–p36.23	593441	7939410	1.31	7.35	Loss	103		4
2_FL	1	p36.23–p36.22	7943593	10118557	2.96	2.17	Gain	20		1
2_FL	1	p36.22–p36.21	10248705	12775335	2.47	2.53	Gain	40		0
2_FL	1	p36.12	21046275	21958754	1.36	0.91	Loss	6	<i>EIF4G3,ECE1,NBPF3,ALPL,RAP1GAP,USP48</i>	1
2_FL	1	p36.11–p35.2	25622672	32040621	1.21	6.42	Loss	91		1
2_FL	1	p35.1–p34.3	32515054	36948596	1.37	4.43	Loss	60		1
2_FL	1	p21.3–p13.2	98737348	113732302	2.91	14.99	Gain	108		2
2_FL	1	q21.1–q21.3	147942929	150345911	2.57	2.40	Gain	79		0
2_FL	1	q21.3–q22	150345973	153861732	3.22	3.52	Gain	120		0
2_FL	1	q31.1	188322363	188428798	1.53	0.11	Loss	1	<i>FAM5C</i>	0
2_FL	2	p21–p15	47167599	61656813	2.70	14.49	Gain	48		4
2_FL	8	p23.3–q24.3	21242	146268947	2.58	146.25	Gain	661		31
2_FL	12	q13.11–q22	44818143	92281807	2.68	47.46	Gain	392		14
2_FL	12	q23.3	103496596	103591140	1.38	0.09	Loss	1	<i>CHST11</i>	0
2_FL	16	p13.3	765	1561817	1.36	1.56	Loss	68		1
2_FL	17	q23.3	59279004	59487220	3.10	0.21	Gain	10	<i>CSH2,GH2,CSH1,CSHL1,GH1,CD79B,SCN4A,C17orf72,ICAM2,ERN1</i>	0
2_FL	19	p13.3–p13.2	6350425	6929013	0.88	0.58	Loss homozygous	15		1
2_FL	19	p13.11	17043939	17248858	1.33	0.20	Loss	7	<i>HAUS8,MYO9B,USE1,OCEL1,NR2F6,USHBP1,BABAM1</i>	0
2_FL	19	p13.11	17252515	17877184	2.36	0.62	Gain	23		0
2_FL	19	p13.11	17880872	18132412	1.37	0.25	Loss	6	<i>CCDC124,KCNN1,ARRDC2,IL12RB1,MAST3,PIK3R2</i>	0
2_FL	19	q13.31–q13.32	49771648	50491034	1.32	0.72	Loss	24		0
2_FL	19	q13.33	55061878	55271562	1.32	0.21	Loss	9	<i>PNKP,AKT1S1,TBC1D17,IL4I1,NUP62,ATF5,SIGLEC11,VRK3,ZNF473</i>	0
2_FL	22	q13.31–q13.33	43466634	49581309	1.48	6.11	Loss	60		3
2_tFL	1	p36.33–p36.23	615321	7939410	1.01	7.32	Loss	100		4
2_tFL	1	p36.23	7943593	8311158	4.43	0.37	Gain	3	<i>PARK7,ERRFI1,SLC45A1</i>	0
2_tFL	1	p36.23–p36.22	8315805	9254560	3.14	0.94	Gain	8	<i>SLC45A1,RERE,ENO1,CA6,SLC2A7,SLC2A5,GPR157,H6PD</i>	1
2_tFL	1	p36.22	9255290	10120199	4.26	0.86	Gain	10	<i>SPSB1,SLC25A33,TMEM201,PIK3CD,CLSTN1,CTNNBIP1,LZIC,NMNAT1,RBP7,UBE4B</i>	0
2_tFL	1	p36.22–p36.21	10234570	13734984	2.96	3.50	Gain	64		0
2_tFL	1	p36.12	21046275	21955982	0.98	0.91	Loss	6	<i>EIF4G3,ECE1,NBPF3,ALPL,RAP1GAP,USP48</i>	1
2_tFL	1	p36.11–p35.2	25575294	32046387	1.05	6.47	Loss	92		1
2_tFL	1	p35.1–p34.3	32498612	36955073	1.06	4.46	Loss	60		1
2_tFL	1	p31.3	61799901	66578027	1.05	4.78	Loss	23		1
2_tFL	1	p31.3–p31.1	67248040	81444601	7.24	14.20	Amp	48		2
2_tFL	1	p31.1–p22.3	81447085	86992584	3.05	5.55	Gain	25		0
2_tFL	1	p22.3	87096179	87279805	3.20	0.18	Gain	2	<i>SEP15,HS2ST1</i>	0
2_tFL	1	q21.1–q21.3	147999203	150406640	3.08	2.41	Gain	80		1
2_tFL	1	q21.3–q22	150406872	153860384	4.23	3.45	Gain	118		3
2_tFL	1	q31.1	188322363	188427763	1.05	0.11	Loss	1	<i>FAM5C</i>	0
2_tFL	2	p25.3–p14	2772	67734431	2.93	67.73	Gain	312		6
2_tFL	3	p14.2	60224233	60334358	1.08	0.11	Loss	1	<i>FHIT</i>	0
2_tFL	3	q27.3–q29	189007811	199322598	2.51	10.31	Gain (subclonal)	61		4
2_tFL	4	q13.1–q22.1	62887665	92476614	1.51	29.59	Loss (subclonal)	162		2
2_tFL	5	p14.3	19754037	19938522	3.10	0.18	Gain	1	<i>CDH18</i>	0
2_tFL	6	q11.1–q23.3	62975943	136945877	1.62	73.97	Loss (subclonal)	263		7
2_tFL	6	q23.3	136947840	137009964	1.01	0.06	Loss homozygous	1	<i>MAP3K5</i>	0
2_tFL	6	q23.3	137010234	137790542	1.62	0.78	Loss (subclonal)	6		0
2_tFL	6	q23.3–q27	137843499	170824447	1.05	32.98	Loss	145		2
2_tFL	8	p23.3–q12.2	151222	61813516	2.70	61.66	Gain	300		10
2_tFL	8	q12.2–q23.3	61814698	112363810	2.31	50.55	Gain (subclonal)	193		6
2_tFL	8	q23.3–q24.21	113026335	129097966	2.75	16.07	Gain	51		5
2_tFL	9	p24.3–q34.3	36431	140211203	1.46	140.17	Loss (subclonal)	753		26
2_tFL	10	p15.3–q11.21	62747	42979712	2.55	42.92	Gain (subclonal)	158		7

2_tFL	10	q11.21--q11.22	42988256	47173619	5.09	4.19 Amp	30		0
2_tFL	10	q11.22--q11.23	47174599	51067065	2.44	3.89 Gain (subclonal)	32		0
2_tFL	10	q11.23	51119363	51405645	4.82	0.29 Amp	4	AGAP7,MSMB,NCOA4,TIMM23	0
2_tFL	10	q11.23--q21.1	51407449	54364014	2.52	2.96 Gain (subclonal)	11		1
2_tFL	10	q21.1--q21.2	54365971	63675778	2.81	9.31 Gain	19		1
2_tFL	10	q21.2--q21.3	63684101	69227511	11.64	5.54 Amp	10	RTKN2,ZNF365,ADO,EGR2,NRBF2,JMJD1C,REEP3,CTNNA3,LRRTM3,DNAJC12	1
2_tFL	11	q12.2	59965964	59992861	1.19	0.03 Loss	2	MS4A5,MS4A1	0
2_tFL	12	p13.33--q13.11	20691	44818143	1.45	44.80 Loss (subclonal)	302		6
2_tFL	12	q13.11--q22	44823390	92281807	3.34	47.46 Gain	392		14
2_tFL	13	q31.1--q34	83104607	114126487	3.06	31.02 Gain	89		9
2_tFL	15	q11.2--q26.3	22214843	100286551	1.50	78.07 Loss (subclonal)	544		22
2_tFL	17	q23.3	59276217	59487457	4.07	0.21 Gain	10	CSH2,GH2,CSH1,CSHL1,GH1,CD79B,SCN4A,C17orf72,ICAM2,ERN1	0
2_tFL	18	p11.32--q23	1543	76116029	1.27	76.11 Loss	261		9
2_tFL	19	p13.3--p13.2	6350425	6929013	0.39	0.58 Loss homozygous	15		1
2_tFL	19	p13.11	17019694	17248858	1.16	0.23 Loss	7	HAUS8,MYO9B,USE1,OCEL1,NR2F6,USHBP1,BABAM1	0
2_tFL	19	p13.11	17252515	17877184	3.01	0.62 Gain	23		0
2_tFL	19	p13.11	17880872	18132412	1.09	0.25 Loss	6	CCDC124,KCNN1,ARRDC2,IL12RB1,MAST3,PIK3R2	0
2_tFL	19	q13.31--q13.32	49780311	50472787	1.08	0.69 Loss	24		0
2_tFL	19	q13.33	55066296	55271562	1.09	0.21 Loss	8	AKT1S1,TBC1D17,IL4I1,NUP62,ATF5,SIGLEC11,VRK3,ZNF473	0
2_tFL	22	q13.31--q13.33	43466634	49565354	1.03	6.10 Loss	60		3
4_FL	1	p36.22--p36.21	10039354	14790404	1.35	4.75 Loss	67		0
4_FL	1	q21.1--q44	142348955	247190999	2.50	104.84 Gain	940		28
4_FL	5	p15.33--q35.3	68520	180722914	2.59	180.65 Gain	840		31
4_FL	6	p25.3--p11.2	94649	57420704	2.76	57.33 Gain	591		8
4_FL	6	p11.2	57420839	58179324	1.51	0.76 Loss	1	PRIM2	0
4_FL	6	p11.2--q11.1	58190043	61944387	2.42	3.75 Gain	0		0
4_FL	6	q11.1--q27	61987967	170761408	1.38	108.77 Loss	413		9
4_FL	7	p22.3--q36.3	52899	158819753	2.52	158.77 Gain	870		34
4_FL	10	q21.2--q21.3	64584156	65488266	2.66	0.90 Gain	3	NRBF2,JMJD1C,REEP3	1
4_FL	11	q13.1--q25	65027479	134449982	2.53	69.42 Gain	555		12
4_FL	17	p13.3--p13.2	514	5774658	1.36	5.77 Loss	114		4
4_FL	17	q21.32--q21.33	42150374	45629759	2.46	3.48 Gain	67		4
4_tFL	1	p36.22--p36.21	10033010	14787734	1.28	4.75 Loss	67		0
4_tFL	1	q21.1--q44	142348955	247190999	2.58	104.84 Gain	940		28
4_tFL	3	q12.3	102957176	103070483	2.80	0.11 Gain	3	CEP97,FAM55C,NFKBIZ	0
4_tFL	5	p15.33--q35.3	68520	180722914	2.58	180.65 Gain	840		31
4_tFL	6	p25.3--p11.2	94649	57420704	2.59	57.33 Gain	591		8
4_tFL	6	p11.2	57656338	58190898	1.50	0.53 Loss	0		0
4_tFL	6	p11.2--p11.1	58248247	58882675	2.55	0.63 Gain	0		0
4_tFL	6	q11.1--q27	58885344	170761408	1.44	111.88 Loss	413		9
4_tFL	8	q24.23	138529051	138590751	3.58	0.06 Gain	0		0
4_tFL	9	p21.3	21958712	21994936	1.02	0.04 Loss homozygous	2	CDKN2A,CDKN2B	0
4_tFL	9	p21.3--p21.2	25160008	26962005	3.34	1.80 Gain	5	TUSC1,LOC100506422,C9orf82,PLAA,IFT74	0
4_tFL	10	q21.2--q21.3	64584156	65488266	2.66	0.90 Gain	3	NRBF2,JMJD1C,REEP3	1
4_tFL	13	q21.31--q34	63017528	114126487	2.44	51.11 Gain (subclonal)	114		9
5_FL	22	q11.1--q13.33	14432516	49581309	2.78	35.15 Gain	422		16
5_tFL	1	p36.33--p36.21	711153	13791415	1.35	13.08 Loss	185		5
5_tFL	2	p25.3--p21	2772	43319632	1.31	43.32 Loss	219		2
5_tFL	2	p21	43342784	44219284	4.36	0.88 Amp	7	THADA,PLEKHH2,LOC728819,DYNC2LI1,ABCG5,ABCG8,LRPPRC	0
5_tFL	2	p21--p16.1	44219339	55263681	2.63	11.04 Gain	45		1
5_tFL	2	p16.1	55268035	55303648	14.60	0.04 Amp	1	C2orf63	0
5_tFL	2	p16.1	55310260	60794324	2.68	5.48 Gain	12		3
5_tFL	2	p16.1--p15	60797388	62137157	19.35	1.34 Amp	12		0
5_tFL	2	p15	62139332	63466058	3.78	1.33 Gain	6	COMMD1,B3GNT2,TMEM17,EHBP1,OTX1,WDCP	0
5_tFL	4	p16.3--q35.2	56707	191261892	1.29	191.21 Loss	728		27

5_tFL	5	q14.3--q23.3	90815042	128695732	1.31	37.88	Loss	101		5
5_tFL	5	q23.3--q31.1	128695860	134837266	2.84	6.14	Gain	52		1
5_tFL	5	q33.3--q35.2	156216010	176480582	1.30	20.26	Loss	98		6
5_tFL	6	p25.3--q11.1	94649	63176412	2.72	63.08	Gain	592		8
5_tFL	6	q23.3--q24.1	137509836	139304400	0.75	1.79	Loss	13	<i>IL22RA2,IFNGR1,OLIG3,TNFAIP3,PERP,KIAA1244,PBOV1,HEBP2,NHSL1,LOC100507462</i>	0
5_tFL	6	q27	165371891	170892918	2.67	5.52	Gain	32		1
5_tFL	8	p23.3--q22.3	21242	104362829	2.66	104.34	Gain	470		15
5_tFL	8	q22.3--q24.3	104365262	146268947	3.42	41.90	Gain	191		16
5_tFL	9	p24.3--p24.1	36431	6460751	2.69	6.42	Gain	36		1
5_tFL	9	p24.1--p22.3	6645056	15970320	2.51	9.33	Gain	14	<i>KDM4C,C9orf123,PTPRD,TYRP1,LURAP1L,MPDZ,NFIB,ZDHHC21,CER1,FREM1</i>	0
5_tFL	9	p22.1--p21.3	19400343	23906120	0.76	4.51	Loss homozygous	27		2
5_tFL	9	p21.1	28284231	29067988	0.73	0.78	Loss homozygous	1	<i>LINGO2</i>	2
5_tFL	9	p21.1--p13.3	30135969	35671548	3.09	5.54	Gain	58		0
5_tFL	9	p13.3	35686625	35950464	1.42	0.26	Loss	13		0
5_tFL	9	p13.3--q12	35951915	69256641	3.31	33.30	Gain	41		1
5_tFL	11	p15.5--q25	188510	134449982	2.67	134.26	Gain	1254		25
5_tFL	13	q11--q21.1	17924937	56996265	2.70	39.07	Gain	190		5
5_tFL	13	q21.1--q21.2	56996566	60280776	3.32	3.28	Gain	3	<i>PCDH17,DIAPH3,TDRD3</i>	0
5_tFL	13	q21.2--q34	60281696	114126487	2.59	53.84	Gain	115		9
5_tFL	15	q13.3	29217243	30001535	1.32	0.78	Loss	4	<i>TRPM1,LOC283710,KLF13,OTUD7A</i>	0
5_tFL	15	q15.1	38013787	38565968	2.68	0.55	Gain	15		0
5_tFL	15	q15.1	39711752	40119732	1.39	0.41	Loss	8	<i>MGA,MAPKBP1,JMJD7,JMJD7-PLA2G4B,PLA2G4B,SPTBN5,EHD4,PLA2G4E</i>	1
5_tFL	15	q15.3--q21.1	42163400	43241439	1.40	1.08	Loss	14		0
5_tFL	15	q21.2	48342303	48564973	2.92	0.22	Gain	3	<i>HDC,GABPB1,USP8</i>	0
5_tFL	15	q21.2	49550541	49678990	1.19	0.13	Loss	1	<i>DMXL2</i>	0
5_tFL	15	q22.31	63100909	63127091	3.95	0.03	Gain	2	<i>MTFMT,OSTBETA</i>	0
5_tFL	15	q22.31	63127479	63425246	2.69	0.30	Gain	9	<i>OSTBETA,RASL12,KBTBD13,UBAP1L,PDCD7,CLPX,CILP,PARP16,IGDCC3</i>	0
5_tFL	15	q22.31	63900662	64080966	2.72	0.18	Gain	2	<i>RAB11A,MEGF11</i>	0
5_tFL	15	q22.31	64670844	64796483	2.72	0.13	Gain	1	<i>SMAD6</i>	0
5_tFL	15	q23	65593381	65715492	2.80	0.12	Gain	2	<i>C15orf61,MAP2K5</i>	0
5_tFL	15	q23	66100386	66149949	3.08	0.05	Gain	1	<i>PIAS1</i>	0
5_tFL	15	q23	66202797	66438234	2.69	0.24	Gain	5	<i>PIAS1,CALML4,CLN6,FEM1B,ITGA11</i>	0
5_tFL	15	q23	66793742	67097110	2.73	0.30	Gain	4	<i>CORO2B,ANP32A,NOX5,SPESP1</i>	0
5_tFL	15	q23	67655743	68088391	2.65	0.43	Gain	0		0
5_tFL	15	q23	68608147	68690952	2.64	0.08	Gain	0		0
5_tFL	15	q23	69790215	69936254	2.77	0.15	Gain	3	<i>THSD4,NR2E3,MYO9A</i>	0
5_tFL	15	q24.1	72013955	72082137	2.83	0.07	Gain	3	<i>LOXL1,STOML1,PML</i>	0
5_tFL	15	q24.1--q24.3	72121790	75812523	1.37	3.69	Loss	51	<i>PML,GOLGA6A,ISLR2,ISLR,STRA6,CCDC33,CYP11A1,SEMA7A,UBL7,ARID3B</i>	1
5_tFL	15	q24.3	75813380	75987374	2.66	0.17	Gain	0		0
5_tFL	15	q25.1	77986304	78245267	2.90	0.26	Loss	5	<i>ST20-MTHFS,ST20,BCL2A1,ZFAND6,FAH</i>	0
5_tFL	15	q25.1--q25.2	79209430	81097261	1.33	1.89	Loss	12	<i>C15orf26,IL16,STARD5,TCM3,MEX3B,EFTUD1,FAM154B,GOLGA6L10,GOLGA6L9,RPS17</i>	0
5_tFL	15	q25.2--q25.3	81265131	85095259	1.34	3.83	Loss	23		1
5_tFL	15	q26.1	87709843	88510380	1.33	0.80	Loss	14		1
5_tFL	15	q26.1	88524553	88637300	3.20	0.11	Gain	5	<i>SEMA4B,CIB1,C15orf58,TTL13,NGRN</i>	0
5_tFL	15	q26.1--q26.3	91214598	98490070	1.31	7.28	Loss	15		1
5_tFL	15	q26.3	99810927	100106129	2.58	0.30	Gain	3	<i>PCSK6,TM2D3,TARSL2</i>	0
5_tFL	17	p13.3	514	1914251	3.63	1.91	Gain	34		3
5_tFL	17	p13.3--p13.2	3380422	3680596	3.97	0.30	Gain	10	<i>TRPV3,TRPV1,SHPK,CTNS,TAX1BP3,TMEM93,P2RX5,ITGAE,GS2,C17orf85</i>	0
5_tFL	17	p13.2--p13.1	3682946	7902899	1.38	4.22	Loss	126		3
5_tFL	17	p13.1	7907010	8243230	3.74	0.34	Gain	17		0
5_tFL	17	p12--p11.2	13743140	16716207	3.50	2.97	Gain	24		1
5_tFL	17	p11.2--q11.2	18520434	24711800	3.57	6.19	Gain	75		3
5_tFL	17	q21.2--q21.31	37222227	40604364	3.72	3.38	Gain	101		0
5_tFL	17	q21.33--q23.2	46461092	57895029	3.55	11.43	Gain	75		4
5_tFL	17	q24.3--q25.1	67882786	68475682	3.51	0.59	Gain	1	<i>SLC39A11</i>	0

5_tFL	17	q25.1--q25.3	71520806	78643088	3.70	7.12 Gain	124		4
6_FL	3	p26.3--p14.1	35333	199322963	2.29	199.29 Gain (subclonal)	1038		31
6_FL	18	p11.32--q23	1543	76116029	2.31	76.11 Gain	261		9
6_FL	21	p11.2--q22.3	9955079	46921373	2.25	36.97 Gain (subclonal)	216		5
6_tFL	1	p36.33--p35.3	554484	28161477	2.98	27.61 Gain	381		9
6_tFL	1	p34.2	41737683	43394115	1.19	1.66 Loss	19		0
6_tFL	1	p33--p13.3	50468014	107555765	1.19	57.09 Loss	257		6
6_tFL	1	p13.2	112642101	112675419	2.92	0.03 Gain	0		0
6_tFL	1	p13.2	112685647	115721406	4.03	3.04 Gain	31		0
6_tFL	1	p13.2--p13.1	115722141	115956354	2.66	0.23 Gain	0		0
6_tFL	1	q32.1	204672701	205116828	3.06	0.44 Gain	9	SRGAP2,IKBKE,RASSF5,EIF2D,DYRK3,MAPKAPK2,IL10,IL19,IL20	0
6_tFL	1	q44	247034179	247190999	3.00	0.16 Gain	4	SH3BP5L,ZNF672,ZNF692,PGBD2	0
6_tFL	2	p25.3--p21	2772	45612546	1.17	45.61 Loss	232		2
6_tFL	2	p21--p16.1	45955995	57518025	1.22	11.56 Loss	46		4
6_tFL	2	p16.1	57522808	60419483	3.17	2.90 Gain	2	VRK2,FANCL	0
6_tFL	2	p16.1--p14	60419635	64107245	4.39	3.69 Gain	21		0
6_tFL	2	p11.2--q31.2	86312760	178494370	3.01	92.18 Gain	385		7
6_tFL	2	q31.2	178495054	180342303	5.12	1.85 Amp	11		0
6_tFL	2	q31.3--q32.1	182682781	182764243	4.26	0.08 Amp	2	PPP1R1C,PDE1A	0
6_tFL	3	p14.2	60429806	60530314	1.11	0.10 Loss	1	FHIT	0
6_tFL	4	p16.3--p16.1	34101	5851055	1.18	5.82 Loss	66		2
6_tFL	4	p16.1--p15.2	5855598	25005103	2.92	19.15 Gain	78		5
6_tFL	4	p15.2	25009755	26439906	1.25	1.43 Loss	7	ANAPC4,SLC34A2,SEL1L3,C4orf52,RBPJ,CCKAR,TBC1D19	0
6_tFL	4	p15.2	26448558	26629150	3.34	0.18 Gain	1	STIM2	0
6_tFL	4	p15.1	29343024	33257915	3.46	3.91 Gain	1	PCDH7	0
6_tFL	4	p14	37423948	40261972	1.26	2.84 Loss	23		1
6_tFL	4	p13--p11	41249260	48788531	1.27	7.54 Loss	33		0
6_tFL	4	q12	52591408	55129585	1.28	2.54 Loss	11		0
6_tFL	4	q12	56285192	59140344	1.41	2.86 Loss	15		0
6_tFL	4	q12--q13.1	59141182	60084277	2.44	0.94 Gain (subclonal)	0		0
6_tFL	4	q13.1	60084829	66082434	1.50	6.00 Loss	3	LPHN3,TECRL,EPHA5	0
6_tFL	4	q21.22	83100752	83243453	1.57	0.14 Loss (subclonal)	0		0
6_tFL	4	q21.22--q21.23	83243498	84982568	1.25	1.74 Loss	16		1
6_tFL	4	q21.23	84982681	85114932	1.69	0.13 Loss (subclonal)	0		0
6_tFL	4	q22.1	89078055	89955588	1.17	0.88 Loss	10	SPP1,PKD2,ABCG2,PPMIK,HERC6,HERC5,PIGY,HERC3,NAP1L5,FAM13A	0
6_tFL	4	q22.1	90254572	91387619	1.12	1.13 Loss	5	TIGD2,GPRIN3,SNCA,MMRN1,FAM190A	0
6_tFL	4	q22.1	91387825	92359308	1.53	0.97 Loss (subclonal)	1	FAM190A	0
6_tFL	4	q22.2--q28.3	94760264	131815411	1.21	37.06 Loss	129		12
6_tFL	4	q28.3	134454740	134638902	1.11	0.18 Loss	0		0
6_tFL	4	q31.3--q32.1	153478187	155113451	0.49	1.64 Loss homozygous	11		0
6_tFL	4	q32.1--q35.2	155519607	191261892	2.89	35.74 Gain	121		3
6_tFL	5	q14.3	85038663	86110864	3.13	1.07 Gain	1	COX7C	0
6_tFL	5	q15	95408905	95805226	1.20	0.40 Loss	1	PCSK1	1
6_tFL	6	p21.32	32491388	32978035	3.20	0.49 Gain	12		0
6_tFL	6	p21.2	36808415	37003438	3.05	0.20 Gain	3	CPNE5,PPIL1,C6orf89	0
6_tFL	7	p22.3--q22.1	52899	101071285	3.10	101.02 Gain	531		18
6_tFL	7	q22.1	101111394	101654461	1.20	0.54 Loss	1	CUX1	0
6_tFL	7	q22.1--q36.3	101654986	158819753	3.13	57.16 Gain	339		16
6_tFL	8	p23.3	21242	1059260	3.13	1.04 Gain	5	OR4F21,ZNF596,FBXO25,C8orf42,ERICH1	0
6_tFL	8	p22	16599436	18519745	3.73	1.92 Gain	15		0
6_tFL	8	p22	18519809	18980218	6.02	0.46 Amp	1	PSD3	0
6_tFL	8	p22	18980530	18999082	3.56	0.02 Gain	0		0
6_tFL	8	p11.21--p11.1	41214243	43905947	6.39	2.69 Amp	24		1
6_tFL	8	p11.1--q11.21	43943193	48599469	2.86	4.66 Gain	1	KIAA0146	0
6_tFL	8	q11.22--q12.1	52109252	55701828	3.61	3.59 Gain	14		0

6_tFL	8	q12.1	55994820	56245378	4.34	0.25 Gain	1	<i>XKR4</i>	0
6_tFL	8	q12.1	56248055	57721111	6.68	1.47 Amp	10	<i>XKR4,TMEM68,TGS1,LYN,RPS20,MOS,PLAG1,CHCHD7,SDR16C5,PENK</i>	0
6_tFL	8	q12.1	57722274	58259080	4.10	0.54 Gain	1	<i>IMPAD1</i>	0
6_tFL	8	q12.3--q13.2	63952402	69134947	4.18	5.18 Gain	28		1
6_tFL	8	q21.2--q21.3	85878377	90358267	4.48	4.48 Gain	19		0
6_tFL	8	q21.3	90358442	92649112	5.34	2.29 Amp	12		0
6_tFL	8	q21.3	92649307	93338298	3.02	0.69 Gain	1	<i>RUNX1T1</i>	0
6_tFL	8	q22.1	94418334	94553718	1.21	0.14 Loss	0		0
6_tFL	8	q22.1	94558092	96726270	4.33	2.17 Gain	17		0
6_tFL	8	q22.1	97165066	97659486	4.00	0.49 Gain	5	<i>GDF6,UQCRB,MTERFD1,PTDSS1,SDC2</i>	0
6_tFL	8	q22.1	97661533	98145269	6.15	0.48 Amp	2	<i>SDC2,PGCP</i>	0
6_tFL	8	q22.1	98145570	98220242	4.04	0.07 Gain	1	<i>PGCP</i>	0
6_tFL	8	q23.1	107505652	108962399	3.99	1.46 Gain	3	<i>OXR1,ABRA,ANGPT1</i>	0
6_tFL	8	q23.1--q23.2	108962614	111319226	6.18	2.36 Amp	11		0
6_tFL	8	q23.2--q23.3	111322431	115585151	2.96	4.26 Gain	1	<i>CSMD3</i>	1
6_tFL	8	q23.3	115591568	117357849	1.17	1.77 Loss	1	<i>TRPS1</i>	0
6_tFL	8	q23.3--q24.12	117376220	120216311	3.03	2.84 Gain	10	<i>EIF3H,UTP23,RAD21,C8orf85,SLC30A8,MED30,EXT1,SAMD12,TNFRSF11B,COLEC10</i>	0
6_tFL	8	q24.12	120228052	120574143	1.21	0.35 Loss	2	<i>MAL2,NOV</i>	0
6_tFL	8	q24.13--q24.21	127210469	127848962	1.20	0.64 Loss	1	<i>FAM84B</i>	0
6_tFL	8	q24.21	127849890	129142628	3.33	1.29 Gain	2	<i>POU5F1B,MYC</i>	4
6_tFL	8	q24.21	129142657	129383064	5.63	0.24 Amp	0		1
6_tFL	8	q24.21	129383948	129969200	3.94	0.59 Gain	0		0
6_tFL	8	q24.21--q24.22	129970051	136294602	1.20	6.32 Loss	17		2
6_tFL	8	q24.23--q24.3	139842825	140253003	1.14	0.41 Loss	1	<i>COL22A1</i>	0
6_tFL	8	q24.3	140254814	143854267	2.86	3.60 Gain	20		2
6_tFL	8	q24.3	143857916	146268947	1.21	2.41 Loss	77		4
6_tFL	9	p24.3--p24.2	36431	3930874	1.08	3.89 Loss	14		0
6_tFL	9	p24.1--p23	5519757	10032538	1.18	4.51 Loss	13		0
6_tFL	9	p22.3	14432045	15169123	1.24	0.74 Loss	4	<i>ZDHHC21,CER1,FREM1,TTC39B</i>	0
6_tFL	9	p22.2--p21.1	17724224	32749617	1.17	15.03 Loss	52		4
6_tFL	9	p13.3--p13.2	34319110	37229559	1.24	2.91 Loss	58		0
6_tFL	9	p13.2--q21.13	37579870	73571800	1.20	35.99 Loss	48		2
6_tFL	9	q21.13	76632190	76664924	1.10	0.03 Loss	1	<i>TRPM6</i>	0
6_tFL	9	q21.13	77599612	77655534	5.76	0.06 Amp	0		0
6_tFL	9	q21.13	78697748	78846853	1.19	0.15 Loss	2	<i>PRUNE2,FOXB2</i>	0
6_tFL	9	q21.13--q21.2	78854467	79629962	2.86	0.78 Gain	3	<i>VPS13A,GNA14,GNAQ</i>	0
6_tFL	9	q21.2	79725157	79980601	2.92	0.26 Gain	1	<i>GNAQ</i>	0
6_tFL	9	q21.31--q21.32	82368317	85050591	1.16	2.68 Loss	6	<i>TLE1,FAM75D4,FAM75D3,FAM75D1,RASEF,FRMD3</i>	0
6_tFL	9	q21.32--q22.33	85907946	98761801	1.19	12.85 Loss	72		6
6_tFL	9	q22.33	98767762	99252224	2.95	0.48 Gain	3	<i>CTSL2,C9orf174,TDRD7</i>	1
6_tFL	9	q22.33--q31.1	99260494	101756406	4.11	2.50 Amp	22		0
6_tFL	9	q31.1	101757089	106447095	1.19	4.69 Loss	26		0
6_tFL	9	q31.1--q31.2	106448604	107676309	5.28	1.23 Amp	9	<i>OR13D1,NIPSNAP3A,NIPSNAP3B,ABCA1,SLC44A1,FSD1L,FKTN,TAL2,TMEM38B</i>	0
6_tFL	9	q31.2	107681781	108769072	1.17	1.09 Loss	1	<i>ZNF462</i>	0
6_tFL	9	q31.2--q31.3	108769698	112217145	4.33	3.45 Amp	19		1
6_tFL	9	q31.3--q33.1	112217644	119653835	1.22	7.44 Loss	53		1
6_tFL	9	q33.1--q33.2	119656299	124704812	4.17	5.05 Gain	36		1
6_tFL	9	q33.2	124705467	125478430	5.28	0.77 Amp	8	<i>RC3H2,ZBTB6,ZBTB26,RABGAP1,GPR21,STRBP,CRB2,DENND1A</i>	2
6_tFL	9	q33.3--q34.11	128077207	131727134	1.22	3.65 Loss	74		2
6_tFL	9	q34.3	137096364	138337594	1.17	1.24 Loss	17		0
6_tFL	9	q34.3	140011350	140174582	1.17	0.16 Loss	1	<i>CACNA1B</i>	0
6_tFL	10	q22.3--q24.32	78232825	103715801	1.27	25.48 Loss	185		5
6_tFL	11	p15.5--q13.1	188510	63227464	2.88	63.04 Gain	635		9
6_tFL	11	q14.1	76747438	80077309	3.05	3.33 Gain	16		1
6_tFL	11	q14.1	80078742	80420865	4.02	0.34 Gain	0		0

6_tFL	11	q14.1	80421879	84834686	3.14	4.41 Gain	8	<i>FAM181B,PRCP,C11orf82,RAB30,PCF11,ANKRD42,CCDC90B,DLG2</i>	0
6_tFL	11	q14.1	84837995	85203217	4.28	0.37 Gain	6		0
6_tFL	11	q14.1	85203531	85289461	5.96	0.09 Amp	1		0
6_tFL	11	q14.1–q14.2	85291901	85392091	4.12	0.10 Gain	2		0
6_tFL	11	q14.2	85396921	86805271	3.08	1.41 Gain	8		0
6_tFL	11	q14.3–q22.1	88590980	100205385	2.88	11.61 Gain	47		3
6_tFL	11	q22.3–q23.2	108882276	113342641	3.01	4.46 Gain	38		2
6_tFL	11	q23.3–q25	118230794	134449982	1.21	16.22 Loss	136		3
6_tFL	14	q21.1–q21.2	37132187	41573082	1.38	4.44 Loss	11		0
6_tFL	14	q32.33	104268612	105270409	2.97	1.00 Gain	21		0
6_tFL	14	q32.33	105708528	106356482	4.37	0.65 Gain	0		0
6_tFL	15	q24.2–q26.3	74045568	100222647	1.22	26.18 Loss	157		7
6_tFL	16	p13.3–p13.13	765	10777075	1.17	10.78 Loss	204		3
6_tFL	16	q12.1	46725075	47418431	1.20	0.69 Loss	5	<i>ABCC12,ABCC11,LONP2,SIAH1,N4BP1</i>	0
6_tFL	17	p13.3–p11.2	514	16319164	1.23	16.32 Loss	260		10
6_tFL	17	p11.2	16322726	16349340	2.84	0.03 Gain	1		0
6_tFL	17	p11.2	16351923	17432785	1.20	1.08 Loss	13	<i>FAM211A</i>	0
6_tFL	18	p11.32–p11.31	1543	3927121	1.18	3.93 Loss	20		0
6_tFL	18	p11.31–p11.23	4306883	7581692	1.13	3.27 Loss	11		0
6_tFL	18	p11.23–p11.21	8041780	10957379	2.83	2.92 Gain	14		0
6_tFL	18	p11.21	12854013	14826834	1.36	1.97 Loss	11		0
6_tFL	18	p11.21–q11.2	15316729	20586952	2.98	5.27 Gain	21		4
6_tFL	18	q11.2–q12.1	21173794	23618309	3.10	2.44 Gain	7	<i>ZNF521,SS18,PSMA8,TAF4B,KCTD1,AQP4,CHST9</i>	0
6_tFL	18	q12.1–q12.2	27192062	33824668	2.99	6.63 Gain	32		1
6_tFL	18	q12.3	37380284	38209827	2.88	0.83 Gain	1		0
6_tFL	18	q12.3	39193740	39633740	2.88	0.44 Gain	0	<i>PIK3C3</i>	0
6_tFL	18	q12.3–q21.2	39972685	47700424	2.92	7.73 Gain	41		1
6_tFL	18	q21.2–q21.32	49221250	56382743	3.08	7.16 Gain	29		1
6_tFL	18	q21.33	57164223	58943607	2.97	1.78 Gain	8	<i>CDH20,RNF152,PIGN,KIAA1468,TNFRSF11A,ZCCHC2,PHLPP1,BCL2</i>	0
6_tFL	21	q11.2–q22.3	13492143	46921373	2.89	33.43 Gain	210		5
6_tFL	22	q11.1–q13.33	14886448	49581309	1.22	34.69 Loss	420		16
8_FL	2	q36.2–q37.3	225525912	240003841	1.78	14.48 Loss (subclonal)	99		3
8_FL	5	p15.33–q35.3	68520	180722914	2.32	180.65 Gain (subclonal)	840		31
8_FL	7	p22.3–p22.2	52899	158819753	2.31	158.77 Gain (subclonal)	870		34
8_FL	8	q23.3	113646896	114258280	2.37	0.61 Gain (subclonal)	1	<i>CSMD3</i>	1
8_FL	8	q23.3–q24.11	116559070	118083857	2.31	1.52 Gain (subclonal)	6		0
8_FL	8	q24.12	119219639	119716899	3.89	0.50 Amp	1		0
8_FL	8	q24.13	125121021	126211585	3.19	1.09 Gain	11	<i>SAMD12</i>	0
8_FL	8	q24.13–q24.21	126212968	129227418	4.87	3.01 Amp	5		4
8_FL	8	q24.21–q24.22	129228127	131890351	3.23	2.66 Gain	4		1
8_FL	8	q24.22–q24.3	131890648	140655995	2.27	8.77 Gain (subclonal)	17		0
8_FL	9	p24.3	78018	140211203	2.27	140.13 Gain (subclonal)	753		26
8_FL	12	p13.33–p13.32	120500	4664536	2.63	4.54 Gain	37		0
8_FL	12	p13.32–q21.31	4664612	80206652	2.30	75.54 Gain (subclonal)	626		20
8_FL	17	p13.3–p12	514	12534536	1.37	12.53 Loss	236		8
8_FL	18	p11.32–q23	1543	76116029	2.30	76.11 Gain (subclonal)	261		9
8_FL	19	p13.3–p13.2	2705548	8223780	1.36	5.52 Loss	145		3
8_FL	19	p13.2	8364960	9751957	2.20	1.39 Gain (subclonal)	33		0
8_FL	19	p13.2–p13.12	11576502	14256257	2.53	2.68 Gain	72		5
8_FL	21	p11.2–q11.2	9758730	46921373	1.76	37.16 Loss (subclonal)	216		5
8_tFL	1	q21.1–q25.2	144308838	177614304	2.50	33.31 Gain	481		14
8_tFL	1	q32.1	197513430	202781620	2.69	5.27 Gain	63		1
8_tFL	2	p25.3–q24.1	26850	157399467	2.74	157.37 Gain	716		11
8_tFL	2	q36.2–q37.3	225523323	239996326	1.26	14.47 Loss	99		3
8_tFL	3	p12.3–q29	77747465	199380503	3.36	121.63 Gain	563		17

8_tFL	4	q13.2--q22.3	70079275	97479378	2.80	27.40 Gain	152		1
8_tFL	4	q22.3	97480485	97630205	4.23	0.15 Gain	0		0
8_tFL	4	q22.3	97633413	97859551	7.57	0.23 Amp	0		0
8_tFL	4	q22.3--q23	97863710	99577022	3.86	1.71 Gain	2	<i>C4orf37,RAP1GDS1</i>	0
8_tFL	4	q23--q24	99585104	104235014	4.99	4.65 Amp	29		1
8_tFL	5	p15.33--q35.3	68520	180508151	2.77	180.44 Gain	835		31
8_tFL	7	p22.3--p22.2	52899	158819753	2.72	158.77 Gain	870		34
8_tFL	8	q23.3	113673156	114255179	3.20	0.58 Gain	1	<i>CSMD3</i>	1
8_tFL	8	q23.3--q24.11	116559070	118086804	2.75	1.53 Gain	6	<i>TRPS1,EIF3H,UTP23,RAD21,C8orf85,SLC30A8</i>	0
8_tFL	8	q24.12	119219639	119727758	5.40	0.51 Amp	1	<i>SAMD12</i>	0
8_tFL	8	q24.12--q24.13	119727919	124890898	2.61	5.16 Gain	26		1
8_tFL	8	q24.13	125121764	126211585	3.58	1.09 Gain	11		0
8_tFL	8	q24.13--q24.21	126212968	129227418	7.09	3.01 Amp	5	<i>NSMCE2,TRIB1,FAM84B,POU5F1B,MYC</i>	4
8_tFL	8	q24.21--q24.22	129228127	131890074	4.18	2.66 Gain	4	<i>GSDMC,FAM49B,ASAP1,ADCY8</i>	1
8_tFL	8	q24.22--q24.3	131890351	146180016	2.52	14.29 Gain	113		8
8_tFL	9	p24.3--p22.3	263651	16242358	2.72	15.98 Gain	48		1
8_tFL	9	p21.1--q12	30416131	68809437	3.49	38.39 Gain	113		1
8_tFL	9	q31.3--q32	113743818	113905003	1.46	0.16 Loss	1	<i>SUSD1</i>	0
8_tFL	9	q33.3--q34.11	126150783	132025814	1.23	5.88 Loss	93		4
8_tFL	12	p13.33--p13.32	62481	4643571	2.47	4.58 Gain (subclonal)	37		0
8_tFL	12	p12.1--q21.31	26206245	80124268	2.75	53.92 Gain	421		14
8_tFL	16	p13.3--p11.2	50690	30432689	2.59	30.38 Gain	381		8
8_tFL	17	p13.3--p12	514	12528539	1.19	12.53 Loss	236		8
8_tFL	17	q21.33--q25.3	45184577	78643088	2.54	33.46 Gain	363		12
8_tFL	18	p11.32--q23	1543	76116029	2.60	76.11 Gain	261		9
8_tFL	19	p13.3--p13.2	2679577	8502621	1.26	5.82 Loss	157		3
8_tFL	19	p13.2	8706618	9751957	2.49	1.05 Gain (subclonal)	24		0
8_tFL	19	p13.2--p13.12	11588503	14327156	3.10	2.74 Gain	72		5
8_tFL	19	p13.12--p13.11	15706852	16121763	2.86	0.41 Gain	9	<i>OR10H3,OR10H5,OR10H1,CYP4F2,CYP4F11,OR10H4,TPM4,RAB8A,HS2D</i>	0
8_tFL	19	q13.2--q13.43	47299114	63789654	2.54	16.49 Gain	531		60
8_tFL	20	p12.3	8284712	8712779	1.48	0.43 Loss	1	<i>PLCB1</i>	0
8_tFL	21	q11.2--q22.3	14326373	46921373	1.43	32.60 Loss	209		5
11_FL	1	p36.32--p36.31	2438154	6830069	1.08	4.39 Loss	41		1
11_FL	2	p22.2--q13	37926735	112347492	2.41	74.42 Gain (subclonal)	381		5
11_FL	2	q13--q22.1	112348785	137956768	1.53	25.61 Loss	110		3
11_FL	2	q22.1--q37.3	137957549	242738117	2.46	104.78 Gain (subclonal)	520		16
11_FL	4	p16.3--p15.1	2269	32777991	1.11	32.78 Loss	151		7
11_FL	4	p14--p13	40692738	42621243	1.02	1.93 Loss	11		0
11_FL	4	p13	44318934	44715492	1.38	0.40 Loss (subclonal)	3	<i>YIPF7,GUF1,GNPDA2</i>	0
11_FL	4	q31.22	146641559	146677742	0.94	0.04 Loss	1	<i>SMAD1</i>	0
11_FL	6	p24.3	10175989	10421650	1.03	0.25 Loss	0		0
11_FL	6	p21.32--p11.1	33233462	58882675	1.49	25.65 Loss (subclonal)	242		5
11_FL	6	q11.1--q27	58885344	170824240	1.07	111.94 Loss	413		9
11_FL	7	p22.3--q36.3	52899	158809998	2.72	158.76 Gain	870		34
11_FL	10	q21.1	56127179	56275656	1.13	0.15 Loss	1	<i>PCDH15</i>	0
11_FL	12	q13.11--q21.1	45916744	70430909	2.72	24.51 Gain	326		11
11_FL	12	q24.23--q24.31	117283938	120936241	1.10	3.65 Loss	46		1
11_FL	13	q11--q21.1	17924937	54777792	1.14	36.85 Loss	185		5
11_FL	13	q21.1	54778512	56201460	0.43	1.42 Loss homozygous	0		0
11_FL	13	q21.1--q31.1	56203402	84424431	1.09	28.22 Loss	35		0
11_FL	13	q31.1--q34	84429112	114126487	1.45	29.70 Loss (subclonal)	88		9
11_FL	14	q11.1--q23.1	18072112	59317037	1.47	41.24 Loss (subclonal)	278		4
11_FL	15	q13.2--q21.3	28036451	52475219	1.05	24.44 Loss	194		8
11_FL	16	p13.3	3571192	3721314	1.14	0.15 Loss	4	<i>SLX4,DNASE1,TRAP1,CREBBP</i>	0
11_FL	16	q23.1	77653349	77779788	2.56	0.13 Gain	1	<i>WWOX</i>	0

11_FL	17	p13.3--p12	514	13449857	1.49	13.45 Loss (subclonal)	239		9
11_FL	19	p13.3	3181135	6417745	1.60	3.24 Loss (subclonal)	84		2
11_FL	19	p13.3--p13.2	6420835	7163256	0.49	0.74 Loss homozygous	17		1
11_FL	19	p13.2	7167771	11023109	1.60	3.86 Loss (subclonal)	110		4
11_tFL	1	p36.32--p36.31	2406272	6830069	1.20	4.42 Loss	42		1
11_tFL	2	p16.1	56624127	57367350	2.91	0.74 Gain	0		0
11_tFL	2	p16.1--p14	57369552	67421596	3.85	10.05 Gain	33		0
11_tFL	3	p23--p22.3	31959517	32130052	1.11	0.17 Loss	3	<i>OSBPL10,ZNF860,GPD1L</i>	0
11_tFL	3	q27.3--q29	188955156	199380503	3.52	10.43 Gain	62		4
11_tFL	5	p14.1	29112217	29131701	1.31	0.02 Loss	0		0
11_tFL	6	q14.1--q27	79283088	170824447	1.21	91.54 Loss	378		7
11_tFL	7	p22.3--q36.3	52899	158810147	3.00	158.76 Gain	870		34
11_tFL	8	q23.3--q24.21	112498759	129199218	4.89	16.70 Amp	51		6
11_tFL	8	q24.21--q24.3	129202441	146268947	2.95	17.07 Gain	117		9
11_tFL	9	p21.3--p21.1	21616157	29477049	0.55	7.86 Loss homozygous	17		2
11_tFL	10	q26.2	128333557	128449569	2.66	0.12 Gain	0		0
11_tFL	11	p15.5--q25	188510	134449982	2.82	134.26 Gain	1254		25
11_tFL	12	q13.11--q21.1	45916744	70430909	3.03	24.51 Gain	326		11
11_tFL	12	q24.11	109450701	109572340	1.15	0.12 Loss	4	<i>RAD9B,PPTC7,TCTN1,HVCN1</i>	0
11_tFL	12	q24.23--q24.31	117288979	120985827	1.18	3.70 Loss	47		1
11_tFL	13	q21.1	54777792	55952351	1.26	1.17 Loss	0		0
11_tFL	13	q21.1	56205133	57098557	1.26	0.89 Loss	5	<i>PRR20A,PRR20B,PRR20C,PRR20D,PRR20E</i>	0
11_tFL	14	q31.3	84077293	84702469	1.22	0.63 Loss	0		0
11_tFL	16	p13.3	3524488	3721314	1.26	0.20 Loss	6	<i>CLUAP1,NLRC3,SLX4,DNASE1,TRAP1,CREBBP</i>	0
11_tFL	19	p13.3--p13.2	6420835	7163256	0.57	0.74 Loss homozygous	17		1
12_FL	1	q21.1--q43	143998992	240165038	2.43	96.17 Gain	861		28
12_FL	2	p11.2	87163333	88912170	1.40	1.75 Loss	12		0
12_tFL	1	p34.3	36305921	36401066	3.93	0.10 Gain	5	<i>TEKT2,ADPRHL2,COL8A2,TRAPPC3,MAP7D1</i>	0
12_tFL	1	q21.1--q44	142719660	247190999	3.10	104.47 Gain	939		28
12_tFL	2	p11.2	87163333	88914734	1.21	1.75 Loss	12		0
12_tFL	8	p23.3--q24.3	21242	146268947	3.09	146.25 Gain	661		31
12_tFL	9	p21.3	21958712	22047218	0.27	0.09 Loss homozygous	2	<i>CDKN2A,CDKN2B</i>	0
12_tFL	19	p13.11	19028998	19238510	1.21	0.21 Loss	11		0
17_FL	1	p36.33--p36.23	51586	8449229	1.60	8.40 Loss	108		5
17_FL	2	p25.3--p16.1	2772	60498580	1.61	60.50 Loss	280		6
17_FL	2	p16.1	60503985	60911866	6.56	0.41 Amp	2	<i>BCL11A,PAPOLG</i>	0
17_FL	2	p16.1--p15	60912514	61510340	21.06	0.60 Amp	7	<i>REL,PUS10,PEX13,KIAA1841,C2orf74,AHSA2,USP34</i>	0
17_FL	2	p15	61512703	61529131	7.41	0.02 Amp	1	<i>USP34</i>	0
17_FL	2	p15	61529183	61563164	16.55	0.03 Amp	2	<i>USP34,XPO1</i>	0
17_FL	2	p15	61565811	62899802	6.75	1.33 Amp	7	<i>XPO1,FAM161A,CCT4,COMMD1,B3GNT2,TMEM17,EHBP1</i>	0
17_FL	2	p15	62903885	63201684	13.66	0.30 Amp	2	<i>EHBP1,OTX1</i>	0
17_FL	2	p15--p14	63202197	64929516	7.64	1.73 Amp	8	<i>WDPCP,MDH1,UGP2,VPS54,PELI1,LGALS1,AFTPH,SERTAD2</i>	0
17_FL	2	p14	64931251	65903456	4.67	0.97 Amp	5	<i>SLC1A4,CEP68,RAB1A,ACTR2,SPRED2</i>	0
17_FL	3	p26.3--p25.2	35333	11752612	1.58	11.72 Loss	55		1
17_FL	3	p25.2	11753256	11926756	3.28	0.17 Gain	1	<i>TAMM41</i>	0
17_FL	3	p25.2--p25.1	12061585	13403413	2.82	1.34 Gain	12		0
17_FL	3	p25.1	13403581	14242747	3.87	0.84 Gain	8	<i>NUP210,HDAC11,FBLN2,WNT7A,CHCHD4,TMEM43,XPC,LSM3</i>	0
17_FL	3	p25.1--p24.2	14244887	23979804	2.61	9.73 Gain	36		1
17_FL	3	q13.31	116564819	117712345	2.53	1.15 Gain	2	<i>GAP43,LSAMP</i>	0
17_FL	4	q21.23--q27	86729459	121031361	2.51	34.30 Gain	126		11
17_FL	5	q14.2--q14.3	82238160	82822305	1.53	0.58 Loss	3	<i>TMEM167A,XRCC4,VCAN</i>	0
17_FL	5	q14.3	82828276	91714851	2.72	8.89 Gain	14		1
17_FL	5	q14.3--q15	91717312	93717023	1.46	2.00 Loss	4	<i>NR2F1,FAM172A,POU5F2,KIAA0825</i>	0
17_FL	5	q15	93902820	94200708	1.43	0.30 Loss	3	<i>KIAA0825,ANKRD32,MCTP1</i>	1
17_FL	5	q15--q21.1	94201137	102273856	2.77	8.07 Gain	24		2

17_FL	5	q21.1	102277763	102400494	1.55	0.12 Loss	1	PAM	0
17_FL	5	q21.2--q21.3	103844607	105245391	1.44	1.40 Loss	0		0
17_FL	5	q21.3--q23.1	108050988	117434692	1.57	9.38 Loss	33		1
17_FL	5	q23.1--q23.2	119118957	125960461	1.56	6.84 Loss	16		0
17_FL	5	q23.2	125960771	126817025	2.91	0.86 Gain	6	PHAX,C5orf48,LMNB1,MARCH3,C5orf63,MEGF10	0
17_FL	5	q23.2--q23.3	126817326	127795471	1.60	0.98 Loss	5	MEGF10,PRRC1,CTXN3,SLC12A2,FBN2	0
17_FL	5	q32--q33.1	147182297	147389198	1.55	0.21 Loss	3	SPINK1,SCGB3A2,C5orf46	0
17_FL	5	q33.1	147391054	147461578	2.76	0.07 Gain	1	SPINK5	0
17_FL	5	q33.1	147464067	147568360	1.58	0.10 Loss	3	SPINK5,SPINK14,SPINK6	0
17_FL	5	q33.1	147569649	151494364	2.79	3.92 Gain	50		4
17_FL	5	q33.2--q34	152534449	162883105	1.73	10.35 Loss (subclonal)	50		3
17_FL	5	q34--q35.3	162885492	180722914	2.44	17.84 Gain (subclonal)	125		7
17_FL	6	p25.3--p21.1	94649	44609851	2.77	44.52 Gain	520		5
17_FL	6	q14.1	77173069	77813182	2.64	0.64 Gain	0		0
17_FL	6	q14.1	77813522	77950530	3.39	0.14 Gain	0		0
17_FL	6	q14.1	77954621	78172987	1.28	0.22 Loss	0		0
17_FL	6	q14.1	78173007	78329805	3.85	0.16 Gain	1	HTR1B	0
17_FL	6	q14.1	78330112	79116626	2.97	0.79 Gain	0		0
17_FL	6	q14.1	79116688	79167788	5.39	0.05 Amp	0		0
17_FL	6	q14.1	79169752	79197675	9.31	0.03 Amp	0		0
17_FL	6	q14.1--q22.31	79199573	125528048	1.46	46.33 Loss	172		3
17_FL	6	q22.31--q22.32	125531900	126751146	2.67	1.22 Gain	7	TPD52L1,HDDC2,HEY2,NCOA7,HINT3,TRMT11,CENPW	0
17_FL	6	q22.32--q27	126760908	170750923	1.51	43.99 Loss	200		4
17_FL	8	p12	33885277	37729664	2.55	3.84 Gain	4	UNC5D,KCNU1,ZNF703,ERLIN2	0
17_FL	8	p12--p11.22	37730758	39516172	3.57	1.79 Gain	24		0
17_FL	8	p11.22--q24.3	39516889	146268947	2.77	106.75 Gain	433		22
17_FL	9	p24.3--p23	78018	9632669	2.66	9.55 Gain	40		1
17_FL	9	p23--p21.3	9633173	24197934	1.61	14.56 Loss	49		2
17_FL	10	q21.1	54948487	55075266	1.34	0.13 Loss	0		0
17_FL	10	q21.1	58691337	58985947	1.30	0.29 Loss	0		0
17_FL	10	q21.3--q25.1	65464126	107804037	1.64	42.34 Loss	322		11
17_FL	11	p15.5--q25	188510	134449982	2.63	134.26 Gain	1254		25
17_FL	12	p13.32--p13.2	3412714	11204755	1.48	7.79 Loss	136		3
17_FL	12	p13.2	11208809	11389266	2.72	0.18 Gain	3	TAS2R42,PRB3,PRB4	0
17_FL	12	p13.2--p13.1	11394381	12849193	1.49	1.45 Loss	12		2
17_FL	12	p13.1	12855526	12975349	2.86	0.12 Gain	2	DDX47,GPRC5A	1
17_FL	12	p13.1	12976711	14748413	1.46	1.77 Loss	9	GPRC5D,HEBP1,KIAA1467,GSG1,EMP1,GRIN2B,ATF7IP,PLBD1,GUCY2C	0
17_FL	12	p12.3	16460765	16662999	1.27	0.20 Loss	1	LMO3	0
17_FL	12	p12.3	16667370	17100139	2.53	0.43 Gain	0		0
17_FL	12	p12.3--p12.2	17100818	20634936	1.38	3.53 Loss	7	RERGL,PIK3C2G,PLCZ1,CAPZA3,PLEKHA5,AEBP2,PDE3A	0
17_FL	12	p12.2	20639015	20659128	2.62	0.02 Gain	1	PDE3A	0
17_FL	12	p12.2--p12.1	20660255	21378710	1.40	0.72 Loss	6	PDE3A,SLCO1C1,SLCO1B3,SLCO1B7,SLCO1B1,SLCO1A2	0
17_FL	12	p12.1	21658138	22373705	1.41	0.72 Loss	5	LDHB,KCNJ8,ABCC9,CMA5,ST8SIA1	0
17_FL	12	q21.1--q21.33	72943979	89107773	2.48	16.16 Gain (subclonal)	45		3
17_FL	12	q21.33	89110139	89985056	1.34	0.87 Loss	3	C12orf12,EPYC,KERA	0
17_FL	12	q21.33	89987152	90506822	2.52	0.52 Gain (subclonal)	2	LUM,DCN	0
17_FL	12	q21.33--q22	90508388	91549103	1.36	1.04 Loss	4	LOC256021,BTG1,CLLU1OS,CLLU1	0
17_FL	12	q22	91550027	92454506	2.83	0.90 Gain	6	C12orf74,PLEKHG7,EEA1,NUDT4,UBE2N,MRPL42	0
17_FL	12	q22	92463162	92626260	1.39	0.16 Loss	2	SOCS2,CRADD	0
17_FL	13	q21.2--q31.1	59520559	80231806	1.53	20.71 Loss	28		0
17_FL	13	q31.3--q34	92498611	114126487	1.59	21.63 Loss	85		2
17_FL	16	q21--q23.1	64641319	76492239	2.43	11.85 Gain (subclonal)	156		4
17_FL	16	q23.1	76493541	77130526	2.85	0.64 Gain	3	VAT1L,CLEC3A,WWOX	0
17_FL	16	q23.1--q24.3	77131129	88815024	2.41	11.68 Gain (subclonal)	101		1
17_FL	17	p13.3--p13.1	514	8012659	1.56	8.01 Loss	194		7

17_FL	17	q21.32--q25.3	44303576	78643088	3.09	34.34	Gain	378		12
17_FL	19	p13.3--p13.2	41898	10438628	2.98	10.40	Gain	299		6
17_FL	20	p12.1	15572156	15629081	2.87	0.06	Gain	1	MACROD2	0
17_FL	20	q11.21	29311247	30793365	2.82	1.48	Gain	28		1
17_tFL	6	p25.3	94649	505558	2.51	0.41	Gain	3	DUSP22,IRF4,EXOC2	0
17_tFL	6	p21.32	32689804	32974073	2.75	0.28	Gain	9	HLA-DQA1,HLA-DQB1,HLA-DQA2,HLA-DQB2,HLA-DOB,TAP2,PSMB8,TAP1,PSMB9	0
17_tFL	8	p23.3--q24.3	21242	146268947	2.50	146.25	Gain	661		31
17_tFL	9	p13.2	37293477	37667675	2.72	0.37	Gain	7	ZCCHC7,GRHPR,ZBTB5,POLR1E,FBXO10,TOMM5,FRMPD1	0
17_tFL	10	q22.3--q23.31	80817120	91427131	1.57	10.61	Loss	60		2
17_tFL	10	q24.1--q24.2	98154050	99737051	1.68	1.58	Loss	26		1
23_FL	6	q22.31--q24.1	124714893	141403016	1.10	16.69	Loss	79		2
23_FL	6	q24.1	141405030	142160177	0.27	0.76	Loss homozygous	0		0
23_FL	6	q24.1--q24.2	142161131	144067548	1.11	1.91	Loss	9	NMBR,VTA1,GPR126,HIVEP2,AIG1,ADAT2,PEX3,FUCA2,PHACTR2	0
23_FL	18	q11.1--q21.33	16783339	58943748	4.10	42.16	Amp	146		9
23_tFL	1	q21.2--q44	148785871	247190999	2.99	98.41	Gain	869		28
23_tFL	2	q14.1	118559039	118597538	1.15	0.04	Loss	1	INSIG2	0
23_tFL	3	p26.3--p26.2	35333	3989902	1.23	3.95	Loss	7	CHL1,CNTN6,CNTN4,IL5RA,TRNT1,CRBN,LRRN1	0
23_tFL	3	p22.1--p14.3	42050121	58327403	1.26	16.28	Loss	242		8
23_tFL	3	p14.2	60416381	60484618	1.29	0.07	Loss	1	FHIT	0
23_tFL	3	q23--q29	144302473	199322963	3.06	55.02	Gain	262		12
23_tFL	5	q33.1--q35.3	151026231	180625577	1.24	29.60	Loss	180		10
23_tFL	6	q21--q23.3	108935586	138103769	1.24	29.17	Loss	138		3
23_tFL	6	q23.3	138104343	138286660	0.55	0.18	Loss homozygous	1	TNFAIP3	0
23_tFL	6	q23.3--q24.1	138291449	142160177	1.20	3.87	Loss	13		0
23_tFL	6	q24.3--q25.2	147688639	152989646	1.26	5.30	Loss	31		0
23_tFL	7	p22.3--q11.22	52899	66417629	2.87	66.36	Gain	302		10
23_tFL	7	q11.22--q36.3	68614618	158819753	3.26	90.21	Gain	568		24
23_tFL	11	q12.3	61732628	61784801	3.11	0.05	Gain	2	SCGB2A1,SCGB1D2	0
23_tFL	11	q14.1--q25	84281642	134449982	2.95	50.17	Gain	338		8
23_tFL	12	p11.23	26840916	27020454	3.15	0.18	Gain	4	ITPR2,ASUN,FGFR1OP2,TM7SF3	0
23_tFL	15	q14--q21.1	34755583	45866033	1.28	11.11	Loss	120		4
23_tFL	18	p11.32	1543	2633292	1.27	2.63	Loss	12		0
23_tFL	18	p11.21--q21.33	15391751	58943748	4.81	43.55	Amp	146		9
23_tFL	22	q11.22	20899640	21388929	1.36	0.49	Loss	5	VPREB1,ZNF280B,ZNF280A,PRAME,GGTLC2	0
29_tFL	1	p11.2--q44	121041748	247190999	2.74	126.15	Gain	940		28
29_tFL	5	q33.1--q34	151563005	162532172	1.22	10.97	Loss	47		3
29_tFL	9	p21.3	21858896	21995382	1.26	0.14	Loss	2	CDKN2A,CDKN2B	0
29_tFL	12	p13.33--q24.33	20691	132288250	2.87	132.27	Gain	995		29
39_FL	1	p36.33--p36.31	802883	6822450	1.55	6.02	Loss	96		4
39_FL	1	p11.2--q44	120928505	247190999	2.37	126.26	Gain	940		28
39_FL	2	p16.1--p15	59800688	61671499	2.49	1.87	Gain	10	BCL11A,PAPOLG,REL,PUS10,PEX13,KIAA1841,C2orf74,AHSA2,USP34,XPO1	0
39_FL	6	q23.3	137330176	138500113	1.49	1.17	Loss	6	IL20RA,IL22RA2,IFNGR1,OLIG3,TNFAIP3,PERP	0
39_FL	12	q11--q21.31	36149979	83518174	2.43	47.37	Gain	387		14
39_FL	17	q21.33--q25.3	45421698	78643088	2.35	33.22	Gain	358		12
39_FL	21	p11.2--q22.3	9758730	46921373	2.34	37.16	Gain	216		5
39_tFL	1	p36.33--p36.31	1145909	6981328	1.16	5.84	Loss	84		1
39_tFL	1	p11.2--q44	121005815	247190999	2.67	126.19	Gain	940		28
39_tFL	2	p16.1--p15	59800688	61671933	3.77	1.87	Gain	10	BCL11A,PAPOLG,REL,PUS10,PEX13,KIAA1841,C2orf74,AHSA2,USP34,XPO1	0
39_tFL	4	p15.32--p15.1	18452134	32767854	1.24	14.32	Loss	22	SLIT2,PACRGL,KCNIP4,GPR125,GBA3,PPARGC1A,DHX15,SOD3,CCDC149,LGI2	2
39_tFL	4	q13.1--q23	63804797	102072927	1.26	38.27	Loss	191		2
39_tFL	4	q25	110198580	112577850	1.24	2.38	Loss	13		2
39_tFL	4	q26--q27	119273283	120694906	1.25	1.42	Loss	10	NDST3,PRSS12,METTL14,SEC24D,SYNPO2,MYOZ2,USP53,C4orf3,FABP2,PDE5A	0
39_tFL	4	q31.21	143476047	143656920	1.26	0.18	Loss	1	INPP4B	0
39_tFL	6	p21.33--p21.32	31573026	32905787	1.29	1.33	Loss	72		1
39_tFL	6	q14.3--q16.1	85399549	97993542	1.27	12.59	Loss	41		0

39_tFL	12	q11--q21.31	36144199	82947204	2.75	46.80 Gain	387		14
39_tFL	12	q21.31	83125919	83518174	2.90	0.39 Gain	0		0
39_tFL	17	q21.33--q25.3	45515869	78643088	2.66	33.13 Gain	357		12
39_tFL	21	p11.2--q22.3	9758730	46921373	2.78	37.16 Gain	216		5
40_FL	2	p25.3	2772	67143328	2.70	67.14 Gain	311		6
40_FL	5	q33.3	156733599	157134118	1.61	0.40 Loss (subclonal)	7	<i>CYFIP2,NIPAL4,ADAM19,SOX30,C5orf52,THG1L,LSM11</i>	0
40_FL	7	p22.3--p22.2	149266	3143731	2.36	2.99 Gain (subclonal)	33		1
40_FL	8	p23.3	21242	146268947	2.36	146.25 Gain	661		31
40_FL	10	p15.3--p12.33	62747	25715582	2.34	25.65 Gain	111		5
40_FL	12	p13.33	31469	71903356	2.39	71.87 Gain	637		17
40_FL	14	q11.2	19600479	105270409	2.39	85.67 Gain	571		62
40_FL	18	q21.33	58945410	76116029	2.40	17.17 Gain	51		0
40_FL	21	p11.2--q11.2	9758730	46921373	2.27	37.16 Gain (subclonal)	216		5
40_tFL	2	p25.3--p14	2772	67151787	2.69	67.15 Gain	311		6
40_tFL	3	q13.11	104571672	104768553	0.88	0.20 Loss	0		0
40_tFL	5	q33.3	156738482	157134118	0.96	0.40 Loss	7	<i>CYFIP2,NIPAL4,ADAM19,SOX30,C5orf52,THG1L,LSM11</i>	0
40_tFL	5	q35.3	176548933	176808981	1.00	0.26 Loss	11		0
40_tFL	6	p25.3--q27	94649	170892918	2.53	170.80 Gain	1004		17
40_tFL	7	p22.3--q36.3	52899	158819753	3.66	158.77 Gain	870		34
40_tFL	8	p23.3--q24.3	21242	146268947	2.55	146.25 Gain	661		31
40_tFL	9	p21.3	20466263	22984940	0.94	2.52 Loss	24		2
40_tFL	11	q13.5--q25	76621605	134449982	2.63	57.83 Gain	362		9
40_tFL	12	p13.33--q21.1	20691	70882381	2.66	70.86 Gain	636		17
40_tFL	13	q11--q34	17924937	114126487	2.52	96.20 Gain	308		14
40_tFL	14	q11.2--q32.33	19495751	106356482	3.19	86.86 Gain	575		62
40_tFL	15	q21.1	42791853	42796158	0.35	0.00 Loss homozygous	1	<i>B2M</i>	0
40_tFL	17	p13.3--p11.2	514	19504358	1.06	19.50 Loss	307		12
40_tFL	17	q11.2--q12	27667593	32834166	1.08	5.17 Loss	69		1
40_tFL	18	p11.32--q23	1543	76116029	2.74	76.11 Gain	261		9
40_tFL	21	p11.2--q22.3	9758730	46921373	2.57	37.16 Gain	216		5

^ Numbering according to the Human Genome hg18 assembly

Note that this table reports the segments of aberrant copy number obtained by the circular binary segmentation (CBS) algorithm; therefore the total number of segments is different from the total number of lesions reported in the text (see also Materials and Methods).

Segments highlighted in grey correspond to subclonal lesions which were retained despite a CN value below the adopted threshold (2.3 and 1.7 for gains and losses, respectively) given the low tumor enrichment of the specimen.

Table V. Complete list of genes found mutated in >10% of tFL, with additional selected genes of relevance within the same pathway

Gene ID [^]	N of M tFL	% of M tFL	N of M with paired FL	N of M FL (in the pair)	Significance according to:			Biological Function/Affected Pathway	N of M unsel FL	% of M unsel FL	p value (tFL vs uFL)
					Gistic*	MutCom	Focal*	MutSigCV			
<i>MLL2</i> ^{^^}	26/39	66.7	10	10	ns	ns	ns	epigenetic remodeling	11/23	47.8	ns
<i>CREBBP</i> ^{^^}	19/39	48.7	8	8	ns	yes	ns	epigenetic remodeling	14/23	60.9	ns
<i>BCL2</i>	19/39	43.6	12	12	ns	yes	ns	apoptosis	11/23	47.8	ns
<i>TNFRSF14</i>	11/39	28.2	6	1	ns	yes	ns	immune evasion	5/23	21.7	ns
<i>EZH2</i>	10/39	25.6	6	4	ns	yes	yes	epigenetic remodeling	1/23***	4.3	ns
<i>STAT6</i>	9/39	23.1	5	3	ns	yes	yes	JAK/STAT6 signaling	1/23	4.3	ns
<i>PIM1</i>	9/39	23.1	8	2	ns	yes	ns	tyrosin kinase/signal transduction	0/23	0.0	0.01
<i>GNA13</i>	8/39	20.5	2	0	ns	yes	yes	G protein; CXCR4/S1P2 signaling	1/23	4.3	ns
<i>MYC</i>	7/39	17.9	4	0	yes	yes	ns	proliferation	0/23	0.0	0.03
<i>HIST1H1E</i>	7/39	17.9	2	1	ns	ns	ns	epigenetic remodeling	2/23	8.7	ns
<i>SYNE1</i> ^{^^}	7/39	17.9	2	0	ns	yes	ns	cytoskeletal organization/spectrin superfamily	0/23	0.0	0.03
<i>ARID1A</i>	7/39	17.9	4	1	ns	yes	ns	epigenetic remodeling	3/23	13.0	ns
<i>FMN2</i>	6/39	15.4	1	0	ns	yes	ns	cytoskeletal organization	0/23	0.0	0.05
<i>FOXO1</i>	6/39	15.4	3	1	ns	ns	ns	PI3K, BCR, CXCR4 signaling	0/23	0.0	0.05
<i>TP53</i>	6/39	15.4	3	0	ns	yes	ns	DNA damage response	3/23	13.0	ns^^
<i>EP300</i> ^{^^}	5/39	12.8	2	1	ns	ns	ns	epigenetic remodeling	1/23	4.3	ns
<i>SGK1</i>	5/39	12.8	2	0	ns	ns	ns	serine-threonine kinase/PI3K signaling pathway, ion channel transport	0/23	0.0	ns
<i>B2M</i>	4/39	10.3	3	0	ns	yes	ns	immune evasion	0/23	0.0	ns**
<i>CCND3</i>	4/39	10.3	1	0	ns	ns	ns	cell cycle regulation	3/23	13.0	ns
<i>FAS</i>	4/39	10.3	3	3	yes	yes	ns	apoptosis	0/23	0.0	ns
<i>IRF8</i>	4/39	10.3	2	1	ns	ns	ns	B cell development/differentiation	1/23	4.3	ns
<i>KLHL6</i>	4/39	10.3	1	0	ns	ns	ns	germinal center responses	1/23	4.3	ns
<i>SOCS1</i>	4/39	10.3	1	0	ns	yes	ns	JAK/STAT6 signaling	0/23	0.0	ns
<i>TAGAP</i>	4/39	10.3	1	0	ns	yes	ns	RhoH GTPase signaling	0/23	0.0	ns
<i>TET2</i>	4/39	10.3	1	1	ns	yes	ns	epigenetic remodeling	0/23	0.0	ns
<i>TNFK</i>	4/39	10.3	1	1	ns	ns	ns	NF-kappaB signaling	0/23	0.0	ns
<i>CARD11</i>	4/39	10.3	1	1	ns	yes	ns	NF-kappaB/BCR signaling	1/23	4.3	ns
<i>PAX5</i>	3/39	7.7	0	0	ns	ns	ns	B cell development/differentiation	1/23	4.3	ns
<i>CD79B</i>	3/39	7.7	3	3	ns	yes	ns	NF-kappaB/BCR signaling	1/23	4.3	ns
<i>EBF1</i>	3/39	7.7	2	0	ns	yes	ns	B cell development/differentiation	1/23	4.3	ns
<i>IRF4</i>	3/39	7.7	3	0	ns	yes	ns	B cell development/differentiation	2/23	8.7	ns
<i>MEF2B</i>	3/39	7.7	1	0	ns	ns	ns	BCL6 regulation (Ying et al., Nat Immunology 2013)	3/23	13.0	ns
<i>MLL3</i>	3/39	7.7	0	0	ns	ns	ns	epigenetic remodeling	0/23	0.0	ns
<i>KDM6B</i>	2/39	5.1	0	0	ns	ns	ns	epigenetic remodeling	0/23	0.0	ns
<i>FBXO11</i>	2/39	5.1	1	0	ns	ns	ns	BCL6 regulation (Duan et al., Nature 2012)	0/23	0.0	ns
<i>TNFAIP3</i>	1/39	2.6	1	0	yes	yes	ns	NF-kappaB	1/23	4.3	ns
<i>PCLO</i> ^{^^}	8/39	20.5	1	1	ns	yes	ns	cytoskeleton organization	2/23	8.7	ns
<i>ABCA13</i> ^{^^}	7/39	17.9	1	0	ns	yes	ns	transmembrane transporter	0/23	0.0	0.03
<i>DMD</i> ^{^^}	7/39	17.9	1	0	ns	ns	ns	bone morphogenetic protien/secreted signaling proteins	0/23	0.0	0.03
<i>FAT4</i> ^{^^}	6/39	15.4	1	0	ns	ns	ns	protocadherin/calcium binding	0/23	0.0	0.05
<i>FBN2</i> ^{^^}	6/39	15.4	5	0	ns	ns	ns	elastic fibers assembly	0/23	0.0	0.05
<i>XIRP2</i> ^{^^}	5/39	12.8	1	1	ns	ns	ns	cohesin complex	1/23	4.3	ns
<i>BMP7</i>	4/39	10.3	3	1	ns	ns	ns	bone morphogenetic protein/TGFbeta and Hippo signaling pathways	0/23	0.0	ns
<i>DTX1</i>	4/39	10.3	2	0	ns	ns	ns	ubiquitin ligase/Notch signaling	0/23	0.0	ns
<i>PCDH7</i>	4/39	10.3	1	0	ns	ns	ns	protocadherin/cell-cell recognition and adhesion	0/23	0.0	ns
<i>TAF1</i>	4/39	10.3	1	0	ns	ns	ns	basal transcription	0/23	0.0	ns
<i>GRID2</i>	4/39	10.3	1	1	ns	ns	ns	glutamate receptor	0/23	0.0	ns
<i>NAV3</i> ^{^^}	4/39	10.3	1	0	ns	ns	ns	nucleotide binding	0/23	0.0	ns

[^] Grey shade indicates genes that are mutated in >10% of cases but were conservatively not included in Figure 6 because not deemed significant by any of the three algorithms (see text), and/or because they do not seem to be expressed in B cells (*ABCA13* and *PCLO*). Genes in red were included in Figure 6 because within the same annotation pathways as significant genes, and/or because of their known functional relevance in B cell development and cancer.

^{^^} Genes with CDS >7000bp.

^{^^^} The TP53 gene was considered tFL-specific because it was invariably absent in the diagnostic FL biopsy of the three informative tFL cases, and because of the broad available literature.

* Blue and red color denote genes identified as significant targets of genomic losses (+/- mutations) and gains (+/- mutations) respectively.

** <0.05 when considering CNAs and SNVs combined.

Table VI. Significant regions of deletion and amplification in tFL according to the GISTIC algorithm

Cytoband	q value	Residual q value	Wide peak boundaries	Total N of genes	Gene names (if less than 20)
Regions of Deletion					
1p36.32	6.05E-09	6.05E-09	chr1:2941054-3782660	12	hsa-mir-551a, <i>DFFB</i> , <i>MEGF6</i> , <i>TP73</i> , <i>KIAA0562</i> , <i>ARHGEF16</i> , <i>WDR8</i> , <i>KIAA0495</i> , <i>LRRC47</i> , <i>PRDM16</i> , <i>TPRG1L</i> , <i>CCDC27</i>
2p25.1	9.62E-02	9.62E-02	chr2:12600517-15312495	3	<i>TRIB2</i> , <i>NAG</i> , <i>FAM84A</i>
4q31.3	2.61E-04	2.61E-04	chr4:153473059-154773142	8	<i>KIAA0922</i> , <i>TRIM2</i> , <i>ARFIP1</i> , <i>FBXW7</i> , <i>MND1</i> , <i>FHDC1</i> , <i>TIGD4</i> , <i>TMEM154</i>
5q33.3	5.94E-02	5.94E-02	chr5:156215994-162532316	33	
6q13	3.52E-04	2.81E-02	chr6:72351905-82626037	29	
6q23.3	2.43E-09	2.43E-09	chr6:138057822-138291448	1	<i>TNFAIP3</i>
8q12.1	5.40E-02	5.40E-02	chr8:59895943-60032285	1	<i>TOX</i>
9p21.3	4.27E-23	4.27E-23	chr9:21958713-21995192	2	<i>CDKN2A</i> , <i>CDKN2B</i>
10q23.31	1.23E-04	1.23E-04	chr10:90642493-90928640	3	<i>ACTA2</i> , <i>FAS</i> , <i>STAMBPL1</i>
13q14.2	1.14E-01	1.14E-01	chr13:28540423-63450802	129	
15q21.1	2.97E-05	2.97E-05	chr15:42156769-42798306	6	<i>B2M</i> , <i>EIF3J</i> , <i>CTDSPL2</i> , <i>SPG11</i> , <i>FRMD5</i> , <i>CASC4</i>
16q12.1	3.29E-02	3.29E-02	chr16:46019838-55595403	57	
17p13.2	1.68E-06	1.68E-06	chr17:3680597-7497729	111	
18p11.32	1.91E-01	1.91E-01	chr18:1-3927943	21	
19p13.3	1.34E-03	1.34E-03	chr19:6345072-6931762	17	
22q11.22	2.23E-01	2.23E-01	chr22:20712333-21392889	5	<i>VPREB1</i> , <i>PRAME</i> , <i>GGTLC2</i> , <i>ZNF280A</i> , <i>ZNF280B</i>
Regions of Amplification					
1q23.2	2.10E-03	2.10E-03	chr1:148974469-157403464	241	
2p16.1	1.99E-03	1.99E-03	chr2:60884247-61145579	3	<i>PEX13</i> , <i>REL</i> , <i>PUS10</i>
3q27.3	4.36E-02	4.36E-02	chr3:188954049-189334637	1	<i>[BCL6]</i>
7p22.1	1.99E-03	1.99E-03	chr7:180195-18005030	92	
8q12.1	2.32E-03	7.39E-02	chr8:56245379-57722273	10	<i>LYN</i> , <i>MOS</i> , <i>PENK</i> , <i>PLAG1</i> , <i>RPS20</i> , <i>CHCHD7</i> , <i>TGS1</i> , <i>XKR4</i> , <i>TMEM68</i> , <i>RDHE2</i>
8q24.21	1.99E-03	1.99E-03	chr8:129081312-129228126	1	<i>[MYC]</i>
9p13.2	1.14E-01	1.14E-01	chr9:37207947-38178523	12	<i>SHB</i> , <i>GRHPR</i> , <i>ZBTB5</i> , <i>FRMPD1</i> , <i>FBXO10</i> , <i>EXOSC3</i> , <i>POLR1E</i> , <i>WDR32</i> , <i>ZCCHC7</i> , <i>MCART1</i> , <i>RG9MTD3</i> , <i>C9orf105</i>
12q13.11	7.13E-03	7.13E-03	chr12:44818144-51745367	124	
15q24.3	1.99E-03	1.99E-03	chr15:1-100338915	558	
18q21.33	5.66E-03	5.66E-03	chr18:57163777-58922894	7	<i>TNFRSF11A</i> , <i>PHLPP</i> , <i>PIGN</i> , <i>CDH20</i> , <i>ZCCHC2</i> , <i>KIAA1468</i> , <i>RNF152</i>
21q22.3	4.63E-02	4.63E-02	chr21:21098338-46944323	195	

Table VII. SNVs identified by whole exome sequencing in genes mutated in >5% of cases

Sample	Gene	Exon^	Strand	Genome_Change^^	cDNA_Change	Protein_Change	Variant_Classification	PPH2_Class	Refseq_prot_Id	Sanger validation	Somatic status
1	ABCA13	18	+	g.chr7:48319056G>T	c.8265G>T	p.W2755C	Missense_Mutation	deleterious	NP_689914	Yes	somatic
6	ABCA13	10	+	g.chr7:48280528T>A	c.1127T>A	p.L376Q	Missense_Mutation	deleterious	NP_689914	nd	somatic
10	ABCA13	38	+	g.chr7:48431556C>T	c.11693C>T	p.T3898I	Missense_Mutation	deleterious	NP_689914	Yes	
25	ABCA13	30	+	g.chr7:48390288G>A	c.10253G>A	p.R3418H	Missense_Mutation	neutral	NP_689914	nd	
38	ABCA13	13	+	g.chr7:48285584A>C	c.1616A>C	p.E539A	Missense_Mutation	deleterious	NP_689914	nd	
48	ABCA13	17	+	g.chr7:48314670G>A	c.5407G>A	p.E1803K	Missense_Mutation	deleterious	NP_689914	nd	
49	ABCA13	30	+	g.chr7:48390287C>T	c.10252C>T	p.R3418C	Missense_Mutation	deleterious	NP_689914	nd	
2	ANKMY1	4	-	g.chr2:241465828C>T	c.721G>A	p.A241T	Missense_Mutation	neutral	NP_057636	Yes	somatic
15	ANKMY1	3	-	g.chr2:241468634T>G	c.506A>C	p.E169A	Missense_Mutation	deleterious	NP_057636	nd	
48	ANKMY1	5,6	-	g.chr2:241463333T>C	c.1534A>G	p.T512A	Missense_Mutation	deleterious	NP_057636	nd	
8	ANKS1A	15	+	g.chr6:35047347G>C	c.2338G>C	p.D780H	Missense_Mutation	neutral	NP_056060	Yes	putative somatic
36	ANKS1A	20	+	g.chr6:35051251G>A	c.2965G>A	p.G989S	Missense_Mutation	deleterious	NP_056060	nd	
2	APC2	13	+	g.chr19:1461961G>A	c.1639_splice	p.E547_splice	Splice_Site		NP_005874	nd	somatic
25	APC2	13	+	g.chr19:1461980G>A	c.1657G>A	p.V553M	Missense_Mutation	neutral	NP_005874	nd	
47	APC2	14	+	g.chr19:1467213G>A	c.3913G>A	p.G1305S	Missense_Mutation	neutral	NP_005874	nd	
18	ARHGEF10	23	+	g.chr8:1876747C>T	c.2852C>T	p.P951L	Missense_Mutation	neutral	NP_055444	nd	
20	ARHGEF10	26	+	g.chr8:1893741C>T	c.3317C>T	p.T1106M	Missense_Mutation	deleterious	NP_055444	nd	
23	ARHGEF10	25	+	g.chr8:1882059C>T	c.3173C>T	p.A1058V	Missense_Mutation	neutral	NP_055444	nd	somatic
7	ARID1A	1	+	g.chr1:27023207delG	c.313delG	p.G105fs	Frame_Shift_Del		NP_006006	Yes	somatic
12	ARID1A	1	+	g.chr1:27023216_27023222del7bp	c.322_328del7bp	p.G108fs	Frame_Shift_Del		NP_006006	Yes	putative somatic
20	ARID1A	1	+	g.chr1:27023539C>A	c.645C>A	p.Y215*	Nonsense_Mutation		NP_006006	Yes	
20	ARID1A	1	+	g.chr1:27023540T>C	c.646T>C	p.Y216H	Missense_Mutation	none	NP_006006	Yes	
24	ARID1A	1	+	g.chr1:27023170_27023171insG	c.276_277insG	p.G92fs	Frame_Shift_Ins		NP_006006	nd	
25	ARID1A	1	+	g.chr1:27023535C>T	c.641C>T	p.S214F	Missense_Mutation	none	NP_006006	Yes	
29	ARID1A	5	+	g.chr1:27087512_27087513insC	c.2086_2087insC	p.S696fs	Frame_Shift_Ins		NP_006006	Yes	putative somatic
39	ARID1A	18	+	g.chr1:27101552_27101553delAA	c.4834_4835delAA	p.K1612fs	Frame_Shift_Del		NP_006006	nd	
5	ARID1B	20	+	g.chr6:157528087C>T	c.5758C>T	p.Q1920*	Nonsense_Mutation		NP_059989	Yes	somatic
4	AUTS2	1	+	g.chr7:69064912T>G	c.273T>G	p.F91L	Missense_Mutation	neutral	NP_056385	nd	putative somatic
10	AUTS2	18	+	g.chr7:70254895C>T	c.2693C>T	p.S898L	Missense_Mutation	deleterious	NP_056385	nd	
27	AUTS2	5	+	g.chr7:69755417G>A	c.724G>A	p.D242N	Missense_Mutation		NP_001120704.1	nd	
5	B2M	1	+	g.chr15:45003781_45003782delCT	c.37_38delCT	p.L13fs	Frame_Shift_Del		NP_004039	Yes	somatic
23	B2M	1	+	g.chr15:45003808C>B	c.64C>T	p.Q22*	Nonsense_Mutation		NP_004039	Yes	somatic
36	B2M	2	+	g.chr15:45007768C>G	c.215C>G	p.S72*	Nonsense_Mutation		NP_004039	Yes	
48	B2M	2	+	g.chr15:45007774_45007775ins7bp	c.221_222ins7bp	p.L74fs	Frame_Shift_Ins		NP_004039	Yes	
2	BCL2	1	-	g.chr18:60985881T>G	c.19A>C	p.T7P	Missense_Mutation	deleterious	NP_000624	nd	somatic
2	BCL2	1	-	g.chr18:60985896C>A	c.4G>T	p.A2S	Missense_Mutation	deleterious	NP_000624	nd	somatic
3	BCL2	1	-	g.chr18:60985896C>G	c.4G>C	p.A2P	Missense_Mutation	deleterious	NP_000624	Yes	somatic
4	BCL2	1	-	g.chr18:60985305G>A	c.595C>T	p.L199F	Missense_Mutation		NP_000624	nd	
4	BCL2	1	-	g.chr18:60985508G>C	c.392C>G	p.A131G	Missense_Mutation	neutral	NP_000624	Yes	
4	BCL2	1	-	g.chr18:60985644G>A	c.256C>T	p.L86F	Missense_Mutation	neutral	NP_000624	nd	
4	BCL2	1	-	g.chr18:60985679G>T	c.221C>A	p.T74N	Missense_Mutation	neutral	NP_000624	nd	
5	BCL2	1	-	g.chr18:60985363C>G	c.537G>C	p.E179D	Missense_Mutation	neutral	NP_000624	Yes	somatic
7	BCL2	1	-	g.chr18:60985362A>G	c.538T>C	p.Y180H	Missense_Mutation	deleterious	NP_000624	Yes	somatic
7	BCL2	1	-	g.chr18:60985793C>T	c.107G>A	p.G36D	Missense_Mutation	neutral	NP_000624	Yes	somatic
20	BCL2	1	-	g.chr18:60985725G>A	c.175C>T	p.P59S	Missense_Mutation	neutral	NP_000624	nd	
23	BCL2	1	-	g.chr18:60985279G>A			Splice_Site		NP_000624	nd	somatic
23	BCL2	1	-	g.chr18:60985292C>T	c.608G>A	p.S203N	Missense_Mutation		NP_000624	nd	somatic
23	BCL2	1	-	g.chr18:60985566A>G	c.334T>C	p.F112L	Missense_Mutation	deleterious	NP_000624	Yes	somatic
23	BCL2	1	-	g.chr18:60985670G>A	c.230C>T	p.A77V	Missense_Mutation	neutral	NP_000624	nd	somatic
23	BCL2	1	-	g.chr18:60985880G>A	c.20C>T	p.T7I	Missense_Mutation	neutral	NP_000624	Yes	somatic
28	BCL2	1	-	g.chr18:60985284A>G	c.616T>C	p.*206R	Nonstop_Mutation		NP_000624	nd	
28	BCL2	1	-	g.chr18:60985727T>C	c.173A>G	p.H58R	Missense_Mutation	neutral	NP_000624	Yes	
28	BCL2	1	-	g.chr18:60985820C>T	c.80G>A	p.G27D	Missense_Mutation	deleterious	NP_000624	Yes	
36	BCL2	1	-	g.chr18:60985492C>G	c.408G>C	p.E136D	Missense_Mutation	deleterious	NP_000624	Yes	
36	BCL2	1	-	g.chr18:60985744C>G	c.156G>C	p.Q52H	Missense_Mutation	deleterious	NP_000624	Yes	
36	BCL2	1	-	g.chr18:60985793C>T	c.107G>A	p.G36D	Missense_Mutation	neutral	NP_000624	Yes	
36	BCL2	1	-	g.chr18:60985866G>A	c.34C>T	p.R12W	Missense_Mutation	deleterious	NP_000624	Yes	
36	BCL2	1	-	g.chr18:60985877C>T	c.23G>A	p.G8E	Missense_Mutation	deleterious	NP_000624	no	
39	BCL2	1	-	g.chr18:60985311C>G	c.589G>C	p.G197R	Missense_Mutation		NP_000624	nd	
39	BCL2	1	-	g.chr18:60985412T>G	c.488A>C	p.N163T	Missense_Mutation	neutral	NP_000624	Yes	
39	BCL2	1	-	g.chr18:60985499A>G	c.401T>C	p.V134A	Missense_Mutation	neutral	NP_000624	Yes	

39	BCL2	1	-	g.chr18:60985508G>A	c.392C>T	p.A131V	Missense_Mutation	neutral	NP_000624	Yes	
39	BCL2	1	-	g.chr18:60985547T>C	c.353A>G	p.Q118R	Missense_Mutation	neutral	NP_000624	Yes	
40	BCL2	1	-	g.chr18:60985644G>A	c.256C>T	p.L86F	Missense_Mutation	neutral	NP_000624	nd	putative somatic
40	BCL2	1	-	g.chr18:60985646G>A	c.254C>T	p.A85V	Missense_Mutation	neutral	NP_000624	Yes	putative somatic
40	BCL2	1	-	g.chr18:60985725G>C	c.175C>G	p.P59A	Missense_Mutation	neutral	NP_000624	Yes	putative somatic
40	BCL2	1	-	g.chr18:60985883C>G	c.17G>C	p.R6T	Missense_Mutation	deleterious	NP_000624	Yes	putative somatic
41	BCL2	1	-	g.chr18:60985671_60985672delCA	c.228_229delTG	p.A76fs	Frame_Shift_Del		NP_000624	Yes	
41	BCL2	1	-	g.chr18:60985674_60985680del7bp	c.220_226del7bp	p.T74fs	Frame_Shift_Del		NP_000624	Yes	
44	BCL2	1	-	g.chr18:60985743G>A	c.157C>T	p.P53S	Missense_Mutation	deleterious	NP_000624	nd	
44	BCL2	1	-	g.chr18:60985764G>A	c.136C>T	p.P46S	Missense_Mutation	neutral	NP_000624	nd	
45	BCL2	1	-	g.chr18:60985307G>A	c.593C>T	p.A198V	Missense_Mutation		NP_000624	nd	
47	BCL2	1	-	g.chr18:60985760C>G	c.140G>C	p.G47A	Missense_Mutation	deleterious	NP_000624	Yes	
49	BCL2	1	-	g.chr18:60985808T>C	c.92A>G	p.D31G	Missense_Mutation	neutral	NP_000624	nd	
49	BCL2	1	-	g.chr18:60985883C>G	c.17G>C	p.R6T	Missense_Mutation	deleterious	NP_000624	nd	
49	BCL2	1	-	g.chr18:60985884T>C	c.16A>G	p.R6G	Missense_Mutation	neutral	NP_000624	nd	
18	BIRC6	6	+	g.chr2:32620612T>C	c.985T>C	p.F329L	Missense_Mutation	deleterious	NP_057336	nd	
18	BIRC6	6	+	g.chr2:32620631T>C	c.1004T>C	p.L335P	Missense_Mutation		NP_057336	nd	
20	BIRC6	55	+	g.chr2:32740142G>A	c.10654G>A	p.V3552I	Missense_Mutation		NP_057336	nd	
43	BIRC6	59	+	g.chr2:32750590A>G	c.11815A>G	p.R3939G	Missense_Mutation		NP_057336	nd	
2	BMP7	1	-	g.chr20:55840986G>A	c.193C>T	p.P65S	Missense_Mutation	neutral	NP_001710	Yes	somatic
11	BMP7	1	-	g.chr20:55841157C>T	c.22G>A	p.A8T	Missense_Mutation	neutral	NP_001710	Yes	somatic
36	BMP7	1	-	g.chr20:55841060C>G	c.119G>C	p.S40T	Missense_Mutation	neutral	NP_001710	nd	
40	BMP7	1	-	g.chr20:55840918G>C	c.261C>G	p.Y87*	Nonsense_Mutation		NP_001710	nd	putative somatic
40	BMP7	1	-	g.chr20:55840929G>T	c.250C>A	p.L84M	Missense_Mutation	neutral	NP_001710	nd	putative somatic
40	BMP7	1	-	g.chr20:55840958C>T	c.221G>A	p.G74D	Missense_Mutation	neutral	NP_001710	nd	putative somatic
17	CACNA1H	33,34	+	g.chr16:1270496delC	c.6564delC	p.S2188fs	Frame_Shift_Del		NP_066921	nd	putative somatic
41	CACNA1H	33,34	+	g.chr16:1270329C>T	c.6397C>T	p.R2133C	Missense_Mutation	neutral	NP_066921	nd	
43	CACNA1H	27,28	+	g.chr16:1265319C>T	c.5117C>T	p.A1706V	Missense_Mutation	deleterious	NP_066921	nd	
6	CADPS	19,20,21	-	g.chr3:62477131T>G	c.2909A>C	p.H970P	Missense_Mutation	deleterious	NP_003707	nd	
10	CADPS	21,22,24	-	g.chr3:62459846A>T	c.3477_splice	p.E1159_splice	Splice_Site		NP_003707	nd	
45	CADPS	21,22,24	-	g.chr3:62459846A>T	c.3477_splice	p.E1159_splice	Splice_Site		NP_003707	nd	
4	CARD11	5	-	g.chr7:2979559C>T	c.688G>A	p.D230N	Missense_Mutation	deleterious	NP_115791	Yes	
44	CARD11	20	-	g.chr7:2954991C>T	c.2719G>A	p.E907K	Missense_Mutation	neutral	NP_115791	Yes	
47	CARD11	4	-	g.chr7:2984142A>T	c.388T>A	p.F130I	Missense_Mutation	deleterious	NP_115791	Yes	
48	CARD11	7	-	g.chr7:2977614T>A	c.1070A>T	p.D357V	Missense_Mutation	deleterious	NP_115791	Yes	
12	CCND3	2,4,5	-	g.chr6:41903731G>A	c.826C>T	p.Q276*	Nonsense_Mutation		NP_001751	nd	putative somatic
37	CCND3	2,4,5	-	g.chr6:41903745_41903746insG	c.811_812insC	p.R271fs	Frame_Shift_Ins		NP_001751	Yes	somatic
38	CCND3	2,4,5	-	g.chr6:41903697A>C	c.860T>G	p.V287G	Missense_Mutation	deleterious	NP_001751	Yes	
44	CCND3	2,4,5	-	g.chr6:41903688A>C	c.869T>G	p.I290R	Missense_Mutation	neutral	NP_001751	Yes	
1	CD58*	3	-	g.chr1:117078587C>T	c.628G>A	p.G210S	Missense_Mutation	neutral	NP_001770	Yes	somatic
25	CD58*	2	-	g.chr1:117087019_117087022del4bp	c.275_278del4bp	p.I92fs	Frame_Shift_Del		NP_001770	nd	
2	CD79B	5,6	-	g.chr17:62006656_62006657insA	c.619_620insT	p.Y207fs	Frame_Shift_Ins		NP_000617	Yes	somatic
18	CD79B	5,6	-	g.chr17:62006604C>G	c.672G>C	p.E224D	Missense_Mutation	deleterious	NP_000617	Yes	
18	CD79B	5,6	-	g.chr17:62006657A>T	c.619T>A	p.Y207N	Missense_Mutation	deleterious	NP_000617	Yes	
24	CD79B	3	-	g.chr17:62007718C>T	c.146G>A	p.S49N	Missense_Mutation	neutral	NP_000617	nd	
7	CDKN2A*	2	-	g.chr9:21971142G>T	c.216C>A	p.C72*	Nonsense_Mutation		NP_000068	Yes	somatic
2	CREBBP	29,30	-	g.chr16:3781323_3781325delAAG	c.5040_5042delCTT	p.L1681del	In_Frame_Del		NP_004371	Yes	somatic
2	CREBBP	26,27	-	g.chr16:3786704A>G	c.4507T>C	p.Y1503H	Missense_Mutation	none	NP_004371	Yes	somatic
3	CREBBP	26,27	-	g.chr16:3786704A>G	c.4507T>C	p.Y1503H	Missense_Mutation	none	NP_004371	nd	somatic
4	CREBBP	26,27	-	g.chr16:3786760A>G	c.4451T>C	p.F1484S	Missense_Mutation	none	NP_004371	Yes	
6	CREBBP	25,26	-	g.chr16:3788617C>T	c.4337G>A	p.R1446H	Missense_Mutation	none	NP_004371	Yes	
7	CREBBP	29,30	-	g.chr16:3781324_3781326delAGG	c.5039_5041delCCT	p.S1680del	In_Frame_Del		NP_004371	nd	somatic
15	CREBBP	26,27	-	g.chr16:3786761A>C	c.4450T>G	p.F1484V	Missense_Mutation	none	NP_004371	nd	
17	CREBBP	26,27	-	g.chr16:3786763A>T	c.4448T>A	p.I1483N	Missense_Mutation	none	NP_004371	Yes	
20	CREBBP	23,24	-	g.chr16:3790421A>T	c.4112T>A	p.V1371D	Missense_Mutation	none	NP_004371	nd	
22	CREBBP	29,30	-	g.chr16:3781324_3781326delAGG	c.5039_5041delCCT	p.S1680del	In_Frame_Del		NP_004371	nd	
23	CREBBP	26,27	-	g.chr16:3786708_3786710delCTC	c.4501_4503delGAG	p.E1501del	In_Frame_Del		NP_004371	Yes	somatic
24	CREBBP	25,26	-	g.chr16:3788647C>A	c.4307G>T	p.S1436I	Missense_Mutation	none	NP_004371	nd	
29	CREBBP	29,30	-	g.chr16:3781324_3781326delAGG	c.5039_5041delCCT	p.S1680del	In_Frame_Del		NP_004371	Yes	
34	CREBBP	25,26	-	g.chr16:3788617C>T	c.4337G>A	p.R1446H	Missense_Mutation	none	NP_004371	nd	
37	CREBBP	25,26	-	g.chr16:3788618G>A	c.4336C>T	p.R1446C	Missense_Mutation	none	NP_004371	nd	somatic
38	CREBBP	25,26	-	g.chr16:3788618G>A	c.4336C>T	p.R1446C	Missense_Mutation	none	NP_004371	nd	
39	CREBBP	25,26	-	g.chr16:3788617C>T	c.4337G>A	p.R1446H	Missense_Mutation	none	NP_004371	Yes	
39	CREBBP	22,23	-	g.chr16:3794893A>T	c.3982_splice	p.R1328_splice	Splice_Site		NP_004371	Yes	

42	CREBBP	13,14	-	g.chr16:3820843G>A	c.2608C>T	p.Q870*	Nonsense_Mutation		NP_004371	nd	
43	CREBBP	26,27	-	g.chr16:3786748G>A	c.4463C>T	p.P1488L	Missense_Mutation	none	NP_004371	nd	
47	CREBBP	15,16	-	g.chr16:3817912T>C	c.3061_splice	p.M1021_splice	Splice_Site		NP_004371	nd	
4	CSMD2	23	-	g.chr1:34166260T>C	c.3470A>G	p.N1157S	Missense_Mutation	neutral	NP_443128	Yes	putative somatic
22	CSMD2	1	-	g.chr1:34631393G>A	c.22C>T	p.P8S	Missense_Mutation	neutral	NP_443128	nd	
25	CSMD2	38	-	g.chr1:34083196G>A	c.5729C>T	p.P1910L	Missense_Mutation	deleterious	NP_443128	nd	
4	DCC	18	+	g.chr18:50923807T>C	c.2818T>C	p.Y940H	Missense_Mutation	neutral	NP_005206	Yes	putative somatic
27	DCC	3	+	g.chr18:50432615G>A	c.614G>A	p.G205E	Missense_Mutation	neutral	NP_005206	nd	
44	DCC	2	+	g.chr18:50278526C>A	c.194C>A	p.S65Y	Missense_Mutation	deleterious	NP_005206	nd	
7	DMD	32	-	g.chrX:32407765C>G	c.4371G>C	p.K1457N	Missense_Mutation	deleterious	NP_003997	nd	somatic
8	DMD	13,14,21,24,25,75	-	g.chrX:31165405T>A	c.10784A>T	p.Q3595L	Missense_Mutation	deleterious	NP_003997	nd	putative somatic
18	DMD	32	-	g.chrX:32407734C>A	c.4402G>T	p.E1468*	Nonsense_Mutation		NP_003997	nd	
20	DMD	40	-	g.chrX:32361293delT	c.5697delA	p.K1899fs	Frame_Shift_Del		NP_003997	Yes	
28	DMD	17,28,78	-	g.chrX:31144764T>C	c.11041A>G	p.R3681G	Missense_Mutation	deleterious	NP_003997	Yes	
42	DMD	9	-	g.chrX:32716063C>A	c.884G>T	p.R295L	Missense_Mutation	deleterious	NP_003997	nd	
49	DMD	44	-	g.chrX:32235154C>A	c.6317G>T	p.W2106L	Missense_Mutation	deleterious	NP_003997	nd	
2	DSC3	8	-	g.chr18:28598745T>G	c.964A>C	p.I322L	Missense_Mutation	neutral	NP_001932	Yes	somatic
41	DSC3	9	-	g.chr18:28598180A>C	c.1120T>G	p.L374V	Missense_Mutation	deleterious	NP_001932	nd	
43	DSC3	11	-	g.chr18:28588069A>C	c.1576T>G	p.S526A	Missense_Mutation	neutral	NP_001932	nd	
1	DTX1	2	+	g.chr12:113515684G>A	c.715G>A	p.G239R	Missense_Mutation	neutral	NP_004407	nd	somatic
18	DTX1	1	+	g.chr12:113496097C>T	c.100C>T	p.H34Y	Missense_Mutation	deleterious	NP_004407	nd	
23	DTX1	1	+	g.chr12:113496076G>A	c.79G>A	p.V27M	Missense_Mutation	deleterious	NP_004407	nd	somatic
23	DTX1	2	+	g.chr12:113515546G>A	c.577G>A	p.G193S	Missense_Mutation	deleterious	NP_004407	nd	somatic
40	DTX1	1	+	g.chr12:113496037G>A	c.40G>A	p.G14S	Missense_Mutation	deleterious	NP_004407	nd	putative somatic
40	DTX1	1	+	g.chr12:113496145A>G	c.148A>G	p.N50D	Missense_Mutation	deleterious	NP_004407	nd	putative somatic
11	EBF1	3	-	g.chr5:158523356C>A	c.350G>T	p.S117I	Missense_Mutation	deleterious	NP_076870	nd	somatic
28	EBF1	10	-	g.chr5:158204520G>A	c.937C>T	p.Q313*	Nonsense_Mutation		NP_076870	nd	
39	EBF1	2	-	g.chr5:158524075G>C	c.198C>G	p.N66K	Missense_Mutation	deleterious	NP_076870	nd	putative somatic
17	EML6	14	+	g.chr2:55096338delG	c.2072delG	p.R691fs	Frame_Shift_Del		NP_001034842	nd	putative somatic
25	EML6	15	+	g.chr2:55098670A>G	c.2272A>G	p.T758A	Missense_Mutation	neutral	NP_001034842	nd	
34	EML6	1	+	g.chr2:54952205G>T	c.7G>T	p.D3Y	Missense_Mutation	deleterious	NP_001034842	nd	
5	EP300	5	+	g.chr22:41525969T>C	c.1244T>C	p.L415P	Missense_Mutation	deleterious	NP_001420	Yes	somatic
5	EP300	27	+	g.chr22:41566470T>G	c.4347T>G	p.H1449Q	Missense_Mutation	none	NP_001420	Yes	somatic
5	EP300	27	+	g.chr22:41566522T>A	c.4399T>A	p.Y1467N	Missense_Mutation	none	NP_001420	Yes	somatic
5	EP300	27	+	g.chr22:41566571A>G	c.4448A>G	p.Y1483C	Missense_Mutation	none	NP_001420	Yes	somatic
7	EP300	30	+	g.chr22:41572350C>T	c.4879C>T	p.R1627W	Missense_Mutation	none	NP_001420	nd	somatic
12	EP300	10	+	g.chr22:41537176C>T	c.2003C>T	p.S668F	Missense_Mutation	deleterious	NP_001420	Yes	putative somatic
27	EP300	1	+	g.chr22:41489097G>A	c.89G>A	p.G30D	Missense_Mutation	deleterious	NP_001420	nd	
49	EP300	3	+	g.chr22:41521925C>G	c.787C>G	p.Q263E	Missense_Mutation	deleterious	NP_001420	nd	
22	EPAS1	8	+	g.chr2:46602873G>A	c.931G>A	p.A311T	Missense_Mutation	neutral	NP_001421	nd	
23	EPAS1	3	+	g.chr2:46583316G>A	c.244G>A	p.E82K	Missense_Mutation	neutral	NP_001421	nd	somatic
43	EPAS1	7	+	g.chr2:46597053G>C	c.867G>C	p.M289I	Missense_Mutation	deleterious	NP_001421	nd	
11	EPHA7	12	-	g.chr6:93967203C>T	c.2149G>A	p.G717R	Missense_Mutation	deleterious	NP_004431	Yes	somatic
26	EPHA7	2	-	g.chr6:94124454_94124457del4bp	c.126_129del4bp	p.Q42fs	Frame_Shift_Del		NP_004431	nd	
36	EPHA7	5	-	g.chr6:94066635T>C	c.1124A>G	p.E375G	Missense_Mutation	deleterious	NP_004431	nd	
6	EZH2	14,15	-	g.chr7:148508727T>G	c.1922A>C	p.Y641S	Missense_Mutation	deleterious	NP_004447	Yes	
22	EZH2	14,15	-	g.chr7:148508727T>A	c.1922A>T	p.Y641F	Missense_Mutation	deleterious	NP_004447	nd	
28	EZH2	14,15	-	g.chr7:148508727T>A	c.1922A>T	p.Y641F	Missense_Mutation	deleterious	NP_004447	nd	
29	EZH2	14,15	-	g.chr7:148508727T>A	c.1922A>T	p.Y641F	Missense_Mutation	deleterious	NP_004447	Yes	
34	EZH2	14,15	-	g.chr7:148508728A>T	c.1921T>A	p.Y641N	Missense_Mutation	deleterious	NP_004447	nd	
35	EZH2	14,15	-	g.chr7:148508728A>T	c.1921T>A	p.Y641N	Missense_Mutation	deleterious	NP_004447	nd	
38	EZH2	14,15	-	g.chr7:148508728A>T	c.1921T>A	p.Y641N	Missense_Mutation	deleterious	NP_004447	nd	
39	EZH2	16,17	-	g.chr7:148506467G>C	c.2030C>G	p.A677G	Missense_Mutation	deleterious	NP_004447	Yes	
44	EZH2	14,15	-	g.chr7:148508727T>G	c.1922A>C	p.Y641S	Missense_Mutation	deleterious	NP_004447	nd	
47	EZH2	14,15	-	g.chr7:148508727T>A	c.1922A>T	p.Y641F	Missense_Mutation	deleterious	NP_004447	nd	
23	FAM129B	4	-	g.chr9:130287433G>A	c.325C>T	p.R109W	Missense_Mutation		NP_073744	nd	somatic
35	FAM129B	18,19	-	g.chr9:130279252C>T	c.857G>A	p.R286H	Missense_Mutation		NP_073744	nd	
41	FAM129B	5	-	g.chr9:130286047G>A	c.500C>T	p.A167V	Missense_Mutation		NP_073744	nd	
27	FAM190A	5	+	g.chr4:91549337C>T	c.1886C>T	p.T629M	Missense_Mutation	deleterious	NP_001138537	nd	
38	FAM190A	1	+	g.chr4:91230369G>A	c.934G>A	p.G312R	Missense_Mutation	deleterious	NP_001138537	nd	
40	FAM190A	3	+	g.chr4:91321254T>A	c.1577T>A	p.L526*	Nonsense_Mutation		NP_001138537	nd	putative somatic
2	FAS	8,9	+	g.chr10:90773895T>A	c.696T>A	p.Y232*	Nonsense_Mutation		NP_000034	Yes	somatic
28	FAS	1	+	g.chr10:90750635T>C	c.2T>C	p.M1T	Missense_Mutation	deleterious	NP_000034	nd	
47	FAS	8,9	+	g.chr10:90774149A>T	c.950A>T	p.D317V	Missense_Mutation	deleterious	NP_000034	Yes	

48	FAS	6,7	+	g.chr10:90771839G>A	c.651_splice	p.P217_splice	Splice_Site		NP_000034	Yes	
11	FAT4	1	+	g.chr4:126238027T>C	c.461T>C	p.V154A	Missense_Mutation	neutral	NP_078858	nd	somatic
20	FAT4	9	+	g.chr4:126371142C>A	c.8971C>A	p.P2991T	Missense_Mutation	deleterious	NP_078858	nd	
27	FAT4	9	+	g.chr4:126372196G>A	c.10025G>A	p.R3342Q	Missense_Mutation	deleterious	NP_078858	nd	
42	FAT4	9	+	g.chr4:126370676G>T	c.8505G>T	p.R2835S	Missense_Mutation	deleterious	NP_078858	nd	
44	FAT4	1	+	g.chr4:126242482A>G	c.4916A>G	p.K1639R	Missense_Mutation	deleterious	NP_078858	nd	
48	FAT4	7	+	g.chr4:126355504G>C	c.7123G>C	p.E2375Q	Missense_Mutation	deleterious	NP_078858	nd	
3	FBN2	28	-	g.chr5:127671763C>A	c.3641G>T	p.G1214V	Missense_Mutation	deleterious	NP_001990	Yes	somatic
5	FBN2	52	-	g.chr5:127624883G>T	c.6573C>A	p.D2191E	Missense_Mutation	deleterious	NP_001990	nd	somatic
11	FBN2	4	-	g.chr5:127863625T>C	c.472A>G	p.T158A	Missense_Mutation	none	NP_001990	nd	somatic
24	FBN2	10	-	g.chr5:127728861C>T	c.1432G>A	p.A478T	Missense_Mutation	neutral	NP_001990	Yes	
42	FBN2	15	-	g.chr5:127710369C>T	c.2047G>A	p.D683N	Missense_Mutation	neutral	NP_001990	Yes	
45	FBN2	20	-	g.chr5:127693022T>G	c.2620A>C	p.N874H	Missense_Mutation	neutral	NP_001990	Yes	
10	FBXO11	5	-	g.chr2:48062012T>G	c.466_splice	p.Y156_splice	Splice_Site		NP_079409	Yes	
11	FBXO11	22	-	g.chr2:48035373A>C	c.2416T>G	p.C806G	Missense_Mutation	deleterious	NP_079409	Yes	somatic
15	FLNB	9	+	g.chr3:58087956A>G	c.1372A>G	p.S458G	Missense_Mutation	neutral	NP_001448	nd	
23	FLNB	16	+	g.chr3:58095850delC	c.2437delC	p.P813fs	Frame_Shift_Del		NP_001448	nd	somatic
26	FLNB	5	+	g.chr3:58080647T>C	c.872T>C	p.V291A	Missense_Mutation	neutral	NP_001448	nd	
3	FMN2	4	+	g.chr1:240351543G>A	c.1967G>A	p.R656H	Missense_Mutation	none	NP_064450	Yes	somatic
10	FMN2	5	+	g.chr1:240371342C>T	c.3230C>T	p.P1077L	Missense_Mutation	none	NP_064450	nd	
20	FMN2	2	+	g.chr1:240286518A>G	c.1655A>G	p.N552S	Missense_Mutation	none	NP_064450	Yes	
20	FMN2	5	+	g.chr1:240370751C>A	c.2639C>A	p.P880H	Missense_Mutation	none	NP_064450	Yes	
20	FMN2	5	+	g.chr1:240370907T>C	c.2795T>C	p.L932P	Missense_Mutation	none	NP_064450	nd	
23	FMN2	1	+	g.chr1:240255566G>C	c.157G>C	p.G53R	Missense_Mutation	none	NP_064450	nd	somatic
24	FMN2	5	+	g.chr1:240370358C>T	c.2246C>T	p.T749M	Missense_Mutation	none	NP_064450	Yes	
28	FMN2	3	+	g.chr1:240341320C>T	c.1882C>T	p.P628S	Missense_Mutation	none	NP_064450	nd	
11	FOXO1	1	-	g.chr13:41240126C>G	c.224G>C	p.S75T	Missense_Mutation	neutral	NP_002006	Yes	somatic
28	FOXO1	1	-	g.chr13:41240286A>G	c.64T>C	p.S22P	Missense_Mutation	deleterious	NP_002006	Yes	
36	FOXO1	2	-	g.chr13:41134980A>T	c.648T>A	p.N216K	Missense_Mutation	deleterious	NP_002006	Yes	
38	FOXO1	1	-	g.chr13:41240286A>G	c.64T>C	p.S22P	Missense_Mutation	deleterious	NP_002006	Yes	
48	FOXO1	1	-	g.chr13:41240279G>A	c.71C>T	p.T24I	Missense_Mutation	deleterious	NP_002006	Yes	
49	FOXO1	1	-	g.chr13:41240294C>G	c.56G>C	p.R19P	Missense_Mutation	deleterious	NP_002006	Yes	
6	GCC2	19	+	g.chr2:109109241C>G	c.4442C>G	p.T1481R	Missense_Mutation	deleterious	NP_852118	nd	
25	GCC2	13	+	g.chr2:109100725A>G	c.3571A>G	p.K1191E	Missense_Mutation		NP_852118	nd	
28	GCC2	6	+	g.chr2:109087929_109087931delAAG	c.2144_2146delAAG	p.E716del	In_Frame_Del		NP_852118	nd	
12	GFRAL	4	+	g.chr6:55214931C>T	c.358C>T	p.R120C	Missense_Mutation	neutral	NP_997293	nd	putative somatic
20	GFRAL	5	+	g.chr6:55216092T>C	c.412T>C	p.C138R	Missense_Mutation	deleterious	NP_997293	nd	
22	GFRAL	6	+	g.chr6:55223895G>A	c.911G>A	p.C304Y	Missense_Mutation	deleterious	NP_997293	nd	
2	GNAI3	3	-	g.chr17:63014369A>T	c.561_splice	p.P187_splice	Splice_Site		NP_006563	Yes	somatic
7	GNAI3	1	-	g.chr17:63052484_63052501del18bp	c.211_228del18bp	p.IHGQDF71del	In_Frame_Del		NP_006563	Yes	somatic
22	GNAI3	4	-	g.chr17:63010820C>T	c.689G>A	p.R230K	Missense_Mutation	neutral	NP_006563	Yes	
22	GNAI3	3	-	g.chr17:63014390A>T	c.542T>A	p.L181*	Nonsense_Mutation		NP_006563	Yes	
22	GNAI3	3	-	g.chr17:63014395A>C	c.537T>G	p.D179E	Missense_Mutation	neutral	NP_006563	Yes	
29	GNAI3	3	-	g.chr17:63014369A>G	c.561_splice	p.P187_splice	Splice_Site		NP_006563	Yes	putative somatic
38	GNAI3	4	-	g.chr17:63010844T>C	c.665A>G	p.D222G	Missense_Mutation	deleterious	NP_006563	Yes	
42	GNAI3	3	-	g.chr17:63014404A>T	c.528T>A	p.Y176*	Nonsense_Mutation		NP_006563	Yes	
42	GNAI3	1	-	g.chr17:63052447A>C	c.265T>G	p.Y89D	Missense_Mutation	deleterious	NP_006563	Yes	
42	GNAI3	1	-	g.chr17:63052460G>T	c.252C>A	p.F84L	Missense_Mutation	deleterious	NP_006563	Yes	
45	GNAI3	4	-	g.chr17:63010844T>C	c.665A>G	p.D222G	Missense_Mutation	deleterious	NP_006563	nd	
17	GNAI3	4	-	g.chr17:63010375T>G	c.1134A>C	p.*378C	Nonstop_Mutation		NP_006563	Yes	
47	GNAI2	5	+	g.chr3:50293622A>G	c.465_splice	p.Y155_splice	Splice_Site		NP_002061	nd	
17	GNAI2	5	+	g.chr3:50293701A>T	c.542A>T	p.K181M	Missense_Mutation	deleterious	NP_002061	nd	putative somatic
40	GNAI2	5	+	g.chr3:50293703A>G	c.544A>G	p.T182A	Missense_Mutation	deleterious	NP_002061	nd	putative somatic
40	GNAI2	5	+	g.chr3:50293712A>G	c.553A>G	p.I185V	Missense_Mutation	neutral	NP_002061	nd	putative somatic
2	GRID2	11	+	g.chr4:94376882C>T	c.1615C>T	p.R539C	Missense_Mutation	deleterious	NP_001501	nd	somatic
15	GRID2	3	+	g.chr4:94006232G>A	c.331G>A	p.A111T	Missense_Mutation	deleterious	NP_001501	nd	
26	GRID2	15	+	g.chr4:94690585G>A	c.2585G>A	p.R862Q	Missense_Mutation	neutral	NP_001501	nd	
35	GRID2	11	+	g.chr4:94377069C>T	c.1802C>T	p.T601M	Missense_Mutation	neutral	NP_001501	nd	
5	HERC1	77	-	g.chr15:63901455C>A	c.14411G>T	p.S4804I	Missense_Mutation	deleterious	NP_003913	nd	somatic
26	HERC1	65	-	g.chr15:63927091A>C			Splice_Site		NP_003913	nd	
43	HERC1	38	-	g.chr15:63964843G>C	c.7897C>G	p.P2633A	Missense_Mutation	deleterious	NP_003913	nd	
1	HIST1H1E	1	+	g.chr6:26157118C>G	c.500C>G	p.A167G	Missense_Mutation	none	NP_005312	nd	somatic
2	HIST1H1E	1	+	g.chr6:26157109C>T	c.491C>T	p.A164V	Missense_Mutation	none	NP_005312	Yes	somatic
11	HIST1H1E	1	+	g.chr6:26156744G>C	c.126G>C	p.E42D	Missense_Mutation	neutral	NP_005312	Yes	somatic

25	HIST1H1E	1	+	g.chr6:26156852C>A	c.234C>A	p.S78R	Missense_Mutation	neutral	NP_005312	nd	
25	HIST1H1E	1	+	g.chr6:26157134C>G	c.516C>G	p.S172R	Missense_Mutation	none	NP_005312	nd	
27	HIST1H1E	1	+	g.chr6:26156926G>A	c.308G>A	p.G103D	Missense_Mutation	deleterious	NP_005312	nd	
36	HIST1H1E	1	+	g.chr6:26156970C>T	c.352C>T	p.P118S	Missense_Mutation	none	NP_005312	nd	
36	HIST1H1E	1	+	g.chr6:26157013C>T	c.395C>T	p.A132V	Missense_Mutation	none	NP_005312	nd	
36	HIST1H1E	1	+	g.chr6:26157109C>T	c.491C>T	p.A164V	Missense_Mutation	none	NP_005312	nd	
49	HIST1H1E	1	+	g.chr6:26156890G>A	c.272G>A	p.G91D	Missense_Mutation	deleterious	NP_005312	nd	
1	HIST1H2AC	1	+	g.chr6:26124663G>A	c.203G>A	p.G68D	Missense_Mutation	deleterious	NP_003503	nd	somatic
2	HIST1H2AC	1	+	g.chr6:26124578T>A	c.118T>A	p.Y40N	Missense_Mutation	deleterious	NP_003503	Yes	somatic
44	HIST1H2AC	1	+	g.chr6:26124563C>G	c.103C>G	p.L35V	Missense_Mutation	neutral	NP_003503	nd	
23	HIST1H3D	1	-	g.chr6:26197090C>T	c.389G>A	p.R130H	Missense_Mutation	neutral	NP_003521	nd	somatic
28	HIST1H3D	1	-	g.chr6:26197384G>C	c.95C>G	p.A32G	Missense_Mutation	neutral	NP_003521	nd	
44	HIST1H3D	1	-	g.chr6:26197219C>G	c.260G>C	p.S87T	Missense_Mutation	neutral	NP_003521	nd	
2	HNRNPU	1	-	g.chr1:245027102G>A	c.508C>T	p.Q170*	Nonsense_Mutation		NP_114032	Yes	somatic
28	HNRNPU	1	-	g.chr1:245027099G>A	c.511C>T	p.Q171*	Nonsense_Mutation		NP_114032	nd	
47	HNRNPU	9	-	g.chr1:245020097G>C	c.1676C>G	p.A559G	Missense_Mutation	deleterious	NP_114032	nd	
8	IRF4	2	+	g.chr6:394888C>A	c.284C>A	p.T95K	Missense_Mutation	deleterious	NP_002451	Yes	putative somatic
11	IRF4	2	+	g.chr6:394899T>C	c.295T>C	p.C99R	Missense_Mutation	deleterious	NP_002451	nd	somatic
17	IRF4	1	+	g.chr6:393222C>T	c.70C>T	p.L24F	Missense_Mutation	deleterious	NP_002451	nd	putative somatic
28	IRF8	8	+	g.chr16:85954883_85954886del4bp	c.1276_1279del4bp	p.V426fs	Frame_Shift_Del		NP_002154	Yes	
35	IRF8	8	+	g.chr16:85954781C>T	c.1174C>T	p.Q392*	Nonsense_Mutation		NP_002154	Yes	
39	IRF8	8	+	g.chr16:85954866_85954872del7bp	c.1259_1265del7bp	p.E420fs	Frame_Shift_Del		NP_002154	Yes	putative somatic
42	IRF8	2	+	g.chr16:85942740T>A	c.319T>A	p.Y107N	Missense_Mutation	deleterious	NP_002154	Yes	
42	IRF8	2	+	g.chr16:85942756T>G	c.335T>G	p.I112S	Missense_Mutation	deleterious	NP_002154	Yes	
11	ITPKB	1	-	g.chr1:226923671C>G	c.1489G>C	p.D497H	Missense_Mutation	deleterious	NP_002212	Yes	somatic
11	ITPKB	1	-	g.chr1:226924292C>T	c.868G>A	p.A290T	Missense_Mutation	neutral	NP_002212	Yes	somatic
40	ITPKB	1	-	g.chr1:226925103C>G	c.57G>C	p.E19D	Missense_Mutation	neutral	NP_002212	nd	putative somatic
24	ITPR1	40	+	g.chr3:4808285G>T	c.5472G>T	p.M1824I	Missense_Mutation	deleterious	NP_001093422	nd	
35	ITPR1	8	+	g.chr3:4687333A>G	c.776A>G	p.K259R	Missense_Mutation	neutral	NP_001093422	nd	
39	ITPR1	41	+	g.chr3:4810375C>T	c.5762C>T	p.A1921V	Missense_Mutation	neutral	NP_001093422	nd	putative somatic
27	KDM6B*	8	+	g.chr17:7751857G>A	c.2251G>A	p.V751I	Missense_Mutation		NP_001073893	nd	
42	KDM6B*	3	+	g.chr17:7749575G>A	c.416G>A	p.S139N	Missense_Mutation	neutral	NP_001073893	nd	
7	KLHL6	7	-	g.chr3:183209942C>T	c.1639G>A	p.E547K	Missense_Mutation	neutral	NP_569713	Yes	somatic
7	KLHL6	4	-	g.chr3:183217590C>T	c.935G>A	p.R312K	Missense_Mutation	deleterious	NP_569713	Yes	somatic
10	KLHL6	7	-	g.chr3:183209954G>A	c.1627C>T	p.Q543*	Nonsense_Mutation		NP_569713	Yes	
10	KLHL6	1	-	g.chr3:183273222C>T	c.220G>A	p.V74I	Missense_Mutation	deleterious	NP_569713	Yes	
12	KLHL6	3	-	g.chr3:183226202_183226203insA	c.553_554insT	p.S185fs	Frame_Shift_Ins		NP_569713	Yes	putative somatic
27	KLHL6	1	-	g.chr3:183273174G>C	c.268C>G	p.L90V	Missense_Mutation	deleterious	NP_569713	Yes	
5	MAST4	28	+	g.chr5:66460745T>C	c.5171T>C	p.V1724A	Missense_Mutation	neutral	NP_055998	nd	somatic
35	MAST4	28	+	g.chr5:66461050_66461052delCTC	c.5476_5478delCTC	p.L1827del	In_Frame_Del		NP_055998	nd	
47	MAST4	28	+	g.chr5:66461005C>G	c.5431C>G	p.L1811V	Missense_Mutation	neutral	NP_055998	nd	
27	MEF2B	7,8	-	g.chr19:19256757T>A	c.977A>T	p.K326M	Missense_Mutation		NP_001139257	Yes	
28	MEF2B	1	-	g.chr19:19261528A>C	c.17T>G	p.I6S	Missense_Mutation		NP_001139257	Yes	
38	MEF2B	1	-	g.chr19:19261535T>C	c.10A>G	p.K4E	Missense_Mutation		NP_001139257	Yes	
1	MLL2	8	-	g.chr12:49446772A>T	c.1038T>A	p.C346*	Nonsense_Mutation		NP_003473	nd	somatic
1	MLL2	5	-	g.chr12:49447861delC	c.573delG	p.R191fs	Frame_Shift_Del		NP_003473	nd	somatic
3	MLL2	14	-	g.chr12:49441836C>T	c.4148G>A	p.C1383Y	Missense_Mutation		NP_003473	nd	somatic
4	MLL2	34	-	g.chr12:49432396G>A	c.8743C>T	p.R2915*	Nonsense_Mutation		NP_003473	Yes	putative somatic
4	MLL2	31	-	g.chr12:49434889G>A	c.6664C>T	p.Q2222*	Nonsense_Mutation		NP_003473	Yes	putative somatic
5	MLL2	53	-	g.chr12:49415877_49415878delTT	c.16469_16470delAA	p.K5490fs	Frame_Shift_Del		NP_003473	Yes	somatic
5	MLL2	5	-	g.chr12:49433328_49433329insA	c.8118_8119insT	p.T2706fs	Frame_Shift_Ins		NP_003473	Yes	somatic
6	MLL2	44	-	g.chr12:49422905C>T	c.14190G>A	p.W4730*	Nonsense_Mutation		NP_003473	Yes	
6	MLL2	32	-	g.chr12:49433247G>A	c.8200C>T	p.R2734*	Nonsense_Mutation		NP_003473	Yes	
7	MLL2	15	-	g.chr12:49440521C>A	c.4289G>T	p.C1430F	Missense_Mutation		NP_003473	nd	somatic
7	MLL2	14	-	g.chr12:49441848_49441849delAT	c.4135_4136delAT	p.M1379fs	Frame_Shift_Del		NP_003473	nd	somatic
8	MLL2	10	-	g.chr12:49445934G>C	c.1532C>G	p.P511R	Missense_Mutation	none	NP_003473	Yes	
11	MLL2	31	-	g.chr12:49433620G>A	c.7933C>T	p.R2645*	Nonsense_Mutation		NP_003473	Yes	somatic
11	MLL2	15	-	g.chr12:49440452T>C	c.4358A>G	p.H1453R	Missense_Mutation		NP_003473	Yes	somatic
12	MLL2	39	-	g.chr12:49425091_49425113del23bp	c.13375_13397del23bp	p.G4459fs	Frame_Shift_Del		NP_003473	Yes	
12	MLL2	31	-	g.chr12:49434525delG	c.7028delC	p.P2343fs	Frame_Shift_Del		NP_003473	Yes	
15	MLL2	39	-	g.chr12:49426469G>A	c.12019C>T	p.Q4007*	Nonsense_Mutation		NP_003473	nd	
15	MLL2	31	-	g.chr12:49434958delA	c.6595delT	p.Y2199fs	Frame_Shift_Del		NP_003473	nd	
17	MLL2	39	-	g.chr12:49426351C>T	c.12137G>A	p.G4046E	Missense_Mutation		NP_003473	Yes	
18	MLL2	48	-	g.chr12:49420245delG	c.15504delC	p.I5168fs	Frame_Shift_Del		NP_003473	nd	

23	MLL2	32	-	g.chr12:49433388G>A	c.8059C>T	p.R2687*	Nonsense_Mutation		NP_003473	Yes	somatic
23	MLL2	7	-	g.chr12:49447105C>T	c.840_splice	p.R280_splice	Splice_Site		NP_003473	Yes	somatic
28	MLL2	16	-	g.chr12:49440191G>A	c.4435C>T	p.Q1479*	Nonsense_Mutation		NP_003473	nd	
28	MLL2	10	-	g.chr12:49445287_49445288insC	c.2178_2179insG	p.L726fs	Frame_Shift_Ins		NP_003473	nd	
29	MLL2	17	-	g.chr12:49439908G>A	c.4633C>T	p.L1545*	Nonsense_Mutation		NP_003473	Yes	
33	MLL2	39	-	g.chr12:49425072_49425073delCA	c.13415_13416delTG	p.V4472fs	Frame_Shift_Del		NP_003473	nd	somatic
33	MLL2	39	-	g.chr12:49426187G>A	c.12301C>T	p.Q4101*	Nonsense_Mutation		NP_003473	nd	somatic
34	MLL2	48	-	g.chr12:49420066T>C	c.15683A>G	p.Y5228C	Missense_Mutation		NP_003473	nd	
34	MLL2	31	-	g.chr12:49434629delC	c.6924delC	p.D2308fs	Frame_Shift_Del		NP_003473	Yes	
34	MLL2	6	-	g.chr12:49447411A>T	c.687T>A	p.C229*	Nonsense_Mutation		NP_003473	nd	
35	MLL2	10	-	g.chr12:49446206delC	c.1260delG	p.G420fs	Frame_Shift_Del		NP_003473	nd	
36	MLL2	41	-	g.chr12:49424459_49424477del19bp	c.13746_13764del19bp	p.G4582fs	Frame_Shift_Del		NP_003473	Yes	
36	MLL2	31	-	g.chr12:49434005delG	c.7548delC	p.P2516fs	Frame_Shift_Del		NP_003473	nd	
37	MLL2	53	-	g.chr12:49415846G>A	c.16501C>T	p.R5501*	Nonsense_Mutation		NP_003473	Yes	somatic
38	MLL2	36	-	g.chr12:49428382C>T	c.10423G>A	p.A3475T	Missense_Mutation		NP_003473	nd	
38	MLL2	34	-	g.chr12:49432694delT	c.8445delA	p.Q2815fs	Frame_Shift_Del		NP_003473	Yes	
38	MLL2	34	-	g.chr12:49432696G>A	c.8443C>T	p.Q2815*	Nonsense_Mutation		NP_003473	nd	
39	MLL2	48	-	g.chr12:49420474C>A	c.15275G>T	p.C5092F	Missense_Mutation		NP_003473	Yes	
39	MLL2	3	-	g.chr12:49448310C>T	c.400_splice	p.G134_splice	Splice_Site		NP_003473	Yes	
41	MLL2	31	-	g.chr12:49433515G>A	c.8038C>T	p.Q2680*	Nonsense_Mutation		NP_003473	nd	
43	MLL2	39	-	g.chr12:49425704G>A	c.12784C>T	p.Q4262*	Nonsense_Mutation		NP_003473	nd	
45	MLL2	51	-	g.chr12:49416497C>T	c.16214G>A	p.R5405H	Missense_Mutation		NP_003473	nd	
45	MLL2	39	-	g.chr12:49425325_49425332del8bp	c.13156_13163del8bp	p.T4386fs	Frame_Shift_Del		NP_003473	nd	
45	MLL2	27	-	g.chr12:49436430T>G	c.5783_splice	p.G1928_splice	Splice_Site		NP_003473	nd	
49	MLL2	13	-	g.chr12:49442444G>A	c.4129C>T	p.Q1377*	Nonsense_Mutation		NP_003473	nd	
3	MLL3	3	-	g.chr7:152027684A>G	c.389_splice	p.S130_splice	Splice_Site		NP_733751	nd	somatic
20	MLL3	38	-	g.chr7:151874148delT	c.8390delA	p.K2797fs	Frame_Shift_Del		NP_733751	nd	
25	MLL3	3	-	g.chr7:152027725G>C	c.350C>G	p.A117G	Missense_Mutation	deleterious	NP_733751	nd	
5	MLLT6	11	+	g.chr17:36873796T>C	c.1763T>C	p.L588P	Missense_Mutation	deleterious	NP_005928	nd	somatic
24	MLLT6	11	+	g.chr17:36873756C>T	c.1723C>T	p.L575F	Missense_Mutation	deleterious	NP_005928	nd	
42	MLLT6	19	+	g.chr17:36880943T>C	c.2954T>C	p.M985T	Missense_Mutation	deleterious	NP_005928	nd	
1	MYC	2	+	g.chr8:128751173C>T	c.710C>T	p.S237F	Missense_Mutation	deleterious	NP_002458	Yes	somatic
1	MYC	2	+	g.chr8:128751223G>C	c.760G>C	p.V254L	Missense_Mutation	neutral	NP_002458	Yes	somatic
1	MYC	2	+	g.chr8:128751256A>T	c.793A>T	p.S265C	Missense_Mutation	deleterious	NP_002458	Yes	somatic
11	MYC	2	+	g.chr8:128750535C>G	c.27C>G	p.N9K	Missense_Mutation	neutral	NP_002458	Yes	somatic
11	MYC	2	+	g.chr8:128750626C>A	c.118C>A	p.L40M	Missense_Mutation	deleterious	NP_002458	Yes	somatic
12	MYC	2	+	g.chr8:128750512C>T	c.4C>T	p.P2S	Missense_Mutation	deleterious	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750526C>G	c.18C>G	p.S6R	Missense_Mutation	neutral	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750543A>C	c.35A>C	p.Y12S	Missense_Mutation	deleterious	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750577C>G	c.69C>G	p.F23L	Missense_Mutation	deleterious	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750653A>G	c.145A>G	p.I49V	Missense_Mutation	deleterious	NP_002458	nd	
12	MYC	2	+	g.chr8:128750683C>A	c.175C>A	p.P59T	Missense_Mutation	deleterious	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750713C>T	c.205C>T	p.L69F	Missense_Mutation	neutral	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750725T>C	c.217T>C	p.S73P	Missense_Mutation	neutral	NP_002458	nd	
12	MYC	2	+	g.chr8:128750810T>A	c.302T>A	p.M101K	Missense_Mutation	deleterious	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750886A>C	c.378A>C	p.K126N	Missense_Mutation	deleterious	NP_002458	Yes	
12	MYC	2	+	g.chr8:128750921T>C	c.413T>C	p.F138S	Missense_Mutation	deleterious	NP_002458	Yes	
12	MYC	2	+	g.chr8:128751015C>A	c.507C>A	p.S169R	Missense_Mutation	neutral	NP_002458	Yes	
12	MYC	3	+	g.chr8:128752698G>C	c.859G>C	p.A287P	Missense_Mutation	neutral	NP_002458	nd	
12	MYC	3	+	g.chr8:128753003C>G	c.1164C>G	p.S388R	Missense_Mutation	deleterious	NP_002458	nd	
28	MYC	2	+	g.chr8:128750525G>A	c.17G>A	p.S6N	Missense_Mutation	neutral	NP_002458	Yes	
29	MYC	2	+	g.chr8:128750495A>G	c.-11_splice		Splice_Site		NP_002458	Yes	putative somatic
29	MYC	2	+	g.chr8:128750680A>G	c.172A>G	p.T58A	Missense_Mutation	deleterious	NP_002458	Yes	putative somatic
29	MYC	2	+	g.chr8:128750921T>C	c.413T>C	p.F138S	Missense_Mutation	deleterious	NP_002458	Yes	putative somatic
40	MYC	2	+	g.chr8:128750677_128750679delCCC	c.169_171delCCC	p.P57del	In_Frame_Del		NP_002458	Yes	putative somatic
40	MYC	2	+	g.chr8:128750681_128750689del9bp	c.173_181del9bp	p.S8_61TPPL>M	In_Frame_Del		NP_002458	Yes	putative somatic
40	MYC	2	+	g.chr8:128750692delT	c.184delT	p.S62fs	Frame_Shift_Del		NP_002458	Yes	putative somatic
40	MYC	2	+	g.chr8:128750696C>A	c.188C>A	p.P63H	Missense_Mutation	deleterious	NP_002458	nd	putative somatic
40	MYC	2	+	g.chr8:128750698_128750702del5bp	c.190_194del5bp	p.S64fs	Frame_Shift_Del		NP_002458	Yes	putative somatic
40	MYC	2	+	g.chr8:128750852T>C	c.344T>C	p.F115S	Missense_Mutation	deleterious	NP_002458	Yes	putative somatic
40	MYC	3	+	g.chr8:128752656G>A	c.817G>A	p.D273N	Missense_Mutation	neutral	NP_002458	nd	putative somatic
49	MYC	2	+	g.chr8:128750752C>G	c.244C>G	p.L82V	Missense_Mutation	neutral	NP_002458	Yes	
49	MYC	2	+	g.chr8:128750844C>A	c.336C>A	p.N112K	Missense_Mutation	deleterious	NP_002458	Yes	
49	MYC	2	+	g.chr8:128751020G>C	c.557G>C	p.C186S	Missense_Mutation	deleterious	NP_002458	Yes	

49	MYC	3	+	g.chr8:128753003C>G	c.1164C>G	p.S388R	Missense_Mutation	deleterious	NP_002458	nd	
2	NAV3	15	+	g.chr12:78513682C>G	c.3706C>G	p.P1236A	Missense_Mutation	deleterious	NP_055718	nd	somatic
25	NAV3	33	+	g.chr12:78583791T>A	c.6083T>A	p.F2028Y	Missense_Mutation	deleterious	NP_055718	nd	
42	NAV3	13	+	g.chr12:78510678T>G	c.2763T>G	p.N921K	Missense_Mutation	deleterious	NP_055718	nd	
47	NAV3	1	+	g.chr12:78225308G>A	c.67G>A	p.A23T	Missense_Mutation	deleterious	NP_055718	nd	
47	NAV3	16	+	g.chr12:78515974C>T	c.4004C>T	p.S1335L	Missense_Mutation	deleterious	NP_055718	nd	
8	NEBL	12	-	g.chr10:21134220T>G	c.1194A>C	p.L398F	Missense_Mutation	deleterious	NP_006384	nd	putative somatic
35	NEBL	25	-	g.chr10:21098829delT	c.2519_splice	p.D840_splice	Splice_Site		NP_006384	nd	
36	NEBL	19	-	g.chr10:21112153T>C	c.1946A>G	p.Q649R	Missense_Mutation	deleterious	NP_006384	nd	
11	NOC4L	14	+	g.chr12:132636707A>G	c.1396A>G	p.S466G	Missense_Mutation	neutral	NP_076983	nd	somatic
26	NOC4L	4	+	g.chr12:132631856G>A	c.376G>A	p.V126M	Missense_Mutation	deleterious	NP_076983	nd	
42	NOC4L	12	+	g.chr12:132636159C>T	c.1204C>T	p.R402W	Missense_Mutation	deleterious	NP_076983	nd	
17	NOTCH1	34	-	g.chr9:139390617G>A	c.7574C>T	p.P2525L	Missense_Mutation	deleterious	NP_060087.3	nd	suspect SNP
20	NOTCH1	34	-	g.chr9:139391133C>T	c.7058G>A	p.S2353N	Missense_Mutation	neutral	NP_060087.3	Yes	
45	NOTCH1	20	-	g.chr9:139402689C>T	c.3320G>A	p.R1107Q	Missense_Mutation	neutral	NP_060087.3	nd	
8	NTRK1	15,16	+	g.chr1:156849951T>C	c.2205_splice	p.E735_splice	Splice_Site		NP_002520	Yes	putative somatic
10	NTRK1	14,15	+	g.chr1:156849143G>C	c.2035G>C	p.D679H	Missense_Mutation	deleterious	NP_002520	nd	
45	NTRK1	4,5	+	g.chr1:156836698G>A			Splice_Site		NP_002520	nd	
1	OR5R1	1	-	g.chr11:56185318G>T	c.391C>A	p.H131N	Missense_Mutation	deleterious	NP_001004744	nd	somatic
35	OR5R1	1	-	g.chr11:56184918G>T	c.791C>A	p.S264*	Nonsense_Mutation		NP_001004744	nd	
39	OR5R1	1	-	g.chr11:56184966A>T	c.743T>A	p.V248D	Missense_Mutation	deleterious	NP_001004744	nd	putative somatic
5	PARP14	12	+	g.chr3:122433232delA	c.3956delA	p.E1319fs	Frame_Shift_Del		NP_060024	nd	somatic
43	PARP14	6	+	g.chr3:122418534C>T	c.1133C>T	p.P378L	Missense_Mutation		NP_060024	nd	
45	PARP14	15	+	g.chr3:122439225G>T	c.4931G>T	p.R1644I	Missense_Mutation		NP_060024	nd	
10	PAX5	2	-	g.chr9:37020768_37020769insC	c.76_77insG	p.V26fs	Frame_Shift_Ins		NP_057953	nd	
27	PAX5	1	-	g.chr9:37033983C>G	c.46G>C	p.G16R	Missense_Mutation	neutral	NP_057953	nd	
38	PAX5	1	-	g.chr9:37033995T>C	c.34A>G	p.T12A	Missense_Mutation	neutral	NP_057953	nd	
2	PCDH7	1	+	g.chr4:30725197G>A	c.2153G>A	p.G718E	Missense_Mutation	deleterious	NP_002580	nd	somatic
10	PCDH7	1	+	g.chr4:30724664_30724666delGGA	c.1620_1622delGGA	p.E541del	In_Frame_Del		NP_002580	nd	
18	PCDH7	1	+	g.chr4:30725806A>C	c.2762A>C	p.D921A	Missense_Mutation	deleterious	NP_002580	nd	
41	PCDH7	1	+	g.chr4:30725919C>T	c.2875C>T	p.P959S	Missense_Mutation	deleterious	NP_002580	nd	
6	PCLO	5	-	g.chr7:82583626C>G	c.6643G>C	p.D2215H	Missense_Mutation	none	NP_149015	nd	
16	PCLO	5	-	g.chr7:82586109T>C	c.4160A>G	p.D1387G	Missense_Mutation	none	NP_149015	nd	
18	PCLO	7	-	g.chr7:825444851C>T	c.12451G>A	p.A4151T	Missense_Mutation	none	NP_149015	nd	
18	PCLO	5	-	g.chr7:82585508G>C	c.4761C>G	p.S1587R	Missense_Mutation	none	NP_149015	nd	
18	PCLO	2	-	g.chr7:82785699C>A	c.258G>T	p.E86D	Missense_Mutation	deleterious	NP_149015	nd	
20	PCLO	20	-	g.chr7:82451796A>C	c.14806T>G	p.*4936E	Nonstop_Mutation		NP_149015	nd	
20	PCLO	18	-	g.chr7:82455920G>T	c.14400C>A	p.N4800K	Missense_Mutation	none	NP_149015	nd	
20	PCLO	12	-	g.chr7:82475953delA	c.-13_splice		Splice_Site		NP_149015	nd	
20	PCLO	5	-	g.chr7:82582944delT	c.7325delA	p.K2442fs	Frame_Shift_Del		NP_149015	nd	
20	PCLO	3	-	g.chr7:82763844_82763845insT	c.3021_3022insA	p.K1007fs	Frame_Shift_Ins		NP_149015	nd	
20	PCLO	2	-	g.chr7:82785217_82785218insT	c.739_740insA	p.I247fs	Frame_Shift_Ins		NP_149015	nd	
22	PCLO	7	-	g.chr7:82544057G>T	c.13245C>A	p.D4415E	Missense_Mutation	none	NP_149015	nd	
37	PCLO	5	-	g.chr7:82585297C>A	c.4972G>T	p.V1658F	Missense_Mutation	none	NP_149015	nd	somatic
41	PCLO	5	-	g.chr7:82585021_82585022insT	c.5247_5248insA	p.K1749fs	Frame_Shift_Ins		NP_149015	nd	
44	PCLO	6	-	g.chr7:82580569C>T	c.9335G>A	p.S3112N	Missense_Mutation	none	NP_149015	nd	
23	PIKFYVE	4	+	g.chr2:209142398T>A	c.579T>A	p.N193K	Missense_Mutation	deleterious	NP_055855	nd	somatic
28	PIKFYVE	3	+	g.chr2:209141493A>C	c.380A>C	p.Q127P	Missense_Mutation	deleterious	NP_055855	nd	
35	PIKFYVE	18	+	g.chr2:209189689G>T	c.2386G>T	p.D796Y	Missense_Mutation	deleterious	NP_055855	nd	
2	PIM1	1	+	g.chr6:37138423G>C	c.72G>C	p.K24N	Missense_Mutation	neutral	NP_002639	Yes	somatic
5	PIM1	1	+	g.chr6:37138423G>C	c.72G>C	p.K24N	Missense_Mutation	neutral	NP_002639	Yes	somatic
11	PIM1	2	+	g.chr6:37138580C>A	c.114C>A	p.Y38*	Nonsense_Mutation		NP_002639	Yes	somatic
11	PIM1	4	+	g.chr6:37138956G>C	c.296G>C	p.G99A	Missense_Mutation	neutral	NP_002639	nd	somatic
11	PIM1	4	+	g.chr6:37138960C>G	c.300C>G	p.F100L	Missense_Mutation	neutral	NP_002639	nd	somatic
11	PIM1	4	+	g.chr6:37139210C>T	c.550C>T	p.L184F	Missense_Mutation	deleterious	NP_002639	nd	somatic
18	PIM1	3	+	g.chr6:37138805C>G	c.238C>G	p.L80V	Missense_Mutation	neutral	NP_002639	Yes	
20	PIM1	1	+	g.chr6:37138423G>C	c.72G>C	p.K24N	Missense_Mutation	neutral	NP_002639	Yes	
23	PIM1	4	+	g.chr6:37139072G>C	c.412G>C	p.A138P	Missense_Mutation	neutral	NP_002639	Yes	somatic
36	PIM1	1	+	g.chr6:37138355C>T	c.4C>T	p.L2F	Missense_Mutation	deleterious	NP_002639	nd	
36	PIM1	1	+	g.chr6:37138379G>A	c.28G>A	p.A10T	Missense_Mutation	neutral	NP_002639	nd	
36	PIM1	1	+	g.chr6:37138419C>T	c.68C>T	p.T23I	Missense_Mutation	neutral	NP_002639	Yes	
36	PIM1	2	+	g.chr6:37138545C>T			Splice_Site		NP_002639	Yes	
36	PIM1	2	+	g.chr6:37138549G>C	c.83G>C	p.G28A	Missense_Mutation	neutral	NP_002639	Yes	
36	PIM1	2	+	g.chr6:37138563C>T	c.97C>T	p.P33S	Missense_Mutation	neutral	NP_002639	nd	

36	PIM1	2	+	g.chr6:37138582A>G	c.116A>G	p.Q39R	Missense_Mutation	neutral	NP_002639	nd	
36	PIM1	2	+	g.chr6:37138642C>G	c.176C>G	p.S59C	Missense_Mutation	neutral	NP_002639	nd	
36	PIM1	3	+	g.chr6:37138804G>C	c.237G>C	p.E79D	Missense_Mutation	neutral	NP_002639	nd	
36	PIM1	4	+	g.chr6:37139238T>G	c.578T>G	p.L193R	Missense_Mutation	deleterious	NP_002639	Yes	
36	PIM1	5	+	g.chr6:37140920G>C	c.756G>C	p.Q252H	Missense_Mutation	neutral	NP_002639	nd	
36	PIM1	6	+	g.chr6:37141795G>A	c.870G>A	p.M290I	Missense_Mutation	deleterious	NP_002639	nd	
40	PIM1	1	+	g.chr6:37138419C>G	c.68C>G	p.T23S	Missense_Mutation	neutral	NP_002639	Yes	
47	PIM1	1	+	g.chr6:37138423G>C	c.72G>C	p.K24N	Missense_Mutation	neutral	NP_002639	Yes	
10	PLD1	17,18	-	g.chr3:171392349G>T	c.2170C>A	p.Q724K	Missense_Mutation	neutral	NP_002653	nd	
23	PLD1	25,26	-	g.chr3:171321022A>T	c.3071T>A	p.V1024E	Missense_Mutation	neutral	NP_002653	nd	somatic
44	PLD1	10	-	g.chr3:171426554G>T	c.1136C>A	p.T379K	Missense_Mutation	deleterious	NP_002653	nd	
2	POU2AF1	1	-	g.chr11:111249884T>G			Splice_Site		NP_006226	Yes	somatic
36	POU2AF1	1	-	g.chr11:111249902T>G	c.1A>C	p.M1L	Missense_Mutation	deleterious	NP_006226	nd	
41	RP1L1	3	-	g.chr8:10469822G>A	c.1786C>T	p.Q596*	Nonsense_Mutation		NP_849188	nd	
41	RP1L1	3	-	g.chr8:10470839T>G	c.769A>C	p.T257P	Missense_Mutation	deleterious	NP_849188	nd	
47	RP1L1	3	-	g.chr8:10469900C>T	c.1708G>A	p.E570K	Missense_Mutation	neutral	NP_849188	nd	
2	SDK1	39	+	g.chr7:4259845C>T	c.5644C>T	p.R1882W	Missense_Mutation	deleterious	NP_689957	Yes	somatic
20	SDK1	38	+	g.chr7:4249662delG	c.5407delG	p.G1803fs	Frame_Shift_Del		NP_689957	nd	
41	SDK1	37	+	g.chr7:4247765C>T	c.5249C>T	p.T1750M	Missense_Mutation	deleterious	NP_689957	nd	
11	SGK1	5,7	-	g.chr6:134494491A>G	c.338T>C	p.L113P	Missense_Mutation	deleterious	NP_005618	Yes	somatic
11	SGK1	4,6	-	g.chr6:134494640A>G	c.293T>C	p.F98S	Missense_Mutation	deleterious	NP_005618	Yes	somatic
11	SGK1	4,6	-	g.chr6:134494655G>C	c.278C>G	p.A93G	Missense_Mutation	deleterious	NP_005618	nd	somatic
11	SGK1	4,6	-	g.chr6:134494671A>C	c.262T>G	p.S88A	Missense_Mutation	neutral	NP_005618	nd	somatic
11	SGK1	2,4	-	g.chr6:134495653_134495676del24bp	c.125_148del24bp	p.IANNSYAC42del	In_Frame_Del		NP_005618	nd	somatic
11	SGK1	1	-	g.chr6:134495706C>G	c.95G>C	p.R32T	Missense_Mutation	deleterious	NP_005618	nd	somatic
20	SGK1	1	-	g.chr6:134495897T>A	c.49A>T	p.M17L	Missense_Mutation	deleterious	NP_005618	nd	
27	SGK1	2,4	-	g.chr6:134495668T>C	c.133A>G	p.N45D	Missense_Mutation	neutral	NP_005618	nd	
28	SGK1	2,4	-	g.chr6:134495654G>C	c.147C>G	p.C49W	Missense_Mutation	deleterious	NP_005618	nd	
40	SGK1	2,4	-	g.chr6:134495938A>G	c.8T>C	p.V3A	Missense_Mutation	neutral	NP_005618	nd	putative somatic
2	SLC1A7	7	-	g.chr1:53558439A>G	c.818T>C	p.V273A	Missense_Mutation	neutral	NP_006662	Yes	somatic
38	SLC1A7	8	-	g.chr1:53556482C>T			Splice_Site		NP_006662	nd	
47	SLC1A7	10	-	g.chr1:53554649T>C	c.1364A>G	p.D455G	Missense_Mutation	deleterious	NP_006662	nd	
11	SOCS1	1	-	g.chr16:11349022T>G	c.314A>C	p.D105A	Missense_Mutation	deleterious	NP_003736	Yes	somatic
24	SOCS1	1	-	g.chr16:11349320G>A	c.16C>T	p.Q6*	Nonsense_Mutation		NP_003736	Yes	
27	SOCS1	1	-	g.chr16:11348795G>A	c.541C>T	p.R181C	Missense_Mutation	deleterious	NP_003736	Yes	
27	SOCS1	1	-	g.chr16:11348802G>C	c.534C>G	p.C178W	Missense_Mutation	deleterious	NP_003736	Yes	
49	SOCS1	1	-	g.chr16:11349109G>A	c.227C>T	p.A76V	Missense_Mutation	neutral	NP_003736	Yes	
49	SOCS1	1	-	g.chr16:11349139C>G	c.197G>C	p.R66P	Missense_Mutation	neutral	NP_003736	Yes	
49	SOCS1	1	-	g.chr16:11349152C>G	c.184G>C	p.A62P	Missense_Mutation	neutral	NP_003736	Yes	
3	STAT6	11	-	g.chr12:57496661T>C	c.1256A>G	p.D419G	Missense_Mutation	neutral	NP_003144	Yes	somatic
5	STAT6	11	-	g.chr12:57496661T>C	c.1256A>G	p.D419G	Missense_Mutation	neutral	NP_003144	nd	somatic
6	STAT6	11	-	g.chr12:57496661T>C	c.1256A>G	p.D419G	Missense_Mutation	neutral	NP_003144	nd	
15	STAT6	11	-	g.chr12:57496661T>C	c.1256A>G	p.D419G	Missense_Mutation	neutral	NP_003144	Yes	
22	STAT6	11	-	g.chr12:57496662C>G	c.1255G>C	p.D419H	Missense_Mutation	deleterious	NP_003144	Yes	
24	STAT6	13	-	g.chr12:57493818T>C	c.1568A>G	p.D523G	Missense_Mutation	deleterious	NP_003144	Yes	
28	STAT6	8	-	g.chr12:57499079G>T	c.856C>A	p.Q286K	Missense_Mutation	deleterious	NP_003144	Yes	
38	STAT6	11	-	g.chr12:57496658T>C	c.1259A>G	p.N420S	Missense_Mutation	neutral	NP_003144	Yes	
43	STAT6	11	-	g.chr12:57496662C>A	c.1255G>T	p.D419Y	Missense_Mutation	neutral	NP_003144	Yes	
8	SYNE1	138,139	-	g.chr6:152461107T>G	c.25436A>C	p.E8479A	Missense_Mutation	neutral	NP_892006	nd	putative somatic
20	SYNE1	39,40	-	g.chr6:152737857T>C	c.5715A>G	p.I1905M	Missense_Mutation	neutral	NP_892006	nd	
40	SYNE1	39,40	-	g.chr6:152737940C>T	c.5632G>A	p.A1878T	Missense_Mutation	neutral	NP_892006	nd	putative somatic
42	SYNE1	85	-	g.chr6:152638007G>A	c.16687C>T	p.R5563W	Missense_Mutation	deleterious	NP_892006	nd	
43	SYNE1	118	-	g.chr6:152540205T>A	c.21977A>T	p.Q7326L	Missense_Mutation	neutral	NP_892006	nd	
43	SYNE1	58,59	-	g.chr6:152690653G>A	c.9604C>T	p.R3202C	Missense_Mutation	deleterious	NP_892006	nd	
47	SYNE1	101	-	g.chr6:152576741A>T	c.19245T>A	p.C6415*	Nonsense_Mutation		NP_892006	nd	
49	SYNE1	110	-	g.chr6:152555079T>A	c.20549A>T	p.D6850V	Missense_Mutation	deleterious	NP_892006	nd	
5	SYNJ2	17	+	g.chr6:158497754A>C	c.2389A>C	p.T797P	Missense_Mutation	neutral	NP_003889	nd	somatic
10	SYNJ2	23	+	g.chr6:158508017delT	c.3339delT	p.P1113fs	Frame_Shift_Del		NP_003889	nd	
16	SYNJ2	16	+	g.chr6:158495639G>A	c.2161G>A	p.V721I	Missense_Mutation	neutral	NP_003889	nd	
4	TAF1	21	+	g.chrX:70613247C>T	c.3145C>T	p.R1049C	Missense_Mutation	deleterious	NP_620278	nd	putative somatic
15	TAF1	20	+	g.chrX:70612764G>A	c.2968G>A	p.G990R	Missense_Mutation	deleterious	NP_620278	nd	
35	TAF1	13	+	g.chrX:70603852T>G	c.1985T>G	p.M662R	Missense_Mutation	deleterious	NP_620278	nd	
45	TAF1	13	+	g.chrX:70603930G>A	c.2063G>A	p.G688E	Missense_Mutation	neutral	NP_620278	nd	
45	TAF1	19	+	g.chrX:70612483T>G	c.2843T>G	p.L948R	Missense_Mutation	deleterious	NP_620278	nd	

1	TAGAP	4	-	g.chr6:159463180T>C	c.245A>G	p.K82R	Missense_Mutation	deleterious	NP_473455	Yes	somatic
10	TAGAP	4,9	-	g.chr6:159457833G>A	c.1222C>T	p.R408W	Missense_Mutation	deleterious	NP_473455	nd	
11	TAGAP	3	-	g.chr6:159464691A>G	c.30_splice	p.G11_splice	Splice_Site		NP_473455	nd	somatic
42	TAGAP	4,9	-	g.chr6:159457620T>C	c.1435A>G	p.S479G	Missense_Mutation	neutral	NP_473455	nd	
2	TBC1D4	1	-	g.chr13:76055882G>A	c.22C>T	p.Q8*	Nonsense_Mutation		NP_055647	Yes	somatic
28	TBC1D4	1	-	g.chr13:76055558C>G	c.346G>C	p.V116L	Missense_Mutation	deleterious	NP_055647	nd	
42	TBC1D4	10	-	g.chr13:75900392A>C	c.1974T>G	p.D658E	Missense_Mutation	deleterious	NP_055647	nd	
8	TBCD	35	+	g.chr17:80895232G>A	c.3277G>A	p.A1093T	Missense_Mutation	neutral	NP_005984	Yes	putative somatic
15	TBCD	14	+	g.chr17:80828105G>A	c.1324G>A	p.A442T	Missense_Mutation	neutral	NP_005984	nd	
49	TBCD	8	+	g.chr17:80755631A>T	c.772_splice	p.A258_splice	Splice_Site		NP_005984	nd	
18	TET2*	5	+	g.chr4:106180823C>T	c.3851C>T	p.S1284F	Missense_Mutation	deleterious	NP_001120680	Yes	
20	TET2*	1	+	g.chr4:106157836C>T	c.2737C>T	p.Q913*	Nonsense_Mutation		NP_001120680	Yes	
43	TET2*	4	+	g.chr4:106164753G>T	c.3621G>T	p.E1207D	Missense_Mutation	deleterious	NP_001120680	Yes	
47	TET2*	1	+	g.chr4:106155400C>A	c.301C>A	p.P101T	Missense_Mutation	deleterious	NP_001120680	Yes	
11	THBS1	2	+	g.chr15:39874489T>A	c.163T>A	p.S55T	Missense_Mutation	neutral	NP_003237	nd	somatic
41	THBS1	2	+	g.chr15:39874747G>A	c.421G>A	p.V141M	Missense_Mutation	deleterious	NP_003237	nd	
48	THBS1	10	+	g.chr15:39881282G>A	c.1768G>A	p.D590N	Missense_Mutation	neutral	NP_003237	nd	
20	TMEM30A	6	-	g.chr6:75965939delT	c.965delA	N322-	Frame_Shift_Del		NP_060717.1	nd	
35	TMEM30A	6	-	g.chr6:75975015G>A	c.385C>T	p.Q129*	Nonsense_Mutation		NP_060717.1	nd	
20	TNFAIP3*	6	+	g.chr6:138200388delG	c.1806delG	p.T602fs	Frame_Shift_Del		NP_006281	nd	
6	TNFRSF14	4	+	g.chr1:2491310delT	c.353delT	p.V118fs	Frame_Shift_Del		NP_003811	Yes	
17	TNFRSF14	3	+	g.chr1:2489823_2489824delGT	c.220_221delGT	p.V74fs	Frame_Shift_Del		NP_003811	Yes	putative somatic
22	TNFRSF14	3	+	g.chr1:2489836G>A	c.233G>A	p.C78Y	Missense_Mutation	deleterious	NP_003811	Yes	
27	TNFRSF14	4	+	g.chr1:2491369T>G	c.412T>G	p.C138G	Missense_Mutation	deleterious	NP_003811	Yes	
29	TNFRSF14	5	+	g.chr1:2492097C>A	c.495C>A	p.C165*	Nonsense_Mutation		NP_003811	Yes	putative somatic
35	TNFRSF14	5	+	g.chr1:2492089C>T	c.487C>T	p.Q163*	Nonsense_Mutation		NP_003811	Yes	
38	TNFRSF14	3	+	g.chr1:2489866A>T	c.263A>T	p.N88I	Missense_Mutation	deleterious	NP_003811	Yes	
41	TNFRSF14	5	+	g.chr1:2492134G>T	c.532G>T	p.E178*	Nonsense_Mutation		NP_003811	Yes	
42	TNFRSF14	4	+	g.chr1:2491292C>T	c.335C>T	p.S112F	Missense_Mutation	deleterious	NP_003811	Yes	
48	TNFRSF14	4	+	g.chr1:2491289G>T	c.332G>T	p.C111F	Missense_Mutation	deleterious	NP_003811	nd	
49	TNFRSF14	5	+	g.chr1:2492154G>A	c.551_splice	p.K184_splice	Splice_Site		NP_003811	nd	
5	TNIF	23	-	g.chr3:170811704G>A	c.2645C>T	p.T882M	Missense_Mutation	neutral	NP_055843	nd	somatic
25	TNIF	27	-	g.chr3:170800163T>C	c.3190A>G	p.T1064A	Missense_Mutation	deleterious	NP_055843	nd	
38	TNIF	24	-	g.chr3:170805171C>T	c.2846G>A	p.R949H	Missense_Mutation	deleterious	NP_055843	nd	
49	TNIF	17	-	g.chr3:170843805G>A	c.1909C>T	p.R637C	Missense_Mutation	deleterious	NP_055843	nd	
3	TP53	7	-	g.chr17:7577025T>A	c.913A>T	p.K305*	Nonsense_Mutation		NP_001119584	Yes	somatic
12	TP53	5	-	g.chr17:7578274T>C	c.575A>G	p.Q192R	Missense_Mutation	deleterious	NP_001119584	Yes	
23	TP53	8	-	g.chr17:7576851A>G	c.993_splice	p.Q331_splice	Splice_Site		NP_001119584	Yes	somatic
40	TP53	4	-	g.chr17:7578532A>T	c.398T>A	p.M133K	Missense_Mutation	neutral	NP_001119584	nd	putative somatic
41	TP53	9	-	g.chr17:7574003G>A	c.1024C>T	p.R342*	Nonsense_Mutation		NP_001119584	Yes	
42	TP53	5	-	g.chr17:7578226T>A	c.623A>T	p.D208V	Missense_Mutation	deleterious	NP_001119584	Yes	
15	TYRP1	4	+	g.chr9:12702424G>C	c.1067G>C	p.R356P	Missense_Mutation	deleterious	NP_000541	nd	
20	TYRP1	7	+	g.chr9:12708973A>C			Splice_Site		NP_000541	nd	
29	TYRP1	4	+	g.chr9:12702400A>T	c.1043A>T	p.Y348F	Missense_Mutation	neutral	NP_000541	nd	
11	XIRP2	8	+	g.chr2:168107995C>A	c.10093C>A	p.Q3365K	Missense_Mutation	neutral	NP_689594	Yes	somatic
18	XIRP2	8	+	g.chr2:168101537T>C	c.3635T>C	p.I1212T	Missense_Mutation	deleterious	NP_689594	nd	
27	XIRP2	8	+	g.chr2:168103831C>T	c.5929C>T	p.L1977F	Missense_Mutation	deleterious	NP_689594	Yes	
27	XIRP2	8	+	g.chr2:168106238G>C	c.8336G>C	p.R2779T	Missense_Mutation	deleterious	NP_689594	Yes	
35	XIRP2	2	+	g.chr2:167992435G>A	c.425G>A	p.R142Q	Missense_Mutation		NP_689594	Yes	
49	XIRP2	8	+	g.chr2:168101947G>C	c.4045G>C	p.G1349R	Missense_Mutation	deleterious	NP_689594	Yes	
29	ZNF462	2	+	g.chr9:109686573C>A	c.380C>A	p.T127N	Missense_Mutation	none	NP_067047	nd	
44	ZNF462	2	+	g.chr9:109691190A>C	c.4997A>C	p.K1666T	Missense_Mutation	deleterious	NP_067047	nd	
45	ZNF462	2	+	g.chr9:109687620C>T	c.1427C>T	p.P476L	Missense_Mutation	none	NP_067047	nd	
26	ZNF845	3	+	g.chr19:53856201G>A	c.2273G>A	p.R758H	Missense_Mutation	neutral	NP_612383	nd	
39	ZNF845	2	+	g.chr19:53848778A>T	c.35A>T	p.D12V	Missense_Mutation	deleterious	NP_612383	nd	putative somatic
48	ZNF845	3	+	g.chr19:53856306A>G	c.2378A>G	p.H793R	Missense_Mutation	deleterious	NP_612383	nd	

^ If multiple isoforms have been described for the gene, the exon number is provided for all isoforms

^^ Numbering according to the Human Genome hg19 assembly

* These genes are included even if mutated in <5% of cases because within focal MCRs of deletion

PPH2, PolyPhen-2 algorithm prediction; nd, not done

REFERENCES

- Al-Tourah, A. J., Gill, K. K., Chhanabhai, M., Hoskins, P. J., Klasa, R. J., Savage, K. J., Sehn, L. H., Shenkier, T. N., Gascoyne, R. D., and Connors, J. M. (2008). Population-based analysis of incidence and outcome of transformed non-Hodgkin's lymphoma. *J Clin Oncol* 26, 5165-5169.
- Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. G., Sabet, H., Tran, T., Yu, X., *et al.* (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403, 503-511.
- Bahler, D. W., and Levy, R. (1992). Clonal evolution of a follicular lymphoma: evidence for antigen selection. *Proc Natl Acad Sci U S A* 89, 6770-6774.
- Bastion, Y., Sebban, C., Berger, F., Felman, P., Salles, G., Dumontet, C., Bryon, P. A., and Coiffier, B. (1997). Incidence, predictive factors, and outcome of lymphoma transformation in follicular lymphoma patients. *J Clin Oncol* 15, 1587-1594.
- Bentley, G., Palutke, M., and Mohamed, A. N. (2005). Variant t(14;18) in malignant lymphoma: a report of seven cases. *Cancer Genet Cytogenet* 157, 12-17.
- Bergsagel, P. L., and Kuehl, W. M. (2013). Degree of focal immunoglobulin heavy chain locus deletion as a measure of B-cell tumor purity. *Leukemia* 27, 2067-2068.
- Beroukhi, R., Getz, G., Nghiemphu, L., Barretina, J., Hsueh, T., Linhart, D., Vivanco, I., Lee, J. C., Huang, J. H., Alexander, S., *et al.* (2007). Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc Natl Acad Sci U S A* 104, 20007-20012.
- Carlotti, E., Wrench, D., Matthews, J., Iqbal, S., Davies, A., Norton, A., Hart, J., Lai, R., Montoto, S., Gribben, J. G., *et al.* (2009). Transformation of follicular lymphoma to diffuse large B-cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone. *Blood* 113, 3553-3557.
- Challa-Malladi, M., Lieu, Y. K., Califano, O., Holmes, A. B., Bhagat, G., Murty, V. V., Dominguez-Sola, D., Pasqualucci, L., and Dalla-Favera, R. (2011). Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell* 20, 728-740.
- Cheung, K. J., Shah, S. P., Steidl, C., Johnson, N., Relander, T., Telenius, A., Lai, B., Murphy, K. P., Lam, W., Al-Tourah, A. J., *et al.* (2009). Genome-wide profiling of follicular lymphoma by array comparative genomic hybridization reveals prognostically significant DNA copy number imbalances. *Blood* 113, 137-148.

Christie, L., Kernohan, N., Levison, D., Sales, M., Cunningham, J., Gillespie, K., Batstone, P., Meiklejohn, D., and Goodlad, J. (2008). C-MYC translocation in t(14;18) positive follicular lymphoma at presentation: An adverse prognostic indicator? *Leuk Lymphoma* 49, 470-476.

Compagno, M., Lim, W. K., Grunn, A., Nandula, S. V., Brahmachary, M., Shen, Q., Bertoni, F., Ponzoni, M., Scandurra, M., Califano, A., *et al.* (2009). Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* 459, 717-721.

Davies, A. J., Rosenwald, A., Wright, G., Lee, A., Last, K. W., Weisenburger, D. D., Chan, W. C., Delabie, J., Braziel, R. M., Campo, E., *et al.* (2007). Transformation of follicular lymphoma to diffuse large B-cell lymphoma proceeds by distinct oncogenic mechanisms. *Br J Haematol* 136, 286-293.

De Jong, D., Voetdijk, B. M., Beverstock, G. C., van Ommen, G. J., Willemze, R., and Kluin, P. M. (1988). Activation of the c-myc oncogene in a precursor-B-cell blast crisis of follicular lymphoma, presenting as composite lymphoma. *N Engl J Med* 318, 1373-1378.

Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., Ritchey, J. K., Young, M. A., Lamprecht, T., McLellan, M. D., *et al.* (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481, 506-510.

Dreyling, M., Ghielmini, M., Marcus, R., Salles, G., Vitolo, U., and Group, E. G. W. (2011). Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 22 Suppl 6*, vi59-63.

Elenitoba-Johnson, K. S., Gascoyne, R. D., Lim, M. S., Chhanabai, M., Jaffe, E. S., and Raffeld, M. (1998). Homozygous deletions at chromosome 9p21 involving p16 and p15 are associated with histologic progression in follicle center lymphoma. *Blood* 91, 4677-4685.

Fabbri, G., Khiabani, H., Holmes, A. B., Wang, J., Messina, M., Mullighan, C. G., Pasqualucci, L., Rabadan, R., and Dalla-Favera, R. (2013). Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J Exp Med* 210, 2273-2288.

Fabbri, G., Rasi, S., Rossi, D., Trifonov, V., Khiabani, H., Ma, J., Grunn, A., Fangazio, M., Capello, D., Monti, S., *et al.* (2011). Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med* 208, 1389-1401.

Fisher, R. I., LeBlanc, M., Press, O. W., Maloney, D. G., Unger, J. M., and Miller, T. P. (2005). New treatment options have changed the survival of patients with follicular lymphoma. *J Clin Oncol* 23, 8447-8452.

Frattoni, V., Trifonov, V., Chan, J. M., Castano, A., Lia, M., Abate, F., Keir, S. T., Ji, A. X., Zoppoli, P., Niola, F., *et al.* (2013). The integrated landscape of driver genomic alterations in glioblastoma. *Nature genetics* 45, 1141-1149.

Green, M. R., Gentles, A. J., Nair, R. V., Irish, J. M., Kihira, S., Liu, C. L., Kela, I., Hopmans, E. S., Myklebust, J. H., Ji, H., *et al.* (2013). Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood* 121, 1604-1611.

Green, T. M., Young, K. H., Visco, C., Xu-Monette, Z. Y., Orazi, A., Go, R. S., Nielsen, O., Gadeberg, O. V., Mourits-Andersen, T., Frederiksen, M., *et al.* (2012). Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol* 30, 3460-3467.

Ho, L., Ronan, J. L., Wu, J., Staahl, B. T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A. I., Ranish, J., and Crabtree, G. R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A* 106, 5181-5186.

Holzelova, E., Vonarbourg, C., Stolzenberg, M. C., Arkwright, P. D., Selz, F., Prieur, A. M., Blanche, S., Bartunkova, J., Vilmer, E., Fischer, A., *et al.* (2004). Autoimmune lymphoproliferative syndrome with somatic Fas mutations. *N Engl J Med* 351, 1409-1418.

Hu, S., Xu-Monette, Z. Y., Tzankov, A., Green, T., Wu, L., Balasubramanyam, A., Liu, W. M., Visco, C., Li, Y., Miranda, R. N., *et al.* (2013). MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. *Blood* 121, 4021-4031; quiz 4250.

Korbel, J. O., and Campbell, P. J. (2013). Criteria for inference of chromothripsis in cancer genomes. *Cell* 152, 1226-1236.

Kuppers, R. (2005). Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* 5, 251-262.

Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S. L., Stewart, C., Mermel, C. H., Roberts, S. A., *et al.* (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214-218.

Lenz, G., Wright, G. W., Emre, N. C., Kohlhammer, H., Dave, S. S., Davis, R. E., Carty, S., Lam, L. T., Shaffer, A. L., Xiao, W., *et al.* (2008). Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U S A* 105, 13520-13525.

Li, H., Kaminski, M. S., Li, Y., Yildiz, M., Ouillet, P., Jones, S., Fox, H., Jacobi, K., Saiya-Cork, K., Bixby, D., *et al.* (2014). Mutations in linker histone genes HIST1H1 B, C, D and E, OCT2 (POU2F2), IRF8 and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood*.

Limpens, J., Stad, R., Vos, C., de Vlaam, C., de Jong, D., van Ommen, G. J., Schuur, E., and Kluin, P. M. (1995). Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals. *Blood* 85, 2528-2536.

Lin, M., Wei, L. J., Sellers, W. R., Lieberfarb, M., Wong, W. H., and Li, C. (2004). dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* 20, 1233-1240.

Lo Coco, F., Gaidano, G., Louie, D. C., Offit, K., Chaganti, R. S., and Dalla-Favera, R. (1993). p53 mutations are associated with histologic transformation of follicular lymphoma. *Blood* 82, 2289-2295.

Lohr, J. G., Stojanov, P., Lawrence, M. S., Auclair, D., Chapuy, B., Sougnez, C., Cruz-Gordillo, P., Knoechel, B., Asmann, Y. W., Slager, S. L., *et al.* (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A* 109, 3879-3884.

Lossos, I. S., Alizadeh, A. A., Diehn, M., Warnke, R., Thorstenson, Y., Oefner, P. J., Brown, P. O., Botstein, D., and Levy, R. (2002). Transformation of follicular lymphoma to diffuse large-cell lymphoma: alternative patterns with increased or decreased expression of c-myc and its regulated genes. *Proc Natl Acad Sci U S A* 99, 8886-8891.

Lossos, I. S., and Gascoyne, R. D. (2011). Transformation of follicular lymphoma. *Best Pract Res Clin Haematol* 24, 147-163.

Mandelbaum, J., Bhagat, G., Tang, H., Mo, T., Brahmachary, M., Shen, Q., Chadburn, A., Rajewsky, K., Tarakhovsky, A., Pasqualucci, L., and Dalla-Favera, R. (2010). BLIMP1 is a tumor suppressor gene frequently disrupted in activated B cell-like diffuse large B cell lymphoma. *Cancer Cell* 18, 568-579.

Montoto, S., Davies, A. J., Matthews, J., Calaminici, M., Norton, A. J., Amess, J., Vinnicombe, S., Waters, R., Rohatiner, A. Z., and Lister, T. A. (2007). Risk and clinical implications of transformation of follicular lymphoma to diffuse large B-cell lymphoma. *J Clin Oncol* 25, 2426-2433.

Morin, R. D., Johnson, N. A., Severson, T. M., Mungall, A. J., An, J., Goya, R., Paul, J. E., Boyle, M., Woolcock, B. W., Kuchenbauer, F., *et al.* (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature genetics* 42, 181-185.

Morin, R. D., Mendez-Lago, M., Mungall, A. J., Goya, R., Mungall, K. L., Corbett, R. D., Johnson, N. A., Severson, T. M., Chiu, R., Field, M., *et al.* (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* 476, 298-303.

Morin, R. D., Mungall, K., Pleasance, E., Mungall, A. J., Goya, R., Huff, R. D., Scott, D. W., Ding, J., Roth, A., Chiu, R., *et al.* (2013). Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* 122, 1256-1265.

- Mottok, A., Renne, C., Seifert, M., Oppermann, E., Bechstein, W., Hansmann, M. L., Kuppers, R., and Brauning, A. (2009). Inactivating SOCS1 mutations are caused by aberrant somatic hypermutation and restricted to a subset of B-cell lymphoma entities. *Blood* 114, 4503-4506.
- Mullighan, C. G., Goorha, S., Radtke, I., Miller, C. B., Coustan-Smith, E., Dalton, J. D., Girtman, K., Mathew, S., Ma, J., Pounds, S. B., *et al.* (2007). Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446, 758-764.
- Mullighan, C. G., Phillips, L. A., Su, X., Ma, J., Miller, C. B., Shurtleff, S. A., and Downing, J. R. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* 322, 1377-1380.
- Ngo, V. N., Young, R. M., Schmitz, R., Jhavar, S., Xiao, W., Lim, K. H., Kohlhammer, H., Xu, W., Yang, Y., Zhao, H., *et al.* (2011). Oncogenically active MYD88 mutations in human lymphoma. *Nature* 470, 115-119.
- O'Keefe, C., McDevitt, M. A., and Maciejewski, J. P. (2010). Copy neutral loss of heterozygosity: a novel chromosomal lesion in myeloid malignancies. *Blood* 115, 2731-2739.
- Okosun, J., Bodor, C., Wang, J., Araf, S., Yang, C. Y., Pan, C., Boller, S., Cittaro, D., Bozek, M., Iqbal, S., *et al.* (2014). Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nature genetics* 46, 176-181.
- Oricchio, E., Nanjangud, G., Wolfe, A. L., Schatz, J. H., Mavrikakis, K. J., Jiang, M., Liu, X., Bruno, J., Heguy, A., Olshen, A. B., *et al.* (2011). The Eph-receptor A7 is a soluble tumor suppressor for follicular lymphoma. *Cell* 147, 554-564.
- Pasqualucci, L., Bhagat, G., Jankovic, M., Compagno, M., Smith, P., Muramatsu, M., Honjo, T., Morse, H. C., 3rd, Nussenzweig, M. C., and Dalla-Favera, R. (2008). AID is required for germinal center-derived lymphomagenesis. *Nature genetics* 40, 108-112.
- Pasqualucci, L., Compagno, M., Houldsworth, J., Monti, S., Grunn, A., Nandula, S. V., Aster, J. C., Murty, V. V., Shipp, M. A., and Dalla-Favera, R. (2006). Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J Exp Med* 203, 311-317.
- Pasqualucci, L., Dominguez-Sola, D., Chiarenza, A., Fabbri, G., Grunn, A., Trifonov, V., Kasper, L. H., Lerach, S., Tang, H., Ma, J., *et al.* (2011a). Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* 471, 189-195.
- Pasqualucci, L., Migliazza, A., Fracchiolla, N., William, C., Neri, A., Baldini, L., Chaganti, R. S., Klein, U., Kuppers, R., Rajewsky, K., and Dalla-Favera, R. (1998). BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci U S A* 95, 11816-11821.

Pasqualucci, L., Neumeister, P., Goossens, T., Nanjangud, G., Chaganti, R. S., Kuppers, R., and Dalla-Favera, R. (2001). Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412, 341-346.

Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V. A., Grunn, A., Messina, M., Elliot, O., *et al.* (2011b). Analysis of the coding genome of diffuse large B-cell lymphoma. *Nature genetics* 43, 830-837.

Pinyol, M., Cobo, F., Bea, S., Jares, P., Nayach, I., Fernandez, P. L., Montserrat, E., Cardesa, A., and Campo, E. (1998). p16(INK4a) gene inactivation by deletions, mutations, and hypermethylation is associated with transformed and aggressive variants of non-Hodgkin's lymphomas. *Blood* 91, 2977-2984.

Pounds, S., Cheng, C., Mullighan, C., Raimondi, S. C., Shurtleff, S., and Downing, J. R. (2009). Reference alignment of SNP microarray signals for copy number analysis of tumors. *Bioinformatics* 25, 315-321.

Ritz, O., Guiter, C., Castellano, F., Dorsch, K., Melzner, J., Jais, J. P., Dubois, G., Gaulard, P., Moller, P., and Leroy, K. (2009). Recurrent mutations of the STAT6 DNA binding domain in primary mediastinal B-cell lymphoma. *Blood* 114, 1236-1242.

Rossi, D., Berra, E., Cerri, M., Deambrogi, C., Barbieri, C., Franceschetti, S., Lunghi, M., Conconi, A., Paulli, M., Matolcsy, A., *et al.* (2006). Aberrant somatic hypermutation in transformation of follicular lymphoma and chronic lymphocytic leukemia to diffuse large B-cell lymphoma. *Haematologica* 91, 1405-1409.

Rossi, D., Deaglio, S., Dominguez-Sola, D., Rasi, S., Vaisitti, T., Agostinelli, C., Spina, V., Bruscaggin, A., Monti, S., Cerri, M., *et al.* (2011). Alteration of BIRC3 and multiple other NF-kappaB pathway genes in splenic marginal zone lymphoma. *Blood* 118, 4930-4934.

Rossi, D., Trifonov, V., Fangazio, M., Bruscaggin, A., Rasi, S., Spina, V., Monti, S., Vaisitti, T., Arruga, F., Fama, R., *et al.* (2012). The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med* 209, 1537-1551.

Roulland, S., Faroudi, M., Mamessier, E., Sungalee, S., Salles, G., and Nadel, B. (2011). Early steps of follicular lymphoma pathogenesis. *Adv Immunol* 111, 1-46.

Saito, M., Novak, U., Piovan, E., Basso, K., Sumazin, P., Schneider, C., Crespo, M., Shen, Q., Bhagat, G., Califano, A., *et al.* (2009). BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* 106, 11294-11299.

Sander, C. A., Yano, T., Clark, H. M., Harris, C., Longo, D. L., Jaffe, E. S., and Raffeld, M. (1993). p53 mutation is associated with progression in follicular lymphomas. *Blood* 82, 1994-2004.

Sawas, A., Diefenbach, C., and O'Connor, O. A. (2011). New therapeutic targets and drugs in non-Hodgkin's lymphoma. *Curr Opin Hematol* 18, 280-287.

Schuetz, J. M., Johnson, N. A., Morin, R. D., Scott, D. W., Tan, K., Ben-Nierah, S., Boyle, M., Slack, G. W., Marra, M. A., Connors, J. M., *et al.* (2012). BCL2 mutations in diffuse large B-cell lymphoma. *Leukemia* 26, 1383-1390.

Sherr, C. J. (2004). Principles of tumor suppression. *Cell* 116, 235-246.

Siegel, R. M., Frederiksen, J. K., Zacharias, D. A., Chan, F. K., Johnson, M., Lynch, D., Tsien, R. Y., and Lenardo, M. J. (2000). Fas preassociation required for apoptosis signaling and dominant inhibition by pathogenic mutations. *Science* 288, 2354-2357.

Solal-Celigny, P., Roy, P., Colombat, P., White, J., Armitage, J. O., Arranz-Saez, R., Au, W. Y., Bellei, M., Brice, P., Caballero, D., *et al.* (2004). Follicular lymphoma international prognostic index. *Blood* 104, 1258-1265.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550.

Swerdlow, S. H., Campo, E., Harris, N. L., Jaffe, E. S., Pileri, S. A., Stein, H., Thiele, J., and Vardiman, J. W. (2008). WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. International Agency for Research on Cancer. (Geneva: World Health Organization).

Tchernitchko, D., Goossens, M., and Wajcman, H. (2004). In silico prediction of the deleterious effect of a mutation: proceed with caution in clinical genetics. *Clin Chem* 50, 1974-1978.

Trifonov, V., Pasqualucci, L., Dalla Favera, R., and Rabadan, R. (2013a). MutComFocal: an integrative approach to identifying recurrent and focal genomic alterations in tumor samples. *BMC Syst Biol* 7, 25.

Trifonov, V., Pasqualucci, L., Tiacci, E., Falini, B., and Rabadan, R. (2013b). SAVI: a statistical algorithm for variant frequency identification. *BMC Syst Biol* 7.

Trinh, D. L., Scott, D. W., Morin, R. D., Mendez-Lago, M., An, J., Jones, S. J., Mungall, A. J., Zhao, Y., Schein, J., Steidl, C., *et al.* (2013). Analysis of FOXO1 mutations in diffuse large B-cell lymphoma. *Blood* 121, 3666-3674.

Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E., and Croce, C. M. (1985). The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science* 229, 1390-1393.

Vakiani, E., Basso, K., Klein, U., Mansukhani, M. M., Narayan, G., Smith, P. M., Murty, V. V., Dalla-Favera, R., Pasqualucci, L., and Bhagat, G. (2008). Genetic and phenotypic analysis of B-

cell post-transplant lymphoproliferative disorders provides insights into disease biology. *Hematol Oncol* 26, 199-211.

van Eijk, M., Defrance, T., Hennino, A., and de Groot, C. (2001). Death-receptor contribution to the germinal-center reaction. *Trends Immunol* 22, 677-682.

Venkatraman, E. S., and Olshen, A. B. (2007). A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* 23, 657-663.

Walter, M. J., Shen, D., Ding, L., Shao, J., Koboldt, D. C., Chen, K., Larson, D. E., McLellan, M. D., Dooling, D., Abbott, R., *et al.* (2012). Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med* 366, 1090-1098.

Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., Wartman, L. D., Lamprecht, T. L., Liu, F., Xia, J., *et al.* (2012). The origin and evolution of mutations in acute myeloid leukemia. *Cell* 150, 264-278.

Ye, J., Ma, N., Madden, T. L., and Ostell, J. M. (2013). IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* 41, W34-40.

Zhang, J., Grubor, V., Love, C. L., Banerjee, A., Richards, K. L., Mieczkowski, P. A., Dunphy, C., Choi, W., Au, W. Y., Srivastava, G., *et al.* (2013). Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* 110, 1398-1403.

ACKNOWLEDGMENTS

I would like to express the sincere appreciation to my advisor and mentor Professor Gianluca Gaidano. He introduced me to the molecular biology and pushed me into pursuing my passion for research. He taught me that good research comes from dedication and sacrifices, but leads to great satisfaction and pride; I can now see that on a daily base.

I would like to express my deepest gratitude to Professor Riccardo Dalla-Favera for granting me the possibility to work in his prestigious laboratory at the Institute for Cancer Genetics at Columbia University and to Professor Laura Pasqualucci for being my mentor, for her patience, motivation, enthusiasm, and knowledge; qualities that, at least in part, I think I was able to acquire.

Besides my mentors, I would like to thank my PhD program board for giving me the opportunity to spend large part of my PhD abroad in a cutting-edge and competitive research infrastructure for the study of lymphoid malignancies biology.

I thank my fellow labmates at the Hematology Division of the University of Eastern Piedmont in Novara as well as at the Institute for Cancer Genetics in New York for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had together in the last years.

My sincere thanks also go to NovarAIL foundation for their financial support; the realization of my abroad research experience would have been much more difficult, if not impossible, without the precious support that they generously granted to me.

A variant of the *LRP4* gene affects the risk of chronic lymphocytic leukaemia transformation to Richter syndrome

Silvia Rasi,^{1*} Valeria Spina,^{1*} Alessio Bruscatto,¹ Tiziana Vaisitti,² Claudio Tripodo,³ Francesco Forconi,⁴ Lorenzo De Paoli,¹ Marco Fangazio,¹ Elisa Sozzi,⁴ Emanuele Cencini,⁴ Luca Laurenti,⁵ Roberto Marasca,⁶ Carlo Visco,⁷ Zijun Y. Xu-Monette,⁷ Valter Gattei,⁸ Ken H. Young,⁷ Fabio Malavasi,² Silvia Deaglio,² Gianluca Gaidano¹ and Davide Rossi¹

¹Division of Haematology, Department of Clinical and Experimental Medicine & IRCAD, Amedeo Avogadro University of Eastern Piedmont, Novara, ²Department of Genetics, Biology and Biochemistry, University of Turin, Turin, ³Department of Human Pathology, University of Palermo, Palermo, ⁴Division of Haematology, University of Siena and AOUS, Siena, ⁵Division of Haematology, Catholic University of the Sacred Heart, Rome, ⁶Division of Haematology, Department of Oncology and Haematology, University of Modena and Reggio Emilia, Modena, Italy, ⁷Department of Pathology and Laboratory Medicine, University of Wisconsin School of Medicine and Public Health, University of Wisconsin Paul P. Carbone Comprehensive Cancer Center, Madison, WI, USA, and ⁸Clinical and Experimental Onco-Haematology Unit, Centro di Riferimento Oncologico, I.R.C.C.S., Aviano, Italy

Received 29 July 2010; accepted for publication 1 October 2010

Correspondence: Davide Rossi, MD, PhD, Division of Haematology, Department of Clinical and Experimental Medicine & BRMA, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy. E-mail: rossidav@med.unipmn.it

*S.R. and V.S. contributed equally to this study.

Summary

Richter syndrome (RS) represents the transformation of chronic lymphocytic leukaemia (CLL) to aggressive lymphoma. Risk factors of CLL transformation to RS are only partly known. We explored the role of the host genetic background as a risk factor for RS occurrence. Forty-five single nucleotide polymorphisms (SNPs) known to be relevant for CLL prognosis were genotyped in a consecutive cohort of 331 CLL, of which 21 had transformed to RS. After correcting for multiple testing and adjusting for previously reported RS risk factors, the *LRP4* rs2306029 TT variant genotype was the sole SNP independently associated with a higher risk of RS transformation (Hazard Ratio: 4.17; $P = 0.001$; $q = 0.047$). The enrichment of *LRP4* TT genotype in RS was confirmed in an independent series ($n = 44$) used for validation purposes. The *LRP4* protein was expressed in CLL ($n = 66$). Bioinformatic analysis scored *LRP4* rs2306029 as a variant with possible deleterious and damaging variant of *LRP4*. *LRP4* genotyping may help the recognition of patients with increased risk of RS at the time of CLL diagnosis.

Keywords: Richter syndrome, chronic lymphocytic leukaemia, diffuse large B cell lymphoma, single nucleotide polymorphism, *LRP4*.

Richter syndrome (RS) represents the transformation of chronic lymphocytic leukaemia (CLL) to aggressive lymphoma, most commonly diffuse large B-cell lymphoma (DLBCL) (Tsimberidou & Keating, 2005; Omoti & Omoti, 2008; Rossi & Gaidano, 2009). Mechanisms and risk factors of

CLL transformation to RS are known only in part (Tsimberidou & Keating, 2005; Omoti & Omoti, 2008; Rossi & Gaidano, 2009). To date, predictors of RS transformation mainly consist of biological features related to the CLL clone, including usage of specific *IGHV* genes, expression of stereo-

typed B cell receptors, CD38 expression, cytogenetics, or telomere length (Aydin *et al*, 2008; Rossi *et al*, 2008a, 2009a,b; Deambrogi *et al*, 2009).

Increasing evidence points to the role of the genetic background of the host in the development and clinical course of CLL. Several genotypic variants have been reported to predispose CLL development (Sellick *et al*, 2007; Di Bernardo *et al*, 2008; Crowther-Swanepoel *et al*, 2010; Slager *et al*, 2010). Furthermore, single nucleotide polymorphisms (SNPs) have emerged as relevant predictors of CLL outcome, both in the context of clinical trials and in consecutive CLL series (Nüchel *et al*, 2007; Gryshchenko *et al*, 2008; Rossi *et al*, 2008b; Sellick *et al*, 2008; Allan *et al*, 2010; Rasi *et al*, 2010).

With respect to RS development, one SNP of the *CD38* gene (rs6449182) has been implicated in the acquisition of the RS clinical phenotype (Aydin *et al*, 2008). Indeed, RS appears to be enriched in cases harbouring the *CD38* rs6449182 minor G allele (Aydin *et al*, 2008). Overall, these data suggest that the interaction between the genetic background of the host and the biological characteristics of the neoplastic clone may affect the clinical course of CLL as well as the risk of RS transformation.

This study aimed to further explore the role of the genetic background of the host as a risk factor for CLL transformation to RS. We report that a nonsynonymous SNP of the *LRP4* gene, a member of the low-density lipoprotein receptor gene family involved in the modulation of Wnt signalling, is a novel risk factor of RS transformation.

Materials and methods

Patients

This study was based on a consecutive series of 331 CLL patients, with typical sIg low/CD5⁺/CD23⁺ phenotype and Matutes score >3 (Matutes *et al*, 1994; Hallek *et al*, 2008). Twenty-one (6.3%) of these patients had transformed to clonally related RS, as documented by the comparison of *IGH-V-D-J* rearrangements and HCDR3 sequences in CLL/RS paired samples. Among CLL undergoing RS transformation, the median time from CLL diagnosis to RS was 29.1 months. For validation purposes, an independent cohort of 44 clonally related RS, as well as 7 cases of clonally unrelated RS, were also collected. RS diagnosis was based on histology of lymph node or extranodal tissue excisional biopsies. All RS cases were classified as DLBCL according to the World Health Organization Classification of Tumours of the Haematopoietic and Lymphoid Tissues (Swerdlow *et al*, 2008). Patients provided informed consent in accordance with local IRB requirements and Declaration of Helsinki.

SNP genotyping

An educated guess approach was utilized for the selection of genes potentially relevant to RS transformation. SNPs were

selected according to the following criteria: (i) reported association with CLL prognosis (Gryshchenko *et al*, 2008; Sellick *et al*, 2008; Allan *et al*, 2010; Rasi *et al*, 2010), (ii) minor allele frequency >5% in Caucasians. Accordingly, 45 SNPs from 45 genes were analysed (Table SI). SNP nomenclature and nucleotide position were defined according to GenBank (<http://www.ncbi.nlm.nih.gov>. Accessed on 14 September 2009). DNA samples for genotyping were extracted from peripheral blood granulocytes collected at CLL diagnosis. Genotyping was performed by SNP minisequencing (ABI Prism SNaPshot Multiplex kit; Applied Biosystems, Foster City, CA, USA), after validation of this approach by DNA direct sequencing of each SNP in a pilot panel of cases ($n = 60$; data not shown). Quality control of genotyping was performed by replicate sample analysis.

Deviation of SNP genotype distribution from Hardy–Weinberg equilibrium (HWE) was tested by chi-square test or Fisher's exact test if appropriate. Seven SNPs were not in HWE (Table SI), and therefore were excluded from subsequent analysis. Linkage disequilibrium (LD) among SNPs was calculated as r^2 values with the use of HAPLOVIEW (<http://www.broad.mit.edu/haploview>. Accessed on 14 September 2009). None of the SNP was in strong LD ($r^2 \geq 0.8$).

Analysis of *LRP4* expression

CLL cells were purified from peripheral blood mononuclear cells (PBMC) preparations by negative selection using anti-CD3, anti-CD16 and anti-CD14 monoclonal antibodies (mAbs; produced and purified in the Turin Laboratory) and magnetic beads (Invitrogen, Milan, Italy) separation. Purified CLL cells were lysed in 1% Nonidet P-40 lysis buffer and whole cell lysates (25 µg/lane) were resolved by 10% sodium dodecyl sulphate polyacrylamide gel electrophoresis and electrophoretically transferred to nitrocellulose membranes (Millipore, Billerica, MA, USA), using a semi-dry apparatus (Hoefer SemyPhor; Pharmacia Biotech, San Francisco, CA, USA). Blots were blocked and incubated with a rabbit polyclonal antibody against LRP4 (Abgent, San Diego, CA, USA), followed by a goat anti-rabbit horseradish peroxidase-conjugated secondary reagent. An anti-actin mAb (Sigma, Milan, Italy) was used to check for homogeneous protein loading in each lane. Blots were developed with an enhanced chemiluminescence (ECL) detection system (Perkin Elmer, Milan, Italy). Densitometric analysis was performed using the public domain NIH IMAGEJ program (version 1.36, available at rsb.info.nih.gov/nihimage/).

For immunophenotypic studies, PBMC preparations were stained using a phycoerythrin (PE)- (for FACS studies) or Alexa568-labeled (for confocal studies) anti-CD19 mAb (produced in the University of Turin Laboratory), fixed and permeabilized using 0.1% saponin. Cells were then saturated using pre-immune goat serum, before incubation with the same rabbit polyclonal antibody against LRP4 (Abgent) used for Western blot. Detection was done with an Alexa-488 conjugated goat anti-rabbit secondary antibody

(Invitrogen). Cells were analysed using a FACSort (BD Biosciences, Milan, Italy), acquiring at least 5000 events per sample using the Cell QUEST software (Becton Dickinson, San Jose, California, USA). For confocal microscopy studies, labelled cells were left to adhere to polylysine-coated slides. Analysis was performed using an Olympus IX71 FV300 confocal microscope and the FLUOVIEW software (Olympus, Milan, Italy).

For immunohistochemical analysis of LRP4 expression, 30 CLL bone marrow biopsies were collected from the archives of the Department of Human Pathology, University of Palermo, Italy. All bone marrow samples were collected at diagnosis, before any treatment was started. Four μm -thick sections were cut from paraffin blocks and put onto slides. Following deparaffinization, sections underwent antigen retrieval by microwave oven treatment in citrate buffer pH 6.0 (Dako, Glostrup, Denmark). Subsequently, endogenous peroxidase activity was quenched by incubating slides with H_2O_2 for 10 min at room temperature. Sections were then incubated with a specific polyclonal antibody anti-human-LRP4 (Sigma Life Sciences, Sigma) overnight at 4°C. Binding of the primary antibody was revealed using a specific biotinylated secondary antibody, the streptavidin-biotin-peroxidase complex system (Novolink Max Polymer Detection System; Novocastra, Newcastle-upon-Tyne, UK), and 3-3'-diaminobenzidine as chromogenic substrate (brown signal). Slides were counterstained with haematoxylin and evaluated under a Leica DM2000 optical microscope (Leica Microsystems Nussloch GmbH, Nussloch, Germany) by an expert hematopathologist. Microphotographs were collected using a Leica DFC320 digital camera (Leica).

In silico analysis

The consequence of LRP4 rs2306029 on protein function was analysed *in silico* according to the PolyPhen (<http://genetics.bwh.harvard.edu/pph>. Accessed on 14 September 2009) and SIFT (<http://blocks.fhcrc.org/sift/SIFT.html>. Accessed on 14 September 2009) algorithms. LRP4 protein sequence information was retrieved from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>. Accessed on 12 April 2010) and the UniProtKB/Swiss-Prot (<http://www.uniprot.org/>. Accessed on 12 April 2010) databases. Alignment of amino acid sequences homologous to human LRP4 was computed using the MEGA4 (<http://www.megasoftware.net>) and CLUSTALW software (<http://www.ebi.ac.uk/Tools/msa/clustalw2>). Weblogo presentation of multiple sequence alignment around LRP4 Ser1554 was generated using the <http://weblogo.berkeley.edu/> software (Accessed on 12 April 2010).

Statistical analysis

The primary endpoint of the study was the cumulative probability of transformation measured from CLL diagnosis to

RS transformation, death or last follow-up. The cumulative probability of transformation was calculated by Kaplan–Meier method (Kaplan & Meier, 1958). The association between SNPs and risk of RS transformation was actuarially assessed by univariate Cox analysis considering the minor allele as acting either in a recessive or in a dominant fashion. False discovery rate (FDR) was used to control for multiple statistical testing (Benjamini & Hochberg, 1995). Accordingly, along with the *P* value, also a *q* value was provided for each association analysis. The *q*-value indicates the probability that a positive association identified by a *P*-value <0.05 might be a false positive association due to chance. Cox proportional hazard regression was used to build a multivariate model for RS transformation (Cox, 1972). Covariates included in multivariate analysis were LRP4 rs2306029 genotype (CT/CC *versus* TT), HCDR3 status (non stereotyped *versus* stereotyped), immunoglobulin variable heavy chain (IGHV) gene usage (non IGHV4-39 *versus* IGHV4-39), and CD38 rs6449182 genotype (CC *versus* CG/GG). In order to account for monotone likelihood, the Firth's bias correction method was applied to Cox regression (Heinze & Schemper, 2001). Unadjusted and adjusted hazard ratios for LRP4 rs2306029 genotype were compared to estimate the proportion of the difference in the risk of RS transformation that was attributable to covariates. Because the hazard ratio derived from Cox analysis measures magnitude of risk rather than a model ability to accurately classify subjects, Harrell *c*-statistics was used to further evaluate the discriminatory value of the prognostic models in terms of RS development (*c*-statistics = 1 indicates perfect discrimination, while *C* statistics = 0.5 equivalent to chance) (Harrell *et al*, 1996). To test whether the difference in the discriminatory value of the prognostic models was significant, we used the *rcorr.p.cens* function in the Design library of R. Recursive-partitioning analysis for censored survival data was performed to identify the factors most influential for RS transformation, and to permit the classification of patients into risk categories. Categorical variables were compared by chi square test and Fisher's exact test when appropriate. Continuous variables were compared by Mann–Whitney test. All statistical tests were two-sided. Statistical significance was defined as *P*-value <0.05. The analysis was performed with the Statistical Package for the Social Sciences (SPSS) software v.18.0 (SPSS, Chicago, IL, USA) and with R statistical package 2.8.1 (<http://www.r-project.org>. Accessed on 15 May 2010).

Results

The LRP4 rs2306029 genotype is an independent predictor of RS transformation

In order to characterize the impact of the genetic background of the host on the risk of RS transformation, 45 SNPs known to associate with CLL progression and/or prognosis were genotyped in the study cohort (*n* = 331) (Table SI). The clinical characteristics at diagnosis of the CLL cohort are shown in Table I. After a median follow-up from diagnosis of

Table I. Clinical features of the CLL cohort.

Characteristics at diagnosis	
Biological features	
IGHV homology $\geq 98\%$	125/326 (38.3%)
Stereotyped HCDR3	77/326 (23.6%)
IGHV4-39 usage	9/326 (2.8%)
FISH stratification	
Normal FISH	89/320 (27.8%)
del13q14	120/320 (37.5%)
+12	55/320 (17.2%)
del11q22-q23	22/320 (6.9%)
del17p13	34/320 (10.6%)
TP53 mutations	32/306 (10.5%)
CD38 $\geq 30\%$	104/326 (31.9%)
Clinical features	
Age (years)	68 (60–74)
Male	175/331 (52.9%)
Binet stage	
A	256/331 (77.3%)
B	41/331 (12.4%)
C	34/331 (10.3%)
Nodal areas ≥ 3	52/331 (15.7%)
Largest node ≥ 3 cm	34/331 (10.3%)
Absolute lymphocyte count ($\times 10^9/l$)	11.5 (6.4–20.0)
Hb (g/l)	136 (128–148)
Platelets ($\times 10^9/l$)	193 (149–235)
LDH (u/l)	360 (305–428)

25th–75th percentiles are reported in parentheses for continuous variables; IGHV, immunoglobulin heavy chain variable gene; HCDR3, immunoglobulin heavy chain complementarity determining region 3; FISH, fluorescence *in situ* hybridization.

75 months, 21/331 (6.3%) CLL transformed to RS, accounting for a 5-year cumulative probability of transformation of 7.1% [95% confidence interval (CI) 3.8–10.4%].

By Cox univariate analysis corrected for multiple comparisons, the sole SNP that displayed a significant association with an increased risk of RS transformation was *LRP4* rs2306029 ($P = 0.001$; $q = 0.047$). CLL who carried the *LRP4* rs2306029 TT variant genotype displayed a higher risk of transformation

[Hazard Ratio (HR): 4.17; 95% CI: 1.75–9.92; transformed/N: 12/79; 5-year risk: 13.5%; 95% CI: 5.3–21.7%] compared to patients carrying the CT/CC genotypes that contained the wild type allele (transformed/N: 9/252; 5-year risk: 4.4%; 95% CI: 1.5–6.9%) (Table II; Fig 1). Although the *CD38* rs6449182 SNP was also significantly associated with RS transformation by Cox univariate analysis ($P = 0.005$), the probability that this association was instead a false positive association due to chance was $>10\%$ ($q = 0.117$).

The *LRP4* rs2306029 genotype appeared to specifically predict RS, whereas it had no impact on other CLL clinical endpoints, including overall survival (HR: 1.24; 95% CI 0.78–1.97; $P = 0.350$) and time to first treatment (HR: 1.00; 95% CI 0.70–1.43; $P = 0.998$).

Other variables at CLL diagnosis associated with an increased risk of RS were CD38 expression $\geq 30\%$ ($P = 0.011$), unfavourable pattern of genetic lesions ($P = 0.004$), IGHV homology $\geq 98\%$ ($P = 0.002$), stereotyped HCDR3 ($P < 0.001$), and usage of *IGHV4-39* ($P < 0.001$) (Table II). By statistical analysis, there was no evidence of an association between the above risk factors and the *LRP4* rs2306029 TT variant genotype ($P > 0.050$ in all instances), indicating that *LRP4* rs2306029 TT is not a surrogate of other RS risk factors (Table III).

By multivariate analysis, the *LRP4* rs2306029 TT genotype was an independent predictor of RS transformation (HR: 3.21; 95% CI: 1.22–8.45; $P = 0.018$), along with stereotyped HCDR3 (HR: 4.45; 95% CI: 1.67–11.86; $P = 0.003$) and *IGHV4-39* usage (HR: 6.13; 95% CI: 1.95–19.19; $P = 0.002$) (Table II). The comparison between unadjusted and adjusted hazard ratio for the *LRP4* rs2306029 TT genotype (Table II) revealed that stereotyped HCDR3 and *IGHV4-39* usage could explain only 24% of the overall risk of RS transformation determined by the *LRP4* rs2306029 TT genotype.

Model of RS prediction based on *LRP4* rs2306029 genotype, HCDR3 status and *IGHV4-39* usage

Recursive-partitioning analysis was used to identify prognostic factors with the most influential predictive significance for

Table II. Univariate and multivariate Cox analysis for cumulative probability of RS transformation.

	Univariate analysis				Multivariate analysis			
	HR	LCI	UCI	P	HR	LCI	UCI	P
<i>LRP4</i> rs2306029 TT	4.17	1.75	9.92	0.001	3.21	1.22	8.45	0.018
<i>CD38</i> rs6449182 CG/GG	3.63	1.46	9.00	0.005				
<i>CD38</i> $\geq 30\%$	3.26	1.31	8.13	0.011				
<i>TP53</i> deletion/mutation	3.81	1.53	9.51	0.004				
IGHV homology $\geq 98\%$	4.61	1.77	12.00	0.002				
Stereotyped HCDR3	7.10	2.87	17.62	<0.001	4.45	1.67	11.86	0.003
<i>IGHV4-39</i> usage	12.79	4.66	35.12	<0.0001	6.13	1.95	19.19	0.002

HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval; IGHV, immunoglobulin heavy chain variable gene; HCDR3, immunoglobulin heavy chain complementarity determining region 3.

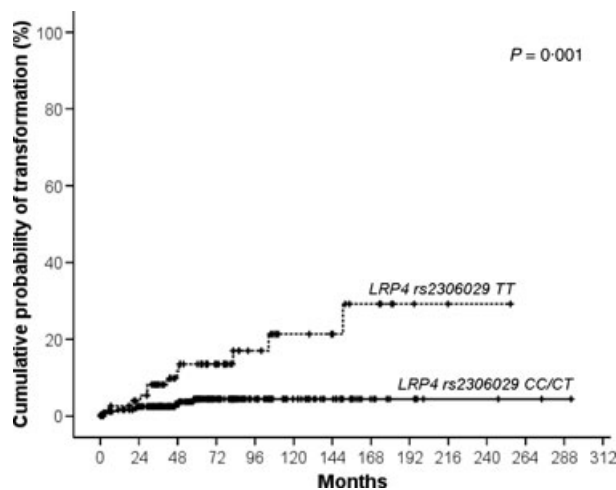


Fig 1. Cumulative probability of transformation to Richter syndrome. Kaplan-Meier plots representing the cumulative probability of transformation to Richter syndrome according to *LRP4* rs2306029 genotype.

cumulative risk of RS transformation and to classify patients into risk categories. A model for RS prediction based on HCDR3 status and *IGHV4-39* usage (Rossi *et al*, 2009a) was applied to the RS cohort of this study. According to this model, RS were stratified into three risk categories (Fig 2A): (i)

CLL utilizing both *IGHV4-39* and stereotyped HCDR3, characterized by a high risk of RS transformation (HR: 43.96; 95% CI: 12.57–153.65; Events/N: 4/5; 5-year risk: 70.0%; $P < 0.001$), (ii) CLL utilizing stereotyped HCDR3 but without *IGHV4-39*, characterized by an intermediate risk of RS transformation (HR: 5.40; 95% CI: 2.05–14.19; Events/N: 10/72; 5-year risk: 13.7%; $P = 0.001$) and (iii) CLL without stereotyped HCDR3, independent of *IGHV4-39* usage, characterized by a low risk of RS transformation (Events/N: 7/429; 5-year risk: 3.5%; reference category).

The introduction of the *LRP4* rs2306029 genotype in this model improved stratification by adding prognostic information within the intermediate risk group of patients characterized by stereotyped HCDR3 but without *IGHV4-39*. In particular, the *LRP4* rs2306029 genotype split intermediate risk patients into two subgroups (Fig 2B): (i) CLL utilizing stereotyped HCDR3 without *IGHV4-39* and harbouring the *LRP4* rs2306029 TT genotype, characterized by a high risk of RS transformation (HR: 14.13; 95% CI: 4.72–42.29; Events/N: 6/16; 5-year risk: 30.4%; $P < 0.001$), (ii) CLL utilizing stereotyped HCDR3 without *IGHV4-39* and harbouring the *LRP4* rs2306029 CT/CC genotype, characterized by a low risk of RS transformation (HR: 2.78; 95% CI: 0.81–9.53; Events/N: 4/56; 5-year risk: 8.5%; $P = 0.103$).

Table III. Biological and clinical characteristics of the CLL series according to *LRP4* rs2306029.

	LRP4 rs2306029						
	TT			CT/CC			
Characteristics	N/total	%	95% CI (%)	N/total	%	95% CI (%)	P
Biological variables							
IGHV homology ≥98	23/77	29.9	19.6–40.0	102/249	41.0	34.8–47.0	0.080
Stereotyped HCDR3	19/75	25.3	15.4–35.1	58/249	23.3	18.0–28.5	0.716
IGHV4-39 usage	4/77	5.2	0.2–10.1	5/249	2.0	0.2–3.7	0.223
FISH stratification							
Normal	21/77	27.3	17.3–37.2	68/243	28.0	22.3–33.6	0.112
del13q14	23/77	29.9	19.6–40.0	97/243	39.9	33.7–46.0	
+12	18/77	23.4	13.9–32.8	37/243	15.2	10.7–19.7	
del11q22-q23	3/77	3.9	0–8.2	19/243	7.8	4.4–11.2	
del17p13	12/77	15.6	7.4–23.6	22/243	9.1	5.4–12.6	
TP53 mutation	7/76	9.2	2.7–15.7	25/230	10.9	6.8–14.8	0.682
CD38 ≥30	20/78	25.6	15.9–35.3	84/248	33.9	27.9–39.7	0.174
Clinical variables							
Age >65 years	44/79	55.7	44.7–66.6	149/252	59.1	53.0–65.2	0.589
Male	45/79	57.0	46.0–67.8	130/252	51.6	45.4–57.7	0.404
Binet stage B–C	18/79	22.8	13.5–32.0	57/252	22.6	17.4–27.7	0.976
Nodal areas ≥3	11/79	13.9	6.2–21.5	41/252	16.3	11.7–20.8	0.617
Lymph node size ≥3 cm	9/79	11.4	4.3–18.4	15/252	9.9	3.0–8.8	0.707
Absolute lymphocyte count >10 × 10 ⁹ /l	23/79	29.1	19.9–39.1	61/252	24.2	18.9–29.5	0.382
Hb <100 g/l	3/79	3.8	0–8.2	14/252	5.6	2.7–8.3	0.771
Platelets (×10 ⁹ /l)	8/79	10.1	3.4–16.7	20/252	7.9	4.6–11.2	0.542
LDH > ULN	12/79	15.2	7.2–23.1	25/252	9.9	6.2–13.6	0.195

IGHV, immunoglobulin heavy chain variable gene; HCDR3, immunoglobulin heavy chain complementarity determining region 3; LDH, lactate dehydrogenase; ULN, upper limit of normal; FISH, fluorescence *in situ* hybridization.

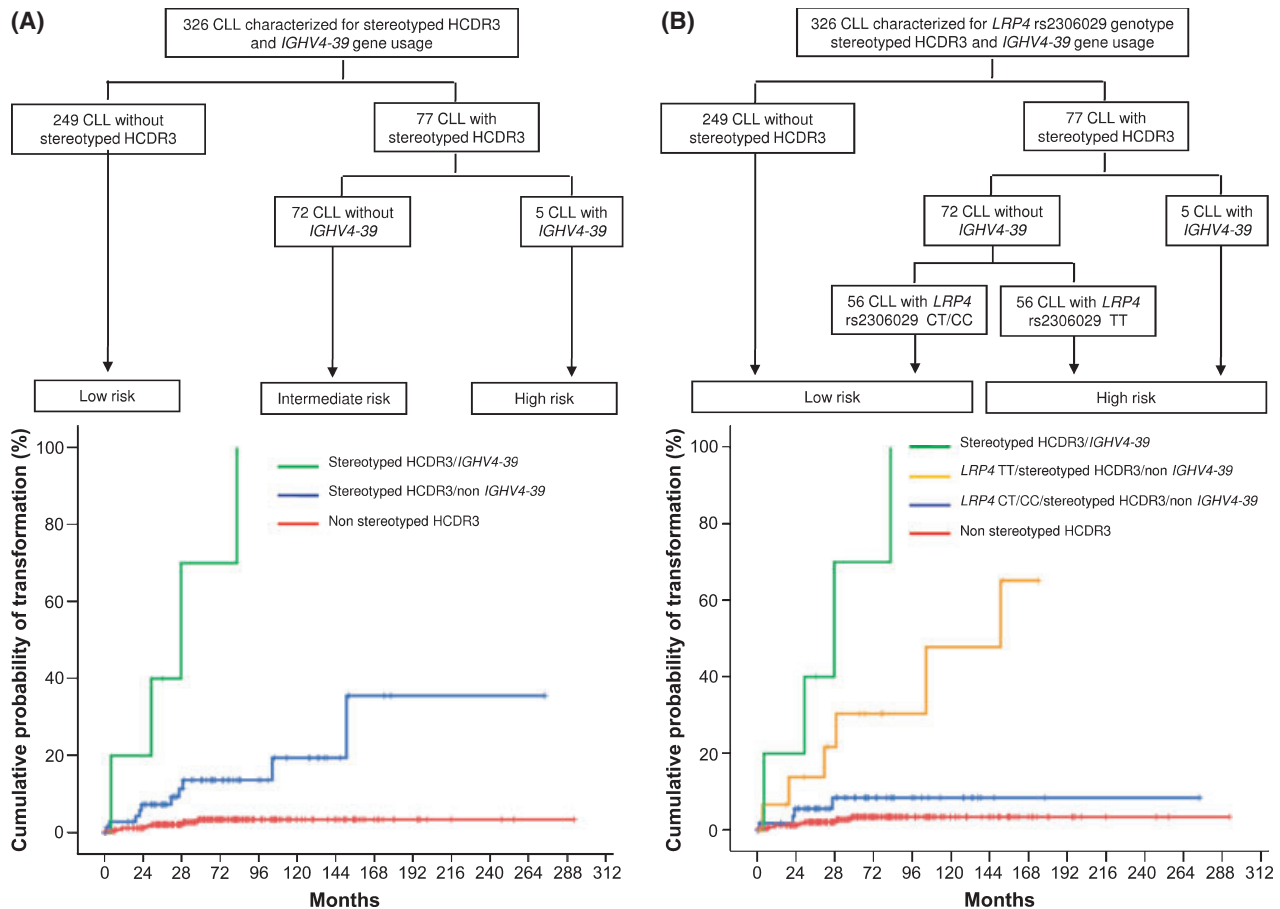


Fig 2. Classification of CLL into risk-of-RS categories and Kaplan-Meier estimates of cumulative probability of transformation to Richter syndrome according to those categories. Recursive-partitioning analysis was used to identify prognostic factors with the most influential predictive significance in a proportional-hazards model for cumulative risk of RS transformation and to classify patients into risk categories. The prognostic factors in the analysis were: (i) HCDR3 status and *IGHV4-39* usage (panel A) and (ii) HCDR3 status, *IGHV4-39* usage and *LRP4* rs2306029 genotype (panel B). The model for RS prediction including HCDR3 status and *IGHV4-39* usage segregated three risk categories (panel A): (i) CLL utilizing both *IGHV4-39* and stereotyped HCDR3 represented the disease category with the highest risk of RS transformation, (ii) CLL utilizing stereotyped HCDR3 but without *IGHV4-39* showed an intermediate risk of RS transformation and (iii) CLL without stereotyped HCDR3, independent of *IGHV4-39* usage, showed the lowest risk of RS transformation. The introduction of *LRP4* rs2306029 in this model improved stratification by adding prognostic information within the intermediate risk group of patients harbouring stereotyped HCDR3 without *IGHV4-39*. In this context, the *LRP4* rs2306029 genotype splits patients into two subgroup (panel B): (i) CLL utilizing stereotyped HCDR3 without *IGHV4-39* and harbouring the *LRP4* rs2306029 TT genotype and (ii) CLL utilizing stereotyped HCDR3 without *IGHV4-39* and harbouring the *LRP4* rs2306029 CT/CC genotype, characterized by a low risk of RS transformation.

We applied *c*-statistics to define the probability of correctly discriminating *a priori* between patients developing RS versus patients not developing RS. Compared to the previous model including HCDR3 status and *IGHV4-39* usage without the *LRP4* rs2306029 genotype, the current model including HCDR3 status, *IGHV4-39* usage along with the *LRP4* rs2306029 genotype significantly improved the probability of correct discrimination *a priori* from 72% (*c*-statistic = 0.72; SD = 0.11) to 85% (*c*-statistic = 0.85; SD = 0.10) (*P* = 0.001).

Validation of *LRP4* rs2306029 TT genotype enrichment in RS

In order to confirm the enrichment of TT genotype in RS, *LRP4* rs2306029 was genotyped in an independent RS series of

clonally related RS (*n* = 44) used for validation purposes and matched to the original RS cohort for age, sex and ethnic origin. The prevalence of the *LRP4* rs2306029 TT genotype did not differ significantly between the original RS series (12/21, 57.1%) and the validation RS series (19/44, 43.2%; *P* = 0.426), thus confirming the enrichment of *LRP4* rs2306029 TT genotype in the context of clonally related RS.

In contrast, the *LRP4* rs2306029 TT genotype was rare (1/7, 14.3%) among cases of clonally unrelated RS.

LRP4 is expressed in CLL cells

LRP4 rs2306029 is a non-synonymous SNP mapping to exon 31 of the *LRP4* gene and leading to Ser1554Gly amino acid substitution within an evolutionarily conserved position in the

context of the LDL-receptor class B beta-propeller motif (Fig 3). This motif is necessary for proper folding and stability of the LRP4 protein (Tomita *et al*, 1998). Bioinformatic analysis according to the PolyPhen algorithm indicated *LRP4* rs2306029 SNP as a possibly deleterious and damaging variant of LRP4 (score = 1.776).

LRP4 protein expression levels and pattern were analysed by Western Blot, flow cytometry confocal microscopy and immunohistochemistry in representative CLL samples. Western blot analysis revealed that CLL cells ($n = 16$) expressed the LRP4 protein with the expected size of *c.* 220 kD (Fig 4A). By flow cytometry performed on PB CLL cells, LRP4 was expressed on all cases analysed ($n = 66$), with a median percentage of expression of 72% (range 30–98%) and a mean fluorescence intensity (MFI) of 51 (range: 14–57) (Fig 4B). By confocal microscopy, LRP4 staining was preferentially localized on the cell membrane of CLL cells and was characterized by dim or marked fluorescence (Fig 4C). By immunohistochemistry, LRP4 was expressed in CLL lymphoid infiltrates of bone marrow biopsies in 19 out of 30 (63.3%) cases. LRP4 expression displayed a variable pattern ranging

from a prominent nuclear staining (8/19, 42.1%) to a mixed nuclear-cytoplasmic (6/18, 31.5%) or cytoplasmic (5/19, 26.3%) reactivity (Fig 5). The *LRP4* rs2306029 genotype did not significantly affect percentage, MFI or subcellular localization of LRP4 expression in CLL cells.

Discussion

This study documents that the *LRP4* rs2306029 genotype is an independent predictor of CLL transformation to RS. LRP4, together with LRP5/6, are members of the multifunctional low-density lipoprotein receptor gene family, that modulate the Wnt/ β -catenin signalling pathway (Kim *et al*, 2008; Ohazama *et al*, 2008; Zhang *et al*, 2008; Choi *et al*, 2009; Dietrich *et al*, 2010). LRP4 acts as an antagonist of the canonical Wnt/ β -catenin signalling pathway, as it displaces the homologous LRP5/6 receptors from the co-receptor complex and prevents the ligands from binding to LRP5/6 (Kim *et al*, 2008; Ohazama *et al*, 2008; Zhang *et al*, 2008; Choi *et al*, 2009; Dietrich *et al*, 2010; Li *et al*, 2010). In this context, the *LRP4* rs2306029 SNP, along with somatically acquired genetic/

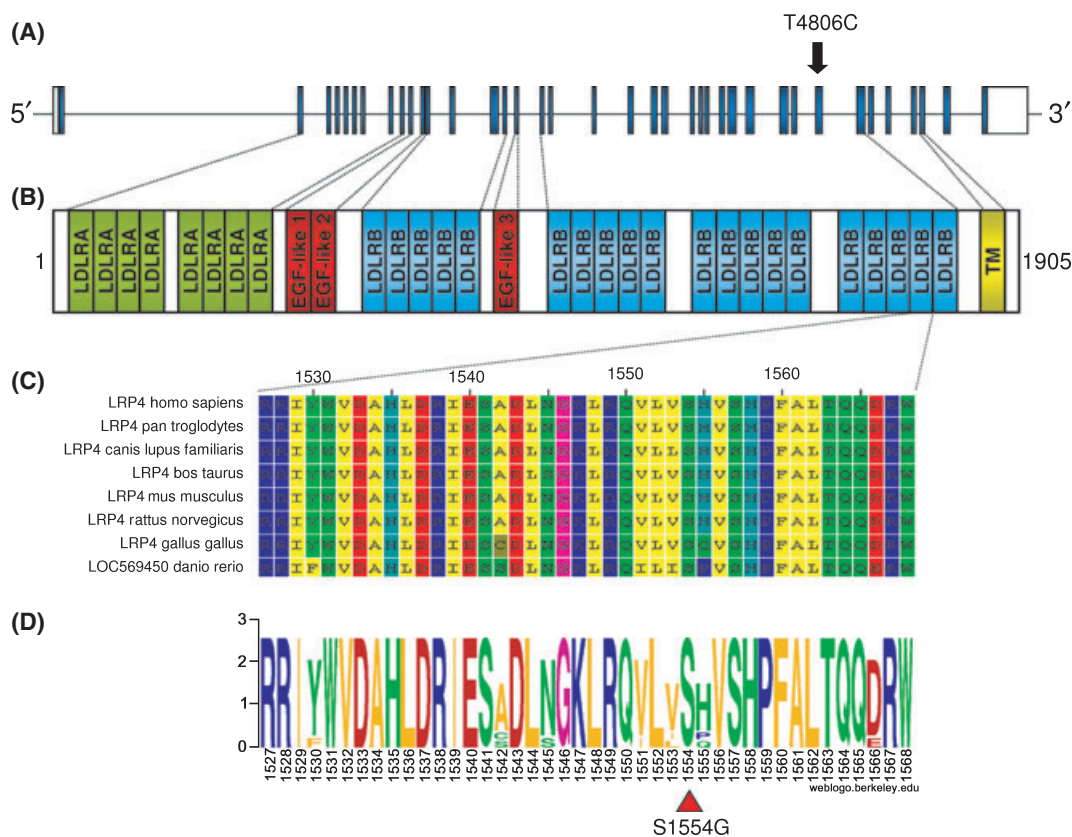


Fig 3. Localization of the *LRP4* rs2306029 SNP within the *LRP4* gene and LRP4 protein. Panels A and B: representation of the *LRP4* gene (panel A) and protein (panel B). *LRP4* rs2306029 is a non-synonymous SNP mapping to exon 31 of the *LRP4* gene (panel A) and leading to Ser1554Gly amino acid substitution within a LDL-receptor class B beta-propeller domain (LDLRB) of the LRP4 protein (panel B). Panel C: alignment of the LRP4 LDL-receptor class B beta-propeller domain amino acid sequences in phylogenesis. Panel D: Weblogo presentation LDL-receptor class B beta-propeller domain amino acid sequences in phylogenesis. The size of each column reflects the conservation of the position, the size of the letter reflects the frequency of each residue within this position. Ser1554 is an evolutionarily conserved amino acid position in the context of the LDL-receptor class B beta-propeller motif.

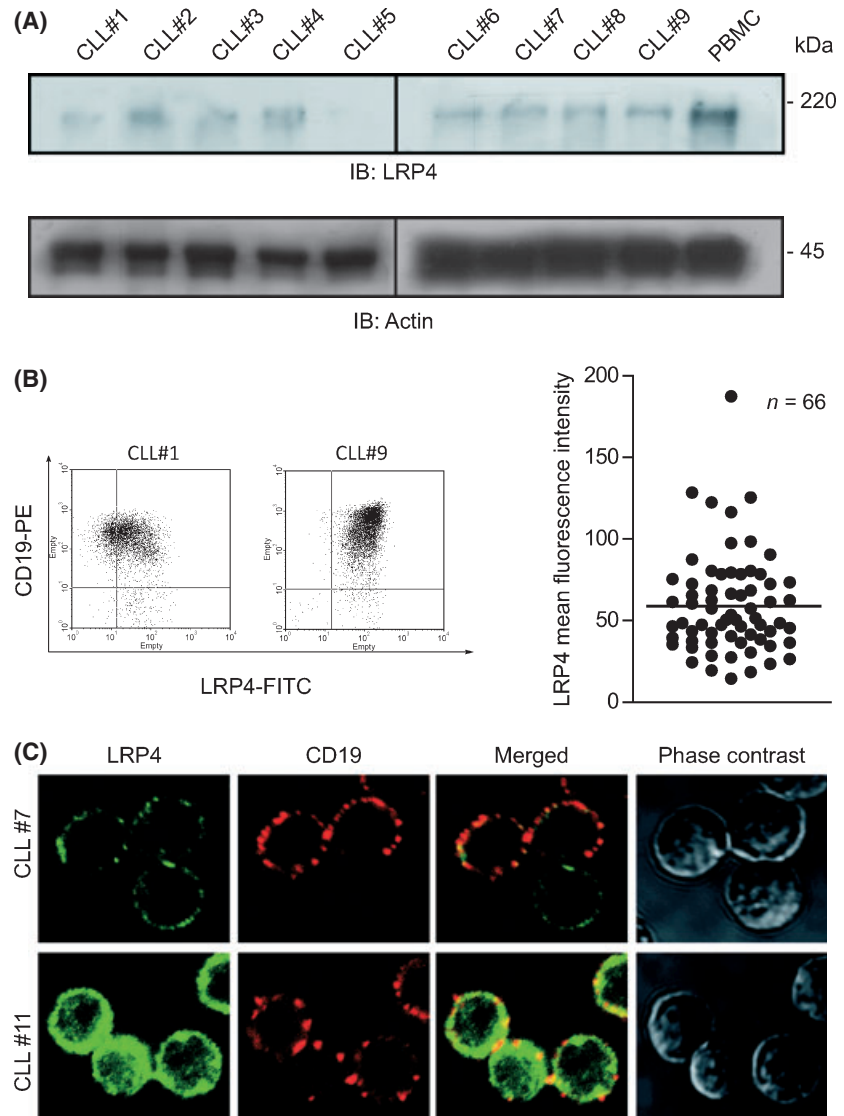


Fig 4. LRP4 expression in CLL. LRP4 was expressed in CLL cells as documented by Western Blot analysis of cell lysates from peripheral blood CD5⁺/CD19⁺ CLL cells (panel A), flow cytometry on peripheral blood mononuclear cells (panel B) and confocal microscopy on CD5⁺/CD19⁺ CLL cells (panel C).

epigenetic lesions of the Wnt/ β -catenin pathway (Liu *et al*, 2006; Chim *et al*, 2008; Sharma & Lichtenstein, 2009; Bennett *et al*, 2010; Gao *et al*, 2010), might contribute to Wnt/ β -

catenin deregulation, that is known to occur in CLL (Lu *et al*, 2004, 2009; Wu *et al*, 2009; Gutierrez *et al*, 2010), as well as in several other cancers (Clevers, 2006).

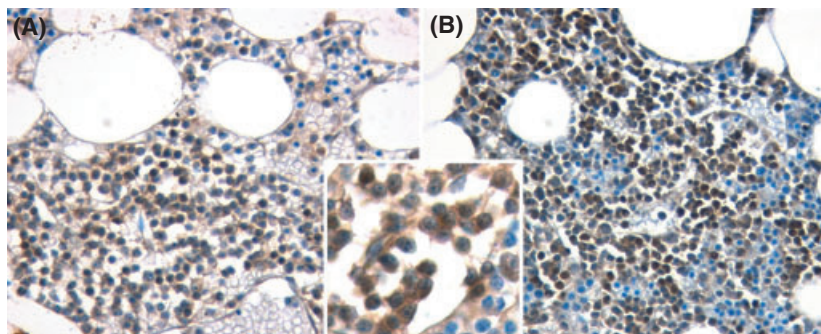


Fig 5. LRP4 expression by immunohistochemistry. Immunohistochemical detection of LRP4 on CLL bone marrow samples highlights LRP4 expression in neoplastic lymphoid infiltrates. By immunohistochemistry, LRP4 shows a variable pattern of expression ranging from a prevalent cytoplasmic localization, with or without nuclear reactivity (panel A and inset), to a prominent nuclear expression (panel B). Strept-ABC method with DAB chromogen (brown signal), original magnifications: (A–B) $\times 200$; inset $\times 400$.

The biological plausibility of the association between *LRP4* rs2306029 genotype and risk of RS transformation is supported by several observations. First, our results document that *LRP4* was expressed by CLL cells in virtually all cases when tested by multiple assays. Second, *LRP4* rs2306029 is a non-synonymous SNP leading to amino acid substitution at position 1554 within an evolutionarily conserved site of the *LRP4* LDL-receptor class B beta-propeller motif, that is necessary for proper folding, stability and functioning of the *LRP4* protein (Tomita *et al*, 1998; Duchesne *et al*, 2006; Simon-Chazottes *et al*, 2006; Drögemüller *et al*, 2007; Dietrich *et al*, 2010; Li *et al*, 2010). Third, bioinformatic analysis performed in this study scores *LRP4* rs2306029 as a variant with possible deleterious and damaging effects on *LRP4* protein function.

Among CLL clinical endpoints, *LRP4* rs2306029 associated only with RS transformation, and not with other types of clinical progression, further reinforcing the notion that predictors of CLL transformation are different, at least in part, from predictors of other CLL clinical outcomes (Aydin *et al*, 2008; Rossi *et al*, 2008a, 2009a,b; Deambrogi *et al*, 2009). Transformation of CLL to RS would thus represent a distinct event compared to progression without histological transformation, requiring a different, though complementary, diagnostic approach for a comprehensive assessment of the risk category of an individual patient.

Recent investigations have disclosed a number of risk factors, both clinical and biological, that might be helpful for RS prediction (Aydin *et al*, 2008; Rossi *et al*, 2008a, 2009a,b; Deambrogi *et al*, 2009). In particular, the concomitant usage of *IGHV4-39* gene with a stereotyped HCDR3 was the strongest predictor of RS transformation in a multi-institutional CLL cohort (Rossi *et al*, 2009a). The *LRP4* rs2306029 genotype can further improve the stratification model for RS prediction. In fact, the algorithm based on *LRP4* rs2306029 genotype, stereotyped HCDR3 and *IGHV4-39* gene usage is highly discriminatory in terms of RS prediction, and enables the segregation of CLL patients into two main risk categories (Fig 2B). The low risk category includes: (i) CLL without stereotyped HCDR3, independent of *IGHV4-39* usage and (ii) CLL utilizing stereotyped HCDR3 without *IGHV4-39* and harbouring the *LRP4* rs2306029 CT/CC genotype. The low risk category shows an expected 5-year probability of transformation of 3–8%, and therefore can be managed according to standard recommendations for CLL. The high-risk category includes: (i) CLL utilizing both *IGHV4-39* and stereotyped HCDR3 and (ii) CLL utilizing stereotyped HCDR3 without *IGHV4-39* but harbouring the *LRP4* rs2306029 TT genotype. This high-risk category shows an expected 5-year probability of transformation of 30–70%.

The most reliable evidence of a true genetic association is replication of the association in multiple populations (Manolio, 2010). Accordingly, a formal replication study in independent series would be required to confirm the association between *LRP4* rs2306029 and RS. In this study, we have addressed this issue by validating the enrichment of the *LRP4* rs2306029 TT genotype in an independent cohort of clonally

related RS. Future validation studies using an actuarial approach are needed to replicate the relevance of the *LRP4* rs2306029 genotype in predicting CLL transformation to RS.

In this study, *CD38* rs6449182, though nearly significant by univariate analysis corrected for multiple comparisons, did not emerge as an independent risk factor of RS transformation by multivariate analysis. The discrepancy between this study and the study by Aydin *et al* (2008) might be explained by differences in design and statistical analysis. First, this study utilized an actuarial approach, while the study by Aydin *et al* (2008) was based on a case-control design. Second, this study explored a large number of SNPs selected for being associated with CLL outcome. In contrast, the study by Aydin *et al* (2008) focused solely on *CD38* rs6449182. Finally, multivariate analysis in this study was adjusted for stereotyped HCDR3 and *IGHV4-39* usage, two covariates that were not known at the time of the study by Aydin *et al* (2008), and that were disclosed as strong independent risk factors for RS development by a subsequent investigation (Rossi *et al*, 2009a,b).

Failure to recognize RS transformation may expose patients to the impact of disease aggressiveness and to the sequelae of multiple lines of ineffective treatment, thus limiting the probabilities of survival. On these bases, a close clinical monitoring directed at early detection of clues of RS development and a careful biopsy policy should be applied to CLL patients at high risk of RS development, including those patients harbouring the *LRP4* rs2306029 TT genotype.

Acknowledgments

This study was supported by FIRB Progetto Futuro in Ricerca and PRIN 2008, Rome, MIUR, Italy; AIRC Special Program Molecular Clinical Oncology – 5 per mille No. 10007, Milan, Italy; Progetto Giovani Ricercatori 2008 and Ricerca Sanitaria Finalizzata, Ministero della Salute, Rome, Italy; Ricerca Sanitaria Finalizzata, Regione Piemonte, Italy; Ricerca Scientifica Applicata, Regione Friuli Venezia Giulia, Trieste ('Linfonet' Project); Novara-AIL Onlus, Novara, Italy; AIL Venezia Section, Pramaggiore Group; V.S. is supported by a fellowship of the Associazione Franca Capurro per Novara Onlus, Novara, Italy.

Authorship and conflicts of interest

D.R. and G.G. designed the study, analysed and interpreted data, performed statistical analysis and drafted the manuscript; S.R., V.S., and A.B. performed SNP genotyping and interpreted data; T.V. performed Western Blot, flow cytometry and confocal experiments and interpreted data; C.T. performed immunohistochemistry and interpreted data; L.D.P., M.F., E.S., and E.C. collected clinical data; L.L., R.M., C.V., and Z.Y.X revised clinically and pathologically the validation series; F.F., V.G., K.H.Y., F.M., and S.D. contributed to data analysis and interpretation.

The authors declare no competing financial interests.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table SI. SNP genotyping in the CLL series.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

- Allan, J.M., Sunter, N.J., Bailey, J.R., Pettitt, A.R., Harris, R.J., Pepper, C., Fegan, C., Hall, A.G., Deignan, L., Bacon, C.M., Pinton, J.C., Houlston, R.S., Broderick, P., Mainou-Fowler, T., Jackson, G.H., Summerfield, G., Evans, P.A., Strefford, J.C., Parker, A., Oscier, D., Pratt, G. & Allsup, D.J. (2010) Variant IRF4/MUM1 associates with CD38 status and treatment-free survival in chronic lymphocytic leukaemia. *Leukemia*, **24**, 877–881.
- Aydin, S., Rossi, D., Bergui, L., D'Arena, G., Ferrero, E., Bonello, L., Omedé, P., Novero, D., Morabito, F., Carbone, A., Gaidano, G., Malavasi, F. & Deaglio, S. (2008) CD38 gene polymorphism and chronic lymphocytic leukemia: a role in transformation to Richter syndrome? *Blood*, **111**, 5646–5653.
- Benjamini, Y. & Hochberg, Y. (1995) Controlling false discovery rate: a practicable and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, **57**, 289–300.
- Bennett, L.B., Taylor, K.H., Arthur, G.L., Rahmatpanah, F.B., Hooshmand, S.I. & Caldwell, C.W. (2010) Epigenetic regulation of WNT signaling in chronic lymphocytic leukemia. *Epigenomics*, **2**, 53–70.
- Chim, C.S., Pang, R. & Liang, R. (2008) Epigenetic dysregulation of the Wnt signalling pathway in chronic lymphocytic leukaemia. *Journal of Clinical Pathology*, **61**, 1214–1219.
- Choi, H.Y., Dieckmann, M., Herz, J. & Niemeier, A. (2009) Lrp4, a novel receptor for Dickkopf 1 and sclerostin, is expressed by osteoblasts and regulates bone growth and turnover in vivo. *PLoS ONE*, **4**, e7930.
- Clevers, H. (2006) Wnt/beta-catenin signaling in development and disease. *Cell*, **127**, 469–480.
- Cox, D.R. (1972) Regression models and life tables. *Journal of the Royal Statistical Society: Series B*, **34**, 187–220.
- Crowther-Swanepoel, D., Broderick, P., Di Bernardo, M.C., Dobbins, S.E., Torres, M., Mansouri, M., Ruiz-Ponte, C., Enjuanes, A., Rosenquist, R., Carracedo, A., Jurlander, J., Campo, E., Juliusson, G., Montserrat, E., Smedby, K.E., Dyer, M.J., Matutes, E., Dearden, C., Sunter, N.J., Hall, A.G., Mainou-Fowler, T., Jackson, G.H., Summerfield, G., Harris, R.J., Pettitt, A.R., Allsup, D.J., Bailey, J.R., Pratt, G., Pepper, C., Fegan, C., Parker, A., Oscier, D., Allan, J.M., Catovsky, D. & Houlston, R.S. (2010) Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. *Nature Genetics*, **42**, 132–136.
- Deambrogi, C., Cresta, S., Cerri, M., Rasi, S., De Paoli, L., Ramponi, A., Gaidano, G. & Rossi, D. (2009) 14q32 Translocations and risk of Richter transformation in chronic lymphocytic leukaemia. *British Journal of Haematology*, **144**, 131–133.
- Di Bernardo, M.C., Crowther-Swanepoel, D., Broderick, P., Webb, E., Sellick, G., Wild, R., Sullivan, K., Vijaykrishnan, J., Wang, Y., Pittman, A.M., Sunter, N.J., Hall, A.G., Dyer, M.J., Matutes, E., Dearden, C., Mainou-Fowler, T., Jackson, G.H., Summerfield, G., Harris, R.J., Pettitt, A.R., Hillmen, P., Allsup, D.J., Bailey, J.R., Pratt, G., Pepper, C., Fegan, C., Allan, J.M., Catovsky, D. & Houlston, R.S. (2008) A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nature Genetics*, **40**, 1204–1210.
- Dietrich, M.F., van der Weyden, L., Prosser, H.M., Bradley, A., Herz, J. & Adams, D.J. (2010) Ectodomains of the LDL receptor-related proteins LRP1b and LRP4 have anchorage independent functions in vivo. *PLoS ONE*, **5**, e9960.
- Drögemüller, C., Leeb, T., Harlizius, B., Tammen, I., Distl, O., Höltershinken, M., Gentile, A., Duchesne, A. & Eggen, A. (2007) Congenital syndactyly in cattle: four novel mutations in the low density lipoprotein receptor-related protein 4 gene (LRP4). *BMC Medical Genetics*, **8**, 5.
- Duchesne, A., Gautier, M., Chadi, S., Grohs, C., Floriot, S., Gallard, Y., Caste, G., Ducos, A. & Eggen, A. (2006) Identification of a doublet missense substitution in the bovine LRP4 gene as a candidate causal mutation for syndactyly in Holstein cattle. *Genomics*, **88**, 610–621.
- Gao, H., Le, Y., Wu, X., Silberstein, L.E., Giese, R.W. & Zhu, Z. (2010) VentX, a novel lymphoid-enhancing factor/T-cell factor-associated transcription repressor, is a putative tumor suppressor. *Cancer Research*, **70**, 202–211.
- Gryshchenko, I., Hofbauer, S., Stoecher, M., Daniel, P.T., Steurer, M., Gaiger, A., Eigenberger, K., Greil, R. & Tinhofer, I. (2008) MDM2 SNP309 is associated with poor outcome in B-cell chronic lymphocytic leukemia. *Journal of Clinical Oncology*, **26**, 2252–2257.
- Gutierrez, Jr, A., Tschumper, R.C., Wu, X., Shanafelt, T.D., Eckel-Passow, J., Huddleston, 3rd, P.M., Slager, S.L., Kay, N.E. & Jelinek, D.F. (2010) LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B cell lymphocytosis. *Blood*, **116**, 2975–83.
- Hallek, M., Cheson, B.D., Catovsky, D., Caligaris-Cappio, F., Dighiero, G., Döhner, H., Hillmen, P., Keating, M.J., Montserrat, E., Rai, K.R. & Kipps, T.J.; International Workshop on Chronic Lymphocytic Leukemia. (2008) Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*, **111**, 5446–5456.
- Harrell, Jr, F.E., Lee, K.L. & Mark, D.B. (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, **15**, 361–387.
- Heinze, G. & Schemper, M. (2001) A solution to the problem of monotone likelihood in Cox regression. *Biometrics*, **57**, 114–119.
- Kaplan, E.L. & Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.
- Kim, N., Stiegler, A.L., Cameron, T.O., Hallock, P.T., Gomez, A.M., Huang, J.H., Hubbard, S.R., Dustin, M.L. & Burden, S.J. (2008) Lrp4 is a receptor for Agrin and forms a complex with MuSK. *Cell*, **135**, 334–342.
- Li, Y., Pawlik, B., Elcioglu, N., Aglan, M., Kayserili, H., Yigit, G., Percin, F., Goodman, F., Nürnberg, G., Cenani, A., Urquhart, J., Chung, B.D., Ismail, S., Amr, K., Aslanger, A.D., Becker, C., Netzer, C., Scambler, P., Eyaad, W., Hamamy, H., Clayton-Smith, J., Hennekam, R., Nürnberg, P., Herz, J., Temtamy, S.A. & Wollnik, B. (2010) LRP4 mutations alter Wnt/beta-catenin signaling and cause limb and kidney malformations in Cenani-Lenz syndrome. *The American Journal of Human Genetics*, **14**, 696–706.
- Liu, T.H., Raval, A., Chen, S.S., Matkovic, J.J., Byrd, J.C. & Plass, C. (2006) CpG island methylation and expression of the secreted frizzled-related protein gene family in chronic lymphocytic leukemia. *Cancer Research*, **66**, 653–658.
- Lu, D., Zhao, Y., Tawatao, R., Cottam, H.B., Sen, M., Leoni, L.M., Kipps, T.J., Corr, M. & Carson, D.A. (2004) Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences USA*, **101**, 3118–3123.
- Lu, D., Liu, J.X., Endo, T., Zhou, H., Yao, S., Willert, K., Schmidt-Wolf, I.G., Kipps, T.J. & Carson, D.A. (2009) Ethacrynic acid exhibits selective toxicity to chronic lymphocytic leukemia cells by inhibition of the Wnt/beta-catenin pathway. *PLoS ONE*, **4**, e8294.
- Manolio, T.A. (2010) Genomewide association studies and assessment of the risk of disease. *New England Journal of Medicine*, **363**, 166–176.
- Matutes, E., Owusu-Ankomah, K., Morilla, R., Garcia Marco, J., Houlihan, A., Que, T.H. & Catovsky, D. (1994) The immunological profile of

- B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia*, **8**, 1640–1645.
- Nückel, H., Frey, U.H., Bau, M., Sellmann, L., Stanelle, J., Dürig, J., Jöckel, K.H., Dührsen, U. & Siffert, W. (2007) Association of a novel regulatory polymorphism (-938C>A) in the BCL2 gene promoter with disease progression and survival in chronic lymphocytic leukemia. *Blood*, **109**, 290–297.
- Ohazama, A., Johnson, E.B., Ota, M.S., Choi, H.Y., Porntaveetus, T., Oommen, S., Itoh, N., Eto, K., Gritli-Linde, A., Herz, J. & Sharpe, P.T. (2008) Lrp4 modulates extracellular integration of cell signaling pathways in development. *PLoS ONE*, **3**, e4092.
- Omoti, C.E. & Omoti, A.E. (2008) Richter syndrome: a review of clinical, ocular, neurological and other manifestations. *British Journal of Haematology*, **142**, 709–716.
- Rasi, S., Forconi, F., Bruscaggin, A., Sozzi, E., Gaidano, G. & Rossi, D. (2010) Impact of the host genetic background on prognosis of chronic lymphocytic leukemia. *Blood*, **115**, 1106–1107.
- Rossi, D. & Gaidano, G. (2009) Richter syndrome: molecular insights and clinical perspectives. *Hematological Oncology*, **27**, 1–10.
- Rossi, D., Cerri, M., Capello, D., Deambrogi, C., Rossi, F.M., Zucchetto, A., De Paoli, L., Cresta, S., Rasi, S., Spina, V., Franceschetti, S., Lunghi, M., Vendramin, C., Bomben, R., Ramponi, A., Monga, G., Conconi, A., Magnani, C., Gattei, V. & Gaidano, G. (2008a) Biological and clinical risk factors of chronic lymphocytic leukaemia transformation to Richter syndrome. *British Journal of Haematology*, **142**, 202–215.
- Rossi, D., Rasi, S., Capello, D. & Gaidano, G. (2008b) Prognostic assessment of BCL2-938C>A polymorphism in chronic lymphocytic leukemia. *Blood*, **111**, 466–468.
- Rossi, D., Spina, V., Cerri, M., Rasi, S., Deambrogi, C., De Paoli, L., Laurenti, L., Maffei, R., Forconi, F., Bertoni, F., Zucca, E., Agostinelli, C., Cabras, A., Lucioni, M., Martini, M., Magni, M., Deaglio, S., Ladetto, M., Nomdedeu, J.F., Besson, C., Ramponi, A., Canzonieri, V., Paulli, M., Marasca, R., Larocca, L.M., Carbone, A., Pileri, S.A., Gattei, V. & Gaidano, G. (2009a) Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clinical Cancer Research*, **15**, 4415–4422.
- Rossi, D., Lobetti Bodoni, C., Genuardi, E., Monitillo, L., Drandi, D., Cerri, M., Deambrogi, C., Ricca, L., Rocci, A., Ferrero, S., Bernocco, E., Capello, D., De Paoli, L., Bergui, L., Boi, M., Omedè, P., Massaia, M., Tarella, C., Passera, R., Boccadoro, M., Gaidano, G. & Ladetto, M. (2009b) Telomere length is an independent predictor of survival, treatment requirement and Richter's syndrome transformation in chronic lymphocytic leukemia. *Leukemia*, **23**, 1062–1072.
- Sellick, G.S., Goldin, L.R., Wild, R.W., Slager, S.L., Ressenti, L., Strom, S.S., Dyer, M.J., Mauro, F.R., Marti, G.E., Fuller, S., Lyttelton, M., Kipps, T.J., Keating, M.J., Call, T.G., Catovsky, D., Caporaso, N. & Houlston, R.S. (2007) A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. *Blood*, **110**, 3326–3333.
- Sellick, G.S., Wade, R., Richards, S., Oscier, D.G., Catovsky, D. & Houlston, R.S. (2008) Scan of 977 nonsynonymous SNPs in CLL4 trial patients for the identification of genetic variants influencing prognosis. *Blood*, **111**, 1625–1633.
- Sharma, S. & Lichtenstein, A. (2009) Aberrant splicing of the E-cadherin transcript is a novel mechanism of gene silencing in chronic lymphocytic leukemia cells. *Blood*, **114**, 4179–4185.
- Simon-Chazottes, D., Tutois, S., Kuehn, M., Evans, M., Bourgade, F., Cook, S., Davisson, M.T. & Guénet, J.L. (2006) Mutations in the gene encoding the low-density lipoprotein receptor LRP4 cause abnormal limb development in the mouse. *Genomics*, **87**, 673–677.
- Slager, S.L., Goldin, L.R., Strom, S.S., Lanasa, M.C., Spector, L.G., Rassenti, L., Leis, J.F., Camp, N.J., Kay, N.E., Vachon, C.M., Glenn, M., Weinberg, J.B., Rabe, K.G., Cunningham, J.M., Achenbach, S.J., Hanson, C.A., Marti, G.E., Call, T.G., Caporaso, N.E. & Cerhan, J.R. (2010) Genetic susceptibility variants for chronic lymphocytic leukemia. *Cancer Epidemiology, Biomarkers & Prevention*, **19**, 1098–1102.
- Swerdlow, S.H., Campo, E., Harris, N.L., Jaffe, E.S., Pileri, S.A., Stein, H., Thiele, J. & Vardiman, J.W. (2008) *World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. IARC, Lyon.
- Tomita, Y., Kim, D.H., Magoori, K., Fujino, T. & Yamamoto, T.T. (1998) A novel low-density lipoprotein receptor-related protein with type II membrane protein-like structure is abundant in heart. *The Journal of Biochemistry*, **124**, 784–789.
- Tsimberidou, A.M. & Keating, M.J. (2005) Richter syndrome: biology, incidence, and therapeutic strategies. *Cancer*, **103**, 216–228.
- Wu, Q.L., Zierold, C. & Ranheim, E.A. (2009) Dysregulation of Frizzled 6 is a critical component of B-cell leukemogenesis in a mouse model of chronic lymphocytic leukemia. *Blood*, **113**, 3031–3039.
- Zhang, B., Luo, S., Wang, Q., Suzuki, T., Xiong, W.C. & Mei, L. (2008) LRP4 serves as a coreceptor of agrin. *Neuron*, **60**, 285–297.

blood

2011 117: 2405-2413
Prepublished online December 14, 2010;
doi:10.1182/blood-2010-07-296244

The host genetic background of DNA repair mechanisms is an independent predictor of survival in diffuse large B-cell lymphoma

Davide Rossi, Silvia Rasi, Alice Di Rocco, Alberto Fabbri, Francesco Forconi, Annunziata Gloghini, Alessio Bruscaggin, Silvia Franceschetti, Marco Fangazio, Lorenzo De Paoli, Riccardo Bruna, Daniela Capello, Annalisa Chiappella, Chiara Lobetti Bodoni, Manuela Giachelia, Maria Chiara Tisi, Enrico M. Pogliani, Francesco Lauria, Marco Ladetto, Stefan Hohaus, Maurizio Martelli, Umberto Vitolo, Antonino Carbone, Robin Foà and Gianluca Gaidano

Updated information and services can be found at:
<http://bloodjournal.hematologylibrary.org/content/117/8/2405.full.html>

Articles on similar topics can be found in the following Blood collections
[Lymphoid Neoplasia](#) (1598 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>



The host genetic background of DNA repair mechanisms is an independent predictor of survival in diffuse large B-cell lymphoma

Davide Rossi,¹ Silvia Rasi,¹ Alice Di Rocco,² Alberto Fabbri,³ Francesco Forconi,³ Annunziata Gloghini,⁴ Alessio Bruscaggin,¹ Silvia Franceschetti,¹ Marco Fangazio,¹ Lorenzo De Paoli,¹ Riccardo Bruna,¹ Daniela Capello,¹ Annalisa Chiappella,⁵ Chiara Lobetti Bodoni,⁶ Manuela Giachelia,⁷ Maria Chiara Tisi,⁷ Enrico M. Pogliani,⁸ Francesco Lauria,³ Marco Ladetto,⁶ Stefan Hohaus,⁷ Maurizio Martelli,² Umberto Vitolo,⁵ Antonino Carbone,⁹ Robin Foà,² and Gianluca Gaidano¹

¹Division of Hematology, Department of Clinical and Experimental Medicine & CRIFF, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; ²Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy; ³Division of Hematology, University of Siena and AOUS, Italy; ⁴Department of Diagnostic Pathology and Laboratory Medicine, Istituto di Ricovero e Cura a Carattere Scientifico-National Cancer Institute, Milan, Italy; ⁵Division of Hematology, Department of Oncology, Azienda Ospedaliero-Universitaria S. Giovanni Battista, Turin, Italy; ⁶Division of Hematology, Department of Experimental Medicine and Oncology, University of Turin, Turin, Italy; ⁷Division of Hematology, Catholic University of the Sacred Heart, Rome, Italy; ⁸Division of Hematology, Department of Clinical and Preventive Medicine, University of Milano Bicocca, Monza, Italy; and ⁹Division of Pathology, Centro di Riferimento Oncologico, Istituto di Ricovero e Cura a Carattere Scientifico-National Cancer Institute, Aviano, Italy

Several drugs used for diffuse large B-cell lymphoma (DLBCL) treatment rely on DNA damage for tumor cell killing. We verified the prognostic impact of the host DNA repair genotype in 2 independent cohorts of DLBCL treated with R-CHOP21 (training cohort, 163 cases; validation cohort, 145 cases). Among 35 single nucleotide polymorphisms analyzed in the training series, *MLH1* rs1799977 was the sole predicting overall survival. DLBCL carrying

the *MLH1* AG/GG genotype displayed an increased death risk (hazard ratio [HR] = 3.23; $P < .001$; $q = 0.009$) compared with patients carrying the AA genotype. Multivariate analysis adjusted for International Prognostic Index identified *MLH1* AG/GG as an independent OS predictor ($P < .001$). The poor prognosis of *MLH1* AG/GG was the result of an increased risk of failing both R-CHOP21 (HR = 2.02; $P = .007$) and platinum-based

second-line (HR = 2.26; $P = .044$) treatment. Survival analysis in the validation series confirmed all outcomes predicted by *MLH1* rs1799977. The effect on OS of *MLH1*, a component of the DNA mismatch repair system, is consistent with its role in regulating the genotoxic effects of doxorubicin and platinum compounds, which are a mainstay of DLBCL first- and second-line treatment. (*Blood*. 2011; 117(8):2405-2413)

Introduction

Diffuse large B-cell lymphoma (DLBCL) is characterized by a heterogeneous clinical course, which may be predicted to some extent by the International Prognostic Index (IPI).^{1,2} In addition to clinical prognosticators, several biologic markers have been proposed as tools for refining outcome stratification in DLBCL.³⁻⁶ Most of these markers rely on features of the tumor clone.³⁻⁶

The genetic background of the host may also be relevant for DLBCL outcome. At present, evidence pointing to a role of the genetic background of the host in driving DLBCL prognosis is restricted to a limited number of studies that have focused on single nucleotide polymorphisms (SNPs) of immune system genes and of genes involved in drug metabolism.⁷⁻¹¹

DNA damage is one of the mainstays of cancer treatment. Several drugs used in DLBCL treatment rely on DNA damage as part of the mechanisms of tumor cell killing.¹²⁻¹⁷ These include drugs of the R-CHOP regimen (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone)¹⁸⁻²³ as well as drugs of second-line regimens containing platinum compounds.^{24,25}

Tumor cells respond to DNA damage induced by chemotherapy by activating 5 major DNA repair pathways: (1) mismatch repair

(MMR),¹²⁻¹⁴ (2) base excision repair,¹²⁻¹⁴ (3) nucleotide excision repair,¹²⁻¹⁴ (4) double strand break repair,¹⁵ and (5) direct reversal.¹⁶

MMR, of which *MLH1* is a major component, recognizes and removes mismatched or unmatched DNA base pairs or insertion-deletion loops.¹²⁻¹⁴ MMR is tightly involved in mediating cytotoxicity of doxorubicin and platinum compounds.^{13,17} DNA adducts produced by doxorubicin and platinum compounds exert a double genotoxic effect because they: (1) induce nucleotide mismatches and (2) interfere with normal MMR activity, thus preventing the repair of nucleotide mismatches from being completed.^{13,17} In cells characterized by MMR competence, the excess of unrepaired DNA triggers apoptosis. In contrast, when MMR is deficient, cells become chemorefractory and proliferate despite DNA damage caused by doxorubicin and platinum compounds.^{13,17} Chemorefractoriness resulting from alterations of DNA repair mechanisms may also involve base excision repair in the case of doxorubicin and alkylating agents,¹²⁻¹⁴ nucleotide excision repair in the case of platinum compounds and alkylating agents,¹²⁻¹⁴ double strand break repair in the case of doxorubicin and etoposide,¹⁵ and direct reversal in the case of alkylating agents.¹⁶

Submitted July 15, 2010; accepted November 28, 2010. Prepublished online as *Blood* First Edition paper, December 14, 2010; DOI 10.1182/blood-2010-07-296244.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2011 by The American Society of Hematology

On these grounds, DNA repair mechanisms and DNA repair capacity represent attractive candidates for explaining interindividual heterogeneity in treatment response of DLBCL. This study aimed at verifying whether SNPs of genes involved in DNA repair pathways may contribute to the prognostic stratification of DLBCL patients treated with R-CHOP.

Methods

Patients and study design

The design of the study was based on a training-validation approach. Accordingly, the patient population consisted of 2 cohorts of DLBCL treated with R-CHOP21 from December 2001 through November 2008. Patients were Italian residents and self-reported to be of white ancestry. Diagnosis of DLBCL was based on the World Health Organization classification of Hematopoietic Tumors.¹ A prior history of lymphoproliferative disorders was ruled out in all cases.

The first cohort (n = 163) was used as learning series because this cohort: (1) was a consecutive series from a single institution (Amedeo Avogadro University of Eastern Piedmont); (2) was provided with a large and homogeneous dataset of prospectively collected variables; (3) was homogeneously treated with the same chemotherapeutic regimen both at DLBCL diagnosis (R-CHOP21) and at relapse/progression (R-DHAP, rituximab, desamethasone, high-dose cytarabine, cisplatin); and (4) was genotyped on DNA from peripheral blood granulocytes in all cases. The following clinical variables were recorded at presentation: date of diagnosis, age, sex, Eastern Cooperative Oncology Group (ECOG) performance status (PS), Ann Arbor stage,²⁶ number of extranodal sites, B symptoms, bulky disease more than 10 cm, liver, and bone marrow involvement; lactate dehydrogenase (LDH, normal range, 200-450 U/L), absolute neutrophil count, hemoglobin and platelet count, bilirubin (normal range, 0.6-1.2 mg/dL), alkaline phosphatase (normal range, 90-360 U/L), albumin (normal range, 34-48 g/L), glomerular filtration rate; cardiac ejection fraction, abnormalities on echocardiography, abnormalities on electrocardiogram; hepatitis C virus, hepatitis B virus, and HIV serology; and number and type of comorbidities according to the Comorbidity Index and Score of Charlson.²⁷ The following variables were recorded during treatment: total number of courses per patient, date and type of response, and date of progression. In addition, the following variables were also recorded for each course of R-CHOP21: date of administration, doses of rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone, delay, toxicity, and dose reduction. The following variables were recorded during follow-up: date of progression, date and type of second-line treatment, date of stem cell transplantation, date of last follow-up visit, date of death, and cause of death.

The second cohort was used as validation series (n = 145). The validation series of DLBCL was composed of a retrospective multi-institutional series from centers based in Rome and participating in the same national lymphoma group (La Sapienza University, Rome, Italy; and Catholic University of the Sacred Heart, Rome, Italy). Inclusion criteria for the validation series were as follows: (1) diagnosis of DLBCL and (2) first-line treatment with R-CHOP21. SNP genotyping was performed on DNA extracted from formalin-fixed, paraffin-embedded tissue biopsies, or peripheral blood granulocytes when available. Second-line regimens used at relapse/progression included platinum compounds in 36 of 39 patients. The following clinical variables were recorded in the validation cohort: date of diagnosis, age, sex, ECOG PS, Ann Arbor stage, number of extranodal sites, bulky disease more than 10 cm, LDH, date of progression, date and type of second-line treatment, date of stem cell transplantation, date of last follow-up visit, and date of death.

Cases with a diagnosis of primary mediastinal large B-cell lymphoma, cases presenting with central nervous system involvement, and cases harboring HIV infection were excluded from both the training and the validation series.

Involved field radiotherapy was part of the initial treatment plan as consolidation on sites of primary bulky disease for both the training and the

validation series. Involved field radiotherapy was administered to 44 of 163 (27.0%) patients in the training series and to 37 of 145 (25.5%) patients in the validation series.

In both training and validation series, response was evaluated at the end of the fourth course and at the end of R-CHOP21 treatment according to the 1999 National Cancer Institute-Sponsored International Working Group guidelines.²⁸ Follow-up visits were scheduled every 3 months for the first 2 years and every 6 months thereafter. In case of nonresponse or progression after R-CHOP21, patients younger than 65 years were offered BEAM (carmustine, etoposide, cytarabine, melphalan) conditioned autologous stem cell transplantation as consolidation after salvage treatment.

Patients provided informed consent in accordance with the Declaration of Helsinki and with approval from the Amedeo Avogadro University of Eastern Piedmont Institutional Review Board.

SNP genotyping

Nonsynonymous SNPs, promoter SNPs, splicing site SNPs, and miRNA binding site SNPs were selected using an educated guess approach based on the following criteria: (1) minor allele frequency more than 5% in whites according to NCBI (<http://www.ncbi.nlm.nih.gov/snp>; accessed August 29, 2009); (2) previously reported functional consequence *in vitro/in vivo*; and/or (3) functional consequences predicted *in silico* according to PupaSuite (<http://pupasuite.bioinfo.cipf.es>; accessed August 29, 2009), PolyPhen (<http://genetics.bwh.harvard.edu/pph>; accessed August 29, 2009), or SIFT (<http://blocks.fhcrc.org/sift/SIFT.html>; accessed August 29, 2009); and/or (4) clinical relevance in settings other than DLBCL.

Sample size calculation was performed according to the assumption of number of patients with the genotype at risk more than or equal to 10 (corresponding to a minor allele frequency of $\geq 6\%$). Based on this assumption, the sample size of the training series (n = 163) would allow detection of at least 30% difference in 3-year overall survival (OS) with a power of 81% ($\alpha = 0.05$).

A total of 35 SNPs from 18 genes were analyzed (supplemental Table 1, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article). These included SNPs affecting: (1) MMR genes (*MLH1* rs1799977 and *MLH1* rs1800734); (2) base excision repair genes (*XRCC1* rs1799782, *XRCC1* rs25487, and *OGG1* rs1052133); (3) nucleotide excision repair genes (*ERCC1* rs3212986, *ERCC2* rs1052555, *ERCC2* rs13181, *ERCC2* rs1799793, *ERCC2* rs238406, *ERCC4* rs1800067, *ERCC4* rs3136038, *ERCC5* rs17655, *ERCC6* rs2228528, *ERCC6* rs2228529, *ERCC6* rs3793784, *XPA* rs1800975, *XPC* rs2227999, *XPC* rs2228000, *XPC* rs2607775, and *XPC* rs2228001); (4) double strand break repair genes (*BRCA1* rs4986850, *BRCA1* rs1799950, *BRCA1* rs799917, *BRCA2* rs144848, *LIG4* rs1805388, *XRCC2* rs3218536, *XRCC3* rs1799794, *XRCC3* rs861539, *XRCC4* rs1805377, *XRCC6* rs5751129, and *XRCC6* rs132788); and (5) direct reversal genes (*MGMT* rs16906252, *MGMT* rs2308321, and *MGMT* rs12917).

SNP genotyping was performed by SNP minisequencing (ABI Prism SNaPshot Multiplex kit, Applied Biosystems), after validation of this approach by DNA direct sequencing of each SNP in a pilot panel of cases (n = 15; data not shown). Quality control of genotyping was performed by replicate sample analysis. The concordance of SNP genotyping performed on granulocytes and on tumor biopsy from the same patients (n = 100) was 100%.

Deviation of SNP genotype distribution from Hardy-Weinberg equilibrium was tested by χ^2 test or Fisher exact test if appropriate. Five SNPs were not in Hardy-Weinberg equilibrium (supplemental Table 1) and therefore were excluded from subsequent analysis. Linkage disequilibrium among SNPs was calculated as r^2 values with the use of HAPLOVIEW (<http://www.broad.mit.edu/haploview/haploview>; accessed August 29, 2009). None of the SNPs was in strong linkage disequilibrium ($r^2 > 0.8$).

Immunohistochemistry

Immunohistochemistry for MLH1 was performed on formalin-fixed, paraffin-embedded DLBCL tissue sections from tissue microarrays (n = 35) and single sections (n = 23) using the mouse monoclonal antibody G168-15 (BD Biosciences Pharmingen) against hMLH1. A score of 0 to 3 for stain

Table 1. Clinical characteristics of the DLBCL cohorts at diagnosis

Clinical characteristic	Training series (n = 163)	Validation series (n = 145)	P
Age > 60 y	104/163 (63.8%)	96/145 (66.2%)	.659
Male:female	92:71	66:79	.056
ECOG PS > 1	21/163 (12.9%)	23/145 (15.9%)	.456
Ann Arbor stage			.301
I	34/163 (20.9%)	25/145 (17.2%)	—
II	44/163 (27.0%)	51/145 (35.2%)	—
III	29/163 (17.8%)	18/145 (12.4%)	—
IV	56/163 (34.4%)	51/145 (35.2%)	—
B symptoms	38/163 (23.3%)		
Bulky	46/163 (28.2%)	37/145 (25.5%)	.593
Extranodal sites > 1	41/163 (25.2%)	28/145 (19.3%)	.220
LDH > ULN	54/163 (33.1%)	66/145 (45.5%)	.931
IPI			.526
Low	80/163 (49.1%)	65/145 (44.8%)	—
Low-intermediate	34/163 (20.9%)	33/145 (22.8%)	—
High-intermediate	25/163 (15.3%)	29/145 (20.0%)	—
High	24/163 (14.7%)	18/145 (12.4%)	—
Non-GC phenotype	53/127 (41.7%)		
Bone marrow function			
ANC, $\times 10^9/L$	4.5 (3.6-5.7)	—	—
Hb, g/dL	12.9 (11.8-14.1)	—	—
Platelets, $\times 10^9/L$	256 (206-320)	—	—
Liver, renal, and cardiac function			
Albumin, g/L	4.1 (3.7-4.4)	—	—
ALP, U/L	174 (147-233)	—	—
Bilirubin, mg/dL	0.7 (0.5-0.9)	—	—
Glomerular filtration rate, mL/min	61 (48-77)	—	—
Cardiac ejection fraction, %	59 (55-64)	—	—
Comorbidities			
Charlson Comorbidity Index	1 (0-2)	—	—

The 25th to 75th percentiles are reported in parentheses for continuous variables.

CG indicates germinal center; ANC, absolute neutrophil count; Hb, hemoglobin; ALP, alkaline phosphatase; and —, not applicable.

intensity was assigned: no staining = 0; weakly positive = 1; moderately positive = 2; and strongly positive = 3. Percentage staining was assessed according to the following scoring system: 0% to 5% = 0; 6% to 20% = 1; 21% to 80% = 2; and 81% to 100% = 3. A combined immunohistochemical score was achieved by multiplying the percentage by the intensity score.²⁹ Both immunostaining and scoring were blinded to *MLH1* rs1799977 genotype. For the tissue microarray construction, hematoxylin and eosin-stained sections from each block were used to define diagnostic areas. Three representative 0.6-mm cores were obtained from each case and inserted in a grid pattern into a recipient paraffin block using a tissue arrayer (Advanced Tissue Arrayer ATA-100; Chemicon International-Millipore).³⁰ Cell of origin was assigned as previously reported.³⁰ Immunohistochemical studies for CD10, BCL-6, and IRF4 were performed on paraffin-embedded tissue sections as previously described.^{30,31}

Statistical analysis

The impact of DNA repair SNPs on R-CHOP efficacy was evaluated by OS. For the purpose of this study, which focused on DNA repair pathways associated with both R-CHOP and second-line regimen drugs, OS was considered as the primary endpoint because it may be influenced by both R-CHOP and second-line treatments. OS was measured from date of DLBCL diagnosis to date of death as a result of any cause (event), or last follow-up (censoring).²⁸ Progression-free survival (PFS) was measured from date of DLBCL diagnosis to date of disease progression or death as a result of any cause (event), or last follow-up (censoring).²⁸ OS after second-line treatment was measured from date of second-line treatment to date of death as a result of any cause (event), transplantation (censoring), or last follow-up (censoring). PFS after second-line treatment was measured from date of second-line treatment to date of progression (event), death as a

result of any cause (event), transplantation (censoring) or last follow-up (censoring). Response was assessed according to published guidelines.²⁸ The association between SNPs and efficacy endpoints was evaluated considering the minor allele of each SNP as acting both in a recessive and in a dominant fashion. For SNPs where 10 or fewer minor allele homozygotes were observed, only the combination of minor allele homozygotes with heterozygotes was analyzed. If this combined frequency was still less than 10, then the SNP was removed from the analysis. Survival analysis was performed using the Kaplan-Meier method.³² Univariate and multivariate analyses were performed by Cox regression.³³ IPI score entered Cox analysis as an ordinal variable with 4 levels (low, low-intermediate, high-intermediate, and high).² False discovery rate was used to control for multiple statistical testing for each SNP as acting both in a recessive and in a dominant fashion.³⁴ Categorical variables were compared by χ^2 test. Continuous variables were compared by Mann-Whitney test. All statistical tests were 2-sided. Statistical significance was defined as *P* value less than .05. The analysis was performed with the Statistical Package for the Social Sciences software Version 17.0, R statistical package 2.8.1 (<http://www.r-project.org>; accessed October 11, 2010).

Results

Clinical characteristics at diagnosis and treatment of the DLBCL training series

All 35 SNPs from 18 genes belonging to 5 different DNA repair pathways were genotyped in the training series of 163 newly diagnosed DLBCL homogeneously treated with R-CHOP21 (supplemental Table 1).

The clinical characteristics at DLBCL diagnosis of the training series are described in Table 1. Overall, 974 courses of R-CHOP21 were administered. The median number of courses per patient was 6 (25th-75th percentile, 5-8) and 120 of 163 (73.6%) patients received at least 6 R-CHOP21 courses. Median relative dose intensity (RDI) of R-CHOP21 drugs was 100% for rituximab, 95.4% for cyclophosphamide, 94.0% for doxorubicin, 100% for vincristine, and 100% for prednisone. Complete remission (CR)/unconfirmed CR after R-CHOP21 was achieved in 127 of 163 (77.9%) patients. PFS at 4 years was 57.9%. OS at 4 years was 72.9%. Forty-six patients failed or progressed after R-CHOP21 and required R-DHAP salvage treatment. Autologous stem cell transplantation consolidation after salvage treatment was performed in 15 patients younger than 65 years.

The *MLH1* rs1799977 genotype is a predictor of OS in DLBCL

Univariate Cox analysis controlled for multiple comparisons by false discovery rate testing identified *MLH1* rs1799977 as the sole SNP predicting OS in the DLBCL training series (supplemental Table 1). DLBCL patients who carried the *MLH1* rs1799977 AG/GG genotype displayed an increased risk of death (hazard ratio [HR] = 3.23; events/N = 22/52; 4-year OS, 55.5%; 95% confidence interval [CI], 40.7%-70.3%) compared with patients who carried the *MLH1* rs1799977 AA genotype (events/N = 18/111; 4-year OS, 80.9%; 95% CI, 72.1%-89.7%; *P* < .001; *q* = .009; Table 2; Figure 1A).

Clinical variables associated with poor OS in the DLBCL training series were ECOG PS > 1 (HR = 2.79; *P* = .007), bulky disease (HR = 2.92; *P* < .001), extranodal sites more than 1 (HR = 2.06; *P* = .027), LDH more than upper limit of normal (ULN; HR = 2.52; *P* = .003), IPI score (HR = 1.54; *P* = .001), and RDI of doxorubicin less than 90% (HR = 2.02; *P* = .028; Table 2; supplemental Figure 1). Multivariate analysis identified the *MLH1* rs1799977 AG/GG genotype (HR = 3.14; *P* < .001) as an

Table 2. Univariate analysis for OS and PFS after R-CHOP21 in the training and validation DLBCL series

	Training series				Validation series			
	HR	LCI	UCI	P	HR	LCI	UCI	P
OS								
<i>MLH1</i> rs1799977 AG/GG	3.23	1.73	6.03	< .001*	2.97	1.42	6.21	.004*
Age > 60 y	2.00	0.95	4.20	.068	3.71	1.29	10.62	.014*
Male	1.59	0.83	3.04	.162	1.69	0.83	3.46	.147
ECOG PS > 1	2.79	1.32	5.92	.007*	1.05	0.41	2.80	.883
Ann Arbor stage III or IV	1.23	0.66	2.30	.515	1.87	0.91	3.87	.088
Bulky	2.92	1.55	5.48	< .001*	1.81	0.86	3.79	.114
Extranodal sites > 1	2.06	1.08	3.92	.027*	2.56	1.22	5.35	.012*
LDH > ULN	2.52	1.35	4.69	.003*	2.77	1.30	5.89	.008*
IPI score	1.54	1.18	2.00	.001*	1.88	1.37	2.58	< .001*
Non-GC phenotype	1.68	0.59	4.36	.352	NA	NA	NA	NA
RDI cyclophosphamide < 90%	1.69	0.91	3.14	.099	NA	NA	NA	NA
RDI doxorubicin < 90%	2.02	1.08	3.78	.028*	NA	NA	NA	NA
Radiotherapy	1.11	0.57	2.16	.749	1.08	0.50	2.37	.830
PFS								
<i>MLH1</i> rs1799977 AG/GG	2.02	1.20	3.38	.007*	1.92	1.09	3.39	.024*
Age > 60 y	1.74	0.97	3.12	.064	2.75	1.28	5.89	.009*
Male	1.35	0.79	2.31	.364	1.34	0.76	2.36	.311
ECOG PS > 1	1.84	0.93	3.64	.081	1.06	0.50	2.85	.863
Ann Arbor stage III or IV	1.91	1.12	3.26	.017*	2.87	1.56	5.30	.001*
Bulky	2.33	1.39	3.91	.001*	1.26	0.67	2.35	.460
Extranodal sites > 1	1.83	1.07	3.15	.028*	2.98	1.64	5.39	< .001*
LDH > ULN	1.91	1.14	3.22	.014*	2.52	1.39	4.55	.002*
IPI score	1.50	1.21	1.86	< .001*	1.85	1.44	2.38	< .001*
Non-GC phenotype	1.05	0.50	2.18	.897	NA	NA	NA	NA
RDI cyclophosphamide < 90%	1.42	0.85	2.37	.183	NA	NA	NA	NA
RDI doxorubicin < 90%	1.47	0.88	2.46	.136	NA	NA	NA	NA
Radiotherapy	.096	0.56	1.72	.963	1.39	0.63	3.08	.406

LCI indicates 95% lower confidence interval; UCI, 95% upper confidence interval; NA, not applicable; and GC, germinal center.

*Significant value.

independent predictor of OS in DLBCL treated with R-CHOP21, along with IPI score (HR = 1.38; $P = .037$) and bulky disease (HR = 2.56; $P = .004$; Table 3).

The *MLH1* rs1799977 genotype is a predictor of failure after R-CHOP21

In the DLBCL training series, the prognostic role of *MLH1* rs1799977 on OS after R-CHOP21 may not be explained by differences in clinical features at presentation or differences in R-CHOP21 feasibility. Indeed, DLBCL patients who carried the *MLH1* rs1799977 AG/GG genotype did not differ from DLBCL patients who carried the *MLH1* rs1799977 AA genotype in terms of age more than 60 years, ECOG PS more than 1, Charlson Comorbidity Index, Ann Arbor stage III or IV, bulky disease, B symptoms, extranodal sites more than 1, LDH more than ULN, IPI 3 to 5, or cell of origin ($P > .05$ in all instances; Table 4). In addition, DLBCL patients who carried the *MLH1* rs1799977 AG/GG genotype did not differ from DLBCL patients who carried the *MLH1* rs1799977 AA genotype in terms of RDI of doxorubicin, RDI of cyclophosphamide, number of R-CHOP21 courses, or R-CHOP21-related mortality ($P > .05$ in all instances; Table 4). Finally, the CR rate after R-CHOP21 (AG/GG 39/of 52, 75.0% vs AA 88 of 111, 79.3%, $P = .539$) and after second-line treatment (AG/GG 3 of 18, 16.7% vs AA 14 of 28, 50.0%, $P = .072$) was not significantly affected by the *MLH1* rs1799977 genotype.

The poor prognosis heralded by *MLH1* rs1799977 AG/GG genotype in DLBCL of the training series was the result of an increased risk of progressing after both first- and second-line

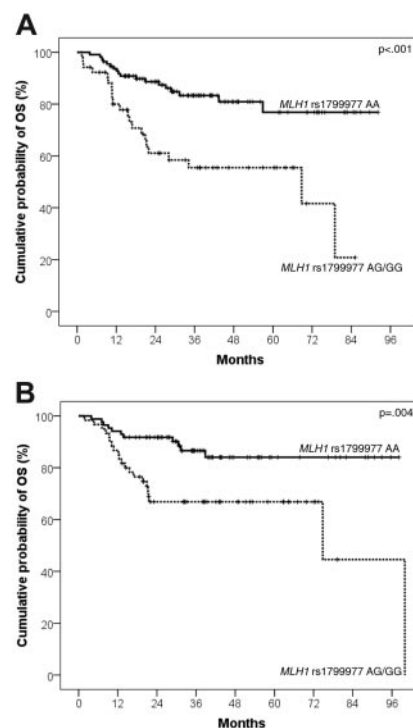


Figure 1. Kaplan-Meier curves for DLBCL survival after R-CHOP21 according to *MLH1* rs1799977 genotype. OS after R-CHOP21 according to *MLH1* rs1799977 genotype in the training series (A) and in the validation series (B).

Table 3. Multivariate analysis for OS and PFS after R-CHOP21 in the training and validation DLBCL series

	Training series				Validation series			
	HR	LCI	UCI	P	HR	LCI	UCI	P
OS								
<i>MLH1</i> rs1799977 AG/GG	3.14	1.67	5.91	< .001*	3.04	1.45	6.37	.003*
Bulky	2.56	1.34	4.89	.004*	1.57	0.74	3.33	.233
IPI	1.38	1.02	1.88	.037*	1.85	1.35	2.53	< .001*
RDI doxorubicin < 90%	1.19	0.58	2.43	.642	NA	NA	NA	NA
PFS								
<i>MLH1</i> rs1799977 AG/GG	1.96	1.17	3.28	.010*	1.98	1.12	3.50	.019*
Bulky	1.96	1.15	3.34	.012*	1.01	0.54	1.91	.954
IPI	1.41	1.12	1.77	.002*	1.86	1.45	2.38	< .001*

LCI indicates 95% lower confidence interval; UCI, 95% upper confidence interval; and NA, not applicable.

*Significant value.

treatment. Indeed, DLBCL patients carrying the *MLH1* rs1799977 AG/GG genotype displayed an increased risk of progression after R-CHOP21 (HR = 2.02; events/N = 26/52; 4-year PFS, 47.5%; 95% CI, 33.0%-62.0%) compared with patients carrying the *MLH1* rs1799977 AA genotype (events/N = 33/111; 4-year PFS, 65.6%; 95% CI, 54.7%-76.5%; $P = .007$; Table 2; Figure 2A).

Clinical variables associated with poor PFS after R-CHOP21 in the DLBCL training series were Ann Arbor stage III or IV (HR = 1.91; $P = .017$), bulky disease (HR = 2.33; $P = .001$), extranodal sites more than 1 (HR = 1.83; $P = .028$), LDH more than ULN (HR = 1.91; $P = .014$), and IPI score (HR = 1.50; $P < .001$; Table 2; supplemental Figure 1).

Multivariate analysis identified the *MLH1* rs1799977 AG/GG genotype (HR = 1.96; $P = .010$) as an independent predictor of PFS after R-CHOP21 in the DLBCL training series, along with IPI score (HR = 1.41; $P = .002$) and bulky disease (HR = 1.96; $P = .012$; Table 3).

The *MLH1* rs1799977 genotype is a predictor of failure after platinum-based second-line regimens

MLH1 rs1799977 AG/GG genotype predicted poor PFS and OS from salvage in patients who failed R-CHOP21 and received R-DHAP as second-line treatment ($n = 46$). Indeed, DLBCL patients carrying the *MLH1* rs1799977 AG/GG genotype displayed an increased risk of progression after R-DHAP (HR = 2.26; 95% CI, 1.02-5.01; events/N = 14/18; 2-year PFS from salvage: 8.6%;

95% CI, 0%-24.0%) compared with patients carrying the *MLH1* rs1799977 AA genotype (events/N = 11/28; 2-year PFS from salvage: 40.7%; 95% CI, 16.6%-64.8%; $P = .044$; Figure 3A).

In addition, DLBCL patients who carried the *MLH1* rs1799977 AG/GG genotype displayed an increased risk of death after R-DHAP (HR = 2.93; 95% CI, 1.14-7.33; events/N = 14/18; 2-year OS from salvage: 16.0%; 95% CI, 0%-35.4%) compared with patients who carried the *MLH1* rs1799977 AA genotype (events/N = 7/28; 2-year OS from salvage: 61.6%; 95% CI, 36.4%-86.8%; $P = .024$; Figure 4A).

Validation of *MLH1* rs1799977 genotype as a predictor of outcome in an independent DLBCL series treated with R-CHOP21

The prognostic value of *MLH1* rs1799977 as a predictor of outcome in DLBCL patients treated with R-CHOP21 was validated in an independent series of DLBCL ($n = 145$). The validation series did not differ from the training series in terms of: (1) clinical features at presentation (Table 1); (2) median follow-up (49 vs 48 months, respectively; $P = .312$); (3) 4-year OS after R-CHOP21 (77.4% vs 72.9%, respectively; $P = .370$); (4) 4-year PFS after R-CHOP21 (64.8% vs 61.1%, respectively; $P = .404$); (5) 2-year OS from salvage (46.0% vs 38.2%, respectively; $P = .920$); (6) 2-year PFS from salvage (25.8% vs 25.3%; $P = .610$); and (7) prevalence of *MLH1* rs1799977 genotypes (AA: 85/145, 58.6%; AG: 51/145, 35.2%; GG: 9/145, 6.2% vs AA:

Table 4. Clinical characteristics of the DLBCL training cohort at diagnosis according to the *MLH1* rs1799977 genotype

Clinical characteristic	<i>MLH1</i> rs1799977		P
	AA	AG/GG	
Age > 60 y	71/111 (64.0%)	33/52 (63.5%)	.950
Male	59/111 (53.2%)	33/52 (63.5%)	.216
ECOG PS > 1	13/111 (11.7%)	8/52 (15.4%)	.514
Charlson Comorbidity Index	1 (0-2)	1 (0-3)	.115
Ann Arbor stage III or IV	56/111 (50.5%)	29/52 (55.8%)	.526
B symptoms	21/111 (18.9%)	17/52 (32.7%)	.052
Bulky	30/111 (27.0%)	16/52 (30.8%)	.620
Extranodal sites > 1	28/111 (25.2%)	13/52 (25.0%)	.975
LDH > ULN	33/111 (29.7%)	21/52 (40.4%)	.177
IPI 3-5	32/111 (28.8%)	17/52 (32.7%)	.616
Non-GC phenotype	35/79 (44.3%)	18/48 (37.5%)	.450
RDI of doxorubicin, %	94.8 (81.8-100)	88.2 (76.1-99.8)	.123
RDI of cyclophosphamide, %	95.8 (87.5-100)	92.1 (80.7-100)	.236
No. of R-CHOP courses	6 (6-8)	6 (4-8)	.372
1-year treatment-related mortality	1.8%	6.0%	.286

The 25th to 75th percentiles are reported in parentheses for continuous variables.
GC indicates germinal center.

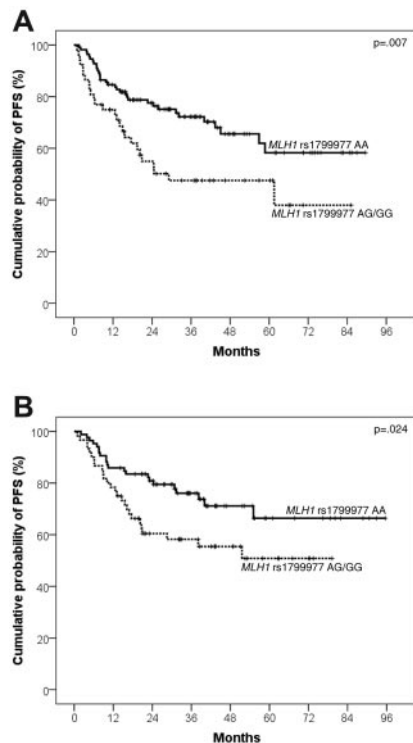


Figure 2. Kaplan-Meier curves for DLBCL progression after R-CHOP21 according to *MLH1* rs179977 genotype. PFS after R-CHOP21 according to *MLH1* rs179977 genotype in the training series (A) and in the validation series (B).

111/163, 68.1%; AG: 43/163, 26.4%; GG: 9/163, 5.5%, respectively; $P = .213$).

Survival analysis in the validation series confirmed that *MLH1* rs179977 is a prognostic factor in DLBCL patients treated with

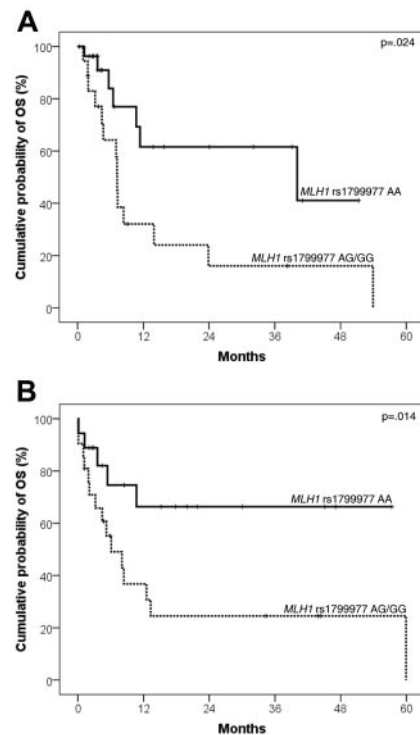


Figure 4. Kaplan-Meier curves for DLBCL survival from salvage treatment according to *MLH1* rs179977 genotype. OS from salvage treatment according to *MLH1* rs179977 genotype in the training series (A) and in the validation series (B).

R-CHOP21. Indeed, by univariate Cox analysis, the *MLH1* rs179977 AG/GG genotype was associated with poor OS (HR = 2.97; events/N = 21/60; 4-year OS, 66.8%; 95% CI, 54.5%-79.1%) compared with the *MLH1* rs179977 AA genotype (events/N = 11/85; 4-year OS, 84.0%; 95% CI, 74.8%-93.2%; $P = .004$; Table 2; Figure 1B).

Other variables associated with poor OS in the validation series were age more than 60 years (HR = 3.71; $P = .014$), extranodal sites more than 1 (HR = 2.56; $P = .012$), LDH more than ULN (HR = 2.77; $P = .008$), and IPI score (HR = 1.88; $P < .001$; Table 2; supplemental Figure 1).

Multivariate analysis selected the *MLH1* rs179977 AG/GG genotype as an independent predictor of OS in the validation series (HR = 3.04; $P = .003$), after adjusting for IPI score (HR = 1.85; $P < .001$) and bulky disease (HR = 1.57; $P = .233$; Table 3).

Analogous to the training series, the *MLH1* rs179977 genotype was associated with poor PFS also in the validation series. Indeed, DLBCL patients carrying the *MLH1* rs179977 AG/GG genotype displayed an increased risk of progression after R-CHOP21 (HR = 1.92; events/N = 26/60; 4-year PFS, 55.4%; 95% CI, 42.1%-68.7%) compared with patients carrying the *MLH1* rs179977 AA genotype (events/N = 22/85; 4-year PFS, 71.1%; 95% CI, 60.0%-82.2%; $P = .024$; Table 2; Figure 2B).

Other variables associated with poor PFS after R-CHOP21 in the validation series were age more than 60 years (HR = 2.75; $P = .009$), Ann Arbor stage III or IV (HR = 2.87; $P = .001$), extranodal sites more than 1 (HR = 2.98; $P < .001$), LDH more than ULN (HR = 2.52; $P = .002$), and IPI score (HR = 1.85; $P < .001$; Table 2; supplemental Figure 1).

Multivariate analysis selected the *MLH1* rs179977 AG/GG genotype as an independent predictor of PFS after R-CHOP21 in the validation series (HR = 1.98; $P = .019$), after adjusting for IPI

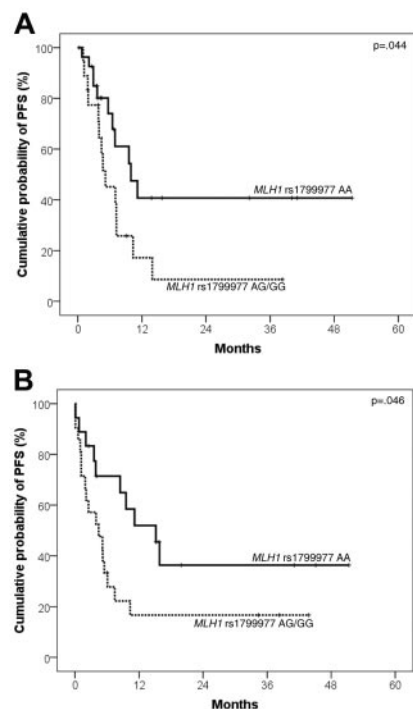
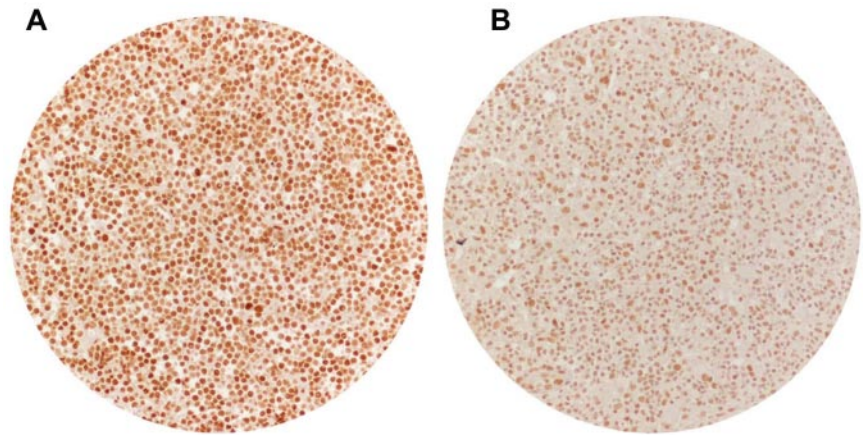


Figure 3. Kaplan-Meier curves for DLBCL PFS from salvage treatment according to *MLH1* rs179977 genotype. PFS from salvage treatment according to *MLH1* rs179977 genotype in the training series (A) and in the validation series (B).

Figure 5. MLH1 expression in DLBCLs. Immunohistochemical analysis of MLH1 protein expression in paraffin sections from representative DLBCL primary biopsies. (A) A representative DLBCL sample harboring the *MLH1* rs1799977 AA genotype and characterized by a strong nuclear positivity for MLH1. (B) A representative DLBCL sample harboring the *MLH1* rs1799977 GG genotype and characterized by a weak nuclear positivity for MLH1. Tissue microarray, immunoperoxidase, hematoxylin counterstain. Images were taken using a Nikon Eclipse 80i microscope (Nikon) with a pan fluor 20×/0.13 objective and Nikon digital sight DS-Fi1 camera equipped with control unit-DS-L2 (Nikon). Images were assembled using Adobe Photoshop 6 (Adobe Systems).



score (HR = 1.86; $P < .001$) and bulky disease (HR = 1.01; $P = .954$; Table 3).

Finally, the *MLH1* rs1799977 AG/GG genotype predicted poor PFS and OS from salvage with second-line regimens also in patients belonging to the validation series who failed R-CHOP21 ($n = 39$). Indeed, patients carrying the AG/GG genotype had a worse PFS from salvage (HR = 2.25; 95% CI, 1.01-4.99; events/N = 17/21; 2-year PFS from salvage: 16.7%; 95% CI, 0.1%-33.3%) compared with patients carrying the AA genotype (events/N = 10/18; 2-year PFS from salvage: 36.4%; 95% CI, 11.2%-61.6%; $P = .046$; Figure 3B). Patients carrying the AG/GG genotype had a worse OS from salvage (HR = 3.01; 95% CI, 1.08-8.39; events/N = 15/21; 2-year OS from salvage: 25.4%; 95% CI, 5.1%-45.7%) compared with patients carrying the AA genotype (events/N = 5/18; 2-year OS from salvage: 66.3%; 95% CI, 41.7%-90.9%; $P = .014$; Figure 4B).

Similarly to the training series, also in the validation series the CR rate after R-CHOP21 (AG/GG 44/60, 73.3% vs AA 71/85, 83.5%, $P = .135$) and after second-line treatment (AG/GG 5/21, 23.8% vs AA 7/18, 38.9%, $P = .377$) was not significantly affected by the *MLH1* rs1799977 genotype.

The *MLH1* rs1799977 genotype correlates with MLH1 protein expression

Immunohistochemical analysis revealed that MLH1 was expressed by all DLBCL patients analyzed ($n = 58$). Percentage of positive tumor cells and intensity of staining for MLH1 were heterogeneous in DLBCL samples (Figure 5). Indeed, MLH1 was expressed by 6% to 20% of tumor cells in 2 of 58 (3.4%) DLBCLs, by 21% to 80% tumor cells in 25 of 58 (43.1%), and by 81% to 100% tumor cells in 31 of 58 (53.4%). In addition, DLBCL showed a weak expression of MLH1 in 5 of 58 (8.6%) cases, a moderate expression in 14 of 58 (24.1%), and a strong expression in 39 of 58 (67.2%). Heterogeneity of MLH1 expression correlated with the *MLH1* rs1799977 genotype. Indeed, the immunohistochemical score for MLH1 expression²⁹ was significantly lower among DLBCL patients harboring the AG/GG genotype (median, 5) compared with cases harboring the AA genotype (median, 9; $P = .034$).

Discussion

The current study on 308 de novo DLBCL documents that the *MLH1* rs1799977 SNP: (1) predicts OS in newly diagnosed

DLBCL patients treated with R-CHOP21, (2) predicts failure after both R-CHOP21 and platinum-based second-line regimens, (3) exerts a prognostic role independent of IPI, and (4) predicts all DLBCL outcomes in series of patients treated at different institutions, as documented by the training-validation approach chosen for study design.

The association between *MLH1* rs1799977 AG/GG genotype and poor DLBCL outcome appears to be biologically plausible based on several lines of evidence. First, the G variant allele of *MLH1* rs1799977 associates with a reduction of MLH1 protein expression, as documented in DLBCL (this study) and other types of cancer.^{35,36} Consistent with the role of *MLH1* in promoting apoptosis triggered by chemotherapy, reduced MLH1 protein expression in cancer cells confers in vitro resistance to doxorubicin and platinum compounds.³⁷⁻³⁹ Second, at variance with other DNA repair genes investigated in this study, *MLH1* not only affects DNA repair mechanisms but is also involved in DNA damage signaling, a process that induces cell cycle arrest and can lead to apoptosis in case of major DNA damages by doxorubicin and platinum compounds.^{12,13,17} Third, the *MLH1* rs1799977 AG/GG genotype harbors prognostic relevance also in disease models other than lymphoma, including acute lymphoblastic leukemia, solid tumors, and inflammatory bowel disease.⁴⁰⁻⁴⁷

The cross-validation approach of the current study documents that the *MLH1* rs1799977 genotype is an independent prognostic factor retaining its value in 2 DLBCL series treated at different institutions using similar therapeutic strategies. This observation suggests that the prognostic value of *MLH1*, although detected retrospectively, appears to be independent of a potential bias because of patient referral or patient management at a single center. Confirmation within the frame of prospective controlled studies will be helpful to fully assess the prognostic role of *MLH1* in DLBCL.

In both the training and validation series, cases of DLBCL with an unfavorable IPI appear to be underrepresented because of the concomitance of the DLCL-04 national trial run by the Italian Lymphoma Intergroup testing the role of autologous stem cell transplantation at diagnosis in young patients with unfavorable IPI (www.clinicaltrials.gov #NCT00499018; accessed July 13, 2010). Despite the underrepresentation of high-risk DLBCL, the impact of *MLH1* rs1799977 genotype is observed also in this DLBCL subgroup (data not shown).

Consistent with the characteristics of our patient series, bulky disease emerged as an independent prognostic factor in the validation series. The prognostic role of bulky disease is not

unexpected in DLBCL cohorts with a preponderance of favorable IPI patients, as clearly documented by the MInT trial.⁴⁸

Although the role of radiotherapy within a combined-modality approach including R-CHOP has not been defined,⁴⁹ in our study, patients presenting with bulky disease were planned to receive involved field radiotherapy to the sites of initial bulk. In both the training and the validation series, involved field radiotherapy administration to sites of initial bulk did not compromise the prognostic effect of *MLH1* rs1799977. Indeed, compared with the *MLH1* rs1799977 AA genotype, the *MLH1* rs1799977 AG/GG genotype was associated with a significantly shorter PFS independent of having or having not received radiotherapy in both the training series (radiotherapy, $P = .037$; no radiotherapy, $P = .022$, respectively) and in the validation series (radiotherapy, $P = .049$; no radiotherapy, $P = .035$).

The impact of *MLH1* rs1799977 genotype on DLBCL outcome may be related to the chance of obtaining CR with induction therapy and/or to the degree of clearance of residual lymphoma cells, which may be responsible of relapse during follow-up. *MLH1* rs1799977 predicted PFS and OS in both the training and validation series. Conversely, the CR rate after R-CHOP21 was not significantly affected by *MLH1* rs1799977. Notably, also *TP53* disruption, a well-known biomarker of tumor cell chemorefractoriness in lymphoid neoplasia, associates with poor survival but not with CR rate in CHOP-treated DLBCL.⁵⁰ These data would suggest that the impact of the *MLH1* rs1799977 genotype on PFS, rather than on CR rate, is the result of a better clearance of residual tumor cells in patients harboring the *MLH1* rs1799977 AA genotype, thus reducing the chance of disease recurrence. It should be noted, however, that the high activity of R-CHOP and the high frequency of CR (127 of 163, 77.9% in the training series; and 115 of 145, 79.3% in the validation series), as well as the sample size of the 2 series might have prevented the detection of a statistical difference in CR rate between *MLH1* rs1799977 genotypes, and that larger studies might be required to fully address this issue.

The results of this study point to *MLH1* rs1799977 as a potential novel prognosticator for predicting treatment failure and survival of DLBCL. With very few exceptions,⁷⁻¹¹ most new prognostic markers of DLBCL have been prompted by investigations of the genetics and of the biology of the lymphoma clone. The example of

the predictive value of the *MLH1* genotype points to the need of focusing also on the genetic background of the host for a comprehensive assessment of DLBCL prognosis.

Acknowledgments

This work was supported by Fondo per gli Investimenti della Ricerca di Base-Programma Futuro in Ricerca 2008, and Programma di Ricerca di Interesse Nazionale 2008, Ministero della Università e della Ricerca, Rome, Italy; Associazione Italiana per la Ricerca sul Cancro, Special Program Molecular Clinical Oncology, 5 × 1000, No. 10007, Milan, Italy; Progetto Giovani Ricercatori 2008, Ricerca Sanitaria Finalizzata, and Progetto Integrato Oncologia-Advanced Molecular Diagnostics project (RFPS-2006-2-339694.1; RFPS-2006-2-339723.2), Ministero della Salute, Rome, Italy; Ricerca Sanitaria Finalizzata, Regione Piemonte, Torino, Italy; Istituto Toscano Tumori 2008 Grant, Florence, Italy; and Novara-Associazione Italiana contro le Leucemie, Linfomi e Mielomi Onlus, Novara.

Authorship

Contribution: D.R. and G.G. designed the study, analyzed and interpreted data, performed statistical analysis, and drafted the manuscript; S.R., A.B., and D.C. performed experimental research and analyzed data; S.F., L.D.P., M.F., and R.B. collected clinical data of the training series; F.F., A.F., A.D.R., A. Chiappella, C.L.B., M.G., and M.C.T. collected clinical data of the validation series; A.G. performed immunohistochemistry and interpreted data; and E.M.P., F.L., M.L., S.H., M.M., U.V., A. Carbone, and R.F. contributed to data analysis and interpretation.

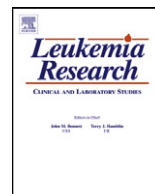
Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Davide Rossi, Division of Hematology, Department of Clinical and Experimental Medicine & CRIF, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy; e-mail: rossidav@med.unipmn.it.

References

1. Swerdlow SH, Campo E, Harris NL, et al. *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: International Agency for Research on Cancer; 2008.
2. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med*. 1993;329(14):987-994.
3. Abramson JS, Shipp MA. Advances in the biology and therapy of diffuse large B-cell lymphoma: moving toward a molecularly targeted approach. *Blood*. 2005;106(4):1164-1174.
4. Lossos IS, Morgensztern D. Prognostic biomarkers in diffuse large B-cell lymphoma. *J Clin Oncol*. 2006;24(6):995-1007.
5. Cheung KJ, Horsman DE, Gascoyne RD. The significance of TP53 in lymphoid malignancies: mutation prevalence, regulation, prognostic impact and potential as a therapeutic target. *Br J Haematol*. 2009;146(3):257-269.
6. Lenz G, Staudt LM. Aggressive lymphomas. *N Engl J Med*. 2010;362(15):1417-1429.
7. Lech-Maranda E, Baseggio L, Bienvenu J, et al. Interleukin-10 gene promoter polymorphisms influence the clinical outcome of diffuse large B-cell lymphoma. *Blood*. 2004;103(9):3529-3534.
8. Habermann TM, Wang SS, Maurer MJ, et al. Host immune gene polymorphisms in combination with clinical and demographic factors predict late survival in diffuse large B-cell lymphoma patients in the pre-rituximab era. *Blood*. 2008;112(7):2694-2702.
9. Kube D, Hua TD, von Bonin F, et al. Effect of interleukin-10 gene polymorphisms on clinical outcome of patients with aggressive non-Hodgkin's lymphoma: an exploratory study. *Clin Cancer Res*. 2008;14(12):3777-3784.
10. Rossi D, Rasi S, Franceschetti S, et al. Analysis of the host pharmacogenetic background for prediction of outcome and toxicity in diffuse large B-cell lymphoma treated with R-CHOP21. *Leukemia*. 2009;23(6):1118-1126.
11. Schoof N, von Bonin F, Zeynalova S, et al. Favorable impact of the interleukin-4 receptor allelic variant I75 on the survival of diffuse large B-cell lymphoma patients demonstrated in a large prospective clinical trial. *Ann Oncol*. 2009;20(9):1548-1554.
12. Kinsella TJ. Coordination of DNA mismatch repair and base excision repair processing of chemotherapy and radiation damage for targeting resistant cancers. *Clin Cancer Res*. 2009;15(6):1853-1859.
13. Martin LP, Hamilton TC, Schilder RJ. Platinum resistance: the role of DNA repair pathways. *Clin Cancer Res*. 2008;14(5):1291-1295.
14. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. *Nature*. 2009;461(7267):1071-1078.
15. Downs JA. Chromatin structure and DNA double-strand break responses in cancer progression and therapy. *Oncogene*. 2007;26(56):7765-7772.
16. egi ME, Sciuscio D, Murat A, Levivier M, Stupp R. Epigenetic deregulation of DNA repair and its potential for therapy. *Clin Cancer Res*. 2009;15(16):5026-5031.
17. Fink D, Aebi S, Howell SB. The role of DNA mismatch repair in drug resistance. *Clin Cancer Res*. 1998;4(1):1-6.
18. Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med*. 2002;346(4):235-242.

19. Feugier P, Van Hoof A, Sebban C, et al. Long-term results of the R-CHOP study in the treatment of elderly patients with diffuse large B-cell lymphoma: a study by the Groupe d'Etude des Lymphomes de l'Adulte. *J Clin Oncol*. 2005; 23(18):4117-4126.
20. Sehn LH, Donaldson J, Chhanabhai M, et al. Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. *J Clin Oncol*. 2005;23(22):5027-5033.
21. Habermann TM, Weller EA, Morrison VA, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. *J Clin Oncol*. 2006; 24(19):3121-3127.
22. Pfreundschuh M, Trümper L, Osterborg A, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol*. 2006;7(5):379-391.
23. Pfreundschuh M, Schubert J, Ziepert M, et al. Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). *Lancet Oncol*. 2008;9(2):105-116.
24. Vellenga E, van Putten WL, van't Veer MB, et al. Rituximab improves the treatment results of DHAP-VIM-DHAP and ASCT in relapsed/progressive aggressive CD20+ NHL: a prospective randomized HOVON trial. *Blood*. 2008; 111(2):537-543.
25. Kewalramani T, Zelenetz AD, Nimer SD, et al. Rituximab and ICE as second-line therapy before autologous stem cell transplantation for relapsed or primary refractory diffuse large B-cell lymphoma. *Blood*. 2004;103(10):3684-3688.
26. Lister TA, Crowther D, Sutcliffe SB, et al. Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. *J Clin Oncol*. 1989; 78(11):1630-1636.
27. Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc*. 1968;16(5):622-626.
28. Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas: NCI Sponsored International Working Group. *J Clin Oncol*. 1999;17(4):1244-1253.
29. Mackay HJ, Cameron D, Rahilly M, et al. Reduced MLH1 expression in breast tumors after primary chemotherapy predicts disease-free survival. *J Clin Oncol*. 2000;18(1):87-93.
30. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004; 103(1):275-282.
31. Carbone A, Gioghini A, Libra M, et al. A spindle cell variant of diffuse large B-cell lymphoma possesses genotypic and phenotypic markers characteristic of a germinal center B-cell origin. *Mod Pathol*. 2006;19(2):299-306.
32. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Am Stat Assoc*. 1958;53:457-481.
33. Cox DR. Regression models and life tables. *J R Stat Assoc*. 1972;34:187-220.
34. Benjamini Y, Hochberg Y. Controlling false discovery rate: a practicable and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57:289-300.
35. Curia MC, Palmirotta R, Aceto G, et al. Unbalanced germ-line expression of hMLH1 and hMSH2 alleles in hereditary nonpolyposis colorectal cancer. *Cancer Res*. 1999;59(15):3570-3575.
36. Kim JC, Roh SA, Koo KH, et al. Genotyping possible polymorphic variants of human mismatch repair genes in healthy Korean individuals and sporadic colorectal cancer patients. *Fam Cancer*. 2004;3(2):129-137.
37. Fink D, Nebel S, Aebi S, et al. The role of DNA mismatch repair in platinum drug resistance. *Cancer Res*. 1996;56(21):4881-4886.
38. Aebi S, Fink D, Gordon R, et al. Resistance to cytotoxic drugs in DNA mismatch repair-deficient cells. *Clin Cancer Res*. 1997;3(10):1763-1767.
39. Fedier A, Schwarz VA, Walt H, Carpin RD, Haller U, Fink D. Resistance to topoisomerase poisons due to loss of DNA mismatch repair. *Int J Cancer*. 2001;93(4):571-576.
40. Mathonnet G, Krajcinovic M, Labuda D, Sinnott D. Role of DNA mismatch repair genetic polymorphisms in the risk of childhood acute lymphoblastic leukaemia. *Br J Haematol*. 2003;123(1):45-48.
41. Bagnoli S, Putignano AL, Melean G, et al. Susceptibility to refractory ulcerative colitis is associated with polymorphism in the hMLH1 mismatch repair gene. *Inflamm Bowel Dis*. 2004;10(6):705-708.
42. Smith TR, Levine EA, Freimanis RI, et al. Polygenic model of DNA repair genetic polymorphisms in human breast cancer risk. *Carcinogenesis*. 2008;29(11):2132-2138.
43. Campbell PT, Curtin K, Ulrich CM, et al. Mismatch repair polymorphisms and risk of colon cancer, tumour microsatellite instability and interactions with lifestyle factors. *Gut*. 2009;58(5):661-667.
44. Nejda N, Iglesias D, Moreno Azcoita M, Medina Arana V, González-Aguilera JJ, Fernández-Peralta AM. A MLH1 polymorphism that increases cancer risk is associated with better outcome in sporadic colorectal cancer. *Cancer Genet Cytogenet*. 2009;193(2):71-77.
45. Vietri MT, Riegler G, De Paola M, et al. I219V polymorphism in hMLH1 gene in patients affected with ulcerative colitis. *Genet Test Mol Biomarkers*. 2009;13(2):193-197.
46. Langeberg WJ, Kwon EM, Koopmeiners JS, Ostrander EA, Stanford JL. Population-based study of the association of variants in mismatch repair genes with prostate cancer risk and outcomes. *Cancer Epidemiol Biomarkers Prev*. 2010;19(1):258-264.
47. Picelli S, Zajac P, Zhou XL, et al. Common variants in human CRC genes as low-risk alleles. *Eur J Cancer*. 2010;46(6):1041-1048.
48. Pfreundschuh M, Ho AD, Cavallin-Stahl E, et al. Prognostic significance of maximum tumour (bulk) diameter in young patients with good-prognosis diffuse large-B-cell lymphoma treated with CHOP-like chemotherapy with or without rituximab: an exploratory analysis of the MabThera International Trial Group (MInT) study. *Lancet Oncol*. 2008;9(5):435-444.
49. Pfreundschuh M. How I treat elderly patients with diffuse large B-cell lymphoma. *Blood*. 2010; 116(24):5103-5110.
50. Young KH, Leroy K, Møller MB, et al. Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood*. 2008; 112(8):3088-3098.



Brief communication

Saliva is a reliable and practical source of germline DNA for genome-wide studies in chronic lymphocytic leukemia

Silvia Rasi^a, Alessio Bruscaggin^a, Andrea Rinaldi^b, Stefania Cresta^a, Marco Fangazio^a, Lorenzo De Paoli^a, Sara Monti^a, Ernesto Gargiulo^a, Ivo Kwee^b, Robin Foà^c, Francesco Bertoni^b, Gianluca Gaidano^a, Davide Rossi^{a,*}

^a Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy

^b Laboratory of Experimental Oncology, Oncology Institute of Southern Switzerland-IOSI, Bellinzona, Switzerland

^c Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy

ARTICLE INFO

Article history:

Received 4 May 2011

Received in revised form 7 May 2011

Accepted 17 May 2011

Available online 12 June 2011

Keywords:

Saliva

Germline DNA

SNP array

Next generation sequencing

Chronic lymphocytic leukemia

ABSTRACT

High-throughput genomics requires tumor DNA matched to germline DNA, that cannot be easily obtained in the context of leukemia. Using chronic lymphocytic leukemia as a model, saliva DNA was frequently devoid of tumor DNA also during overt disease, and passed quality controls for SNP-array (77/102, 75.4%) and next generation sequencing (71/102, 69.6%). Compared to saliva, urine provides germline DNA of similar quality but in lower amounts. Saliva DNA was successfully run on SNP 6.0 arrays, and passed quality control call rate thresholds. On these bases, saliva represents a useful source of germline DNA for high-throughput genomic studies of hematologic neoplasia.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Availability of patients' germline DNA has become mandatory in the era of high throughput genomics, such as SNP array and next generation sequencing [1,2]. In fact, identification of tumor-specific somatic mutations requires calling of sequence variants in both tumor and matched germline DNA in order to filter out the vast number of private or unknown germline polymorphisms that, otherwise, could be erroneously scored as tumor-specific mutations [1,2].

In solid cancers, peripheral blood (PB) is commonly utilized as a source of germline DNA [1]. Conversely, in leukemic disorders, PB cells provide a suitable source of germline DNA only when the disease is in profound complete remission (CR) [1]. Consequently, in leukemic patients who do not achieve CR because of refractoriness or because they are not candidate to treatment, difficulties in obtaining germline DNA might hamper identification of tumor-specific genetic abnormalities. In chronic lymphocytic leukemia (CLL), for example, outside the setting of a profound CR, PB fractions

containing granulocytes or T-cells are often contaminated by neoplastic cells that might introduce tumor DNA in the control samples [1,2]. Based on these assumptions, alternative sources of germline DNA should be sought in leukemic disorders.

Skin biopsies have been recently proposed as an alternative source of germline DNA in hematologic malignancies [1]. However, patients' compliance to this invasive procedure may be suboptimal, and surgical trauma might cause leukemic cell contamination of the biopsy. In contrast, saliva and urine can be easily collected without invasive procedures, although their value in genome-wide investigations of hematologic neoplasia has not been formally established.

This study assessed the adequacy of saliva and urine as germline DNA sources for genome-wide studies. CLL was chosen as a model since: (i) most cases are left untreated at diagnosis; and (ii) treatment rarely achieves a profound CR [3], thus preventing the achievement of pure germline DNA in most patients.

2. Materials and methods

2.1. Patients and samples

One ml of saliva was collected before a meal from 82 CLL with overt disease (median B-lymphocyte count $14.4 \times 10^9/L$; range 5.0 – $151.1 \times 10^9/L$) and from 20 CLL in CR. A spot urine sample (50 ml) was collected from 30/82 CLL with overt disease (median B-lymphocyte count $20.0 \times 10^9/L$; range 5.0 – $94.6 \times 10^9/L$) and from 5/20 CLL in CR. Paired blood samples (10 ml) were also collected. PB granulocytes were separated from mononuclear cells by Ficoll-Paque density

* Corresponding author at: Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy. Tel.: +39 0321 660698; fax: +39 0321 620421.

E-mail address: rossidav@med.unipmn.it (D. Rossi).

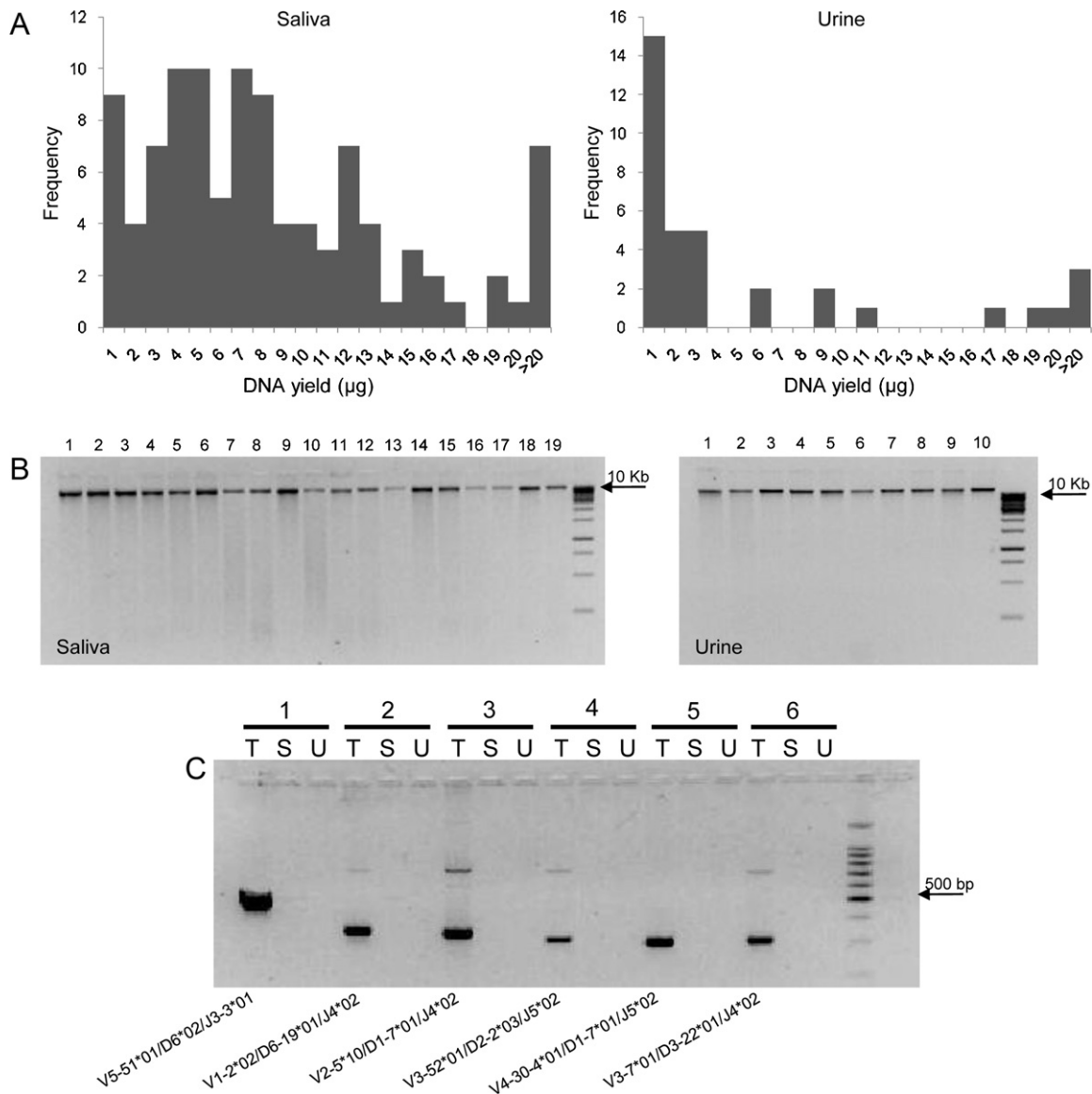


Fig. 1. Quality controls of genomic DNA from saliva and urine. Panel A: yields of genomic DNA from saliva and urine samples. Panel B: molecular weight estimation in representative genomic DNA samples of saliva ($n = 19$) and urine ($n = 10$) by 0.8% agarose gel electrophoresis. High molecular weight standards are shown for comparison. In all samples, a band corresponding to high molecular weight DNA is observed. Panel C: PCR analysis of patient-specific IGHV-D-J rearrangement in genomic DNA from tumor (T), saliva (S) and urine (U). In all samples shown in the figure, a band of IGHV-D-J rearrangement can be readily observed in tumor samples, but not in matched samples of saliva and urine.

gradient centrifugation (GE Healthcare, Milan, Italy). Genomic DNA was purified by cell lysis followed by digestion with proteinase K, "salting out" extraction, and ethanol precipitation. Patients provided informed consent in accordance with local IRB requirements and Declaration of Helsinki.

2.2. DNA quality control

UV spectrum readings (Nanodrop 2000C, Thermo Scientific, Wilmington, DE) were used to determine DNA purity and concentration. DNA molecular weight was estimated by 0.8% agarose gel electrophoresis. Contamination by tumor DNA was assessed by PCR analysis of tumor-specific markers, i.e. IGHV-D-J rearrangements. By dilution experiments, the sensitivity of this strategy was 10^{-3} . Each PCR product visualized by agarose gel electrophoresis was sequenced by ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems) and aligned to ImMunoGeneTics sequence directory (<http://imgt.cines.fr>) in order to characterize the IGHV-D-J rearrangement [4]. The yield of human DNA in saliva and urine was estimated using the Quantifiler Human DNA Quantification kit (Applied Biosystems). Efficiency in PCR amplification of DNA from saliva, urine, and granulocytes was monitored by real time amplification of the *KRT1* control gene. Samples were processed using the Affymetrix Human Mapping GeneChip 6.0 arrays (Affymetrix, Santa Clara, CA, USA) as previously described [5]. Call rates were calculated using Affymetrix Genotyping Console 4.0 with the Birdseed algorithm.

3. Results and discussion

Overall, the median yield of DNA extracted from 1 ml of saliva was 7.4 μ g (range 0.3–176.2 μ g) (Fig. 1A). Purity of saliva DNA was acceptable (median $OD_{260/280}$: 1.8; range: 1.6–2.0). DNA from saliva was of high molecular weight in 99/102 (97.0%) cases (Fig. 1B). Because buccal DNA may contain non-human DNA from oral microflora or from food remnants that may affect large-scale genotyping [6], we quantified the representation of human DNA. The median percentage of human DNA in saliva DNA was 93% (range 70–100%). Contamination of saliva DNA by tumor DNA was restricted to 13/102 (12.7%) cases, including 12/82 (14.6%) CLL with overt disease and 1/20 (5.0%) CLL in CR (Fig. 1C). DNA from saliva was amplifiable for the *KRT1* control gene in all cases as efficiently (median C_T : 24.0; range 20.6–26.8) as DNA from matched PB granulocytes (median C_T : 24.4; range 22.0–28.3) (Fig. 2).

For comparative purposes, contamination by tumor DNA was also assessed in matched PB granulocyte DNA. PB granulocyte DNA

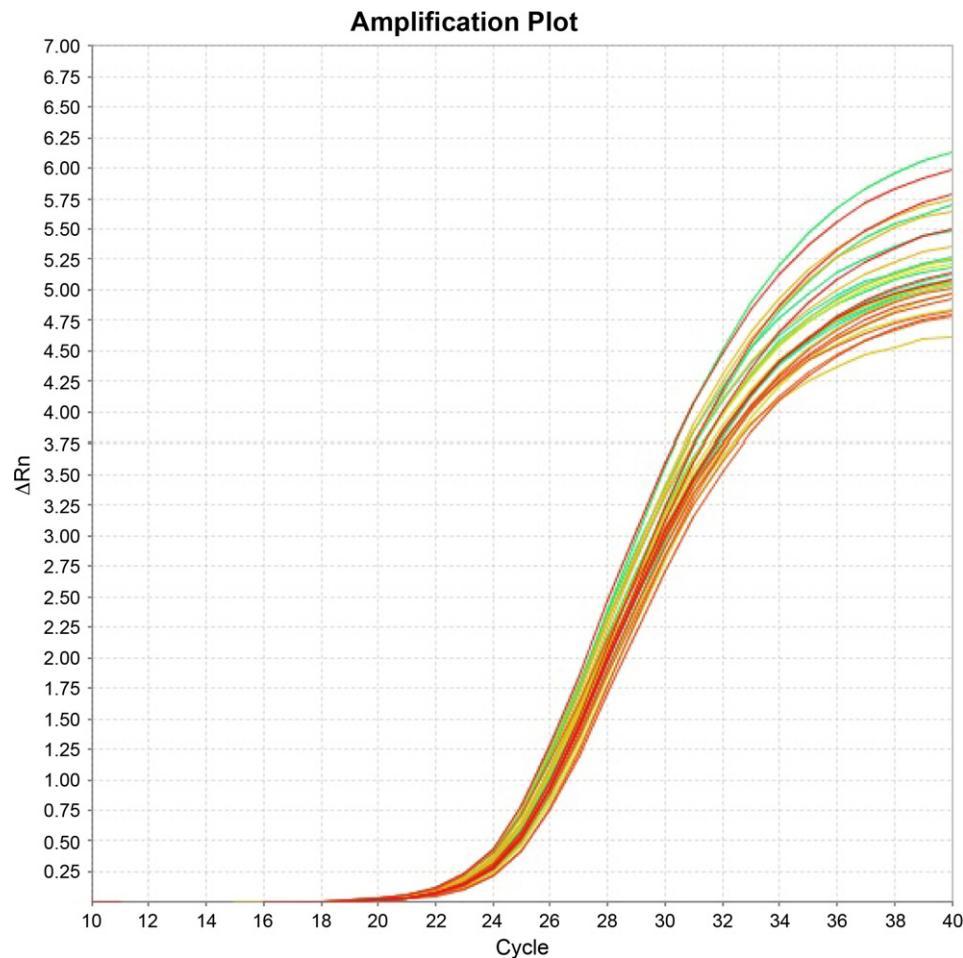


Fig. 2. Amplification plots. PCR amplification suitability of genomic DNA obtained from saliva ($n=12$; yellow) vs. genomic DNA from urine ($n=12$; green) vs. genomic DNA from peripheral blood granulocytes ($n=12$; red). The amplification plots are representative of *KRT1* used as a control gene, and are superimposable in all types of biological samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

was consistently contaminated by tumor DNA in all (82/82, 100%) CLL with overt disease. Importantly, also when PB was collected during post-treatment CR, 50% (10/20) of PB granulocyte DNA was contaminated by tumor DNA.

Overall, 77/102 (75.4%) saliva DNAs presented all the quality requirements necessary to undergo SNP array studies, including total amount >500 ng, purity ($OD_{260/280}$ ratio: 1.6–2.1), high molecular weight, non-human DNA contamination $<20\%$, absence of tumor DNA contamination, and suitability for PCR amplification. An amount of at least $3 \mu\text{g}$ of good quality DNA from saliva for next generation sequencing studies was obtained in 71/102 (69.6%) cases.

The median yield of spot urine DNA was $2.4 \mu\text{g}$ (range $0.1\text{--}34.4 \mu\text{g}$) (Fig. 1A). Purity of urine DNA was acceptable in all cases (median $OD_{260/280}$: 1.8; range: 1.6–2.1). DNA from urine was of high molecular weight in 35/35 (100%) cases (Fig. 1B). The median percentage of human DNA in urine DNA was 96.4% (range 36–100%). Contamination of urine DNA by tumor DNA occurred in 5/35 (14.2%) cases, including 5/30 (16.6%) CLL with overt disease. Contamination was never observed in urine samples collected from CLL ($n=5$) in CR (Fig. 1C). Urine DNA was amplifiable for *KRT1* in all cases as efficiently (median C_T : 23.4; range 22.9–25.8) as DNA from matched PB granulocytes (median C_T : 23.4; range 22.1–23.5) (Fig. 2). An amount of at least 500 ng of good quality urine DNA for SNP array studies was achieved in 29/35 (82.8%) cases. An amount of at least $3 \mu\text{g}$ of good quality DNA for next generation sequencing studies was achieved in 13/35 (37.1%) cases.

Given the low yield of urine DNA after spot sampling, DNA was also extracted after 24 h urine collection in 4 CLL with overt disease. The DNA amount in 24 h urine collections ranged from $39.8 \mu\text{g}$ to $88.5 \mu\text{g}$. DNA was pure, of high molecular weight, not contaminated by tumor DNA, and amplifiable. However, the percentage of human DNA in the total amount of DNA was highly variable (2.2%, 1.5%, 72.3%, 100%), probably reflecting contamination by microbial growth during the 24-h collection.

Suitability of saliva and urine DNA for high throughput genomic studies was assessed by SNP array. DNAs from saliva ($n=5$) and urine ($n=5$) were successfully run on SNP 6.0 arrays, passed the quality control call rate thresholds, and yielded an average call rate (median: 96.5 and 97.42, respectively) appropriate for this analysis.

Overall, our data suggest that DNAs extracted from saliva and spot urine of CLL patients are superimposable in terms of quality. However, in a single spot collection saliva provides a higher amount of germline DNA compared to urine. Urine collection for 24 h, though yielding a higher amount of DNA compared to spot collection, may be contaminated with non-human DNA.

Saliva DNA is in most cases exclusively enriched of human DNA, and in up to 75% of cases it passes quality control filters adopted for genome-wide studies. Overall, these observations document that saliva is a useful source of germline DNA for genome-wide studies applied to cancer patients, and not only hereditary disorders [7]. In addition, given its high yield, saliva might represent an alternative to skin biopsy also for next generation sequencing investigations.

Conflict of interest statement

All authors have no conflict of interest to declare.

Acknowledgments

This study was supported by Progetto FIRB-Programma “Futuro in Ricerca” 2008 and PRIN 2008, MIUR, Rome, Italy; AIRC, Special Program Molecular Clinical Oncology, 5 × 1000, No. 10007, Milan, Italy; Progetto Giovani Ricercatori 2008, Ministero della Salute, Rome, Italy; and Novara-Associazione Italiana contro le Leucemie, Linfomi e Mielomi (AIL) Onlus, Novara; Helmut Horten Foundation, San Salvatore Foundation, Fondazione per la Ricerca e la Cura sui Linfomi (Lugano, CH).

Contributors. S.R. and A.B. contributed equally to this study; S.R., A.B., G.G. and D.R. designed the study, analyzed and interpreted data, and drafted the manuscript; S.C., S.M., and E.G. contributed to data collection; M.F. and L.D.P. provided well characterized biological samples; A.R., I.K. and F.B. performed and interpreted genomic analyses; R.F. contributed to the critical interpretation of results.

References

- [1] Heinrichs S, Li C, Look AT. SNP array analysis in hematologic malignancies: avoiding false discoveries. *Blood* 2010;115:4157–61.
- [2] Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet* 2010;11:685–96.
- [3] Gribben JG, O'Brien S. Update on therapy of chronic lymphocytic leukemia. *J Clin Oncol* 2011;29:544–50.
- [4] Rossi D, Spina V, Cerri M, Rasi S, Deambrogi C, De Paoli L, et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clin Cancer Res* 2009;15:4415–22.
- [5] Mian M, Rinaldi A, Mensah AA, Rossi D, Ladetto M, Forconi F, et al. Del(13q14.3) length matters: an integrated analysis of genomic, FISH and clinical data in 169 chronic lymphocytic leukemia patients with 13q deletion alone or normal karyotype. *Hematol Oncol*; in press.
- [6] Herráez DL, Stoneking M. High fractions of exogenous DNA in human buccal samples reduce the quality of large-scale genotyping. *Anal Biochem* 2008;383:329–31.
- [7] Bahlo M, Stankovich J, Danoy P, Hickey PF, Taylor BV, Browning SR. Australian and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene), Brown MA, Rubio JP. Saliva-derived DNA performs well in large-scale, high-density single-nucleotide polymorphism microarray studies. *Cancer Epidemiol Biomarkers Prev* 2010;19:794–8.

Predictive markers and driving factors behind Richter syndrome development

Expert Rev. Anticancer Ther. 11(3), 433–442 (2011)

Marco Fangazio^{1*},
Lorenzo De Paoli^{1*},
Davide Rossi¹ and
Gianluca Gaidano^{1†}

¹Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy

[†]Author for correspondence:

Tel.: +39 321 660 655

Fax: +39 321 620 421

gaidano@med.unipmn.it

*These authors contributed equally to this manuscript

Transformation of chronic lymphocytic leukemia (CLL) to diffuse large B-cell lymphoma (DLBCL) is known as Richter syndrome (RS). In the entire CLL population, the cumulative prevalence of RS development steadily increases at a rate of 1% per year. Considering conventional predictors of CLL, patient subgroups at high risk of developing RS are characterized by the expression of CD38, absence of del13q14, and a lymph node size >3 cm. Novel risk factors for predicting RS development at CLL diagnosis have been recently identified and include: the host genotype of the CD38 locus and of other genes; telomere length of CLL cells; stereotyped B-cell receptor; and usage of specific immunoglobulin variable genes (*IGHV4-39*). Importantly, although some risk factors predict both CLL progression and transformation to RS, others (CD38 genotype, absence of del13q14, *IGHV4-39* usage, stereotyped B-cell receptor) appear to specifically predict RS. The definition of RS encompasses at least two different conditions: DLBCLs that are clonally related to the pre-existing CLL (accounting for most cases), or DLBCL unrelated to the CLL clone. The transition from CLL to clonally related RS is accompanied by the acquisition of novel genetic alterations that may account for the chemorefractoriness of RS. Genome-wide studies that are currently ongoing are important for identifying novel molecular lesions implicated in RS that might represent a suitable target for future therapeutic strategies.

KEYWORDS: B-cell receptor • CD38 • chronic lymphocytic leukemia • del13q14 • diffuse large B-cell lymphoma • genetic lesions • *IGHV* genes • prognostic factors • Richter syndrome

Definition & frequency of Richter syndrome

Over time, a fraction of chronic lymphocytic leukemia (CLL) patients develop Richter syndrome (RS), representing the clinicopathologic transformation of CLL to an aggressive lymphoma; most commonly diffuse large B-cell lymphoma (DLBCL). Historically, the definition of RS also applies to CLL cases that have developed classical Hodgkin's lymphoma, although this event is very rare. RS is named after Maurice N Richter, who described the syndrome for the first time in the 1920s [1]. Transformation of CLL to RS should be distinguished from other types of CLL progression recognized by the International Workshop on Chronic Lymphocytic Leukemia – National Cancer Institute (IWCLL-NCI) guidelines [2]. Histologic documentation is mandatory for diagnosing RS, which otherwise can only be clinically suspected, but not proven [3,4]. Transformation of CLL to DLBCL must also be distinguished from prolymphocytic transformation, CLL 'acceleration' and the

development of several other B-cell malignancies that may arise at an increased rate in CLL patients [4,5].

In this article, we shall restrict the term RS solely to CLL transformation to DLBCL, since this clinical scenario represents over 95% of all RS cases. Despite this terminological restriction, the clinical definition of RS maintains a certain degree of heterogeneity and may be molecularly distinguished into at least two biologically different entities: transformation of CLL cells to a clonally related DLBCL; and/or the development of a DLBCL that is unrelated to the CLL clone [3,6–11]. Transformation of CLL cells to a clonally related DLBCL is by far the most frequent of the two conditions, accounting for approximately 80–90% of all RS cases. In clonally related RS, the pathogenetic link between CLL cells and the emerging DLBCL clone is apparently obvious, and is substantiated by a gain of novel molecular lesions at a certain time point of the natural evolution of the CLL clone, conceivably corresponding to the time of

clinicopathologic transformation. On the other hand, the mechanisms underlying the development of a DLBCL that is clonally unrelated to CLL cells are not fully understood and may be related, at least in part, to the condition of immune deregulation that frequently characterizes patients affected by CLL.

With the advent of novel immunotherapies for CLL, and especially after the introduction of the anti-CD52 antibody alemtuzumab that causes a profound T-cell depletion, occasional CLL patients have been reported to develop clinically aggressive lymphomas characterized by Epstein–Barr virus (EBV) infection [12,13]. These cases of alemtuzumab-associated aggressive lymphomas are clinically and biologically distinct from RS, and should be considered as a novel type of immunodeficiency-related lymphoma developing after T-cell-depleting therapies in patients already immunocompromised because of the underlying disease and/or because of previous chemotherapy [12,13].

The propensity displayed by CLL toward transformation to DLBCL is also shared by other indolent B-cell lymphoproliferative disorders, namely follicular lymphoma and marginal zone lymphoma. The risk of transformation varies among the different histologies, being highest in the case of follicular lymphoma. Currently, it is unknown whether one or more common molecular pathways recapitulates transformation to DLBCL in all these different clinical settings, or whether transformation from indolent to aggressive B-cell malignancies is due to different genetic and epigenetic lesions depending on the type of initial disease. Differences in the pattern of genetic lesions accompanying transformation of CLL and of follicular lymphoma to DLBCL suggest that the molecular pathways underlying transformation do not fully overlap in these clinical contexts [14].

The incidence of RS varies among available series. In particular, since RS diagnosis requires pathological assessment, heterogeneity in biopsy policies among institutions represents a substantial factor influencing the incidence of RS observed in different CLL series [4]. Notably, whereas re-biopsy is the standard practice at each progression of follicular lymphoma [15], a similar attitude is not commonly followed in the case of CLL, even in patients with substantial enlargement of the lymph nodes. The hematologist's perception of RS as a rare complication of CLL may be due, at least in part, to under-recognition of RS in the current clinical practice.

In a recent report from our group [3], the incidence of RS was defined in a consecutive series of CLL patients who, according to our institution policy, homogeneously underwent biopsy of the index lesion during CLL follow-up or at CLL diagnosis in case of:

- Lymph node size of at least 5 cm;
- Rapid lymph node enlargement (doubling of the largest transversal diameter within a 3-month period);
- Appearance of any suspicious extranodal lesions;
- Appearance of B symptoms;
- Marked elevation of LDH.

These criteria were used for generating the clinical suspicion of RS and for obtaining subsequent confirmation by histology [3]. The strict application of these criteria to a consecutive CLL series revealed that the cumulative incidence of RS at 5 and 10 years exceeds 10 and 15%, respectively [3].

In the same series, the incidence of RS did not markedly increase in parallel with the progressive failure of CLL therapy [3]. This observation may be ascribed to the aggressive biopsy policy adopted, which may favor early recognition of RS before exposing patients to multiple lines of therapy that, being targeted to CLL progression, might be of little efficacy in the instance of transformation to RS.

Richter syndrome is generally regarded as a very late event in the clinical history of CLL. Data from consecutive series, however, document that a significant fraction of RS cases may occur relatively early after CLL diagnosis [3,4]. In some cases, the diagnosis of RS and CLL are concomitant. In one recent study, median time to RS transformation was 23.0 months from CLL diagnosis [3]. In addition, RS was not observed after 82.4 months of follow-up and the rate of transformation remained at 16.2% from that time point onward [3]. These data are consistent with those from other series available in the literature [4] and suggest that, in a fraction of patients, predisposition to RS transformation may be related to biological characteristics that are already intrinsic to the CLL clone at the time of CLL diagnosis. This notion highlights the need to identify patients who are already at risk of RS development at the time of CLL diagnosis.

Diagnosis & prognosis of RS

Classically, RS is pathologically represented by DLBCL. The disease most frequently involves the lymph nodes, but extranodal localizations are also not uncommon and may affect the GI tract, liver, tonsils, skin and bone marrow, among other sites [3,4,9–11,16–20]. In some patients, transformation from CLL to RS does not occur simultaneously at all sites, but at a certain time point might be restricted to a single lesion, either nodal or extranodal. From a practical standpoint, this knowledge mandates that any biopsy aimed at exploring whether RS has occurred should be directed at the index lesion; for example, the lesion displaying the largest diameter by imaging. The PET characteristics of the lesion, in particular the standardized uptake value, may also guide the choice of the site to be biopsied, since sites affected by RS are expected to have standardized uptake values overlapping those of *de novo* DLBCL [21]. The precise value of PET in RS diagnosis and in predicting RS involvement of a given organ, however, is still a matter of investigation [21]. In exceptional cases, when the possibility of a biopsy is prevented by the patient's performance status or by difficult access to the anatomical site of the index lesion, cytological proof of RS by fine-needle aspiration may also be acceptable for reaching a diagnosis of RS.

Richter syndrome cases diagnosed as DLBCL generally consist of confluent sheets of large neoplastic B-lymphocytes, which are clearly distinguishable from the proliferation of small lymphocytes with the scant cytoplasm and clumped chromatin that is typical of CLL. Cases of CLL that lack clear-cut features of

DLBCL, despite presenting with numerous proliferation centers or an increased proportion of prolymphocytes, should not be diagnosed as RS and should not be considered as transformed cases. Finally, RS cases presenting as DLBCL expressing CD5 should be distinguished from *de novo* CD5-positive DLBCL, which represent a molecularly distinct entity [22].

Morphologically, RS may display both immunoblastic and centroblastic features that denote histologic variants of DLBCL occurring *de novo*. Based on the Hans algorithm utilized for immunohistochemical classification of *de novo* DLBCL [23], most (~90%) cases of RS express interferon regulatory factor (IRF)4 and therefore belong to the postgerminal center phenotype, whereas only 10% of RS display the germinal center profile characterized by CD10 and/or BCL6 positivity in the absence of IRF4 [11]. The proportion of RS cases with a germinal center phenotype versus that of cases with a postgerminal center phenotype differs substantially from that detected in *de novo* DLBCL [11,23]. It should be noted, however, that the Hans algorithm was developed for *de novo* DLBCL and that the immunohistochemistry markers that it exploits may not be appropriate for fully revealing the biological heterogeneity of RS [23]. Gene-expression profiling studies of RS cases are currently lacking, but may allow a proper understanding of the histogenesis of RS.

Although the CD5 and CD23 antigens are invariably expressed in CLL cells, their expression is frequently lost at the time of transformation of CLL to DLBCL [11]. CD20 is generally expressed by RS cells and represents an important target for immunotherapy with anti-CD20 monoclonal antibodies, including rituximab [11].

Prognosis of RS is generally considered highly unfavorable. However, a study of 148 patients with biopsy or fine-needle aspiration-proven RS has documented that survival of RS is not uniform among patients, ranging from a few weeks to 15 years, and may be predicted on clinical grounds by the RS score [24]. The RS score is not intended for predicting RS development in CLL patients, but rather predicts RS prognosis once that transformation has occurred [24]. The RS score is a model predicting an individual patient's risk of death and is based on five adverse factors predicting short survival, namely:

- Zubrod performance status greater than 1;
- Elevated LDH level;
- Platelet count no greater than $100 \times 10^9/l$;
- Tumor size of at least 5 cm;
- Two or more prior lines of therapy [24].

Patients are assigned to one of four risk groups based on the number of presenting risk factors: 0 or 1, low risk; 2, low-intermediate risk; 3, high-intermediate risk; 4 or 5, high risk. In the original report of the RS score, survival was 1.12 years in patients with RS score 0–1; 0.90 years in RS score 2; 0.33 years in RS score 3; and 0.14 years in RS score 4–5 [24]. Remarkably, risk factors that are relevant to the International Prognostic Index, namely number of extranodal sites of disease, age and stage, were not relevant to the RS score, confirming the notion that RS and *de novo* DLBCL are very different diseases. By contrast, survival prognostication in

RS takes advantage of parameters reflecting marrow failure, such as thrombocytopenia, and parameters reflecting immune system exhaustion and selective pressure to chemorefractory clones, such as number of prior lines of treatment.

Very recent data from our group indicate that, in addition to the clinical risk factors included in the RS score, survival post-transformation of RS may also be predicted by tumor genotype and by the clonal relationship of the RS phase with the pre-existing CLL clone [25].

Clinical & biological tools for predicting RS development

Early recognition of RS may be clinically useful in order to avoid the exposure of patients to multiple lines of therapy that, being targeted to CLL progression, are of little efficacy on the transformed clone. This notion prompts the need for the close monitoring of CLL patients harboring clinical and/or biological risk factors of RS development. Recent advances in the field have disclosed a number of risk factors, both clinical and biological, which might be helpful for RS prediction [3]. It is remarkable that several of the predictors identified to date are specific for CLL transformation to RS, and are not related to other markers predicting CLL clinical progression according to IWCLL-NCI guidelines [2,3]. According to the data, the existence of predictors for CLL transformation to RS documents that RS transformation and CLL progression without histological transformation are distinct events both clinically and biologically, which require a different, although complementary, diagnostic approach for a comprehensive assessment of the risk category of the individual patient.

The biological risk factors that have been tested for their ability to predict RS include: immunoglobulin heavy-chain variable gene (*IGHV*) homology >98%, *IGHV* gene usage, stereotyped B-cell receptor (BCR), FISH karyotype, *TP53* status, *BCL2* genotype, 14q32 translocations, telomere length, and expression of CD38, ZAP70 and CD49 (FIGURE 1; TABLE 1) [3,26–29]. Of these, the following have been scored as independent risk factors for RS development: absence of del13q14; usage of specific *IGHV* genes, namely *IGHV4-39*; usage of a stereotyped BCR; expression of CD38 (FIGURE 2); genetic background of the host and telomere length [3,26–29].

Among clinical predictors, a lymph node size greater than 3 cm, which is a parameter of solid disease, was the sole clinical risk factor for RS selected by two models of multivariate analysis [3]. Conversely, parameters of leukemic disease, namely lymphocyte count, advanced Rai stage, splenomegaly, and percentage and pattern of bone marrow involvement, were not scored as independent clinical risk factors for RS [3]. As expected, parameters of leukemic disease did predict for short time to progression without transformation, further reinforcing the notion that RS transformation and CLL progression according to NCI guidelines are distinct events in the natural history of CLL (TABLE 1) [3].

As stated earlier, FISH karyotype at CLL diagnosis predicts RS. In particular, the absence of del13q14 confers an independent risk of RS development [3]. Conceivably, pathogenetic differences between CLL with and without del13q14 underlie the different predisposition of these two CLL subsets toward RS transformation.

Table 1. Predictors of chronic lymphocytic leukemia progression and/or of Richter syndrome transformation at the time of chronic lymphocytic leukemia diagnosis.

CLL characteristics	Progression	Transformation
<i>Biological features at diagnosis</i>		
Telomere length	Yes	Yes
CD38 $\geq 30\%$	Yes	Yes
del17p13	Yes	No
del11q22–q23	Yes	No
IGHV homology $\geq 98\%$	Yes	No
Absence of del13q14	No	Yes
IGHV4-39 usage	No	Yes
Stereotyped VH CDR3	No	Yes
CD38 genotype	No	Yes
LRP4 genotype	No	Yes
<i>Clinical features at diagnosis</i>		
Largest node ≥ 3 cm	Yes	Yes
Rai stage III–IV	Yes	No
Splenomegaly	Yes	No

CLL: Chronic lymphocytic leukemia; IGHV: Immunoglobulin variable gene; VH CDR3: Heavy chain complementarity-determining region 3.
Data taken from [3,28,29,50,51].

In fact, del13q14 is thought to play an important role in disease pathogenesis and its presence activates molecular pathways that are distinct from those of CLL without del13q14 [30–32]. The concept that propensity to DLBCL transformation may be affected by the genetic profile of the pre-existing indolent phase is not unique to CLL among B-cell malignancies. In fact, transformation to DLBCL of at least another indolent B-cell disorder, for example MALT lymphoma, has been documented to depend upon differences in disease pathogenesis of the pre-existing indolent B-cell clone [33]. Other FISH abnormalities of CLL, including del17p13, del11q23, +12 or 14q32 translocations, do not appear to be associated with an increased risk of RS [3,34].

A sizeable fraction of CLLs (30%) show a high degree of structural similarity in the BCR because of homology of the expressed heavy chain complementarity determining region 3 (VH CDR3) [35–38]. This phenomenon is known as stereotyped BCR, and strongly suggests the role of antigen stimulation in disease development [39]. Recently, we have shown that carrying a stereotyped BCR at CLL diagnosis may help in the identification of patients at risk of RS transformation [28]. Comparing the immunogenetic features of 69 cases of RS with those of 714 cases of CLL, the prevalence of stereotyped VH CDR3 was significantly higher in RS than in nontransformed CLL (49.3 vs 21.3%) [28]. Of note, applying an actuarial approach to this cohort, stereotyped VH CDR3 at the time of CLL diagnosis was revealed to be an independent predictor of RS transformation [28]. The risk of RS transformation carried by stereotyped BCR was independent of the mutation status of *IGHV* genes [28].

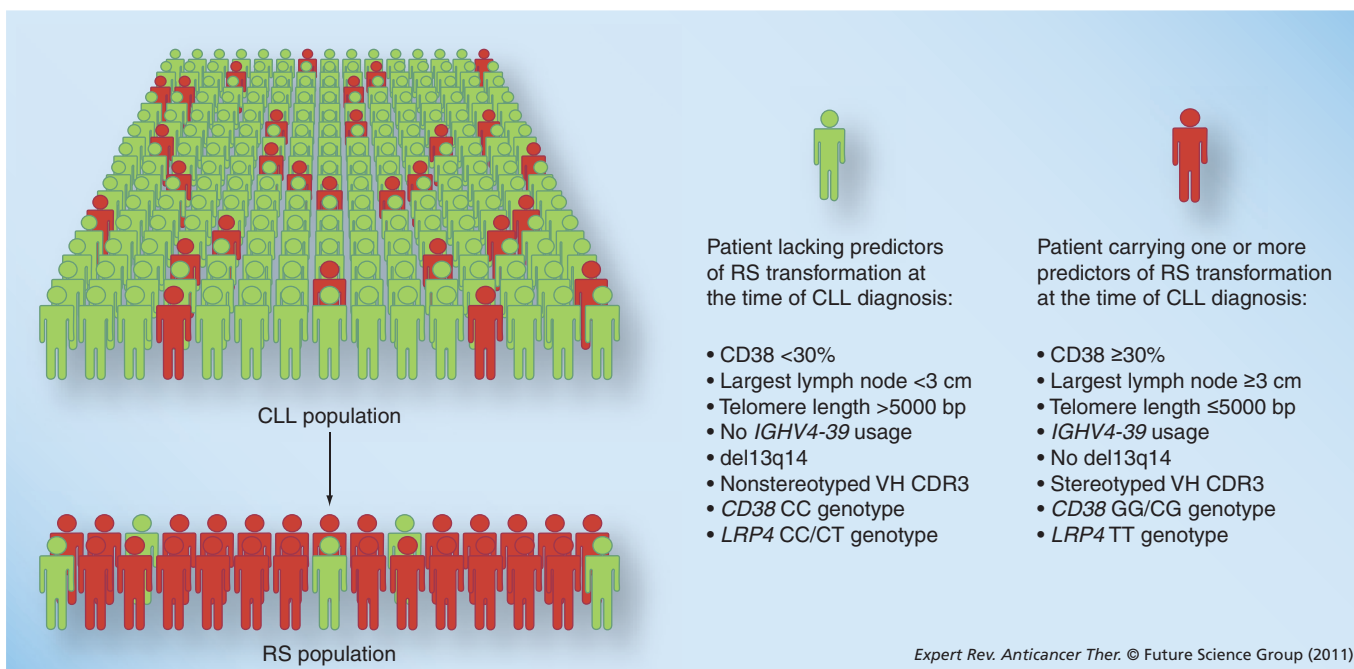


Figure 1. Clinical and molecular risk factors that, at the time of chronic lymphocytic leukemia diagnosis, predict Richter syndrome transformation. At the time of CLL diagnosis, several risk factors predict an increased risk of RS transformation, including: expression of *CD38*; lymph node size; telomere length; absence of del13q14; immunogenetic features of the CLL clone; GG/CG genotypes of the *CD38* rs6449182 SNP; and TT genotype of the *LRP4* rs2306029 SNP.
CLL: Chronic lymphocytic leukemia; IGHV: Immunoglobulin variable gene; RS: Richter syndrome; VH CDR3: Heavy chain complementarity-determining region 3.

Among CLLs carrying a stereotyped BCR, patients utilizing a specific *IGHV* gene, for example *IGHV4-39*, in a stereotyped fashion (so-called BCR 'subset 8') had the highest risk of RS transformation (17-fold higher than cases without a stereotyped BCR; 5-year risk: 68 vs 4%) [28]. In multivariate analysis, usage of *IGHV4-39* along with stereotyped VH CDR3 was an independent risk factor of RS development [28]. This observation, beside being of clinical interest, points to the role of antigen stimulation in RS development. The role of antigen stimulation in favoring RS development is further sustained by the observation that the BCR expressed by CLLs belonging to subset 8 is characterized by the highest affinity for non-muscle myosin II, an autoantigen that might be involved in CLL promotion [40].

Both CD38 expression and lymph node size have been shown to be independent predictors of RS [3]. From a pathogenetic standpoint, this observation suggests that molecular circuits in CLL lymph nodes may be important for determining the risk of RS in individual patients and may directly contribute to CLL transformation to DLBCL. In fact, a strict association between CD38 expression and lymph node enlargement in CLL is suggested by several lines of evidence [41–44]. First, CD38 expression is higher in CLL cells of lymph nodes than in CLL cells of the peripheral blood or bone marrow [41–43]. Second, CD38 expression is required for CLL trafficking toward lymphoid tissues [44]. Third, CD38 expression associates with predominantly nodal disease [41–43]. In addition, higher levels of CD38 expression in CLL lymph nodes associate with lymphoma-like tumor distribution and diffuse bone marrow infiltration pattern [43]. These same features have been reported to predict RS [3]. On this basis, the lymph node may thus provide an optimal microenvironment for proliferation and blastic transformation of CD38-expressing CLL cells. In fact, interactions with other cells and cytokines that take place in the lymph node may provide adequate stimulation for CD38 upregulation and signaling in CLL cells [45–47]. Intriguingly, *in vitro* activation of CD38 signaling induces transformation of CLL cells to plasmablasts that are reminiscent of the blasts appearing during RS evolution [45].

Expression of *CD38* is regulated at multiple levels [45]. The 5' end of intron 1 of the *CD38* gene is involved in the induction of *CD38* expression by transcription factors in myeloid cells [48]. Within the same region maps a well-characterized single nucleotide polymorphism (SNP), leading to a C>G variation at position 184 (*CD38* rs6449182) [49]. The presence of the minor G allele, either in a heterozygous or a homozygous condition, strongly associates with RS [50]. Compared with CC homozygotes, GG homozygous CLL patients had a 30.6% increase in the relative risk of developing RS, whereas GC heterozygotes showed an

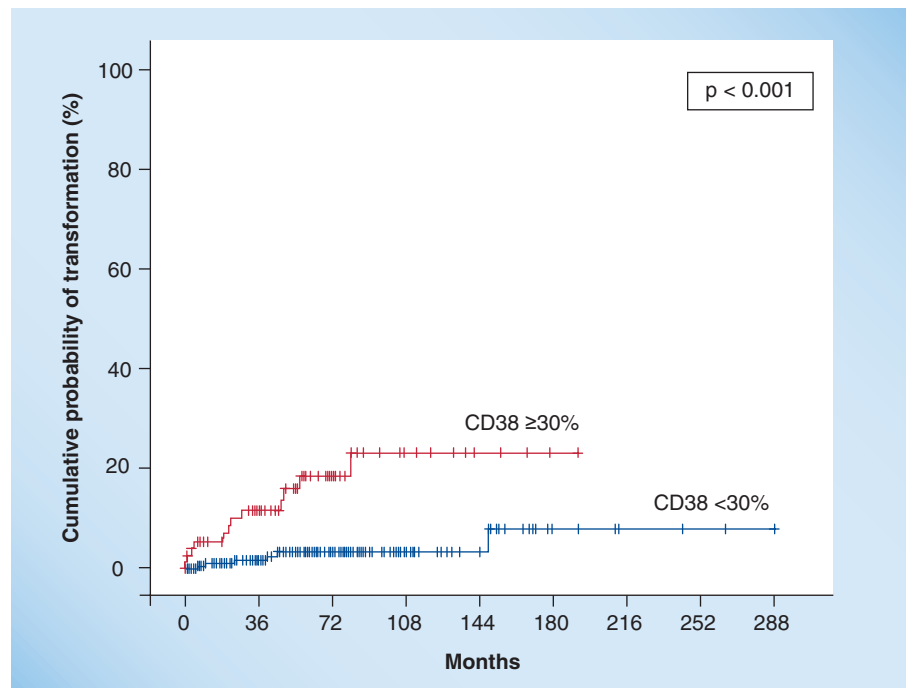


Figure 2. Prediction of Richter syndrome transformation according to CD38 expression at chronic lymphocytic leukemia diagnosis. Kaplan–Meier curves showing the value of CD38 in predicting Richter syndrome transformation in the chronic lymphocytic leukemia series (n = 360). Adapted from the Division of Hematology of the Amedeo Avogadro University of Eastern Piedmont (years: 2000–2010).

intermediate probability of 12.4% [50]. At 5 years, GG homozygotes demonstrated an increase in probability of developing RS that approximately doubled that associated with GC heterozygotes [50]. These observations document that the host genetic background may be of relevance in predicting RS development among individuals affected by CLL [50].

The impact of *CD38* SNPs on the risk of RS has prompted further studies aimed at defining the role of the host genetic background in predisposing CLL patients to RS development. Analysis of a large number of SNPs has revealed that the genetic background of the *LRP4* gene (rs2306029) is also an important predictor of RS transformation [51]. In particular, patients that are homozygotes for the minor allele of *LRP4* rs2306029 display a significantly higher 5-year cumulative probability of transformation to aggressive lymphoma compared with patients carrying the CT/CC genotypes that contain the wild-type allele [51]. Notably, *LRP4* encodes for a protein involved in the Wnt/β-catenin signaling pathway that is known to be activated in CLL cells [52]. The fact that alterations of the *LRP4* protein are pathogenetic in a clinical context other than CLL, points to a possible role of the interaction between *LRP4* and Wnt/β-catenin also in RS pathogenesis [53].

When assessed at the time of diagnosis, the telomere length of CLL cells has also been shown to be an independent predictor of RS transformation [29]. As already demonstrated for other neoplasms, telomere length recapitulates the cell proliferative history and short telomere length – denoting an aggressive clinical phenotype and chemorefractoriness [54].

Table 2. Risk of Richter syndrome transformation at 5 years according to different Richter syndrome biological predictors.

RS predictor	5-year risk of RS transformation (%)	Ref.
Stereotyped VH CDR3	14.4	[28]
<i>IGHV4-39</i> usage	35.4	[28]
Telomere length <5000 bp	18.9	[29]
<i>CD38</i> GG/CG genotype	14.2	[50]
<i>LRP4</i> TT genotype + stereotyped VH CDR3	30.4	[51]

RS: Richter syndrome; VH CDR3: Heavy chain complementarity-determining region 3.

All studies on RS predictors that are available to date did not differentiate between clonally related and clonally unrelated cases [3,28,29,50,51]. At present, therefore, it is not possible to define whether the RS predictors identified until now are useful only for predicting clonally related RS, which represent the majority of cases in all series investigated, or whether they might also be informative for prediction of clonally unrelated RS. Future studies are needed to address this issue.

Another unresolved issue concerns the relative weight associated with each predictor of RS that has been identified. Since RS predictors emerged from different studies that included different series of patients, it is not currently possible to define the contribution of each predictor to the risk of RS development. Available evidence, however, suggests that different predictors may vary in relevance in determining the risk of RS (TABLE 2).

Molecular pathogenesis of RS

The molecular mechanisms leading to the transformation of CLL to RS are largely unknown. In classic RS represented by DLBCL, immunogenotypic studies have revealed that the majority (~80%), although not the totality, of RS cases are clonally related to the pre-existing CLL clones [3,4,6–11]. Cases of RS that are clonally related to the pre-existing CLL phase conceivably stem from the progressive accumulation of molecular lesions, either genetic and/or epigenetic, that favor the emergence and predominance of a large cell population causing clinical aggressiveness.

In all available RS series, a fraction (10–20%) of RS cases are unrelated to the pre-existing CLL clone [3,4,6–11,25]. The precise pathogenesis of clonally unrelated RS is unclear and requires extensive investigation. Accumulation of molecular lesions in the CLL clone cannot be considered a relevant factor in the development of clonally unrelated RS. Rather, these RS cases may be favored by alterations of the host genetic background and immunologic function, or by derangements in the lymph node micro-environment induced by CLL cells. Emergence of DLBCLs that are clonally unrelated to CLLs may be part of a more generalized phenomenon that, for reasons that are currently obscure, render CLL patients prone to developing a second lymphoid malignancy or, more generally, a second cancer [5].

Scant knowledge is currently available concerning the genetic lesions that associate with transformation from CLL to clonally related DLBCL. Acquisition of *TP53* mutations and/or 17p13 deletion is a very frequent molecular alteration in clonally related RS [3,25,55,56]. Interestingly, acquisition of *TP53* disruption by mutation and/or deletion is also common in other types of transformation from indolent to aggressive B-cell malignancies [57,58], suggesting that this molecular pathway is a general feature of clinical progression of mature B-cell neoplasia. Curiously, despite their involvement in RS pathogenesis, *TP53* mutations at CLL diagnosis do not increase the subsequent risk of RS development [3].

In *de novo* DLBCLs arising in different clinical contexts, aberrant somatic hypermutation (ASHM) of the *c-MYC*, *RhoH/ITF*, *PAX5* and *PIM1* proto-oncogenes represents the most frequent genetic lesion currently known [59–61]. ASHM in DLBCL is thought to be caused by malfunctioning of activation-induced cytidine deaminase (AID), an enzyme that is essential for physiological somatic hypermutation of normal B cells [62,63]. Although AID levels increase at the time of CLL transformation to RS [9,64], investigations of ASHM in RS have yielded controversial results [14,64]. The apparently low incidence of ASHM in RS distinguishes CLL transformation to DLBCL from follicular lymphoma transformation to DLBCL, which, unlike RS, is frequently accompanied by a high load of ASHM occurring at the time of histologic transformation [14].

Cytogenetic studies have not revealed recurrent abnormalities selectively clustering with RS [65–69]. The report of trisomy 12 in RS is not disease specific, since trisomy 12 occurs in several other B-cell malignancies, including CLL and several types of lymphoma [70]. Consistent with the role of absence of del13q14 as a risk factor for RS, loss of del13q14 has been reported in RS and may be due to the emergence of a subclone lacking del13q14 [3]. By contrast with other types of B-cell lymphoma, 14q32 translocations involving immunoglobulin genes do not contribute to the molecular mechanisms that are operative at the time of transformation from CLL to DLBCL [34]. Consistently, translocations of *BCL1*, *BCL2* and *BCL6* are not found in RS [3]. Lack of involvement of 14q32 translocations in RS development suggests that RS pathogenesis differs from that of other aggressive B-cell non-Hodgkin's lymphomas, including *de novo* DLBCL, in which proto-oncogene deregulation by juxtaposition to immunoglobulin heavy or light chain loci is a relatively common mechanism [34,71]. Molecular cytogenetic studies have shown that RS associates with a higher degree of genomic complexity compared with CLL, although none of the genetic lesions reported to date appear to be specific for RS [72,73].

The role of EBV infection has been suggested by some studies as a potentially relevant factor for RS pathogenesis. The observation that the overwhelming majority of RS do not carry EBV infection in the malignant cells, however, does not favor this hypothesis [3]. The presence of EBV sequences has been documented in some, although not all, cases of RS originating in patients previously treated with fludarabine for their pre-existent CLL [74,75]. EBV infection in these cases has been thought to be related to the immune deregulation caused by purine analogues and/or alemtuzumab.

As detailed earlier, available studies of RS pathogenesis indicate that RS development may not be easily explained by genetic lesions identified in other B-cell disorders, namely CLL and *de novo* DLBCL. This lack of knowledge points to the need for novel studies aimed at clarifying the molecular basis of RS. Genome-wide studies that are currently in progress will hopefully reveal novel molecular alterations and gene-expression abnormalities involved in RS development. Elucidation of the genetic complexity of RS might also be useful for the future design of modalities of targeted therapy against this disorder.

Expert commentary

Advancements during the last few years have revitalized the interest on RS and on our ability to predict the development of this clinical condition. A limited but growing number of molecular and phenotypic markers have been identified that may facilitate the identification of CLL patients who are at risk of RS development at some point during the natural history of their disease [3,28,29,50,51]. These biological predictors of RS include the usage of specific *IGHV* genes and of peculiar subsets of stereotyped BCR, expression and genotype of *CD38*, genotype of *LRP4*, pattern of cytogenetic alterations detectable by FISH, and telomere length [3,28,29,50,51]. Of note, transformation to RS and progression of CLL without transformation are distinct clinical entities with different predictors (TABLE 1) [3,28,29,39,50,51].

The availability of biological predictors of RS transformation suggests the adoption of a close surveillance and of a dedicated biopsy policy for patients who are judged to be at high risk. A close surveillance and an aggressive biopsy policy may anticipate the timing of RS recognition, which is clinically relevant since late diagnosis is an adverse factor in post-transformation RS prognosis [24]. In the individual RS patient, the status of some RS predictors can be obtained from diagnostic procedures that have entered the clinical practice of CLL patients at the time of diagnosis. The assessment of other RS predictors, however, requires the use of more refined techniques that are available only in highly specialized laboratories [28]. Genome-wide technologies may purport novel and specific markers of

RS development that associate with a high predictive value for CLL patients, and may refine or complement available tools for RS prediction.

Five-year view

Despite advancements in the understanding of the disease and in the identification of clinical and biological predictors, many questions remain unanswered in RS. First, although the term RS currently applies to DLBCLs that are both clonally related and clonally unrelated to the pre-existing CLL phase [3,4], it may very well be that the two conditions are distinct disorders with different prognosis and treatment requirements. Preliminary data from our group do suggest that this scenario is likely [25]. Second, the lack of knowledge of RS molecular lesions, both genetic and epigenetic, that are causative in determining the transformation from CLL to RS, hampers the possibility of targeted therapies against RS, and prompts genome-wide studies aimed at elucidating RS pathogenesis. Finally, the issue of early diagnosis of RS in CLL patients needs to be expanded and refined. This task requires the consolidation of RS risk factors in CLL patients ideally in a prospective fashion and the definition of noninvasive diagnostic tools for generating a suspicion of RS with high sensitivity and specificity. The design of a diagnostic algorithm, utilizing biological and clinical risk factors of RS development coupled with nuclear imaging of suspected RS lesions, would be highly desirable for offering CLL patients a tool that allows recognition of RS transformation at an early stage.

Financial & competing interests disclosure

Work by the authors of this article was supported by PRIN 2008 (Rome, Italy); AIRC 5 × 1000 (No. 10007; Milan, Italy); FIRB Progetto Futuro in Ricerca and PRIN 2008 (MIUR, Rome, Italy); Progetto Giovani Ricercatori and Ricerca Finalizzata, Ministero della Salute (Rome, Italy); Ricerca Sanitaria Finalizzata, Regione Piemonte (Torino, Italy); and by Novara-AIL Onlus. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Key issues

- During the clinical history of chronic lymphocytic leukemia (CLL), the cumulative incidence of Richter syndrome (RS) exceeds 15% at 10 years from diagnosis.
- RS tends to be a relatively early event during CLL clinical history (median time to transformation: 23.0 months from CLL diagnosis).
- The diagnosis of RS requires histological proof.
- Most cases of RS are represented by diffuse large B-cell lymphomas that are clonally related to the CLL phase.
- At the time of CLL diagnosis, biological risk factors of CLL that are independent predictors of RS transformation include: expression of *CD38*; absence of *del13q14*; telomere length; stereotyped B-cell receptor; *IGHV4-39* usage; and host genotype of *CD38* and *LRP4*.
- At the time of CLL diagnosis, clinical risk factors of CLL that are independent predictors of Richter transformation include lymph node size of at least 3 cm.
- Although some risk factors predict both CLL progression and transformation to RS, others (absence of *del13q14*; *IGHV4-39* usage; stereotyped B-cell receptor; *CD38* genotype) appear to selectively predict RS.
- The transition from CLL to clonally related RS is accompanied by the acquisition of novel genetic alterations that might account for the disease chemorefractoriness.
- RS should be distinguished from cases of aggressive lymphomas that may occur in CLL patients after therapies inducing markedly severe T-cell depletion and that represent a type of immunodeficiency-related lymphoma.

References

Papers of special note have been highlighted as:

• of interest

•• of considerable interest

- 1 Richter MN. Generalized reticular cell sarcoma of lymph nodes associated with lymphatic leukaemia. *Am. J. Pathol.* 4(4), 285–292 (1928).
- 2 Hallek M, Cheson BD, Catovsky D *et al.* International Workshop on Chronic Lymphocytic Leukemia. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute Working Group 1996 guidelines. *Blood* 111(12), 5446–5456 (2008).
- 3 Rossi D, Cerri M, Capello D, Deambrogi C *et al.* Biological and clinical risk factors of chronic lymphocytic leukaemia transformation to Richter syndrome. *Br. J. Haematol.* 142(2), 202–215 (2008).
- **Retrospective study identifying clinical and biological predictors of Richter syndrome (RS) transformation at the time of chronic lymphocytic leukemia (CLL) diagnosis.**
- 4 Tsimberidou A-M, Keating MJ. Richter syndrome. Biology, incidence, and therapeutic strategies. *Cancer* 103(2), 216–228 (2005).
- **Comprehensive review of the clinical and biological features of RS.**
- 5 Maddocks-Christianson K, Slager SL, Zent CS *et al.* Risk factors for development of a second malignancy in patients with chronic lymphocytic leukaemia. *Br. J. Haematol.* 139(3), 398–404 (2007).
- 6 Cherepakhin V, Baird SM, Meisenholder GW, Kipps TJ. Common clonal origin of chronic lymphocytic leukemia and high grade lymphoma of Richter's syndrome. *Blood* 82(10), 3141–3147 (1993).
- 7 Matolcsy A, Inghirami G, Knowles DM. Molecular genetic demonstration of the diverse evolution of Richter's syndrome (chronic lymphocytic leukemia and subsequent large cell lymphoma). *Blood* 83(5), 1363–1372 (1994).
- 8 Nakamura N, Kuze T, Hashimoto Y *et al.* Analysis of the immunoglobulin heavy chain gene of secondary diffuse large B-cell lymphoma that subsequently developed in four cases with B-cell chronic lymphocytic leukemia or lymphoplasmacytoid lymphoma (Richter syndrome). *Pathol. Int.* 50(8), 636–643 (2000).
- 9 Timár B, Fülöp Z, Csernus B *et al.* Relationship between the mutational status of VH genes and pathogenesis of diffuse large B-cell lymphoma in Richter's syndrome. *Leukemia* 18(2), 326–330 (2004).
- 10 Smit LA, van Maldegem F, Langerak AW *et al.* Antigen receptors and somatic hypermutation in B-cell chronic lymphocytic leukemia with Richter's transformation. *Haematologica* 91(7), 903–911 (2006).
- 11 Mao Z, Quintanilla-Martinez L, Raffeld M *et al.* IgVH mutational status and clonality analysis of Richter's transformation: diffuse large B-cell lymphoma and Hodgkin lymphoma in association with B-cell chronic lymphocytic leukemia (B-CLL) represent 2 different pathways of disease evolution. *Am. J. Surg. Pathol.* 31(10), 1605–1614 (2007).
- 12 O'Brien SM, Kantarjian HM, Thomas DA *et al.* Alemtuzumab as treatment for residual disease after chemotherapy in patients with chronic lymphocytic leukemia. *Cancer* 98(12), 2657–2663 (2003).
- 13 Janssens A, Berth M, De Paepe P *et al.* EBV negative Richter's syndrome from a coexistent clone after salvage treatment with alemtuzumab in a CLL patient. *Am. J. Hematol.* 81(9), 706–712 (2006).
- 14 Rossi D, Berra E, Cerri M *et al.* Aberrant somatic hypermutation in transformation of follicular lymphoma and chronic lymphocytic leukemia to diffuse large B-cell lymphoma. *Haematologica* 91(10), 1405–1409 (2006).
- 15 Montoto S, Davies AJ, Matthews J *et al.* Risk and clinical implications of transformation of follicular lymphoma to diffuse large B-cell lymphoma. *J. Clin. Oncol.* 25(17), 2426–2433 (2007).
- 16 Armitage JO, Dick FR, Corder MP. Diffuse histiocytic lymphoma complicating chronic lymphocytic leukemia. *Cancer* 41(2), 422–427 (1978).
- 17 Ott MM, Ott G, Roblick U *et al.* Localized gastric non-Hodgkin's lymphoma of high-grade malignancy in patients with pre-existing chronic lymphocytic leukemia or immunocytoma. *Leukemia* 9(4), 609–614 (1995).
- 18 Ratnavel RC, Dunn-Walters DK, Boursier L *et al.* B-cell lymphoma associated with chronic lymphocytic leukaemia: two cases with contrasting aggressive and indolent behaviour. *Br. J. Dermatol.* 140(4), 708–714 (1999).
- 19 Parrens M, Sawan B, Dubus P *et al.* Primary digestive Richter's syndrome. *Mod. Pathol.* 14(5), 452–457 (2001).
- 20 Omoti CE, Omoti AE. Richter syndrome: a review of clinical, ocular, neurological and other manifestations. *Br. J. Haematol.* 142(5), 709–716 (2008).
- 21 Bruzzi JF, Macapinlac H, Tsimberidou AM *et al.* Detection of Richter's transformation of chronic lymphocytic leukemia by PET/CT. *J. Nucl. Med.* 47(8), 1267–1273 (2006).
- **Pilot investigation of the sensitivity and specificity of ¹⁸FDG-PET in diagnosing RS transformation in CLL patients.**
- 22 Matolcsy A, Chadburn A, Knowles DM. *De novo* CD5-positive and Richter's syndrome-associated diffuse large B cell lymphomas are genotypically distinct. *Am. J. Pathol.* 147(1), 207–216 (1995).
- 23 Hans CP, Weisenburger DD, Greiner TC *et al.* Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* 103(1), 275–282 (2004).
- 24 Tsimberidou A-M, O'Brien S, Khouri I *et al.* Clinical outcomes and prognostic factors in patients with Richter's syndrome treated with chemotherapy or chemoimmunotherapy with or without stem-cell transplantation. *J. Clin. Oncol.* 24(15), 2343–2351 (2006).
- **Defines the RS score that allows prediction of the risk of death in an individual patient affected by RS.**
- 25 Rossi D, Spina V, Deambrogi C *et al.* The genetics of Richter syndrome identifies disease heterogeneity and is an independent predictor of survival post transformation. Presented at: *15th Congress of the European Hematology Association*. Barcelona, Spain, 10–13 June 2010.
- 26 Rossi D, Rasi S, Capello D, Gaidano G. Prognostic assessment of BCL2–938C>A polymorphism in chronic lymphocytic leukemia. *Blood* 111(1), 466–468 (2008).
- 27 Rossi D, Zucchetto A, Rossi FM *et al.* CD49d expression is an independent risk factor of progressive disease in early stage chronic lymphocytic leukemia. *Haematologica* 93(10), 1575–1579 (2008).
- 28 Rossi D, Spina V, Cerri M *et al.* Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clin. Cancer Res.* 15(13), 4415–4422 (2009).

- **Large study documenting the relevance of stereotyped B-cell receptor for predicting RS development at the time of CLL diagnosis.**
- 29 Rossi D, Lobetti Bodoni C, Genuardi E *et al.* Telomere length is an independent predictor of survival, treatment requirement and Richter's syndrome transformation in chronic lymphocytic leukemia. *Leukemia* 23(6), 1062–1072 (2009).
- **Comprehensive analysis of telomere length as a predictor of CLL outcome, including transformation to RS.**
- 30 Calin GA, Ferracin M, Cimmino A *et al.* A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.* 353(17), 1793–1801 (2005).
- 31 Kienle DL, Korz C, Hosch B *et al.* Evidence for distinct pathomechanisms in genetic subgroups of chronic lymphocytic leukemia revealed by quantitative expression analysis of cell cycle, activation, and apoptosis-associated genes. *J. Clin. Oncol.* 23(16), 3780–3792 (2005).
- 32 Klein U, Lia M, Crespo M *et al.* The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell* 17(1), 28–40 (2010).
- **Elegant investigation formally documenting the role of del13q14 in the development of CLL.**
- 33 Starostik P, Patzner J, Greiner A *et al.* Gastric marginal zone B-cell lymphomas of MALT type develop along 2 distinct pathogenetic pathways. *Blood* 99(1), 3–9 (2002).
- 34 Deambrogi C, Cresta S, Cerri M *et al.* 14q32 translocations and risk of Richter's transformation in chronic lymphocytic leukaemia. *Br. J. Haematol.* 144(1), 131–133 (2009).
- 35 Stamatopoulos K, Belessi C, Moreno C *et al.* Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: pathogenetic implications and clinical correlations. *Blood* 109(1), 259–270 (2007).
- 36 Bomben R, Dal Bo M, Capello D *et al.* Molecular and clinical features of chronic lymphocytic leukaemia with stereotyped B cell receptors: results from an Italian multicenter study. *Br. J. Haematol.* 144(4), 492–506 (2009).
- 37 Sutton LA, Kostareli E, Hadzidimitriou A *et al.* Extensive intraclonal diversification in a subgroup of chronic lymphocytic leukemia patients with stereotyped IGHV4–34 receptors: implications for ongoing interactions with antigen. *Blood* 114(20), 4460–4468 (2009).
- 38 Darzentas N, Hadzidimitriou A, Murray F *et al.* A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. *Leukemia* 24(1), 125–132 (2010).
- 39 Caligaris-Cappio F, Ghia P. Novel insights in chronic lymphocytic leukemia: are we getting closer to understanding the pathogenesis of the disease? *J. Clin. Oncol.* 26(27), 4497–4503 (2008).
- 40 Chu CC, CATERA R, Zhang L *et al.* Many chronic lymphocytic leukemia antibodies recognize apoptotic cells with exposed nonmuscle myosin heavy chain IIA: implications for patient outcome and cell of origin. *Blood* 115(19), 3907–3915 (2010).
- 41 Del Poeta G, Maurillo L, Venditti A *et al.* Clinical significance of CD38 expression in chronic lymphocytic leukemia. *Blood* 98(9), 2633–2639 (2001).
- 42 Ibrahim S, Keating M, Do KA *et al.* CD38 expression as an important prognostic factor in B-cell chronic lymphocytic leukemia. *Blood* 98(1), 181–186 (2001).
- 43 Jaksic O, Paro MM, Kardum Skelin I, Kusec R, Pejisa V, Jaksic B. CD38 on B-cell chronic lymphocytic leukemia cells has higher expression in lymph nodes than in peripheral blood or bone marrow. *Blood* 103(5), 1968–1969 (2004).
- 44 Deaglio S, Vaisitti T, Aydin S *et al.* CD38 and ZAP-70 are functionally linked and mark CLL cells with high migratory potential. *Blood* 110(12), 4012–4021 (2007).
- 45 Deaglio S, Capobianco A, Bergui L *et al.* CD38 is a signaling molecule in B-cell chronic lymphocytic leukemia cells. *Blood* 102(6), 2146–2155 (2003).
- 46 Deaglio S, Vaisitti T, Aydin S, Ferrero E, Malavasi F. In-tandem insights from basic science combined with clinical research: CD38 as both marker and key component of the pathogenetic network underlying chronic lymphocytic leukemia. *Blood* 108(4), 1135–1144 (2006).
- 47 Patten PE, Buggins AG, Richards J *et al.* CD38 expression in chronic lymphocytic leukemia is regulated by the tumor microenvironment. *Blood* 111(10), 5173–5181 (2008).
- 48 Kishimoto H, Hoshino S, Ohori M *et al.* Molecular mechanisms of human CD38 gene expression by retinoic acid. Identification of retinoic acid response element in the first intron. *J. Biol. Chem.* 273(25), 15429–15434 (1998).
- 49 Ferrero E, Saccucci F, Malavasi F. The human CD38 gene: polymorphism, CpG island, and linkage to the CD157 (BST-1) gene. *Immunogenetics* 49(7–8), 597–604 (1999).
- 50 Aydin S, Rossi D, Bergui L *et al.* CD38 gene polymorphism and chronic lymphocytic leukemia: a role in Richter syndrome? *Blood* 111(12), 5646–5653 (2008).
- **Documenting the relevance of the host genotype as a risk factor for RS development.**
- 51 Rasi S, Spina V, Brusca A *et al.* A variant of the LRP4 gene affects the risk of chronic lymphocytic leukaemia transformation to Richter syndrome. *Br. J. Haematol.* 152(3), 284–294 (2011).
- 52 Lu D, Zhao Y, Tawatao R *et al.* Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. *Proc. Natl Acad. Sci. USA* 101(9), 3118–3123 (2004).
- 53 Li Y, Pawlik B, Elcioglu N *et al.* LRP4 mutations alter Wnt/ β -catenin signaling and cause limb and kidney malformations in Cenani–Lenz syndrome. *Am. J. Hum. Genet.* 86(5), 696–706 (2010).
- 54 Kim NW, Piatyszek MA, Prowse KR *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* 266(5193), 2011–2015 (1994).
- 55 Gaidano G, Ballerini P, Gong JZ *et al.* p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. *Proc. Natl Acad. Sci. USA* 88(12), 5413–5417 (1991).
- 56 Lee JN, Giles F, Huh YO *et al.* Molecular differences between small and large cells in patients with chronic lymphocytic leukemia. *Eur. J. Haematol.* 71(4), 235–242 (2003).
- 57 Lo Coco F, Gaidano G, Louie DC, Offit K, Chaganti RS, Dalla-Favera R. p53 mutations are associated with histologic transformation of follicular lymphoma. *Blood* 82(8), 2289–2295 (1993).
- 58 Lossos IS. Higher-grade transformation of follicular lymphoma – a continuous enigma. *Leukemia* 19(8), 1331–1333 (2005).
- 59 Pasqualucci L, Neumeister P, Goossens T *et al.* Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412(6844), 341–346 (2001).

- 60 Gaidano G, Pasqualucci L, Capello D *et al.* Aberrant somatic hypermutation in multiple subtypes of AIDS-associated non-Hodgkin lymphoma. *Blood* 102(5), 1833–1841 (2003).
- 61 Libra M, Capello D, Gloghini A *et al.* Analysis of aberrant somatic hypermutation (SHM) in non-Hodgkin lymphomas of patients with HCV infection. *J. Pathol.* 206(1), 87–91 (2005).
- 62 Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. *Nat. Rev. Immunol.* 8(1), 22–33 (2008).
- 63 Pasqualucci L, Bhagat G, Jankovic M *et al.* AID is required for germinal center-derived lymphomagenesis. *Nat. Genetics* 40(1), 108–112 (2008).
- 64 Reiniger L, Bödör C, Bognár Á *et al.* Richter's and prolymphocytic transformation of chronic lymphocytic leukemia are associated with high mRNA expression of activation-induced cytidine deaminase and aberrant somatic hypermutation. *Leukemia* 20(6), 1089–1095 (2006).
- 65 Han T, Sadamori N, Ozer H *et al.* Cytogenetic studies in 77 patients with chronic lymphocytic leukemia: correlations with clinical, immunologic, and phenotypic data. *J. Clin. Oncol.* 2(10), 1121–1132 (1984).
- 66 Hébert J, Jonveaux P, d'Agay MF, Berger R. Cytogenetic studies in patients with Richter's syndrome. *Cancer Genet. Cytogenet.* 73(1), 65–68 (1994).
- 67 Brynes RK, McCourty A, Sun NC, Koo CH. Trisomy 12 in Richter's transformation of chronic lymphocytic leukemia. *Am. J. Clin. Pathol.* 104(2), 199–203 (1995).
- 68 Caraway NP, Du Y, Zhang H-Z, Hayes K, Glassman AB, Katz RL. Numeric chromosomal abnormalities in small lymphocytic and transformed large cell lymphomas detected by fluorescence *in situ* hybridization of fine-needle aspiration biopsies. *Cancer* 90(2), 126–132 (2000).
- 69 Santulli B, Kazmierczak B, Napolitano R *et al.* A 12q13 translocation involving the *HMGI-C* gene in Richter transformation of a chronic lymphocytic leukemia. *Cancer Genet. Cytogenet.* 119(1), 70–73 (2000).
- 70 Zenz T, Mertens D, Döhner H, Stilgenbauer S. Molecular diagnostics in chronic lymphocytic leukemia – pathogenetic and clinical implications. *Leuk. Lymphoma* 49(5), 864–873 (2008).
- 71 Lenz G, Staudt LM. Aggressive lymphomas. *N. Engl. J. Med.* 362(15), 1417–1429 (2010).
- 72 Beà S, López-Guillermo A, Ribas M *et al.* Genetic imbalances in progressed B-cell chronic lymphocytic leukemia and transformed large-cell lymphoma (Richter's syndrome). *Am. J. Pathol.* 161(3), 957–968 (2002).
- 73 Scandurra M, Rossi D, Deambrogi C *et al.* Genomic profiling of Richter's syndrome: recurrent lesions and differences with *de novo* diffuse large B-cell lymphomas. *Hematol. Oncol.* 28(2), 62–67 (2010).
- 74 Cohen Y, Da'as N, Libster D, Amir G, Berrebi A, Polliack A. Large-cell transformation of chronic lymphocytic leukemia and follicular lymphoma during or soon after treatment with fludarabine-rituximab containing regimens: natural history or related therapy complication. *Eur. J. Hematol.* 68(2), 80–83 (2002).
- 75 Thornton PD, Bellas C, Santon A *et al.* Richter's transformation of chronic lymphocytic leukemia. The possible role of fludarabine and the Epstein–Barr virus in its pathogenesis. *Leukemia Res.* 29(4), 389–395 (2005).

Analysis of the chronic lymphocytic leukemia coding genome: role of *NOTCH1* mutational activation

Giulia Fabbri,¹ Silvia Rasi,⁵ Davide Rossi,⁵ Vladimir Trifonov,² Hossein Khiabanian,² Jing Ma,⁶ Adina Grunn,¹ Marco Fangazio,⁵ Daniela Capello,⁵ Sara Monti,⁵ Stefania Cresta,⁵ Ernesto Gargiulo,⁵ Francesco Forconi,⁷ Anna Guarini,⁸ Luca Arcaini,⁹ Marco Paulli,¹⁰ Luca Laurenti,¹¹ Luigi M. Larocca,¹² Roberto Marasca,¹³ Valter Gattei,¹⁴ David Oscier,¹⁵ Francesco Bertoni,¹⁶ Charles G. Mullighan,⁶ Robin Foà,⁸ Laura Pasqualucci,^{1,3} Raul Rabadan,² Riccardo Dalla-Favera,^{1,3,4} and Gianluca Gaidano⁵

¹Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, ²Department of Biomedical Informatics and Center for Computational Biology and Bioinformatics, ³Department of Pathology and Cell Biology, and ⁴Department of Genetics and Development, Columbia University, New York, NY 10032

⁵Division of Hematology, Department of Clinical and Experimental Medicine and Interdisciplinary Research Center on Autoimmune Diseases, Amedeo Avogadro University of Eastern Piedmont, 28100 Novara, Italy

⁶Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN 38105

⁷Division of Hematology, University of Siena, 53100 Siena, Italy

⁸Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, 00161 Rome, Italy

⁹Division of Hematology, Department of Hematology Oncology, University of Pavia Medical School, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy

¹⁰Institute of Pathology, University of Pavia, 27100 Pavia, Italy

¹¹Institute of Hematology and ¹²Institute of Pathology, Catholic University of the Sacred Heart, 00618 Rome, Italy

¹³Division of Hematology, Department of Oncology and Hematology, University of Modena and Reggio Emilia, 41124 Modena, Italy

¹⁴Clinical and Experimental Onco-Hematology, Centro di Riferimento Oncologico, Istituto di Ricovero e Cura a Carattere Scientifico, 33081 Aviano, Italy

¹⁵Department of Hematology, Royal Bournemouth Hospital, Bournemouth BH7 7DW, UK

¹⁶Oncology Institute of Southern Switzerland, CH-6500 Bellinzona, Switzerland

The pathogenesis of chronic lymphocytic leukemia (CLL), the most common leukemia in adults, is still largely unknown. The full spectrum of genetic lesions that are present in the CLL genome, and therefore the number and identity of dysregulated cellular pathways, have not been identified. By combining next-generation sequencing and copy number analysis, we show here that the typical CLL coding genome contains <20 clonally represented gene alterations/case, including predominantly nonsilent mutations, and fewer copy number aberrations. These analyses led to the discovery of several genes not previously known to be altered in CLL. Although most of these genes were affected at low frequency in an expanded CLL screening cohort, mutational activation of *NOTCH1*, observed in 8.3% of CLL at diagnosis, was detected at significantly higher frequency during disease progression toward Richter transformation (31.0%), as well as in chemorefractory CLL (20.8%). Consistent with the association of *NOTCH1* mutations with clinically aggressive forms of the disease, *NOTCH1* activation at CLL diagnosis emerged as an independent predictor of poor survival. These results provide initial data on the complexity of the CLL coding genome and identify a dysregulated pathway of diagnostic and therapeutic relevance.

Chronic lymphocytic leukemia (CLL) is the most common adult leukemia in the Western world (Chiorazzi et al., 2005; Tam and Keating, 2010; Zenz et al., 2010b). CLL is characterized

by a marked degree of clinical heterogeneity, ranging from patients that harbor a highly stable disease with a nearly normal life expectancy,

CORRESPONDENCE

Riccardo Dalla-Favera:
rd10@columbia.edu
OR

Gianluca Gaidano:
gaidano@med.unipmn.it

Abbreviations used: BIR, baculovirus IAP repeat; CLL, chronic lymphocytic leukemia; CNA, copy number alteration; DLBCL, diffuse large B cell lymphoma; RS, Richter syndrome; SNP, single-nucleotide polymorphism; T-ALL, T cell acute lymphoblastic leukemia; WES, whole-exome sequencing.

G. Fabbri and S. Rasi contributed equally to this paper.

L. Pasqualucci, R. Rabadan, R. Dalla-Favera, and G. Gaidano contributed equally to this paper.

© 2011 Fabbri et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

to patients with a rapidly progressive disease that are destined to shortly succumb (Chiorazzi et al., 2005; Tam and Keating, 2010; Zenz et al., 2010b). The variable course of CLL is driven, at least in part, by heterogeneity in the leukemia biology (Tam and Keating, 2010; Zenz et al., 2010b). A fraction of CLL is prone to transformation to diffuse large B cell lymphoma (DLBCL), a condition known as Richter syndrome (RS) that is characterized by a rapidly deteriorating clinical course with an extremely dismal outcome (Tsimberidou and Keating, 2005; Rossi and Gaidano, 2009).

The molecular pathogenesis of CLL is not fully understood. In contrast to other B cell malignancies, CLL is not associated with recurrent balanced chromosomal translocations (Klein and Dalla-Favera, 2010). Deletion of 13q14 affecting the *DLEU2/MIR15-16* cluster is conceivably the initiating lesion in nearly 50% of CLL cases, followed by the accumulation of additional alterations during tumor progression, namely del17p13 and del11q22-q23 affecting *TP53* and *ATM*, respectively (Döhner et al., 2000; Calin et al., 2002; Klein et al., 2010; Zenz et al., 2010b). Rare *NOTCH1* mutations have been sporadically reported (Sportoletti et al., 2010).

These known genetic lesions, however, do not fully recapitulate the molecular pathogenesis of CLL and do not entirely explain the marked clinical heterogeneity of the disease or the development of severe complications, such as RS transformation and chemorefractoriness, which, despite recent advances, still represent unmet clinical needs (Rossi and Gaidano, 2009; Tam and Keating, 2010; Zenz et al., 2010b; Fangazio et al., 2011). In the instance of CLL transformation, *TP53* disruption and *MYC* alterations are observed in up to 50% of RS cases, but the genetic lesions of the remaining 50% are virtually unknown (Rossi et al., 2011). Also, *TP53* disruption is observed in 40% of CLL cases refractory to therapy, but the molecular basis of the remaining 60% CLL is unknown (Zenz et al., 2010a).

Genome-wide methods aimed at the characterization of the entire spectrum of genetic lesions present in the CLL genome may be useful to provide further insights into CLL pathogenesis, and might contribute to elucidate the molecular basis of CLL clinical evolution, including RS transformation and development of chemorefractoriness. On these grounds, we have exploited an integrated approach based on next generation whole-exome sequencing (WES) and genome-wide high-density single-nucleotide polymorphism (SNP) analysis to investigate the CLL coding genome. We report that CLL is characterized by a limited number of structural genetic alterations compared with other B cell malignancies, and carries a restricted number of mutations in genes potentially relevant for CLL pathogenesis.

Among the genes found mutated in CLL, *NOTCH1* emerged as a highly recurrent target of genetic lesions in specific phases of the disease. *NOTCH1* encodes a class I transmembrane protein functioning as a ligand-activated transcription factor and playing an important role in cell differentiation, proliferation, and apoptosis (Grabher et al., 2006; Aster et al., 2008, 2011; Paganin and Ferrando, 2011). Upon ligand-binding,

NOTCH1 undergoes multiple proteolytic cleavages that allow its intracellular portion to translocate to the nucleus, thus leading to transcriptional activation of multiple target genes, including *MYC* (Palomero et al., 2006). In T cell acute lymphoblastic leukemia (T-ALL), activating mutations of *NOTCH1* are the predominant genetic alteration, accounting for up to 60% of the cases (Weng et al., 2004). The present study documents a high frequency of *NOTCH1* mutations in aggressive clinical phases of CLL, exemplified by RS transformation and chemorefractoriness, and identifies *NOTCH1* alterations as an independent predictor of poor prognosis.

RESULTS

Mutational load by exome sequencing analysis

To identify somatic gene alterations associated with CLL pathogenesis and to gain insights into the overall complexity of the CLL coding genome, we performed exome capture and next generation sequencing analysis of paired tumor and normal DNAs from 5 newly diagnosed, previously untreated CLL patients (Table S1, discovery panel; and Fig. S1). We also analyzed by direct sequencing the coding exons of *NOTCH1*, as this gene has been previously implicated in the disease (Sportoletti et al., 2010), but was not represented in the capture array.

Validation of the candidate variants by Sanger sequencing in the same samples confirmed the presence of 40 somatic, non-silent mutations involving 39 distinct genes, with *NOTCH1* representing the only recurrently mutated one ($n = 2/5$ cases; Table S2, Table S3, and Supplemental Results). Most of the mutations detected by this approach appeared to be represented in a predominant clone, based on the presence of a clear peak in the chromatogram. The overall mutation load was homogeneous across the five cases (mean = 8, range = 6–10 lesions/case) and included 36 single nucleotide substitutions ($n = 32$ aa changes and 4 premature stop codons) and 4 frameshift insertions/deletions (Fig. 1 a and Table S3). Although assessed on a limited number of events, the mutational spectrum of CLL was analogous to that reported in other cancer types (e.g., colorectal, pancreatic, and brain tumors; Greenman et al., 2007; Jones et al., 2008; Parsons et al., 2011), with a predominance of transitions over transversions ($n = 25:11$, ratio 2.2) a preferential targeting of G and C nucleotides (69.4 vs. 30.6%, affecting A/T nucleotides), and a bias for mutations at 5'-CpG-3' dinucleotides ($P < 0.001$; Fig. S2).

Genome-wide SNP analysis

Analysis of the same 5 CLL samples (and paired normal DNAs) by the Affymetrix SNP 6.0 platform revealed a total of 12 somatic copy number alterations (CNA), with an average of 2.4 lesions per case (range: 1–5; Table S4). 10 of these regions harbored genetic elements, including protein coding genes and/or noncoding RNA genes, and 5 of them encompassed <3 genes that, because of their focal involvement, represent the likely targets of the aberration (Fig. 1 b). Consistent with previous studies (Kujawski et al., 2008; Grubor et al., 2009), the majority of the observed lesions were represented

by deletions (11/12, 91.7%), ranging in size from 1.3 kb to loss of a large portion of a chromosomal arm, as in CLL1 that carried a del11q. Deletion of 13q14 was the only recurrent lesion, while trisomy 12 (CLL4) represented the only region of gain (Fig. S3).

Importantly, the SNP array approach successfully identified all four lesions previously detected by FISH analysis and known to associate with CLL, including two del13q14 (CLL2 and CLL5), one trisomy 12 (CLL4), and one del11q (CLL1; Fig. S3). In both instances, del13q14 encompassed the *DLEU2/MIR15-16* locus along with 61–68 additional genetic elements, including the *RB1* gene. The del11q spanned more than 100 genes, including the tumor suppressor *ATM* and the negative regulator of the noncanonical NF- κ B pathway *BIRC3*.

Overall complexity of the CLL coding genome

The integration between WES and copy number data provided a total of 52 genetic lesions, ranging from 7 to 13 lesions per case (average = 10.4) and mostly represented by point mutations (Fig. 1 c). Even after considering that some mutations may have been missed because of limited probe density or sequence coverage of the SNP and exome sequencing platforms, this mutation burden is 5–20 times lower than that observed in solid tumor genomes (Jones et al., 2008; Pleasance et al., 2010a; Pleasance et al., 2010b), in multiple

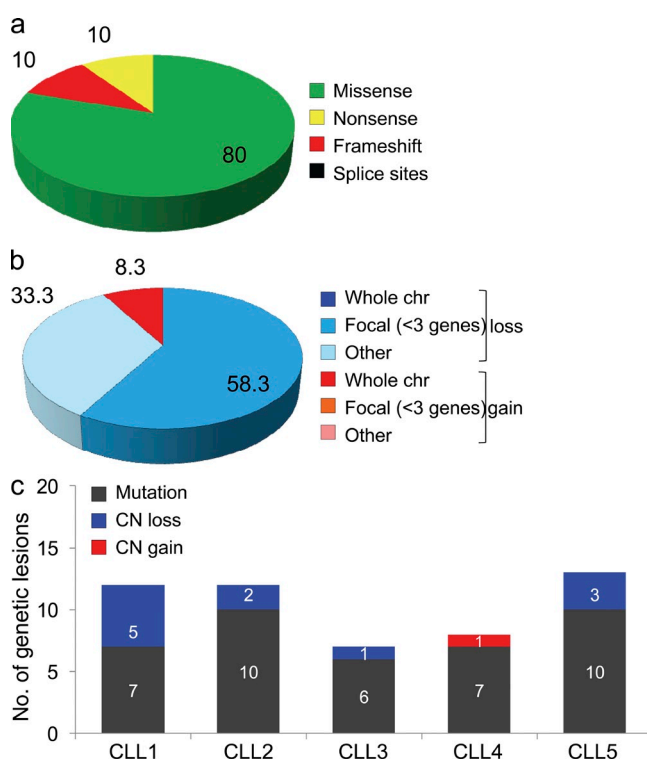


Figure 1. Overall load of genetic lesions in CLL. (a) Frequency of Sanger-validated somatic nonsilent mutations identified through WES in the five CLL discovery cases. (b) Overall frequency of somatically acquired CNA according to aberration type in the 5 CLL discovery cases. (c) Combined load of somatically acquired nonsilent mutations and CNA identified in the CLL discovery panel. CN, copy number.

myeloma (Chapman et al., 2011), and in DLBCL (unpublished data). Conversely, this load is comparable to that reported in acute myeloid leukemia (Mardis et al., 2009) and in pediatric medulloblastoma (Parsons et al., 2011). This observation, together with the low number of CNAs in general, and whole chromosome alterations in particular, suggests that the CLL genome is relatively stable and homogeneous in terms of mutational load, ploidy, and large chromosomal rearrangements.

Identification of frequently mutated genes

The genes found to be mutated by the aforementioned approach include 32 genes listed in the Catalogue of Somatic Mutations in Cancer database and previously found to be mutated in other cancer genomes (Table S3; Forbes et al., 2011). Although the functional significance of the missense mutations affecting these genes is largely unknown, 22 out of 32 (68.8%) are expected to alter the function of the encoded protein, based on the Polyphen-2 prediction algorithm (Table S3; Adzhubei et al., 2010). The Gene Ontology database (<http://www.geneontology.org/>) revealed that one third of the identified mutated genes are involved in transcriptional regulation and chromatin remodeling processes ($n = 13/39$, 33.3%).

One criterion to assess the pathogenic relevance of the observed genetic lesions is to examine their recurrence in the disease. In the discovery panel, *NOTCH1* was the only gene mutated in more than one patient (2/5 cases); however, the lack of recurrence in other mutated genes may be ascribed to the limited number of cases investigated. We therefore expanded the mutational analysis to an independent screening panel of 48 CLL cases with available matched germline DNA (see Table S5 for the detailed characterization of the CLL screening panel). This panel was representative of both the *IGHV* mutated ($n = 24$) and unmutated ($n = 24$) disease categories. The entire coding region and splice sites of the genes found mutated in the discovery panel were studied in the screening phase. This analysis, combined with the results obtained in the 5 cases belonging to the discovery panel, revealed the presence of recurrent mutations ($n \geq 2$) in 4 genes, i.e., *TGM7*, *BIRC3*, *PLEKHG5* (all mutated in $\leq 5\%$ of cases), and *NOTCH1* (mutated in 15.1% of cases; Fig. 2).

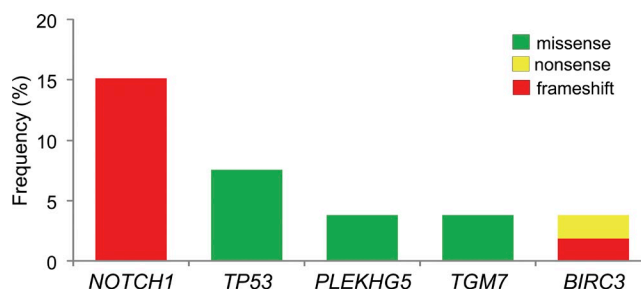
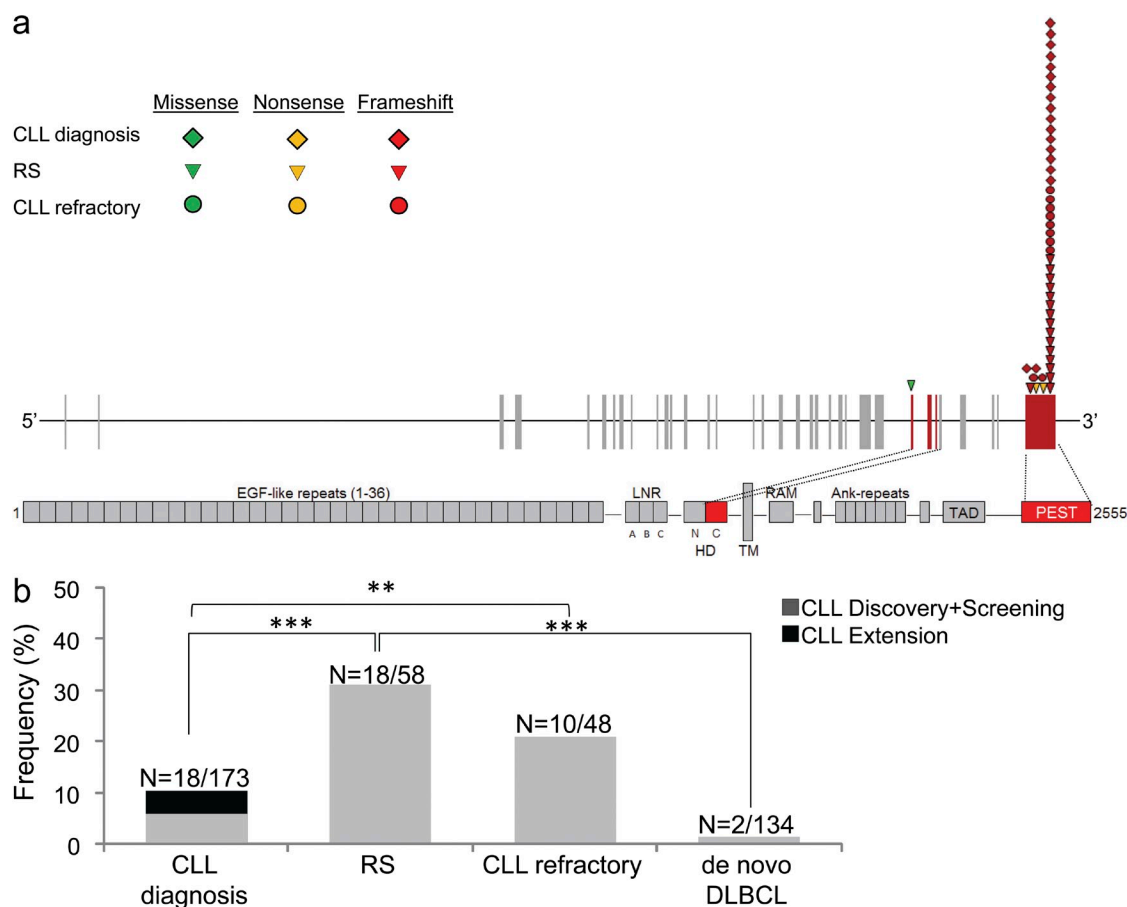


Figure 2. Recurrent mutations in CLL. Percentage of CLL primary cases harboring mutations in candidate genes after targeted resequencing of an expanded screening dataset ($n = 48$ CLL cases). The final number of mutated cases over total analyzed (including the 5 discovery CLL) for each gene is as follows: *NOTCH1*, 8/53; *TP53*, 4/53; *PLEKHG5*, 2/53; *TGM7*, 2/53; *BIRC3*, 2/53.

Although the role of the two missense mutations found in the *TGM7* gene, which encodes a transglutaminase of unclear specific function, remains to be determined, other mutations provide preliminary evidence of functional significance. *BIRC3*, a negative regulator of the noncanonical NF- κ B pathway that is affected by chromosomal translocations in MALT lymphoma and by focal deletions in multiple myeloma (Zhou et al., 2005; Annunziata et al., 2007; Keats et al., 2007; Rosebeck et al., 2011), was biallelically inactivated in CLL1. In this patient, a copy number loss was observed together with a frameshift deletion located between the third baculovirus IAP repeat (BIR) and the caspase activation and recruitment domain (CARD), leading to removal of the RING domain, which is required for the inhibitory function of BIRC3 (Gyrd-Hansen and Meier, 2010). A second case (CLL45) harbored a nonsense mutation (E424X) truncating the corresponding

The recurrence of *NOTCH1* mutations (15.1%) appeared potentially relevant, given the well-established involvement of this gene in the pathogenesis of T-ALL and its previously reported association with CLL at low frequency (Weng et al., 2004;



1392

Sulis et al., 2008; Sportoletti et al., 2010). This finding, together with the preliminary observation of *NOTCH1* mutations in a few cases of RS (see following paragraph), prompted a systematic analysis of *NOTCH1* in different clinical phases of CLL.

***NOTCH1* mutations at CLL diagnosis occur at low frequency and predict poor outcome**

To define the frequency of *NOTCH1* mutations at CLL diagnosis, we analyzed a consecutive cohort of CLL patients ($n = 120$) sampled at diagnosis and followed at the Amedeo Avogadro University of Eastern Piedmont (extension panel). This cohort was representative of the main clinical characteristics of the disease at diagnosis (Table S6). Specifically, *IGHV* homology $\geq 98\%$ occurred in 39/120 (32.5%) patients, del13q14 was observed in 61/112 (54.5%) cases, normal FISH in 32/112 (28.6%), trisomy 12 in 17/112 (15.2%), del11q22-q23 in 8/112 (7.1%), and *TP53* disruption by mutation and/or deletion in 13/117 (11.1%; Table S6). After a median follow up of 70.3 mo, the 5-yr treatment free survival (TFS) probability of the CLL cohort was 60.0% and the 5-yr overall survival (OS) probability was 83.3%.

At CLL diagnosis, mutations of *NOTCH1* were identified in 10/120 (8.3%) patients, all of which displayed a recurrent 2-bp frameshift deletion (Δ CT7544-7545, P2515fs) that is predicted to cause *NOTCH1*-impaired degradation through the truncation of the C-terminal PEST domain (Fig. 3 a and Table S7; Weng et al., 2004). *NOTCH1* mutations clustered with CLL harboring unfavorable features, including unmutated *IGHV* genes (7/39, 17.9% vs. 3/81, 3.7% in *IGHV* mutated CLL; $P = 0.013$), and *TP53* disruption (3/13, 23.1% vs. 7/104, 6.7% in *TP53* wild-type CLL; $P = 0.047$). Consistently, CLL harboring *NOTCH1* mutations were characterized by a significantly shorter time to progression to a disease requiring treatment (median TFS, 7.9 mo) compared with CLL cases harboring wild-type *NOTCH1* alleles (median

TFS, 128.5 mo; $P < 0.001$; Fig. 4). Also, CLL harboring *NOTCH1* mutations were characterized by a significantly shorter OS (median OS, 93.8 mo; 5-yr OS, 53.3%) compared with wild-type CLL (median OS not reached; 5-yr OS, 86.0%; $P = 0.001$; Fig. 4). By multivariate Cox analysis, *NOTCH1* mutations were a predictor of poor OS (HR, 10.1; 95% CI, 2.6–39.0; $P = 0.001$) independent of confounding covariates, including *TP53* disruption (HR, 4.6; 95% CI, 1.6–13.1; $P = 0.004$) and unmutated *IGHV* genes (HR, 6.2; 95% CI, 1.4–25.9; $P = 0.012$).

Collectively, these results confirm previous findings on the low frequency and prognostic relevance of *NOTCH1* mutations at CLL diagnosis (Di Ianni et al., 2009; Sportoletti et al., 2010), but identify *NOTCH1* mutations as a predictor of unfavorable outcome independent of other biological features of CLL, namely *IGHV* mutation status and *TP53* disruption.

Frequent activating mutations of *NOTCH1* in RS

The association between *NOTCH1* mutations at CLL diagnosis and unfavorable outcome prompted the investigation of this gene in patients whose disease subsequently acquired a high-risk clinical phenotype. As a model of CLL clinical progression, we initially chose RS transformation because: (a) it represents the most aggressive phase of the clinical spectrum of CLL, as also documented by the fact that RS is rapidly lethal in most patients; (b) the molecular lesions driving CLL transformation to RS are not fully understood (Tsimberidou and Keating, 2005; Rossi and Gaidano, 2009; Fangazio et al., 2011; Rossi et al., 2011).

Based on this rationale, *NOTCH1* mutations were analyzed in the diagnostic biopsy of 58 cases of histologically proven RS, all represented by DLBCL transformed from a previous CLL phase. The clinical and biological characteristics of the RS cases are summarized in Table S8. A total of 19

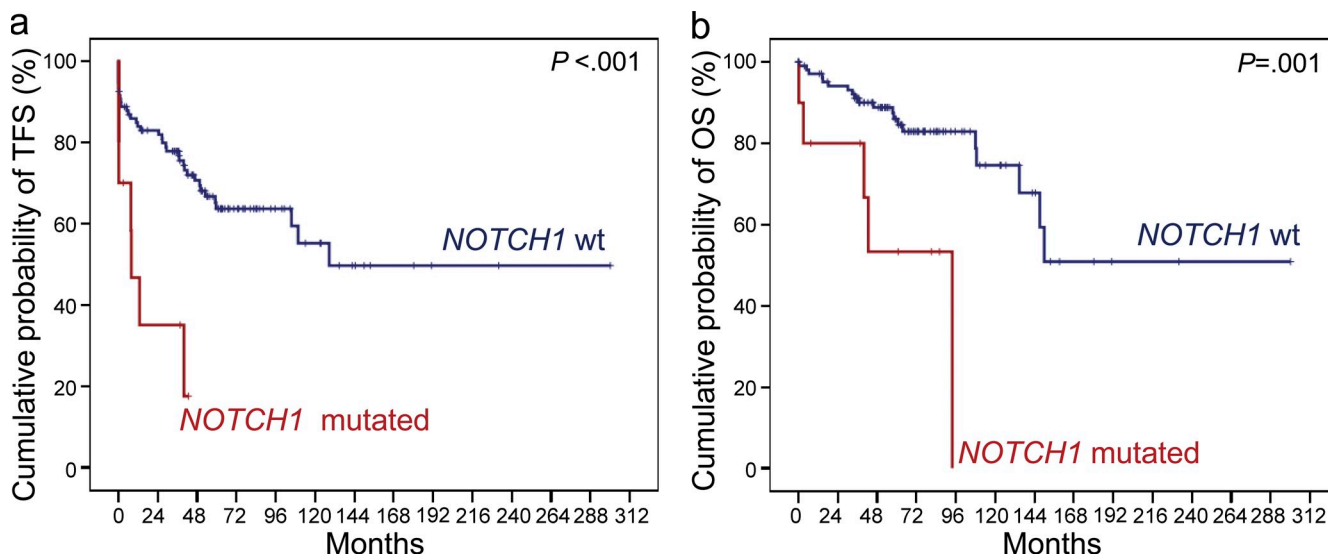


Figure 4. Kaplan-Meier estimates of treatment-free survival (TFS) and overall survival (OS), according to the mutational status of *NOTCH1*.

mutations were identified in 18/58 (31.0%) cases (Table S7). The frequency of *NOTCH1* mutations in RS was significantly higher compared with the frequency observed at CLL diagnosis in the consecutive series ($P < 0.001$; Fig. 3b). *NOTCH1* mutations in RS were represented by a recurrent 2-bp frameshift deletion (Δ CT7544-7545, P2515fs; $n = 15$), one frameshift insertion, and two nonsense mutations (Fig. 3a). On the basis of their distribution along the *NOTCH1* protein, all frameshift and nonsense mutations of *NOTCH1* are predicted to cause *NOTCH1*-impaired degradation through the truncation of the C-terminal PEST domain (Table S7 and Fig. 3a). One single mutation (G5164A, V1722M) affected the heterodimerization domain of *NOTCH1* (Fig. 3a).

To establish the timing of acquisition of *NOTCH1* mutations during the clinical history of RS patients, we investigated paired sequential samples collected at the time of CLL diagnosis. 15 patients harboring 16 *NOTCH1* mutations and belonging to the RS series were analyzed by this approach. Of the 16 *NOTCH1* mutations observed in the RS phase, 5 (31.3%) were not detectable in the paired CLL phase within the sensitivity threshold of Sanger sequencing (Fig. 5a). Conversely, 6/16 (37.5%) mutations detected in the RS phase were already present at the time of CLL diagnosis, where their clonal representation appeared to be similar to that observed in RS, as estimated by Sanger sequencing (Fig. 5a). Interestingly, 5/16 (31.2%) *NOTCH1* mutations identified in the RS

phase were already present at subclonal levels in the paired CLL phase. To obtain a more quantitative estimate of the clonal representation of *NOTCH1* mutated versus germline alleles, the CLL and RS phases from case RS63 were subjected to ultradeep next generation sequencing (sequencing depth $\sim 5,000$ per sample) using the Roche 454 technology. This approach showed that the Δ CT7544-7545 (P2515fs) mutation, occurring in 58.6% (3,304/5,642) of the sequencing reads obtained from the RS phase, was restricted to 5.3% (239/4524) of the reads obtained at the time of CLL diagnosis (Fig. 5b, Fig. S4, and Fig. S5), indicating an 11-fold increase during the 30-mo interval between CLL diagnosis and RS development. These data suggest that clonally represented mutations of RS might already be present years before RS transformation, and that the clone harboring the mutation is progressively selected during the CLL clinical history ending into RS transformation.

One RS patient (RS86) carrying a *NOTCH1* mutation disrupting the PEST domain (Δ CT7544-7545, P2515fs) in the CLL phase and in a first RS biopsy, subsequently acquired a second *NOTCH1* mutation (G5164A, V1722M) affecting the heterodimerization domain when RS further progressed from nodal localization to peripheral blood invasion. The detection of PEST and heterodimerization domain mutations occurring in the same patient is of relevance, as mutations in these two distinct domains have been shown to act synergistically

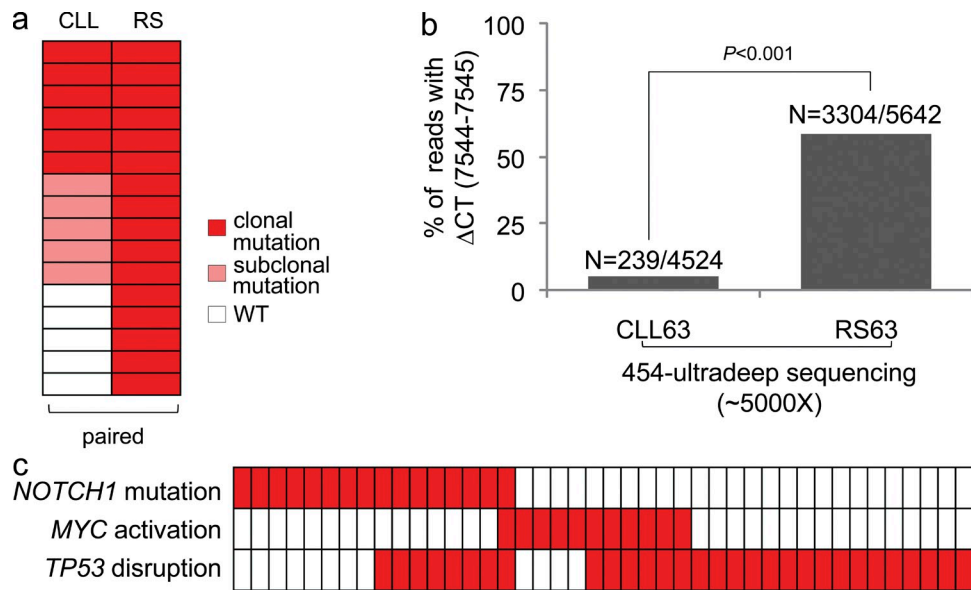


Figure 5. Timing of *NOTCH1* mutations in RS. (a) *NOTCH1* mutations in some cases of RS are not identifiable in the preceding CLL phase, whereas in other patients may already be present before transformation, at a clonal/subclonal level. In the heatmap, rows correspond to *NOTCH1* mutations and columns represent sequential phases of the disease (paired CLL diagnostic samples and RS samples), color-coded based on the mutational status. (b) The *NOTCH1* mutation detected in RS case RS63 is present at subclonal levels at the time of CLL diagnosis (sample CLL63). The graph shows a quantitative estimate of the balance between wild-type and mutated (Δ CT7544-7545) *NOTCH1* alleles assessed by ultradeep 454 sequencing in the paired CLL-RS samples (CLL63 and RS63). 30 mo elapsed between collection of the CLL sample (diagnosis of CLL) and collection of the RS sample (obtained at diagnosis of RS). Numbers on top indicate the actual number of sequencing reads carrying the mutations over total analyzed (for details, see Materials and methods). (c) Mutations of *NOTCH1* are largely mutually exclusive with *MYC* activating events (translocations and copy number gains) in RS. In the heatmap, rows correspond to the indicated genes, and columns represent individual RS samples, color-coded based on the gene status (white, wild-type; red, mutated or activated or disrupted).

in promoting *NOTCH1* signaling activation (Weng et al., 2004). The progressive accumulation of multiple *NOTCH1* mutations during clinical history further reinforces the notion that *NOTCH1* activation might be relevant for CLL progression.

The occurrence of *NOTCH1* activation in RS was then compared with the distribution of other genetic alterations that are recurrently associated with this disease. *MYC*, one of the target genes and major downstream effectors of *NOTCH1* (Palomero et al., 2006; Li et al., 2011), was targeted by translocations in 9/51 (17.6%) RS cases and by gene amplification in 2 additional cases. *NOTCH1* mutations and *MYC* abnormalities distributed in a mutually exclusive fashion (Fig. 5 c), although at borderline statistical significance ($P = 0.076$), conceivably because of the limited number of cases included in the analysis. On these bases, by combining *NOTCH1* mutations and *MYC* abnormalities, 26/51 (51.0%) RS harbor genetic lesions ultimately leading to *MYC* activation. Among CLL, *MYC* abnormalities were extremely rare (1/120; 0.8%), in agreement with the notion that they are generally acquired at RS transformation (Huh et al., 2008; Rossi et al., 2011). The sole case of CLL carrying a *MYC* alteration was devoid of *NOTCH1* mutations.

TP53 was disrupted by mutations and/or deletion of the locus in 30/54 (55.5%) RS. *TP53* disruption frequently paired with *NOTCH1* mutations (8/16, 50%) and *MYC* abnormalities (7/11, 63.6%) in the same RS sample, thus contributing to a dual-hit genetic mechanism of transformation (Fig. 5 c).

All cases of RS included in this study were histologically classified as DLBCL, a pathological entity characterized by high molecular heterogeneity that may arise either de novo or upon transformation from an indolent B cell malignancy, as in the case of CLL transformation to RS (Swerdlow, 2008). To assess the specificity of *NOTCH1* mutations for RS, 134 cases of de novo DLBCL were also investigated. Mutations of *NOTCH1* were very rare in this category, being restricted to 2/134 (1.5%) cases (Table S7 and Fig. 3 b).

The frequent recurrence of *NOTCH1* mutations in RS prompted the investigation of other *NOTCH1*-related genes, namely *NOTCH2* and *FBXW7* (O'Neil et al., 2007; Thompson et al., 2007; Lee et al., 2009). Mutations of *NOTCH2* were detected in 2/32 (6.2%) cases of RS, and included one amino acid change (K2121N), that was shown to be acquired at transformation, and a premature stop codon (R2400*); in this case, no material was available from the corresponding diagnostic CLL sample). *FBXW7* mutations scores were consistently negative (0/32).

Overall, these data document that mutations activating the Notch pathway are a frequent event in the pathogenesis of RS. In the context of the clinicopathological spectrum of DLBCL subtypes, *NOTCH1* mutations appear to preferentially associate with transformed cases.

Frequent activating mutations of *NOTCH1* in chemorefractory CLL

Because RS is a highly chemorefractory condition (Rossi and Gaidano, 2009; Tsimberidou and Keating, 2005), we

hypothesized that *NOTCH1* mutations might also provide a marker of chemorefractoriness in patients who clinically progressed because of treatment failure without transformation to RS. To address this issue, *NOTCH1* mutations were investigated in 48 cases of progressive and chemorefractory CLL in which RS had been ruled out based on clinical picture, lymph node biopsy, and/or ^{18}F FDG-PET studies (Table S9). Samples from these cases were obtained at the time of progression immediately before starting the treatment to which the patient eventually failed to respond. Mutations were identified in 10/48 (20.8%) cases of chemorefractory CLL (Table S7), including 8/35 (22.8%) fludarabine-refractory and 2/13 (15.3%) alkylator-refractory cases. The frequency of *NOTCH1* mutations in chemorefractory CLL was significantly higher compared with the frequency observed at CLL diagnosis in the consecutive series ($P = 0.033$; Fig. 3 b). All mutations led to a shift in the reading frame and, as observed in RS, were mostly (8/10) represented by the 2-bp frameshift deletion $\Delta\text{CT7544-7545}$ (P2515fs). On the basis of their distribution along the *NOTCH1* protein, mutations observed in chemorefractory CLL are all predicted to increase *NOTCH1* stability via truncation of the C-terminal PEST domain (Fig. 3 a; Weng et al., 2004).

In CLL, the molecular mechanisms of chemorefractoriness are represented by *TP53* disruption in up to 40% of cases, indicating that other currently unknown mechanisms might be operative in the remaining 60% of cases. In the current series of chemorefractory CLL, *TP53* was disrupted by mutations and/or deletions in 19/48 (39.6%) cases. *NOTCH1* mutations occurred in 6/29 (20.7%) *TP53* wild-type chemorefractory CLL, whereas they overlapped with *TP53* disruption in 4/19 cases (21.1%).

Based on these data, *NOTCH1* mutations represent a frequent genetic lesion of CLL associated with progressive disease failing treatment.

DISCUSSION

The goal of this study was to characterize the nature and frequency of somatic genetic alterations affecting the CLL coding genome by integrating the analysis of structural and sequence mutations of DNA. This combined approach has allowed the definition of the degree of complexity characterizing the CLL coding genome, with implications for the mechanisms leading to genetic alterations in this disease. This analysis has also revealed the involvement of the *NOTCH1* pathway in disease progression and transformation, with diagnostic and therapeutic implications.

Although analysis of larger cohorts of patients will be required to conclusively assess the precise number of genetic lesions that are present in the CLL genome, the combined set of gene alterations (i.e., copy number changes and mutations) identified in the CLL discovery panel provides initial information about the order of magnitude of the lesions associated with this disease. Given the presence of ~ 8 nonsilent mutations and ~ 2 CNAs per CLL case, our results suggest that the

coding genome of CLL harbors at least 10 potentially relevant genetic aberrations. As an additional ~40% of nonsilent mutations may have escaped detection by the WES approach used, because of relatively low depth of coverage, the total number of gene alterations present in a CLL genome may be <20. This estimate refers to those lesions that are clonally represented in the CLL genome at diagnosis, and which may have thus contributed to the malignant transformation process. As such, this preliminary CLL genome picture may serve as an initial database for the determination of frequency in additional panel of cases.

The order of magnitude of lesions detected in the coding genome of CLL appears considerably lower than that reported for common epithelial cancers. Among hematologic malignancies, the complexity of the CLL genome is also on the lower side, markedly smaller than that of DLBCL (<100) and multiple myeloma (~50), and comparable to that of some acute leukemias (~10; Ley et al., 2008; Mardis et al., 2009; Chapman et al., 2011; unpublished data). The predominance of point mutations over CNAs is notable, as it is unusual compared with most cancer types (Mullighan et al., 2007; Lenz et al., 2008; Parsons et al., 2011). This observation, together with the well-documented rarity of reciprocal balanced chromosomal translocations in CLL (Klein and Dalla-Favera, 2010), is consistent with previous models suggesting a derivation from a B cell counterpart in which immunoglobulin gene remodeling mechanisms are not active and those controlling ploidy remain intact.

On the other hand, similar to most cancer types are: (a) the observed predominance of transitions over transversions; (b) the preferential targeting of C:G and G:C base pairs; and (c) the significant bias toward alterations at CpG dinucleotides. The distribution of the mutations along the entire span of transcribed regions does not seem to reflect the abnormal and/or ectopic activity of members belonging to the APO-BEC family of deaminases, including activation induced deaminase (Albesiano et al., 2003; Delker et al., 2009), which targets the 5' portion of genes within 2 kb from the transcription initiation site. Instead, the observed pattern of alterations is more consistent with a derivation from endogenous biochemical processes, such as the spontaneous deamination of 5-methylcytosine residues (Parsons et al., 2011).

The screening of the mutations observed in the five cases of the discovery panel into an extended panel of cases has shown that only very few of these mutations display some minimal degree of recurrence, even when the entire gene is analyzed. This outcome appears different from that obtained in other malignancies, e.g., in DLBCL, where an analogous approach has led to the identification of a few highly recurrent genetic lesions (Pasqualucci et al., 2011). This result may reflect the fact that the most significant alterations involved in CLL pathogenesis have already been identified, including del13q14 involving the *DLEU2/MIR15-16* cluster, trisomy 12, and del11q. Alternatively, the rarity in recurrence may reflect the relationship of the observed lesions in functional pathways that are not presently recognized. One example in

support of this notion is the observation of three lesions (*BIRC3*, *MYD88*, and *PLEKHG5*) each detected at very low frequency, yet all linked to perturbation of the NF- κ B activity, a well-documented oncogenic pathway (Staudt, 2010). Each of the genetic alterations observed in this initial study represents the basis for extensive analysis of additional lesions in their respective cellular pathways.

A novel finding emerging from this study is the association between somatic mutations of *NOTCH1* and different aspects of disease progression. The mutations observed in CLL have consequences on *NOTCH1* function that have been documented in T-ALL, where their oncogenic activity has been extensively demonstrated in vitro and in vivo (Weng et al., 2004; O'Neil et al., 2006). In fact, the presence of *NOTCH1* mutations in rare cases of CLL at diagnosis has been previously reported and is confirmed in our study (Di Ianni et al., 2009; Sportoletti et al., 2010). Conversely, the association of these mutations at a significantly higher frequency with both RS and CLL investigated at the time of chemorefractoriness strongly suggests that *NOTCH1* mutations are selected during disease progression. At CLL diagnosis, *NOTCH1* mutations identify a disease subgroup enriched for poor risk genetic features, and characterized by an aggressive clinical phenotype. Future investigations involving large and prospective CLL cohorts are required to show how *NOTCH1* mutations interact with other high-risk genetic features.

In the case of RS, our data cannot conclusively distinguish whether *NOTCH1* mutations are already present at subclonal levels early during CLL development and are subsequently selected (as demonstrated for a few cases), or whether they are acquired de novo during disease transformation to RS. A conclusive demonstration of the precise timing of *NOTCH1* mutations in RS awaits studies aimed at tracking *NOTCH1* mutations with high sensitivity techniques in disease phases preceding RS development (Campbell et al., 2008). Also, the recurrent presence of *NOTCH1* mutations in chemorefractory CLL might be caused either by selection of more malignant clones during disease progression or by a role of these mutations in determining drug resistance. This issue may be addressed by dedicated studies of *NOTCH1* mutations in prospective cohorts of CLL, ideally in the context of controlled clinical trials.

Concerning the pathogenesis of RS, our results show that *NOTCH1* mutations are largely mutually exclusive with *MYC* oncogenic activation. This finding is consistent with the observation that *NOTCH1* directly stimulates *MYC* transcription and suggests that activation of oncogenic *MYC* may be one common final pathway selected for tumorigenesis (Palomero et al., 2006). Conversely, both *NOTCH1* mutational activation and *MYC* deregulation often coexist with inactivation of the *TP53* tumor suppressor gene, an event that is commonly observed in association with *MYC* activation in tumors, which may be selected to prevent the apoptotic effects and the response to genomic instability both induced by *MYC* overexpression (Eischen et al., 1999; Schmitt et al., 2002; Dominguez-Sola et al., 2007; Rossi et al., 2011).

From a clinical standpoint, *NOTCH1* mutations cluster with CLL subgroups that are currently scored at extremely high risk because of *TP53* disruption, refractoriness to fludarabine, or transformation to RS. In the context of these highly unfavorable clinical settings, the appearance or selection of *NOTCH1* mutations during disease progression has diagnostic and therapeutic implications. The detection of these mutations at the subclonal level using high sensitivity methods, such as the ultradeep next generation sequencing shown in Fig. 5 b, Fig. S4, and Fig. S5, may provide an objective and measurable biomarker for the early detection of the risk of progression or chemorefractoriness and therefore influence the clinical management of CLL. The validity of this attractive hypothesis needs to be explored by dedicated studies aimed at assessing the role of *NOTCH1* as a predictor of RS and/or chemorefractoriness at the time of CLL diagnosis. Finally, *NOTCH1* represents a well-established therapeutic target with some drugs already available, such as those inhibiting its enzymatic conversion to active transcription factor, and others under active development (Real et al., 2009). Our results add CLL to the cadre of common diseases in which these drugs should be tested for their efficacy alone or in combination with available therapeutic regimens.

MATERIALS AND METHODS

Patients. Samples from 5 newly diagnosed and previously untreated CLL patients (discovery panel) were obtained from the Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, as frozen peripheral blood mononuclear cells isolated by Ficoll-Paque gradient centrifugation. Diagnosis of CLL was based on IWCLL-NCI Working Group criteria and confirmed by a flow cytometry score >3 (Hallek et al., 2008). In all cases, the fraction of tumor cells corresponded to >80%, as assessed by FACS analysis of CD19/CD5 expression. Matched normal DNA was obtained from peripheral blood granulocytes that had been shown to be devoid of tumor cells by PCR analysis of patient-specific IGHV-D-J rearrangements. The clinical and biological characteristics of cases belonging to the discovery panel are summarized in Table S1. Patients were representative of the two major immunogenetic subgroups of CLL (*IGHV* mutated, two cases; *IGHV* unmutated, three cases), and harbored common CLL-associated genetic abnormalities.

The screening panel used to assess the recurrence of mutations affecting genes that were identified through WES was composed of 48 newly diagnosed and previously untreated CLL patients provided by Amedeo Avogadro University of Eastern Piedmont. In all cases, the fraction of tumor cells corresponded to >80%, as assessed by FACS analysis of CD19/CD5 expression. Matched normal DNA was obtained for all cases from purified granulocytes or saliva and was confirmed to be devoid of tumor-derived material by PCR analysis of tumor-specific IGHV-D-J rearrangements. The clinical and biological characteristics of cases belonging to the screening panel are summarized in Table S5. Patients were representative of the two major immunogenetic subgroups of CLL (*IGHV* mutated, 24 cases; *IGHV* unmutated, 24 cases).

The prevalence of *NOTCH1* mutations was further assessed in 3 clinical CLL panels representative of different disease phases, and including a consecutive series of newly diagnosed and previously untreated CLL ($n = 120$, extension panel), a cohort of chemorefractory CLL ($n = 48$), and a cohort of CLL transformed to RS ($n = 58$). Diagnosis of chemorefractoriness was according to guidelines (Hallek et al., 2008). Diagnosis of RS was based on the histology of lymph node or extranodal tissue excisional biopsies. After institutional pathological review, all RS cases were classified as DLBCL according to the World Health Organization Classification of Tumours of the Hematopoietic and Lymphoid Tissues (Swerdlow, 2008). In all RS cases, molecular

studies were performed on the biopsy used for RS diagnosis. Biological samples from the paired CLL phase were available in 45 (77.5%) cases.

The clinical and biological characteristics of the consecutive series of newly diagnosed and previously untreated CLL ($n = 120$) are summarized in Table S6. After a median follow up from diagnosis for living patients of 70.3 mo, the 5 yr-treatment free survival probability of the CLL cohort was 60.0% and the 5-yr overall survival probability was 83.3%. The clinical and biological characteristics of the chemorefractory CLL cohort ($n = 48$) are summarized in Table S9. The median overall survival from chemorefractoriness was 1.6 yr. The clinical and biological characteristics of the RS cohort ($n = 58$) are summarized in Table S8. The median overall survival from RS transformation was 1.1 yr.

The study was approved by the Institutional Review Board of Columbia University and by the Ethical Committee of the Azienda Ospedaliera Maggiore della Carità di Novara, Amedeo Avogadro University of Eastern Piedmont.

IGHV mutational status. *IGHV* mutational status was performed as previously reported (Rossi et al., 2009). PCR products were directly sequenced with the ABI PRISM BigDye Terminator v1.1 Ready Reaction Cycle Sequencing kit (Applied Biosystems) using the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Sequences were aligned to the ImmunoGeneTics sequence directory and considered mutated if homology to the corresponding germline gene was <98% (Damle et al., 1999; Hamblin et al., 1999).

FISH analysis. The following probes were used for FISH analysis of CLL and RS cases: LSID13S319, CEP12, LSIp53, LSIATM, LSI IGH/BCL2, LSI IGH/CCND1, LSI BCL6, LSI IGH/MYC/CEP8, *MYC* break-apart, LSI N-MYC, CEP2, CEP3, CEP11, CEP18, CEP19 (Abbott); a BCL-3 split signal probe (Dako); 6q21/ α -satellite (Kreatech Biotechnology); BAC clones 373L24-REL, 440P05-BCL11A, RP11-177O8-BIRC3. The detailed protocol used for FISH studies is described elsewhere (Rossi et al., 2009).

Whole-exome capture and massively parallel sequencing. Whole-exome capture and next generation sequencing of the five CLL tumor/normal DNA samples were performed using the NimbleGen Sequence Capture 2.1M Human Exome Array and the 454 Genome Sequencer FLX instrument, according to the manufacturer's protocol. The array features oligonucleotides for ~180,000 coding exons and 551 miRNA exons, corresponding to ~85% of nonrepetitive sequences in the CCDS database (total size of target region, 34 Mb). The resulting target-enriched pool was amplified and subjected to high-throughput sequencing on the Genome Sequencer FLX Instrument of 454 Life Sciences using 2 PicoTiterPlate runs per sample (~1 million reads/plate, at 400-bp reads). All procedures were performed at Roche and 454 Life Sciences.

454 sequencing data analysis. The computational pipeline used for the analysis of the deep sequencing data is reported in Fig. S1.

The sequencing reads obtained from the Roche FLX454 Sequencer were aligned to the human reference genome (National Center for Biotechnology Information build 36; hg18 assembly) with the Genome Sequencer Reference Mapper, version 2.3 (Roche). This application discards sequencing reads aligning to multiple positions of the reference sequence. Furthermore, because duplicate reads of a single DNA origin are a known potential artifact of the emulsion PCR step, groups of reads mapping to the same basepair of the reference sequence are counted as a single read in the subsequent computation of variants.

The alignments of the uniquely mapped reads to the reference genome were used to define the genetic variants in the CLL tumor and normal DNAs as single base-pair substitutions, small insertions, and small deletions (indels). The Mapper application declares the presence of a variant allele with high-confidence, if there are at least three nonduplicate reads harboring the variant (two in one sense and one in the opposite). Variant candidates in tumor samples were identified with the GSMapper software, using high confidence variant detection method. Germline variants were filtered using a less conservative

approach (presence in ≥ 1 read) and by comparison with common SNPs reported in the National Center for Biotechnology Information database. Using this observation, one can predict the sensitivity of the high-confidence criteria to detect heterozygous variant alleles in a sample containing 100% cancer genetic material as the fraction of sites covered with a depth of at least 6. Assuming that the two alleles of a heterozygous variant are covered with equal probability, a more precise prediction for the sensitivity to detect such variants is given by the high-confidence coverage (HCC) computed by the formula:

$$HCC = \frac{\sum_{k=m}^{\infty} \frac{N_k}{2^k N_T}}{\sum_{n=3}^k \binom{k}{n}}$$

where N_k corresponds to the number of basepairs covered by k sequencing reads and N_T corresponds to the total number of basepairs covered.

Variants obtained by the high-confidence criterion of the Mapper application were filtered further for nonsilent exonic mutations and mutations affecting a nucleotide within 4 nt of a consensus splice site. Furthermore, germline variants reported in publicly available databases (dbSNP, Ensembl Database, and UCSC Genome Browser) and/or identified in the paired normal DNAs were removed. Finally, variants caused by known systematic errors of the 454 sequencing technology (small indels affecting homopolymer repeats) and variants below the sensitivity of direct Sanger sequencing (i.e., observed in $<20\%$ of the reads), were also discarded.

Validation of candidate somatic mutations by Sanger sequencing.

Candidate nonsilent somatic mutations were subjected to validation by conventional Sanger-based resequencing analysis of PCR products obtained from both normal and tumor high molecular weight genomic DNA using primers specific for the exon encompassing the variant. The sequences surrounding the genomic locations of the candidate tumor-specific nonsilent mutations identified through WES were obtained from the UCSC Human Genome database (validation phase). The validation screen allowed us to eliminate additional private polymorphisms that had been missed by the next generation sequencing approach, presumably because of low coverage of the normal sample, as well as false positives caused by errors introduced during the sequencing procedure and/or by mapping errors, as it may occur in the case of repetitive regions or highly homologous regions in the genome. Confirmed nonsilent mutations were tested for their functional consequences in silico by using the PolyPhen-2 (polymorphism phenotyping) algorithm, which is based on structure/sequence conservation.

Mutation screening of identified genes. The complete coding sequences and exon/intron junctions of genes identified through the WES approach, including *NOTCH1* (under GenBank/EMBL/DBJ accession no. NM_017617.2), were analyzed in a larger dataset (screening panel, $n = 48$ CLL cases) by PCR amplification and direct sequencing of whole genome amplified DNA obtained from two independent aliquots of high molecular weight genomic DNA (12 ng each) using the REPLI-g Mini kit (QIAGEN).

Sequences for all annotated exons and flanking introns were retrieved from the UCSC Human Genome database, using the corresponding mRNA accession no. as a reference. PCR primers, located ≥ 50 bp upstream or downstream to target exon boundaries, were either derived from previously published studies or designed in the Primer 3 program, and filtered using UCSC In Silico PCR to exclude pairs yielding more than a single product. All PCR primers and conditions are available upon request. The oligonucleotides used for the validation phase produced PCR products and chromatograms evaluable in $\sim 99.3\%$ of the target regions in the 5 CLL cases belonging to the discovery panel. The vast majority of the oligonucleotides used for the screening phase ($\sim 98\%$) yielded PCR products and high quality sequencing in 40 or more of the 48 screening samples. The *NOTCH1* gene was also screened in an independent panel of 120 CLL diagnostic samples, 58 RS cases, and 48 chemorefractory CLL. *NOTCH2* (accession no. NM_024408.3) and *FBXW7* (accession no. NM_033632.2) were screened in 32 RS cases. All PCR primers and conditions are available upon request.

Purified amplicons were sequenced by conventional Sanger method (Genewiz, Inc.) and compared with the corresponding germline sequences using the Mutation Surveyor Version 2.41 software package (SoftGenetics) after automated and/or manual curation. All variants were sequenced from both strands on independent PCR products obtained from high molecular weight genomic DNA. Synonymous mutations, previously reported polymorphisms, and changes present in the matched normal DNA, which was available for all samples, were removed from the analysis. To rule out the possibility of contamination across samples, DNA of selected RS cases harboring the highly recurrent $\Delta CT(7544-7545)$ *NOTCH1* mutation was re-extracted from the original tissue biopsy and retested in an independent laboratory, which confirmed the presence of the mutation.

Copy number analysis by high-density SNP array analysis. Genome-wide DNA profiles were obtained from high molecular weight tumor and normal genomic DNA of the five CLL cases of the discovery panel using the Genome-Wide Human SNP Array 6.0 (Affymetrix). In brief, DNA was restriction enzyme digested, ligated, PCR-amplified, purified, labeled, fragmented, and hybridized to the arrays according to the manufacturer's instructions. Data analysis was performed as previously described (Mullighan et al., 2007; Pounds et al., 2009).

RNA extraction and gene expression profile analysis. Adequate material for extraction of RNA was available for three of the five CLL discovery cases, which were characterized by gene expression profile analysis. Total RNA extraction was performed by TRIzol (Invitrogen) following the manufacturer's instruction. Gene expression profile analysis of normal mature B cells (naive, memory, centroblasts, and centrocytes) and primary CLL cases was performed using Affymetrix HG-U133Plus 2.0 arrays. The cRNA was labeled and hybridized to the array according to the manufacturer's protocol. The data were analyzed with Microarray Suite version 5.0 (MAS 5.0) using Affymetrix default analysis settings and global scaling as normalization method. Probe sets that had a mean (μ) < 50 , or a coefficient of variation (CV) < 0.3 , as assessed in a panel of 259 independent B cell-related gene expression profiles analyzed on the same array, were filtered out and considered not informative. Genes found mutated in the WES analysis were considered expressed if the corresponding probes were called "Present" in the CLL case carrying the mutation and/or in $> 90\%$ of CLL cases from an independent panel hybridized to the same platform ($n = 16$).

Backtracking of RS-associated *NOTCH1* mutation in the paired CLL diagnostic sample through 454-ultradeep sequencing.

Oligonucleotides containing the *NOTCH1* specific sequences, along with the 10-bp MID tag and amplicon library A and B sequencing adapters, were used to amplify the 349-bp target region of *NOTCH1* exon 34 carrying the recurrent *NOTCH1* mutation $\Delta CT(7544-7545)$, P2515fs (see Results) in the CLL and RS phases from case 63. PCR primers and conditions are available upon request. Purified PCR products were quantitated, diluted, pooled, and amplified by emulsion PCR. The obtained amplicon library was subjected to ultradeep sequencing on the Genome Sequencer Junior instrument (454 Life Sciences), according to the manufacturer's instructions. The obtained sequencing reads were mapped to the reference sequence by Amplicon Variant Analyzer (Roche).

To identify subclonal variants in the CLL phase, we used the BackTracker algorithm, developed in the Rabadan laboratory at Columbia University. We used the long-read component of BWA aligner (Li and Durbin, 2010) to map the reads to the *NOTCH1* reference genome (hg18). Next, using in-house developed software for variant detection, we compiled a list of substitutions, insertions, and deletions. Positions with Phred score below 30 were removed. For deletions, we assigned the mean Phred score of the neighboring mapped positions. Excluding variants not seen in reads mapped to both forward and reverse strands, we fitted a negative binomial or Luria-Delbruck distribution to variant depths, for which we applied a nonlinear least square fit to values larger than one. Unlike the approach presented by Campbell et al. (Campbell et al., 2008), homo-polymers were not eliminated

(*NOTCH1* mutation was located in a homo-polymer region), no artificial depth cut-offs were imposed, and higher Phred score values were considered. Probabilities for candidate variants were corrected to take into account multiple positions (Bonferroni or Sidac). Three threshold values were established for substitutions, deletions, and insertions (Fig. 5 b, Fig. S4, and Fig. S5). In the CLL phase, we found the C deletion at position 7544 at 5.3% frequency and the T deletion at 7545 at 6.6% frequency, both passing the cutoffs.

Functional categories and pathway analysis of the mutated genes.

The genes found to be mutated in CLL were assigned to functional categories or to annotated pathways using the gene ontology Gene Ontology database (<http://www.geneontology.org/>), the publicly available bioinformatic tool DAVID 2008 (Database for Annotation, Visualization and Integrated Discovery; <http://david.abcc.ncifcrf.gov/>), and the Molecular Signatures Database from the Broad Institute (MSigDB; <http://www.broadinstitute.org/gsea/msigdb/index.jsp>).

Analysis of clinical correlations. OS of the consecutive series of newly diagnosed and previously untreated CLL was measured from date of diagnosis to date of death or last follow up (censoring). Treatment free survival (TFS) was measured from date of CLL diagnosis to date of progressive and symptomatic disease requiring treatment according to IWCLL-NCI Working Group guidelines, death or last follow up (censoring; Hallek et al., 2008). Refractory disease was defined as treatment failure (stable disease or progressive disease during treatment) or disease progression within 6 mo from antileukemic therapy (Hallek et al., 2008). OS of the chemorefractory CLL cohort was measured from date of chemorefractoriness to date of death or last follow up (censoring). RS survival was measured from date of RS diagnosis to date of death or last follow up (censoring). The date of RS diagnosis was defined as the date of the diagnostic biopsy. Survival analysis was performed by the Kaplan-Meier method. Multivariate analysis was performed by the Cox regression model. All statistical tests were two-sided. Statistical significance was defined as P value <0.05. The analysis was performed with the Statistical Package for the Social Sciences (SPSS) software v.18.0 (Chicago, IL).

Accession nos. The WES data and the SNP array data from the five discovery CLL cases will be deposited in dbGaP under accession no. phs000364.v1.p1.

Online supplemental material. Fig. S1 shows the approach used for WES analysis. Fig. S2 shows the mutation spectrum in the five CLL discovery cases. Fig. S3 shows the known CLL-associated copy number aberrations identified by SNP array analysis in the 5 CLL discovery cases. Figs. S4 and S5 show the results of the ultra-deep 454 DNA sequencing analysis of *NOTCH1* in one CLL/RS pair. Table S1 summarizes the features of the 5 CLL discovery cases. Table S2 shows the results of the 454 sequencing and mapping after whole-exome capture in the five CLL discovery cases. Table S3 lists the validated somatic mutations identified by WES in the CLL discovery panel. Table S4 lists the tumor-acquired copy number changes identified in the five CLL discovery cases. Tables S5, S6, S8, and S9 summarize the clinical and biological characteristics of the CLL screening panel, CLL extension panel, RS cohort, and chemorefractory CLL cohort, respectively. Table S7 lists the *NOTCH1* mutations found in the study. Online supplemental material is available at <http://www.jem.org/cgi/content/full/jem.20110921/DC1>.

We would like to thank T. Palomero, E. Tzilianos, V. Miljkovic, and the Genomics Technologies Shared Resource of the Herbert Irving Comprehensive Cancer Center at Columbia University for help with the SNP array hybridization, and B. Boese at 454 Life Sciences for assistance with the whole-exome capture and sequencing procedure. We would also like to thank M. Iacono at Roche Italy for assistance with the Roche 454 Junior backtrack experiments.

This study was supported by National Institutes of Health grant CA-37295 (to R. Della-Favera); the Italian Association for Cancer Research (AIRC), Special Program Molecular Clinical Oncology, AIRC 5 × 1000, no. 10007, Milan, Italy (to G. Gaidano and to R. Foà); Progetto FIRB-Programma "Futuro in Ricerca" 2008 (to D. Rossi) and PRIN 2008 (to G. Gaidano), MIUR, Rome, Italy; Progetto Giovani Ricercatori 2008, Ministero della Salute, Rome, Italy (to D. Rossi); Novara-AIL Onlus, Novara (to

G. Gaidano); Helmut Horten Foundation and San Salvatore Foundation (to F. Bertoni). The work of V. Trifonov, H. Khabanian, and R. Rabadan is supported by the Northeast Biodefense Center (U54-AI057158), the National Institutes of Health (U54 CA121852-05), and the National Library of Medicine (1R01LM010140-01). S. Monti and S. Cresta are being supported by fellowships from Novara-AIL Onlus. L. Pasqualucci is on leave from the Institute of Hematology, University of Perugia Medical School.

The authors have no competing financial interests.

Submitted: 6 May 2011

Accepted: 26 May 2011

REFERENCES

- Adzhubei, I.A., S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, and S.R. Sunyaev. 2010. A method and server for predicting damaging missense mutations. *Nat. Methods*. 7:248–249. doi:10.1038/nmeth0410-248
- Albesiano, E., B.T. Messmer, R.N. Damle, S.L. Allen, K.R. Rai, and N. Chiorazzi. 2003. Activation-induced cytidine deaminase in chronic lymphocytic leukemia B cells: expression as multiple forms in a dynamic, variably sized fraction of the clone. *Blood*. 102:3333–3339. doi:10.1182/blood-2003-05-1585
- Annunziata, C.M., R.E. Davis, Y. Demchenko, W. Bellamy, A. Gabrea, F. Zhan, G. Lenz, I. Hanamura, G. Wright, W. Xiao, et al. 2007. Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma. *Cancer Cell*. 12:115–130. doi:10.1016/j.ccr.2007.07.004
- Aster, J.C., W.S. Pear, and S.C. Blacklow. 2008. Notch signaling in leukemia. *Annu. Rev. Pathol.* 3:587–613. doi:10.1146/annurev.pathmechdis.3.121806.154300
- Aster, J.C., S.C. Blacklow, and W.S. Pear. 2011. Notch signalling in T-cell lymphoblastic leukaemia/lymphoma and other haematological malignancies. *J. Pathol.* 223:262–273. doi:10.1002/path.2789
- Calin, G.A., C.D. Dumitru, M. Shimizu, R. Bichi, S. Zupo, E. Noch, H. Alder, S. Rattan, M. Keating, K. Rai, et al. 2002. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA*. 99:15524–15529. doi:10.1073/pnas.242606799
- Campbell, P.J., E.D. Pleasance, P.J. Stephens, E. Dicks, R. Rance, I. Goodhead, G.A. Follows, A.R. Green, P.A. Futreal, and M.R. Stratton. 2008. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc. Natl. Acad. Sci. USA*. 105:13081–13086. doi:10.1073/pnas.0801523105
- Chapman, M.A., M.S. Lawrence, J.J. Keats, K. Cibulskis, C. Sougnez, A.C. Schinzel, C.L. Harview, J.P. Brunet, G.J. Ahmann, M. Adli, et al. 2011. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 471:467–472. doi:10.1038/nature09837
- Chiorazzi, N., K.R. Rai, and M. Ferrarini. 2005. Chronic lymphocytic leukemia. *N. Engl. J. Med.* 352:804–815. doi:10.1056/NEJMra041720
- Damle, R.N., T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S.L. Allen, A. Buchbinder, D. Budman, K. Dittmar, J. Kolitz, et al. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 94:1840–1847.
- Delker, R.K., S.D. Fugmann, and F.N. Papavasiliou. 2009. A coming-of-age story: activation-induced cytidine deaminase turns 10. *Nat. Immunol.* 10:1147–1153. doi:10.1038/ni.1799
- Di Ianni, M., S. Baldoni, E. Rosati, R. Ciurnelli, L. Cavalli, M.F. Martelli, P. Marconi, I. Screpanti, and F. Falzetti. 2009. A new genetic lesion in B-CLL: a *NOTCH1* PEST domain mutation. *Br. J. Haematol.* 146:689–691. doi:10.1111/j.1365-2141.2009.07816.x
- Döhner, H., S. Stilgenbauer, A. Benner, E. Leupolt, A. Kröber, L. Bullinger, K. Döhner, M. Bentz, and P. Lichter. 2000. Genomic aberrations and survival in chronic lymphocytic leukemia. *N. Engl. J. Med.* 343:1910–1916. doi:10.1056/NEJM200012283432602
- Dominguez-Sola, D., C.Y. Ying, C. Grandori, L. Ruggiero, B. Chen, M. Li, D.A. Galloway, W. Gu, J. Gautier, and R. Dalla-Favera. 2007. Non-transcriptional control of DNA replication by c-Myc. *Nature*. 448:445–451. doi:10.1038/nature05953

- Eischen, C.M., J.D. Weber, M.F. Roussel, C.J. Sherr, and J.L. Cleveland. 1999. Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis. *Genes Dev.* 13:2658–2669. doi:10.1101/gad.13.20.2658
- Fangazio, M., L. De Paoli, D. Rossi, and G. Gaidano. 2011. Predictive markers and driving factors behind Richter syndrome development. *Expert Rev. Anticancer Ther.* 11:433–442. doi:10.1586/era.10.237
- Forbes, S.A., N. Bindal, S. Bamford, C. Cole, C.Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, et al. 2011. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* 39:D945–D950. doi:10.1093/nar/gkq929
- Grabher, C., H. von Boehmer, and A.T. Look. 2006. Notch 1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukaemia. *Nat. Rev. Cancer.* 6:347–359. doi:10.1038/nrc1880
- Greenman, C., P. Stephens, R. Smith, G.L. Dalgliesh, C. Hunter, G. Bignell, H. Davies, J. Teague, A. Butler, C. Stevens, et al. 2007. Patterns of somatic mutation in human cancer genomes. *Nature.* 446:153–158. doi:10.1038/nature05610
- Grubor, V., A. Krasnitz, J.E. Troge, J.L. Meth, B. Lakshmi, J.T. Kendall, B. Yamrom, G. Alex, D. Pai, N. Navin, et al. 2009. Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). *Blood.* 113:1294–1303. doi:10.1182/blood-2008-05-158865
- Gyrd-Hansen, M., and P. Meier. 2010. IAPs: from caspase inhibitors to modulators of NF-kappaB, inflammation and cancer. *Nat. Rev. Cancer.* 10:561–574. doi:10.1038/nrc2889
- Hallek, M., B.D. Cheson, D. Catovsky, F. Caligaris-Cappio, G. Dighiero, H. Döhner, P. Hillmen, M.J. Keating, E. Montserrat, K.R. Rai, and T.J. Kipps; International Workshop on Chronic Lymphocytic Leukemia. 2008. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood.* 111:5446–5456. doi:10.1182/blood-2007-06-093906
- Hamblin, T.J., Z. Davis, A. Gardiner, D.G. Oscier, and F.K. Stevenson. 1999. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood.* 94:1848–1854.
- Huh, Y.O., K.I. Lin, F. Vega, E. Schlette, C.C. Yin, M.J. Keating, R. Luthra, L.J. Medeiros, and L.V. Abruzzo. 2008. MYC translocation in chronic lymphocytic leukaemia is associated with increased prolymphocytes and a poor prognosis. *Br. J. Haematol.* 142:36–44. doi:10.1111/j.1365-2141.2008.07152.x
- Jones, S., X. Zhang, D.W. Parsons, J.C. Lin, R.J. Leary, P. Angenendt, P. Mankoo, H. Carter, H. Kamiyama, A. Jimeno, et al. 2008. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science.* 321:1801–1806. doi:10.1126/science.1164368
- Keats, J.J., R. Fonseca, M. Chesi, R. Schop, A. Baker, W.J. Chng, S. Van Wier, R. Tiedemann, C.X. Shi, M. Sebag, et al. 2007. Promiscuous mutations activate the noncanonical NF-kappaB pathway in multiple myeloma. *Cancer Cell.* 12:131–144. doi:10.1016/j.ccr.2007.07.003
- Klein, U., and R. Dalla-Favera. 2010. New insights into the pathogenesis of chronic lymphocytic leukemia. *Semin. Cancer Biol.* 20:377–383. doi:10.1016/j.semcancer.2010.10.012
- Klein, U., M. Lia, M. Crespo, R. Siegel, Q. Shen, T. Mo, A. Ambesi-Impiomato, A. Califano, A. Migliazza, G. Bhagat, and R. Dalla-Favera. 2010. The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell.* 17:28–40. doi:10.1016/j.ccr.2009.11.019
- Kujawski, L., P. Ouillette, H. Erba, C. Saddler, A. Jakubowiak, M. Kaminski, K. Shedden, and S.N. Malek. 2008. Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. *Blood.* 112:1993–2003. doi:10.1182/blood-2007-07-099432
- Lee, S.Y., K. Kumano, K. Nakazaki, M. Sanada, A. Matsumoto, G. Yamamoto, Y. Nannya, R. Suzuki, S. Ota, Y. Ota, et al. 2009. Gain-of-function mutations and copy number increases of Notch2 in diffuse large B-cell lymphoma. *Cancer Sci.* 100:920–926. doi:10.1111/j.1349-7006.2009.01130.x
- Lenz, G., G.W. Wright, N.C. Emre, H. Kohlhammer, S.S. Dave, R.E. Davis, S. Carty, L.T. Lam, A.L. Shaffer, W. Xiao, et al. 2008. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl. Acad. Sci. USA.* 105:13520–13525. doi:10.1073/pnas.0804295105
- Ley, T.J., E.R. Mardis, L. Ding, B. Fulton, M.D. McLellan, K. Chen, D. Dooling, B.H. Dunford-Shore, S. McGrath, M. Hickenbotham, et al. 2008. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature.* 456:66–72. doi:10.1038/nature07485
- Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 26:589–595. doi:10.1093/bioinformatics/btp698
- Li, X., T. Sanda, A.T. Look, C.D. Novina, and H. von Boehmer. 2011. Repression of tumor suppressor miR-451 is essential for NOTCH1-induced oncogenesis in T-ALL. *J. Exp. Med.* 208:663–675. doi:10.1084/jem.20102384
- Mardis, E.R., L. Ding, D.J. Dooling, D.E. Larson, M.D. McLellan, K. Chen, D.C. Koboldt, R.S. Fulton, K.D. Delehaanty, S.D. McGrath, et al. 2009. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* 361:1058–1066. doi:10.1056/NEJMoa0903840
- Maystadt, I., R. Rezsöházy, M. Barkats, S. Duque, P. Vannuffel, S. Remacle, B. Lambert, M. Najimi, E. Sokal, A. Munnich, et al. 2007. The nuclear factor kappaB-activator gene PLEKHG5 is mutated in a form of autosomal recessive lower motor neuron disease with childhood onset. *Am. J. Hum. Genet.* 81:67–76. doi:10.1086/518900
- Mullighan, C.G., S. Goorha, I. Radtke, C.B. Miller, E. Coustan-Smith, J.D. Dalton, K. Girtman, S. Mathew, J. Ma, S.B. Pounds, et al. 2007. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature.* 446:758–764. doi:10.1038/nature05690
- Ngo, V.N., R.M. Young, R. Schmitz, S. Jhavar, W. Xiao, K.-H. Lim, H. Kohlhammer, W. Xu, Y. Yang, H. Zhao, et al. 2011. Oncogenically active MYD88 mutations in human lymphoma. *Nature.* 470:115–119. doi:10.1038/nature09671
- O'Neil, J., J. Calvo, K. McKenna, V. Krishnamoorthy, J.C. Aster, C.H. Bassing, F.W. Alt, M. Kelliher, and A.T. Look. 2006. Activating Notch1 mutations in mouse models of T-ALL. *Blood.* 107:781–785. doi:10.1182/blood-2005-06-2553
- O'Neil, J., J. Grim, P. Strack, S. Rao, D. Tibbitts, C. Winter, J. Hardwick, M. Welcker, J.P. Meijerink, R. Pieters, et al. 2007. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to γ -secretase inhibitors. *J. Exp. Med.* 204:1813–1824. doi:10.1084/jem.20070876
- Paganin, M., and A. Ferrando. 2011. Molecular pathogenesis and targeted therapies for NOTCH1-induced T-cell acute lymphoblastic leukemia. *Blood Rev.* 25:83–90. doi:10.1016/j.blre.2010.09.004
- Palomero, T., W.K. Lim, D.T. Odom, M.L. Sulis, P.J. Real, A. Margolin, K.C. Barnes, J. O'Neil, D. Neuberg, A.P. Weng, et al. 2006. NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proc. Natl. Acad. Sci. USA.* 103:18261–18266. doi:10.1073/pnas.0606108103
- Parsons, D.W., M. Li, X. Zhang, S. Jones, R.J. Leary, J.C. Lin, S.M. Boca, H. Carter, J. Samayoa, C. Bettegowda, et al. 2011. The genetic landscape of the childhood cancer medulloblastoma. *Science.* 331:435–439. doi:10.1126/science.1198056
- Pasqualucci, L., D. Dominguez-Sola, A. Chiarenza, G. Fabbri, A. Grunn, V. Trifonov, L.H. Kasper, S. Lerach, H. Tang, J. Ma, et al. 2011. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature.* 471:189–195. doi:10.1038/nature09730
- Pleasant, E.D., R.K. Cheetham, P.J. Stephens, D.J. McBride, S.J. Humphray, C.D. Greenman, I. Varela, M.L. Lin, G.R. Ordóñez, G.R. Bignell, et al. 2010a. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature.* 463:191–196. doi:10.1038/nature08658
- Pleasant, E.D., P.J. Stephens, S. O'Meara, D.J. McBride, A. Meynert, D. Jones, M.-L. Lin, D. Beare, K.W. Lau, C. Greenman, et al. 2010b. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature.* 463:184–190. doi:10.1038/nature08629
- Pounds, S., C. Cheng, C. Mullighan, S.C. Raimondi, S. Shurtliff, and J.R. Downing. 2009. Reference alignment of SNP microarray signals for copy number analysis of tumors. *Bioinformatics.* 25:315–321. doi:10.1093/bioinformatics/btm624

- Real, P.J., V. Tosello, T. Palomero, M. Castillo, E. Hernando, E. de Stanchina, M.L. Sulis, K. Barnes, C. Sawai, I. Homminga, et al. 2009. Gamma-secretase inhibitors reverse glucocorticoid resistance in T cell acute lymphoblastic leukemia. *Nat. Med.* 15:50–58. doi:10.1038/nm.1900
- Rosebeck, S., L. Madden, X. Jin, S. Gu, I.J. Apel, A. Appert, R.A. Hamoudi, H. Noels, X. Sagaert, P. Van Loo, et al. 2011. Cleavage of NIK by the API2-MALT1 fusion oncoprotein leads to noncanonical NF-kappaB activation. *Science*. 331:468–472. doi:10.1126/science.1198946
- Rossi, D., and G. Gaidano. 2009. Richter syndrome: molecular insights and clinical perspectives. *Hematol. Oncol.* 27:1–10. doi:10.1002/hon.880
- Rossi, D., V. Spina, M. Cerri, S. Rasi, C. Deambrogi, L. De Paoli, L. Laurenti, R. Maffei, F. Forconi, F. Bertoni, et al. 2009. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clin. Cancer Res.* 15:4415–4422. doi:10.1158/1078-0432.CCR-08-3266
- Rossi, D., V. Spina, C. Deambrogi, S. Rasi, L. Laurenti, K. Stamatopoulos, L. Arcaini, M. Lucioni, G.B. Rocque, Z.Y. Xu-Monette, et al. 2011. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood*. 117:3391–3401. doi:10.1182/blood-2010-09-302174
- Schmitt, C.A., J.S. Fridman, M. Yang, E. Baranov, R.M. Hoffman, and S.W. Lowe. 2002. Dissecting p53 tumor suppressor functions in vivo. *Cancer Cell*. 1:289–298. doi:10.1016/S1535-6108(02)00047-8
- Sportoletti, P., S. Baldoni, L. Cavalli, B. Del Papa, E. Bonifacio, R. Ciurnelli, A.S. Bell, A. Di Tommaso, E. Rosati, B. Crescenzi, et al. 2010. NOTCH1 PEST domain mutation is an adverse prognostic factor in B-CLL. *Br. J. Haematol.* 151:404–406. doi:10.1111/j.1365-2141.2010.08368.x
- Staudt, L.M. 2010. Oncogenic activation of NF-kappaB. *Cold Spring Harb. Perspect. Biol.* 2:a000109. doi:10.1101/cshperspect.a000109
- Sulis, M.L., O. Williams, T. Palomero, V. Tosello, S. Pallikuppam, P.J. Real, K. Barnes, L. Zuurbier, J.P. Meijerink, and A.A. Ferrando. 2008. NOTCH1 extracellular juxtamembrane expansion mutations in T-ALL. *Blood*. 112:733–740. doi:10.1182/blood-2007-12-130096
- Swerdlow, S.H., E. Campo, N.L. Harris, E.S. Jaffe, S.A. Pileri, H. Stein, J. Thiele, and J.W. Vardiman, editors. 2008. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. International Agency for Research on Cancer (IARC), Lyon. 439 pp.
- Tam, C.S., and M.J. Keating. 2010. Chemoimmunotherapy of chronic lymphocytic leukemia. *Nat Rev Clin Oncol.* 7:521–532. doi:10.1038/nrclinonc.2010.101
- Thompson, B.J., S. Buonamici, M.L. Sulis, T. Palomero, T. Vilimas, G. Basso, A. Ferrando, and I. Aifantis. 2007. The SCF^{FBW7} ubiquitin ligase complex as a tumor suppressor in T cell leukemia. *J. Exp. Med.* 204:1825–1835. doi:10.1084/jem.20070872
- Tsimberidou, A.M., and M.J. Keating. 2005. Richter syndrome: biology, incidence, and therapeutic strategies. *Cancer*. 103:216–228. doi:10.1002/cncr.20773
- Weng, A.P., A.A. Ferrando, W. Lee, J.P. Morris IV, L.B. Silverman, C. Sanchez-Irizarry, S.C. Blacklow, A.T. Look, and J.C. Aster. 2004. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science*. 306:269–271. doi:10.1126/science.1102160
- Zenz, T., B. Eichhorst, R. Busch, T. Denzel, S. Häbe, D. Winkler, A. Bühler, J. Edelmann, M. Bergmann, G. Hopfinger, et al. 2010a. TP53 mutation and survival in chronic lymphocytic leukemia. *J. Clin. Oncol.* 28:4473–4479. doi:10.1200/JCO.2009.27.8762
- Zenz, T., D. Mertens, R. Küppers, H. Döhner, and S. Stilgenbauer. 2010b. From pathogenesis to treatment of chronic lymphocytic leukaemia. *Nat. Rev. Cancer*. 10:37–50.
- Zenz, T., D. Vollmer, M. Trbusek, J. Smardova, A. Benner, T. Soussi, H. Helfrich, M. Heuberger, P. Hoth, M. Fuge, et al. 2010c. TP53 mutation profile in chronic lymphocytic leukemia: evidence for a disease specific profile from a comprehensive analysis of 268 mutations. *Leukemia*. 24:2071–2079.
- Zhou, H., M.Q. Du, and V.M. Dixit. 2005. Constitutive NF-kappaB activation by the t(11;18)(q21;q21) product in MALT lymphoma is linked to deregulated ubiquitin ligase activity. *Cancer Cell*. 7:425–431. doi:10.1016/j.ccr.2005.04.012

SUPPLEMENTAL MATERIAL

Fabbri et al., <http://www.jem.org/cgi/content/full/jem.20110921/DC1>

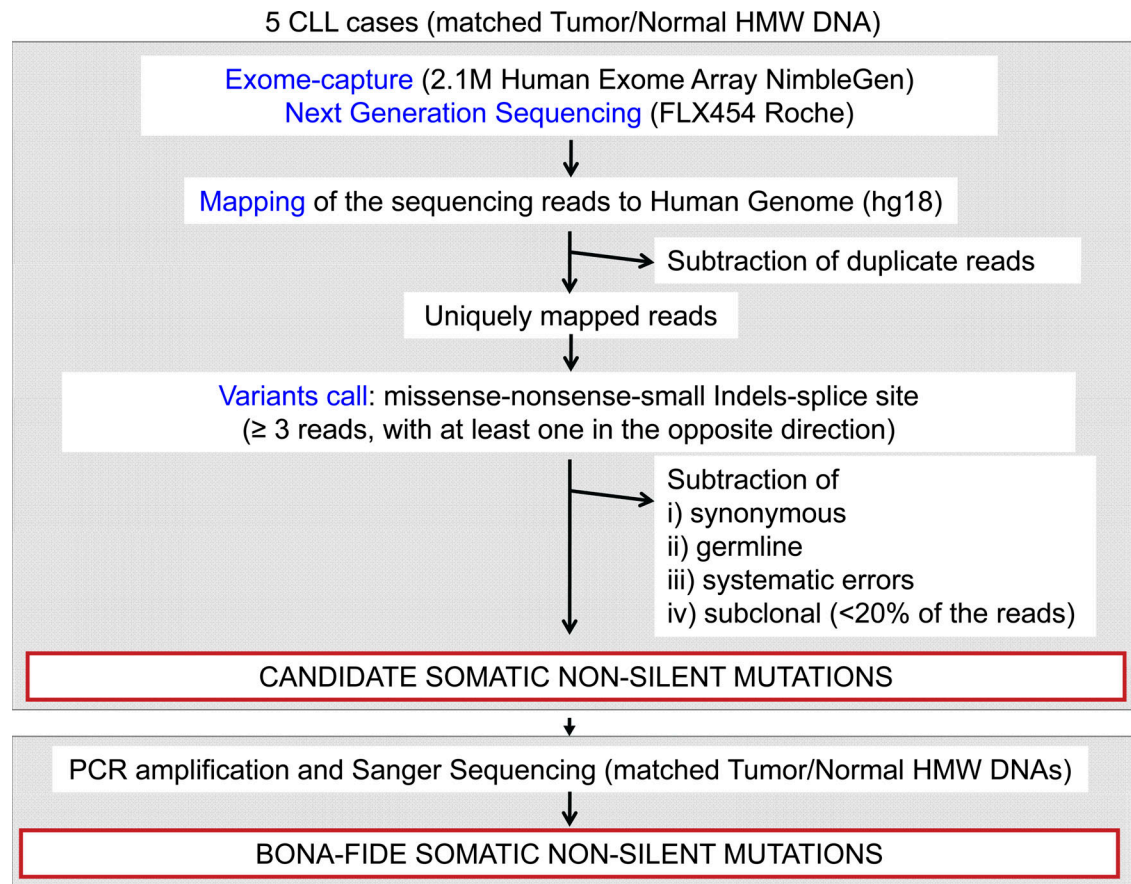


Figure S1. WES approach. Computational pipeline used for the analysis of 454 WES data and the validation of candidate sequence variants.

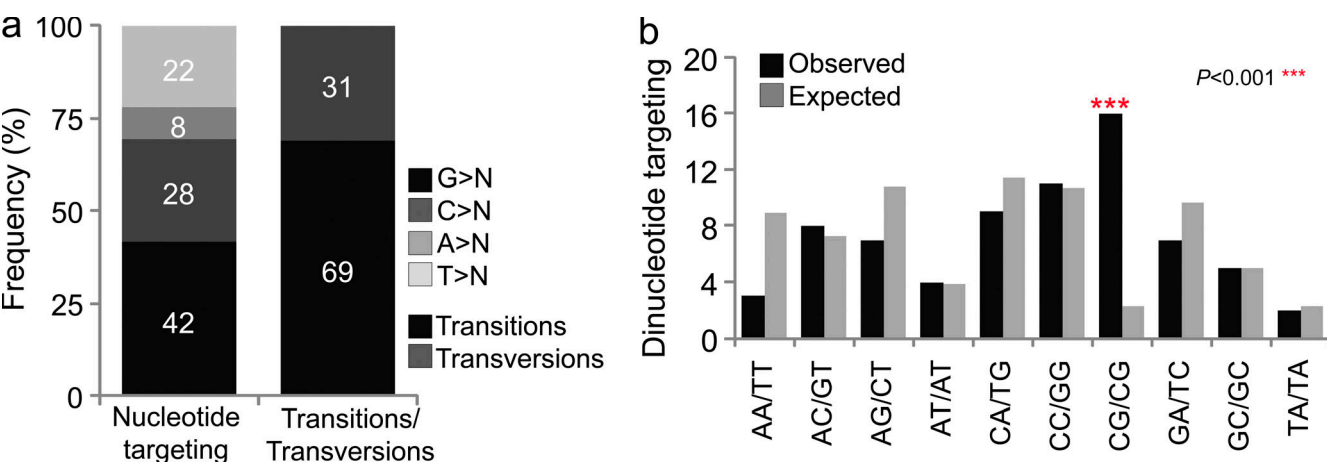


Figure S2. Mutation spectrum in CLL. (a) Nucleotide substitution pattern of nonsynonymous mutations identified in the exomes of the 5 CLL discovery cases. (b) Number of mutations observed at specific dinucleotides (black bars). Gray bars indicate the expected number of events calculated based on the sequence composition of the target exomic region (see Materials and methods for details).

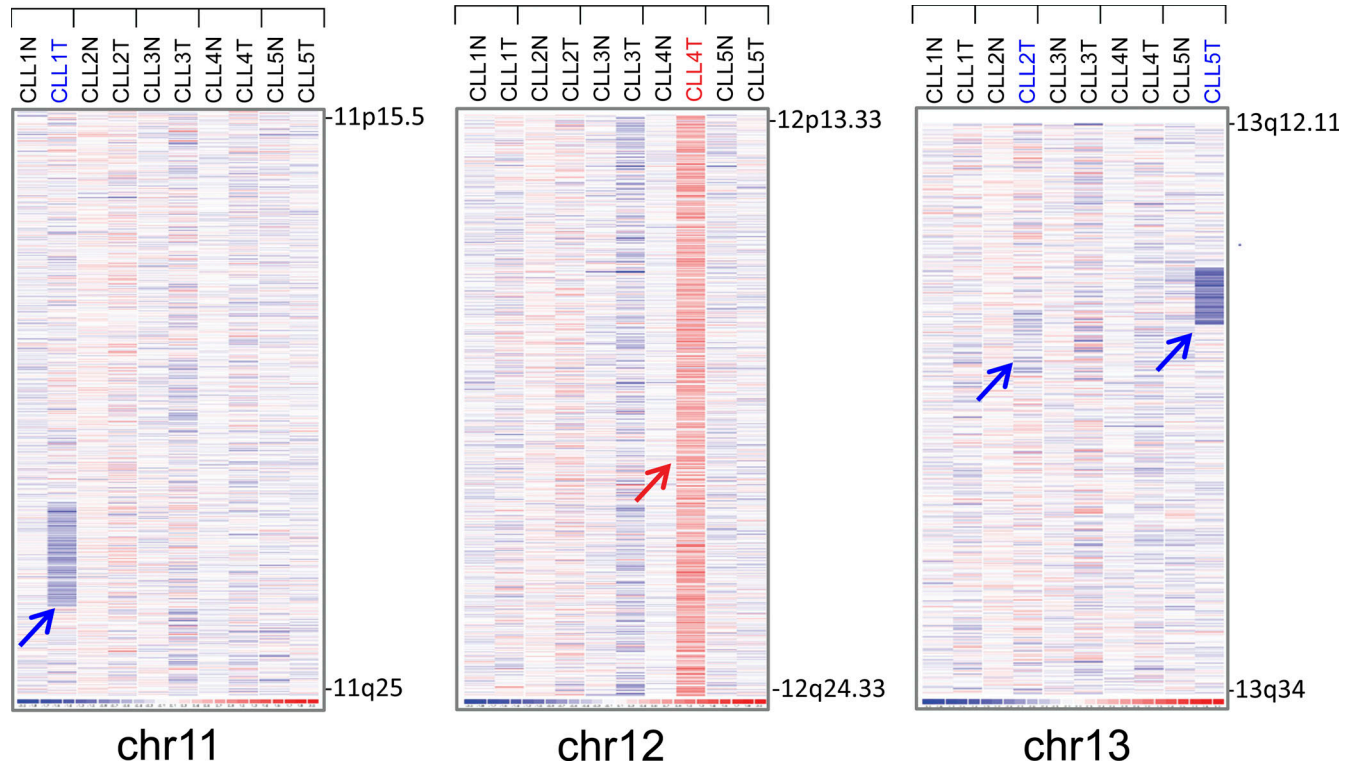


Figure S3. Genome-wide SNP array analysis identifies known CLL-associated CNAs. Heatmap showing median-smoothed $\log_2$ copy number ratio of three chromosomal regions that were found to carry genomic aberrations by SNP array analysis, in concordance with the results of FISH analysis. Arrows point to the aberrant interval. In the red-blue scale, white corresponds to a normal (diploid) copy number log ratio, blue is deletion, and red is gain. CNA, copy number aberrations.

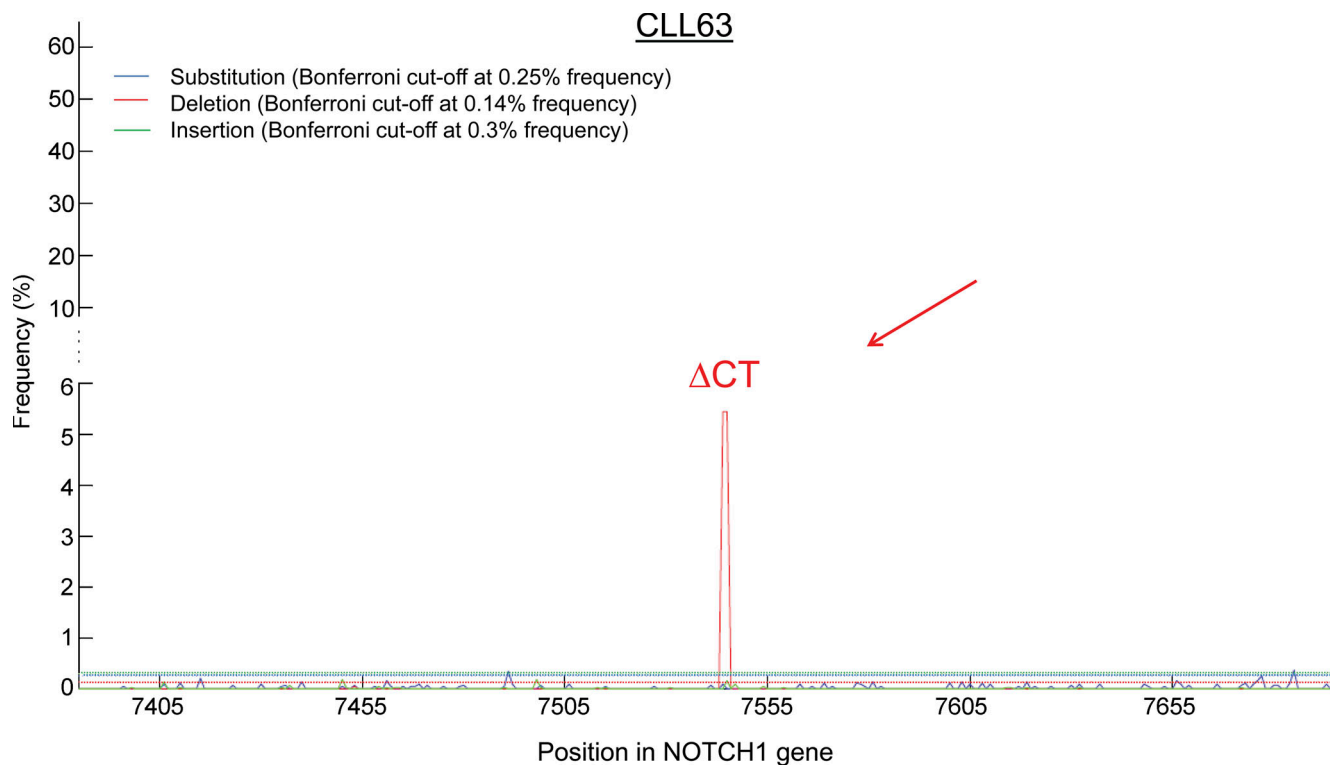


Figure S4. Detection of a subclonal *NOTCH1* mutation in the CLL phase of the CLL63/RS63 pair by ultradeep 454 sequencing. 30 mo elapsed between collection of the CLL sample (diagnosis of CLL) and collection of the RS sample (diagnosis of RS). The CT deletion at position 7544–7545 (available from EMBL/GenBank/DDBJ under accession no. NM_017617.2) was detected at 5.3% frequency, passing the cut-offs established by Bonferroni-corrected probabilities derived from fitting a negative binomial or Luria-Delbruck distribution to variant depths (see Materials and methods). The change in the y-axis scale is devised to compare the cut-off frequencies to the mutation frequencies.

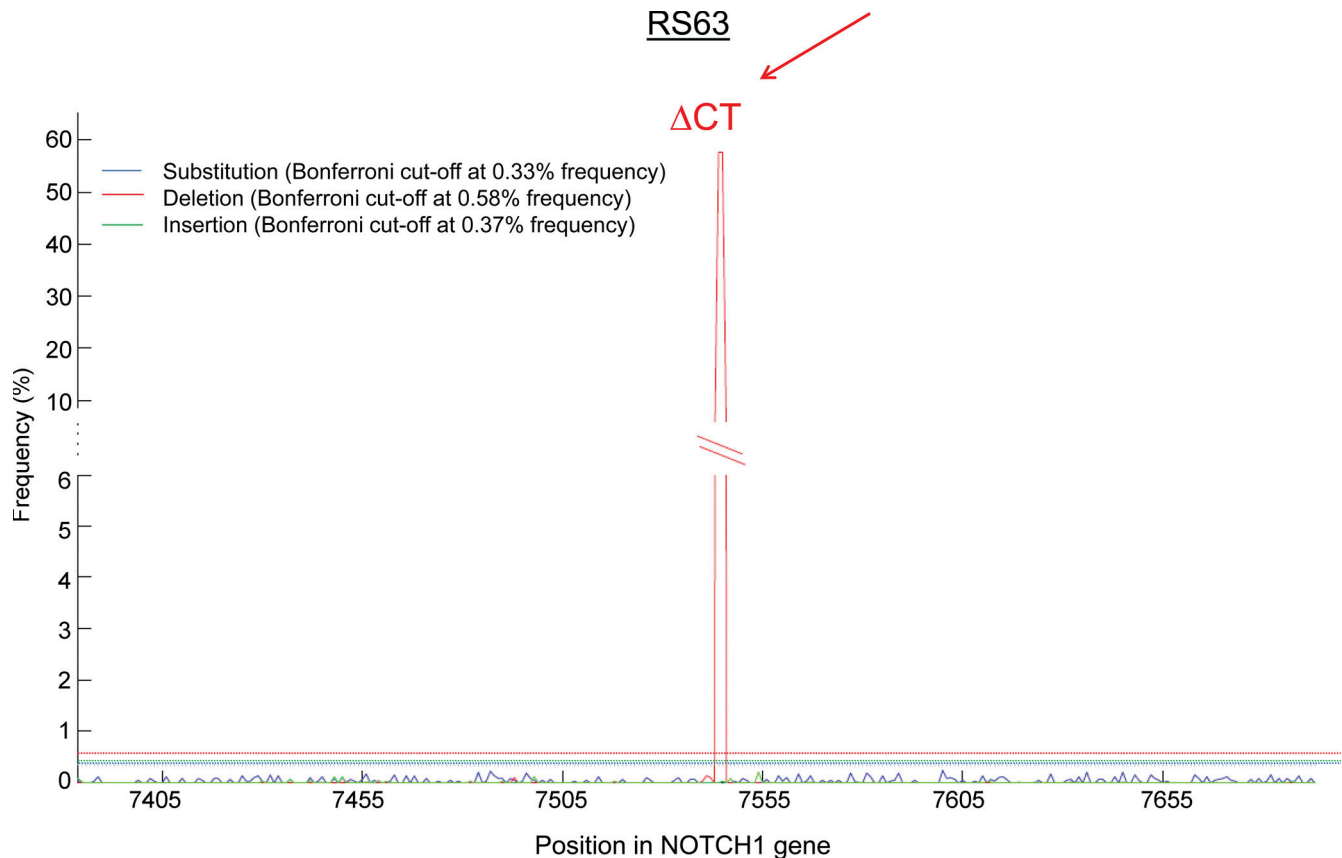


Figure S5. Detection of a *NOTCH1* mutation in the RS phase of the CLL63/RS63 pair through ultradeep 454 sequencing. 30 mo elapsed between collection of the CLL sample (diagnosis of CLL) and collection of the RS sample (diagnosis of RS). The CT deletion at position 7544–7545 (accession no. NM_017617.2) was detected at 58.6% frequency. The change in the y-axis scale is devised to compare the Bonferroni cut-off frequencies (see Materials and methods) to the mutation frequency.

SUPPLEMENTAL RESULTS

Identification and validation of WES candidate somatic mutations

The WES approach provided ~ 2.5 millions reads/case (average length 305 bp), with a mean depth of $10.1\times$ (range: 8.6 to 11.8 per case), and $\sim 79\%$ of the target sequence being covered by at least 5 reads (range: 74.6 to 86.3%; Table S2). Based on comparison with the heterozygous SNP call rate obtained in the same cases by analysis of Affymetrix SNP 6.0 array data, the sensitivity for the identification of heterozygous somatic mutations was estimated to be $\sim 60\%$. This level of resolution allowed us to provide an initial estimate of the order of magnitude of the mutation load in this disease, and to identify genes that are recurrently mutated at high frequency (probability of detecting mutations in genes that are affected at $\sim 30\%$ prevalence = 63%).

Overall, 276 variants were predicted by the WES analysis to be specifically associated with the tumor sample in the 5 patients analyzed. Of these, 38 were confirmed to be somatic in origin by Sanger sequencing of the corresponding paired tumor and normal genomic DNA, whereas 170 were also found in the matched normal genomic DNA, thus representing germline polymorphisms that were not previously annotated and escaped detection during the high-throughput sequencing analysis, possibly because of the relatively low mean depth of coverage. The remaining 68 variants were absent in both tumor and normal genomic DNA when tested by Sanger sequencing.

Table S1. Features of the 5 CLL discovery cases analyzed by Whole Exome Sequencing

Characteristic	CLL patient ID				
	CLL1	CLL2	CLL3	CLL4	CLL5
Age at diagnosis, years	67	71	46	69	63
Sex	F	F	M	F	M
Rai stage	0	IV	II	0	II
Lymphocytes, x10 ⁹ /Liter	38	124	36	16.7	68.5
Hb, g/dl	13.4	9.5	14.4	13.1	12
Platelets, x10 ⁹ /Liter	249	159	166	155	135
B2M, mg/Liter	2.7	6.3	2.2	3.3	3.2
LDH, U/Liter	551	500	340	480	372
Initial management	Wait and see	Treatment	Wait and see	Wait and see	Wait and see
Time to first treatment (months)	86	Therapy started at diagnosis	33	6	9
Transformation to DLBCL	yes	no	no	no	no
Survival (months)	166	55	107	216	108
Status	Death	Alive	Alive	Death	Alive
CD19+/CD5+ cells (%)	90	80	90	90	95
IGHV mutational status	Unmutated	Unmutated	Unmutated	Mutated	Mutated
V gene	1-69*01	3-30*04	1-69*01	4-34*01	3-23*01
D gene	3-3*01	5-24*01	3-16*02	3-22*01	5-5*01
J gene	6*02	6*02	6*02	3*02	4*02
IGHV homology (%)	100	100	100	94.38	92.34
Stereotyped HCDR3	Stereotyped	Not stereotyped	Not stereotyped	Not stereotyped	Not stereotyped
FISH ^a	del11q22-q23	del13q14 ^c	Normal	+12, 14q translocation	del13q14
ZAP-70 expression ^b	positive	negative	positive	negative	negative
CD38 expression ^b	negative	negative	positive	positive	negative
TP53 mutational status	Unmutated	Unmutated	Unmutated	Unmutated	Unmutated

^a based on FISH analysis for del13q, del11q, trisomy12, del17p, and 14q translocation. ^b Evaluated by FACS analysis. ZAP-70 expression was considered positive if the percentage of positive leukemic cells was above 20%, CD38 expression was considered positive if the percentage of positive leukemic cells was above 30%.

^c present in 18% of nuclei.

Table S2. Results of Titanium 454 sequencing and mapping after whole exome capture

Sample ID	Target Region Coverage (%)			Mean Depth ^b	N of mapped reads (%)	Unique reads in region (%)	N of variant reads (HC)		
	1X	≥5X	HCC ^a				Total	Known	Novel ^c
CLL1-T	95.8	74.6	67.2	8.6	2374973 (98.78)	1556726 (68.6)	9730	9163	567
CLL1-N	95.9	77.2	70	9.1	2490456 (99.06)	1679964 (70.4)	10162	9531	631
CLL2-T	96.5	86.3	81	11.8	2684394 (99.35)	2008402 (77.8)	12457	11675	782
CLL2-N	95.8	80.8	75.3	11.2	2643134 (99.18)	1870633 (73.9)	11392	10652	740
CLL3-T	95.7	76.5	69.6	9.1	2399542 (99.17)	1664769 (72.6)	10459	9744	715
CLL3-N	95.8	78.4	73.1	11.1	2531945 (99.30)	1849789 (76.2)	10925	10240	685
CLL4-T	95.7	75.8	68.8	9.1	2157095 (99.14)	1560948 (75.7)	10201	9603	598
CLL4-N	96.2	78.6	71	9	2389230 (99.10)	1633385 (71.7)	10956	10244	712
CLL5-T	96.4	82.2	74.9	10.3	2440574 (98.95)	1746590 (75.3)	12459	11556	903
CLL5-N	96.5	83.7	77.3	11.6	2695359 (99.27)	1969017 (76.3)	12367	11520	847

^a High Confidence Coverage, i.e. % of the target region where coverage depth was estimated to be sufficient for detecting both alleles of heterozygous variants. This value can be taken as an estimate of the mutation-detection sensitivity (see methods).

^b Mean number of sequence reads covering the target exome.

^c N of variant reads not reported in dbSNP database, before cross-comparison with paired normal.

Abbreviations: HC, high confidence.

Table S3. Validated somatic mutations identified by whole exome sequencing in the CLL discovery panel

Sample ID	Gene ID	Mutation type	Chromosomal position ^a	Base change	Aminoacid change	read%	CovT	CovN	Affected domain	Polyphen-2 Prediction	Score	Expression ^b	COSMIC v52 (N of Mutated cases)
CLL1	ACTN1	missense	chr14:68422007	C>T	A425T	67%	7	6	spectrin-2	benign	0	A	1
CLL1	AP1G2	nonsense	chr14:23102859	G>A	R380*	30%	13	9	nd	altering	N/A	P	1
CLL1	BBS2	nonsense	chr16:55105893	G>T	Y106*	43%	7	15	nd	altering	N/A	P	1
CLL1	BIRC3	frameshift deletion	chr11:101706960	Δ 32bases	E368fs	30%	11	9	interdomain (BIR3 and CARD)	altering	N/A	P	4
CLL1	FRMPD1	missense	chr9:37736153	C>T	P1375L	75%	4	4	nd	benign	0.025	A	7
CLL1	FSCN3	nonsense	chr7:127025782	C>T	R340*	50%	8	7	actin-crosslinking domain	altering	N/A	A	1
CLL1	GOLM1	missense	chr9:87882296	G>A	R54C	42%	12	8	luminal	altering	0.948	A	-
CLL2	ABCA7	missense	chr19:996094	G>A	G437S	40%	15	11	extracellular	altering	0.98	P (377.6)	3 (8) ^e
CLL2	ALPK3	missense	chr15:83202310	G>A	G1315S	33%	14	6	nd	altering	0.949	A (5.3)	7
CLL2	DNAH2	missense	chr17:7618928	T>C ^c	L1543P	20%	16	17	interdomain (TPR1-TPR2)	altering	0.999	NA	7
CLL2	GPRASP1	missense	chrX:101798114	A>T	I873F	23%	23	12	Glu-rich region	altering	0.569	P (1220.8)	2
CLL2	INHBC	missense	chr12:56115006	G>A	G24S	47%	16	10	propeptide (AA19-236)	benign	0.019	P (73.7)	-
CLL2	NOTCH1	frameshift deletion	chr9:138510742	Δ 28 bases ^d	P2415fs	-	-	-	PEST domain	altering	N/A	NA	628 (812) ^e
CLL2	SCGN	missense	chr6:25809474	G>A ^c	V255M	18%	18	11	calcium binding	benign	0.002	A	3
CLL2	WHSC1	missense	chr4:1932599	G>A	E1099K	33%	13	27	SET domain	altering	0.982	P (290.5)	6
CLL2	YBX1	missense	chr1:42935456	G>A	A226T	43%	7	6	nd	benign	0	P (7900.4)	4
CLL2	ZNF300	missense	chr5:150255232	G>C	Q588E	50%	8	3	ZF N.12 (AA 577-599)	altering	0.984	A (15.3)	1
CLL3	EMX2	nonsense	chr10:119297581	C>T	Q203*	38%	14	11	homeobox DNA binding	altering	N/A	NA	-
CLL3	MPP2	missense	chr17:39312768	C>T	E447K	33%	9	14	guanylate kinase-like	altering	0.418	A (93.5)	-
CLL3	NOTCH1	frameshift deletion	chr9:138510462	ΔCT ^d	P2515fs	-	-	-	PEST domain	altering	N/A	NA	628 (812) ^e
CLL3	OR4D11	missense	chr11:59028199	C>A	T192N	37%	10	20	extracellular	altering	0.924	NA	1
CLL3	SNX9	frameshift deletion	chr6:158250968	Δ ACTG	F290fs	44%	9	11	PX domain	altering	N/A	P (829.5)	2
CLL3	TGM7	missense	chr15:41366909	G>T	N242K	33%	9	25	nd	altering	0.978	A (9.5)	2
CLL4	ASPHD1	missense	chr16:29820259	C>T	R156C	33%	9	2	luminal	altering	0.974	A (97.3)	-
CLL4	FAM123B	missense	chrX:63329239	T>A	E218V	54%	13	9	nd	altering	0.903	A (12.1)	110
CLL4	HIST1H2AM	missense	chr6:27968622	G>C	N95K	33%	12	14	kistone-fold	altering	0.983	P (180.8)	2
CLL4	MOG	missense	chr6:29735146	C>T	R54C	71%	7	9	Ig-like V-type, EC	altering	0.999	A (30.2)	1
CLL4	OR5AN1	missense	chr11:58888757	T>G	S84A	45%	11	12	extracellular	altering	0.508	NA	1
CLL4	PLEKHG5	missense	chr11:6453687	A>T	Y525N	33%	12	8	nd	altering	1	A (25.3)	1
CLL4	SLC1A1	missense	chr9:4575404	A>G	D474G	56%	9	10	extracellular	benign	0.012	A (64.4)	-
CLL5	ACTL6A	missense	chr3:180770679	C>T	P78L	75%	7	7	nd	altering	0.979	P	1
CLL5	ADCY8	missense	chr8:131862024	T>C	R1184G	75%	8	0	cytoplasmic	benign	0.174	NA	2
CLL5	GDF2	missense	chr10:48033768	G>A	T369M	60%	7	0	chain (AA320-429)	altering	0.999	A	2
CLL5	HEXIM1	missense	chr17:40582696	C>T	A119V	56%	12	0	nd	benign	0	P	1
CLL5	MYD88	missense	chr3:38157645	T>C	L265P	50%	9	10	TIR domain	altering	0.999	P	92 (93) ^e
CLL5	NHSL2	missense	chrX:71275588	G>A	G489S	100%	15	0	nd	altering	0.906	A	-
CLL5	PUM2	missense	chr2:20317179	T>C	N983S	67%	12	11	Pumilio 4	benign	0	P	1
CLL5	SORCS1	missense	chr10:108327217	G>C	P1153R	44%	9	0	cytoplasmic	benign	0.001	P	1
CLL5	ST18	missense	chr8:53240291	T>C	T418A	50%	8	0	ZF N.2 (AA 409-439)	altering	0.988	NA	4
CLL5	ZNF644	missense	chr1:91176521	T>G	Y993S	55%	12	13	ZF N.6 (AA 963-987)	altering	0.995	P	2

^a Numbering according to NCBI Build 36.1 (hg18).^b Expression refers to the mutated case or to an independent panel of 16 CLL samples. A gene is considered expressed if the probe is called Present in the mutated case or if it has >90% of Present calls in the 16 CLL cases panel. The value reported in brackets corresponds to the signal value from the most highly expressed probe in the U133 Plus 2.0 platform.^c Subclonal mutation based on Sanger sequencing.^d Mutation identified through Sanger targeted re-sequencing of the mutational hotspots on NOTCH1 (HD-JME-PEST).^e When the number of mutations exceeds the number of mutated cases, its value is reported in brackets.

Abbreviations: read%, fraction of sequence reads showing the variant nucleotide; CovT, coverage tumor; CovN, coverage normal; nd, not determined; N/A, not available; NA, probe not informative or not present; A, Absent call; P, Present call; HD, heterodimerization domain; JME, juxtamembrane domain.

Table S4. Segments of tumor-acquired copy number changes identified in the CLL discovery panel

Sample ID	Chr	Cytoband	Start ^a	End ^a	Segment size(kb)	Copy number	N of genes in segment	First 10 genes in segment	N of miRNA in segment	First 10 miRNA in segment
CLL1	4	q21.1--q21.23	77636110	86139515	8503.405	1.43	44	SHROOM3 ANKRD56 SEPT11 CCNI CCNG2 CXCL13 CNOT6L MRPL1 FRAS1 ANXA3	1	hsa-mir-575
CLL1	4	q21.23	86587986	86867769	279.783	1.42	1	ARHGAP24	0	-
CLL1	6	p12.1	55703661	55723285	19.624	1.03	0	-		
CLL1	11	q21--q23.3	94306466	116378519	22072.053	1.4	124	CWC15 KDM4D KDM4DL SFRS2B ENDOD1 SESN3 FAM76B CEP57 MTMR2 MAML2	2	hsa-mir-34b hsa-mir-34c
CLL1	15	q26.1	89423285	89433097	9.812	1.06	0	-		
CLL2	10	q26.13	123263324	123337486	74.162	1.53	1	FGFR2	0	-
CLL2	13	q14.2--q21.2	47171346	60112584	12941.238	1.63	62	SUCLA2 NUDT15 MED4 ITM2B RB1 LPAR6 RCBTB2 CYSLTR2 FNDC3A MLNR	3	hsa-mir-16-1 hsa-mir-15a hsa-mir-1297
CLL3	14	q23.2	63771822	63773115	1.293	0.78	1	ESR2	0	-
CLL4	12	p13.33--q24.33	20691	132288250	132267.559	2.99	1085	IQSEC3 SLC6A12 SLC6A13 KDM5A CCDC77 B4GALNT3 NINJ2 WNK1 RAD52 ERC1	28	hsa-mir-3649 hsa-mir-200c hsa-mir-141 hsa-mir-1244-3 hsa-mir-613 hsa-mir-614 hsa-mir-3974 hsa-mir-920 hsa-mir-4302 hsa-mir-4698
CLL5	8	q24.3	140701232	141264442	563.210	1.06	2	KCNK9 TRAPPC9	0	-
CLL5	13	q14.11--q14.3	40457346	50287238	9829.892	1.06	69	ELF1 WBP4 KBTBD6 KBTBD7 MTRF1 NAA16 OR7E37P C13orf15 KIAA0564 DGKH	2	hsa-mir-16-1 hsa-mir-15a
CLL5	18	p11.32	2664799	2723565	58.766	0.96	1	SMCHD1		

^a Numbering according to NCBI Build 36.1 (hg18).

Table S5. Features of the 48 CLL cases belonging to the screening panel

Case ID	Age at diagnosis, years	Sex	Rai stage	Lymphocytesx 10 ⁹ /L	Hb g/dL	Platelets x10 ⁹ /L	Initial management	Time to first treatment (months)	Transformation to DLBCL	Survival (months)	Status	CD38 ≥30%	ZAP70 ≥20%	V gene	D gene	J gene	IGHV homology (%)	IGHV M status	Stereotyped HCDR3	Trisomy 12	del17p13	del11q22-q23	del13q14	TP53 M
CLL9	60	M	0	18.9	17.0	219	Wait and see	Untreated	No	232	alive	Neg	Neg	2-5*01	6-13*01	4*02	95.53	M	No	Neg	Neg	Neg	Pos	UM
CLL10	64	M	I	37.9	13.5	193	Wait and see	Untreated	No	101	alive	Neg	Neg	4-30-4*01	1-7*01	6*03	90.90	M	No	Neg	Neg	Neg	Pos	UM
CLL11	64	F	0	12.2	13.1	166	Wait and see	Untreated	No	62	alive	Neg	Neg	4-61*01	5-24*01	5*02	92.85	M	No	Neg	Neg	Neg	Pos	UM
CLL12	67	M	0	15.2	15.4	257	Wait and see	Untreated	No	65	alive	Neg	Neg	3-33*01	6-13*01	4*02	92.27	M	No	Neg	Neg	Neg	Pos	UM
CLL13	71	M	0	16.6	15.5	126	Wait and see	Untreated	No	76	alive	Neg	Neg	4-61*02	3-10*01	4*02	93.78	M	No	Neg	Neg	Neg	Pos	UM
CLL14	73	F	I	23.1	14.9	164	Wait and see	27	No	76	alive	Neg	Neg	4-59*08	1-26*01	4*02	95.71	M	No	Neg	Neg	Neg	Neg	UM
CLL15	68	F	IV	170.5	6.9	101	Treatment	therapy started at diagnosis	No	49	alive	Neg	Neg	4-4*01	3-10*01	4*02	96.27	M	No	Neg	Neg	Neg	Pos	UM
CLL16	82	F	IV	12.5	13.2	92	Wait and see	25	No	30	death	Neg	Pos	4-39*01	1-1*01	4*02	95.53	M	No	Neg	Neg	Neg	Pos	UM
CLL17	68	F	0	19.7	13.4	179	Wait and see	Untreated	No	191	alive	Neg	Neg	4-39*01	6-19*01	6*02	96.56	M	No	Neg	Neg	Neg	Pos	UM
CLL18	64	M	I	16.0	16.2	168	Wait and see	Untreated	No	55	alive	Neg	Pos	4-34*01	2-2*01	4*02	93.89	M	No	Neg	Neg	Neg	Pos	UM
CLL19	74	F	0	16.9	12.3	371	Wait and see	Untreated	No	51	alive	Neg	Neg	4-34*01	6-6*01	4*02	94.81	M	No	Neg	Neg	Neg	Pos	UM
CLL20	59	M	0	15.3	14.3	242	Wait and see	110	No	123	alive	Neg	Neg	4-34*01	5-5*01	4*02	96.14	M	No	Neg	Neg	Neg	Pos	UM
CLL21	55	F	0	14.8	15.7	268	Wait and see	Untreated	No	124	alive	Neg	Neg	4-34*01	2-15*01	6*02	96.19	M	Yes	Neg	Neg	Neg	Neg	UM
CLL22	55	M	0	17.8	13.6	194	Wait and see	Untreated	No	300	alive	Neg	Neg	3-7*01	3-22*01	4*02	92.85	M	No	Neg	Neg	Neg	Pos	UM
CLL23	83	M	0	16.9	13.8	191	Wait and see	Untreated	No	57	death	Neg	Neg	3-74*01	2-2*01	4*02	93.40	M	No	Neg	Pos	Neg	Neg	M
CLL24	59	F	0	17.2	14.2	214	Wait and see	Untreated	No	53	alive	Neg	Neg	3-7*02	3-10*01	4*02	93.75	M	No	Neg	Neg	Neg	Pos	UM
CLL25	82	F	I	38.7	11.8	156	Wait and see	27	No	85	alive	Pos	Neg	3-48*03	na	5*02	93.75	M	No	Pos	Neg	Neg	Neg	UM
CLL26	73	F	0	16.8	12.6	238	Wait and see	129	No	159	alive	Neg	Neg	3-33*03	na	4*02	94.33	M	No	Neg	Neg	Neg	Neg	UM
CLL27	44	M	0	14.2	14.3	230	Wait and see	Untreated	No	88	alive	Neg	Neg	3-33*01	5-12*01	6*03	95.13	M	No	Neg	Neg	Neg	Pos	UM
CLL28	74	M	0	12.5	15.3	176	Wait and see	Untreated	No	84	alive	Neg	Pos	3-23*01	6-6*01	4*02	94.79	M	No	Neg	Neg	Neg	Neg	UM
CLL29	65	F	0	24.4	13.5	264	Wait and see	Untreated	No	100	alive	Neg	Neg	3-15*01	1-1*01	4*03	93.42	M	No	Neg	Neg	Neg	Pos	UM
CLL30	73	M	0	21.1	14.0	184	Wait and see	106	No	127	alive	Neg	Neg	3-23*01	5-24*01	4*02	93.69	M	No	Neg	Neg	Neg	Pos	UM
CLL31	81	F	I	13.7	12.8	161	Wait and see	12	No	111	alive	Neg	Neg	3-23*01	3-10*01	6*02	93.75	M	No	Pos	Neg	Neg	Pos	UM
CLL32	50	M	0	13.4	13.2	235	Wait and see	Untreated	No	114	alive	Pos	Pos	3-21*01	3-10*01	6*02	94.66	M	Yes	Neg	Neg	Neg	Neg	UM
CLL33	59	M	IV	15.6	7.1	45	Treatment	therapy started at diagnosis	No	77	alive	Neg	Neg	3-48*03	na	4*02	98.49	UM	No	Neg	Neg	Neg	Neg	M
CLL34	81	F	0	16.9	13.2	278	Wait and see	Untreated	No	80	alive	Neg	Neg	3-52*02	3-22*01	4*02	100.00	UM	No	Neg	Neg	Neg	Neg	UM
CLL35	77	M	IV	39.0	6.6	93	Treatment	therapy started at diagnosis	No	59	alive	Pos	Neg	4-39*01	5-24*01	4*02	100.00	UM	No	Neg	Neg	Neg	Neg	UM
CLL36	72	M	0	16.5	13.6	188	Wait and see	47	No	53	alive	Pos	Neg	5-51*01	3-3*01	4*02	100.00	UM	No	Neg	Neg	Neg	Pos	UM
CLL37	72	F	0	17.5	12.6	255	Wait and see	50	No	90	alive	Pos	Neg	1-69*02	5-12*01	4*02	100.00	UM	No	Pos	Neg	Neg	Pos	UM
CLL38	80	F	0	12.9	15.9	313	Wait and see	Untreated	No	38	alive	Pos	Neg	1-69*01	3-22*01	5*02	100.00	UM	Yes	Pos	Neg	Neg	Neg	UM
CLL39	67	M	I	21.5	13.5	126	Wait and see	29	No	73	alive	Pos	Neg	3-30-3*03	3-3*01	5*02	100.00	UM	No	Neg	Neg	Pos	Pos	UM
CLL40	71	F	0	81.0	12.0	191	Wait and see	8	No	86	alive	Pos	Neg	1-2*02	3-16*01	4*02	100.00	UM	No	Neg	Neg	Neg	Neg	UM
CLL41	79	M	I	49.4	12.3	109	Treatment	therapy started at diagnosis	No	18	death	Neg	Pos	4-30-4*01	3-22*01	5*02	100.00	UM	Yes	Neg	Pos	Neg	Pos	M
CLL42	48	M	0	13.8	14.8	170	Wait and see	40	No	63	alive	Neg	Pos	3-33*01	3-10*01	6*02	100.00	UM	No	Neg	Neg	Pos	Pos	UM
CLL43	62	F	I	16.9	12.0	120	Wait and see	Untreated	No	143	alive	Neg	Pos	3-20*01	2-2*03	6*03	99.60	UM	Yes	Pos	Neg	Neg	Neg	UM
CLL44	73	M	I	15.9	12.2	217	Wait and see	40	No	59	alive	Neg	Pos	1-69*01	3-16*01	6*02	99.24	UM	No	Neg	Neg	Neg	Pos	UM
CLL45	51	M	IV	23.0	14.2	82	Treatment	therapy started at diagnosis	No	147	death	Pos	Pos	5-51*01	3-22*01	4*02	100.00	UM	Yes	Neg	Neg	Pos	Pos	UM
CLL46	64	F	0	15.6	14.1	202	Wait and see	Untreated	No	96	alive	Pos	Pos	1-69*06	3-16*02	3*02	100.00	UM	Yes	Neg	Neg	Neg	Pos	UM
CLL47	46	M	I	24.4	14.3	391	Wait and see	6	No	69	alive	Pos	Pos	3-21*01	na	6*02	98.50	UM	Yes	Neg	Neg	Neg	Neg	UM
CLL48	91	M	I	70.8	11.5	132	Wait and see	Untreated	No	38	alive	Neg	Pos	5-51*01	5-18*01	6*02	100.00	UM	No	Pos	Neg	Neg	Neg	UM
CLL49	77	M	I	19.7	13.8	301	Wait and see	7	No	77	alive	Neg	Pos	3-48*01	3-10*01	4*02	100.00	UM	No	Pos	Neg	Neg	Neg	UM
CLL50	74	F	0	15.0	14.8	165	Wait and see	60	No	70	alive	Neg	Pos	3-30-3*01	3-3*01	3*02	100.00	UM	Yes	Pos	Neg	Neg	Neg	UM
CLL51	66	M	0	16.2	14.4	212	Wait and see	59	No	60	alive	Neg	Pos	1-69*01	3-10*01	6*02	100.00	UM	No	Neg	Neg	Neg	Neg	UM
CLL52	63	F	I	141.8	11.3	226	Treatment	therapy started at diagnosis	No	94	death	Neg	Pos	1-69*01	6-19*01	4*02	99.21	UM	No	Pos	Neg	Neg	Neg	UM
CLL53	58	F	I	20.5	13.1	183	Wait and see	14	No	88	alive	Pos	Pos	3-23*01	3-9*01	4*02	100.00	UM	No	Neg	Neg	Neg	Pos	UM
CLL54	62	M	I	19.5	16.0	255	Wait and see	5	No	46	alive	Pos	Pos	1-69*01	3-3*01	6*02	100.00	UM	No	Neg	Neg	Neg	Neg	UM
CLL55	65	M	I	13.7	13.5	188	Wait and see	13	No	81	alive	Pos	Pos	7-4-1*01	6-19*01	4*02	100.00	UM	Yes	Neg	Neg	Neg	Neg	UM
CLL56	62	M	IV	129.9	10.8	59	Treatment	therapy started at diagnosis	No	61	alive	Pos	Pos	3-30-3*01	3-3*02	6*02	100.00	UM	Yes	Neg	Neg	Neg	Neg	M

Table S6. Clinical and biological characteristics of the consecutive series of newly diagnosed and previously untreated CLL^a

	Number	%
Clinical features at CLL diagnosis		
Age >70	54/120	45.0
Male	73/120	60.8
Rai stage		
0	78/120	65.0
I	25/120	20.8
II	6/120	5.0
III	1/120	0.8
IV	10/120	8.3
Lymphocyte count (x 10 ⁹ /L) ^b	18.8	(9.5-28.8)
Hb <11 g/dL	7/120	5.8
Platelets <100 x 10 ⁹ /L	8/120	6.7
LDH > ULN	13/120	10.8
Beta-2-microglobulin		
≤2.5 mg/L	78/120	65.0
2.6-3.5 mg/L	28/120	23.3
>3.5 mg/L	14/120	11.7
<i>IGHV</i> homology ≥98%	39/120	32.5
CD38 ≥30%	36/120	30.0
ZAP70 ≥20%	37/120	30.8
del17p13 and/or <i>TP53</i> mutation	13/117	11.1
del11q22-q23	8/112	7.1
+12	17/112	15.2
del13q14	61/112	54.5
Normal FISH	32/112	28.6

^a ULN, upper limit of normal; *IGHV*, immunoglobulin heavy chain variable region gene; FISH, fluorescence in situ hybridization.

^b Median and 25th-75th percentiles are reported.

Table S7. *NOTCH1* mutations in CLL, RS and de novo DLBCL

Case ID	Nucleotide change ^a	AA change ^a	Protein domain	Reported in T-ALL (COSMICv52)
CLL at diagnosis				
CLL2	Δ 28 bases (7247-7274)	P2415fs	PEST	Yes
CLL3	ΔCT (7544-7545) ^c	P2515fs	PEST	Yes
CLL34	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL40	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL52	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL55	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL56	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL38	+CCCC(7023-7024)	S2342fs	PEST	Yes ^f
CLL200	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL227	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL229	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL261	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL270	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL275	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL277	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL280	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL290	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL293	ΔCT (7544-7545)	P2515fs	PEST	Yes
Paired CLL/Richter				
RS-57	ΔCT (7544-7545) ^b	P2515fs	PEST	Yes
RS-62	ΔCT (7544-7545) ^d	P2515fs	PEST	Yes
RS-63	ΔCT (7544-7545) ^c	P2515fs	PEST	Yes
RS-64	ΔCT (7544-7545) ^c	P2515fs	PEST	Yes
RS-66	C7510T ^d	Q2504*	PEST	Yes
RS-72	ΔCT (7544-7545) ^c	P2515fs	PEST	Yes
RS-74	ΔCT (7544-7545) ^d	P2515fs	PEST	Yes
RS-78	ΔCT (7544-7545) ^e	P2515fs	PEST	Yes
RS-81	dupl.50bp (7295-7344) ^b	R2434fs	PEST	No
RS-85	C7321T ^b	Q2441*	PEST	Yes
RS-86-A	ΔCT (7544-7545) ^d	P2515fs	PEST	Yes
RS-86-B	G5164A ^b	V1722M	HD	Yes
RS-88	ΔCT (7544-7545) ^c	P2515fs	PEST	Yes
RS-91	ΔCT (7544-7545) ^d	P2515fs	PEST	Yes
RS-92	ΔCT (7544-7545) ^d	P2515fs	PEST	Yes
RS-98	ΔCT (7544-7545) ^e	P2515fs	PEST	Yes
RS-100	ΔCT (7544-7545) ^e	P2515fs	PEST	Yes
RS-103	ΔCT (7544-7545) ^c	P2515fs	PEST	Yes
RS-114	ΔCT (7544-7545) ^b	P2515fs	PEST	Yes
Chemorefractory CLL				
CLL-121-F-R	ΔC7446	F2482fs	PEST	Yes
CLL-123-A-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL-126-F-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL-129-A-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL-130-F-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL-137-F-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL-141-F-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL-146-F-R	ΔC7392	A2464fs	PEST	Yes ^f
CLL-152-F-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
CLL-158-F-R	ΔCT (7544-7545)	P2515fs	PEST	Yes
de novo DLBCL				
2132	ΔCT (7544-7545)	P2515fs	PEST	Yes
2100	C7246T	Q2415*	PEST	No

^a Numbering according to GenBank accession No. NM_017617.2 (mRNA) and NP_060087.2 (protein) respectively.

^b Mutation not detectable at CLL diagnosis within the sensitivity threshold of Sanger sequencing.

^c Mutation represented at a subclonal level at CLL diagnosis according to Sanger sequencing.

^d Mutation represented at a clonal level at CLL diagnosis according to Sanger sequencing.

^e paired CLL phase not available.

^f Different nucleotide change causing frameshift at the same aminoacid position of the NOTCH1 protein.

Abbreviations: F-R, fludarabine-refractoriness; A-R, alkylator-refractoriness.

Table S8. Clinical and biological characteristics of the RS cohort^a

	Number	%
Clinical features at RS diagnosis		
Age >60 years	40/58	69.0
Male	37/58	63.8
ECOG PS >1	27/58	46.5
Ann Arbor stage III-IV	55/58	94.8
Rai stage		
0	0/58	0
I	21/58	36.2
II	7/58	12.1
III	7/58	12.1
IV	23/58	39.7
B symptoms	26/58	44.8
Extranodal sites >1	20/58	34.4
Tumor size >5 cm	32/58	55.1
Platelets <100 x 10 ⁹ /L	24/58	41.3
LDH >1.5 ULN	34/58	58.6
Prior CLL therapies >1	22/58	37.9
IPI >2	36/58	62.0
RS score >1	35/58	60.3
Pathologic features at RS diagnosis		
Non-GC phenotype	55/58	94.8
EBV infection	3/58	5.1
Genetic features		
<i>TP53</i> disruption	30/54	55.5
<i>MYC</i> aberrations	11/51	21.5
<i>IGHV</i> homology ≥98%	37/56	66.1
Stereotyped VH CDR3	22/56	39.3

^a ULN, upper limit of normal; IPI, international prognostic index; GC, germinal center; *IGHV*, immunoglobulin heavy chain variable region gene; VH CDR3, immunoglobulin heavy chain complementarity-determining region 3.

Table S9. Clinical and biological characteristics of the chemorefractory CLL cohort^a

	Number	%
Clinical features at diagnosis		
Age >70	33/48	68.8
Male	31/48	64.6
Rai stage		
0	0/48	0
I	8/48	16.7
II	8/48	16.7
III	3/48	6.3
IV	29/48	60.4
Lymphocyte count (x 10 ⁹ /L) ^b	43.0	(14.4-99.1)
Hb <11 g/dL	23/48	47.9
Platelets <100 x 10 ⁹ /L	29/48	60.4
LDH > ULN	24/48	50.0
Beta-2-microglobulin		
≤2.5 mg/L	6/48	12.5
2.6-3.5 mg/L	13/48	27.1
>3.5 mg/L	29/48	60.4
Number of prior therapies		
1	20/48	41.7
2	16/48	33.3
>2	12/48	25.0
Treatment regimen at refractoriness		
FCR	12/48	25.0
FA	5/48	10.4
FR	2/48	4.2
FC	4/48	8.3
F	12/48	25.0
R-COP	13/48	27.1
<i>IGHV</i> homology ≥98%	30/48	62.5
CD38 ≥30%	24/48	50.0
ZAP70 ≥20%	24/48	50.0
del17p13 and/or <i>TP53</i> mutation	19/48	39.6
del11q22-q23	13/48	27.1
+12	16/48	33.3
del13q14	24/48	50.0
Normal FISH	5/48	10.4

^a ULN, upper limit of normal; FCR, fludarabine, cyclophosphamide, rituximab; FA, fludarabine, alemtuzumab; FR, fludarabine, rituximab; FC, fludarabine, cyclophosphamide; F, fludarabine; R-COP, rituximab, cyclophosphamide, vincristine, prednisone; *IGHV*, immunoglobulin heavy chain variable region gene; FISH, fluorescence in situ hybridization.

^b Median and 25th-75th percentiles are reported.

Molecular history of Richter syndrome: origin from a cell already present at the time of chronic lymphocytic leukemia diagnosis

Davide Rossi^{1*}, Valeria Spina^{1*}, Francesco Forconi², Daniela Capello¹, Marco Fangazio¹, Silvia Rasi¹, Maurizio Martini³, Valter Gattei⁴, Antonio Ramponi⁵, Luigi M. Larocca³, Francesco Bertoni⁶ and Gianluca Gaidano¹

¹Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy

²Division of Hematology, University of Siena and AOUS, Siena, Italy

³Institute of Pathology, Catholic University of the Sacred Heart, Rome, Italy

⁴Centro di Riferimento Oncologico, Aviano, Italy

⁵Division of Pathology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy

⁶Oncology Institute of Southern Switzerland IOSI, Bellinzona, CH Switzerland

Richter syndrome (RS) represents the transformation of chronic lymphocytic leukemia to aggressive lymphoma. We explored intraclonal diversification (ID) of immunoglobulin genes in order to (i) follow the evolutionary history of the RS clone (ii) compare the role of ID in clonally related RS vs. clonally unrelated cases. Most (10/11, 90.9%) clonally related RS stem from the predominant clone observed at CLL diagnosis. One single RS had a transformation pattern compatible with sequential evolution from a secondary CLL subclone. Once RS transformation had occurred, all secondary CLL subclones disappeared and were substituted by the dominant RS clone with its own descendants. These observations suggest that genetic lesions associated with RS transformation are acquired by a cell belonging to the original CLL clone, rather than being progressively accumulated by later CLL subclones. Accordingly, most (9/11, 81.1%) clonally related RS harbored a genetic lesion disrupting *TP53* that was already present, though at subclonal levels, in 5/11 (45.5%) samples of the paired CLL phase. A fraction of clonally related RS switched off ID (4/11, 36.4%) or reduced the levels of ID (5/11, 45.4%) at transformation. Conversely, all clonally unrelated RS harbored ID and were characterized by a significantly higher mutation frequency compared to clonally related RS (median: 1.18×10^{-3} vs. 0.13×10^{-3} ; $p = 0.002$). These data indicate that (i) clonally related RS stems from a cell that is already present within the initial CLL clone and (ii) clonally unrelated and clonally related RS are biologically distinct disorders also in terms of antigen affinity maturation.

Key words: chronic lymphocytic leukemia, Richter syndrome, immunoglobulin gene, intraclonal diversification, clonal evolution
Additional Supporting Information may be found in the online version of this article.

*D.R. and V.S. contributed equally to this study.

Grant sponsor: Progetto FIRB Programma “Futuro in Ricerca” 2008, PRIN 2008 and PRIN 2009. MIUR, Rome, Italy; AIRC, Special Program Molecular Clinical Oncology, 5 x 1000, No. 10007, Milan, Italy; Progetto Giovani Ricercatori 2008, Ministero della Salute, Rome, Italy; Ricerca Sanitaria Finalizzata, Regione Piemonte, Torino, Italy; Oncosuisse grant OCS-02296-08-2008; Helmut Horten Foundation, San Salvatore Foundation, Fondazione per la Ricerca e la Cura sui Linfomi, Lugano, Switzerland; Istituto Toscano Tumori (ITT) 2008 Grant, Florence, Italy; Novara-AIL Onlus, Novara, Italy; Siena-AIL, Siena, Italy.

DOI: 10.1002/ijc.26322

History: Received 5 May 2011; Accepted 14 Jul 2011; Online 27 Jul 2011

Correspondence to: Davide Rossi, Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy, Tel: +39 0321 660698, Fax: +39 0321 620421, E-mail: rossidav@med.unipmn.it

The definition of Richter syndrome (RS) encompasses two biologically different conditions: (i) transformation of chronic lymphocytic leukemia (CLL) to clonally related diffuse large B-cell lymphoma (DLBCL) and (ii) development of a DLBCL unrelated to the CLL clone.^{1–4} The role of antigen in promoting CLL transformation to RS is supported by the recent observation of a high prevalence of stereotyped B-cell receptors in CLL transforming to RS, and by the documentation of biased usage of immunoglobulin heavy variable (*IGHV*) genes, including *IGHV4-39*.^{5,6}

A common assumption in the context of transformation from indolent to aggressive B-cell malignancy holds that transformation reflects the final stage of evolution of the indolent B-cell clone.⁷ An alternative scenario, however, might be the divergent evolution of both the indolent and the aggressive phase from a common precursor cell.^{7,8} In RS, the precise timing of onset of the RS clone during CLL history is unknown. Also, although antigen stimulation may be important for the growth of indolent B-cell malignancies, it might no longer be required by the subsequent aggressive malignancy.⁹ With respect to RS, antigen stimulation contributes to CLL predisposition to DLBCL,⁶ but it is unclear whether antigen retains any role once that transformation has occurred.

Table 1. ID of *IGHV* genes in clonally related and clonally unrelated RS^a

Case	Diagnosis	<i>IGHV</i>	<i>IGHD</i>	<i>IGHJ</i>	Identity (%)	HCDR3 subset	CM	UCM	R	S	ID status	MF × 10 ⁻³
Clonally related RS												
7839	CLL phase	1-69	3-3	4	100	Not stereotyped	1	4	5	0	CID	0.69
	DLBCL phase	1-69	3-3	4	100	Not stereotyped	2	2	2	2	CID	0.28
458	CLL phase	1-69	3-16	3	100	6	1	7	8	0	CID	1.11
	DLBCL phase	1-69	3-16	3	100	6	0	0	0	0	No ID	0.00
4233	CLL phase	7-4-1	3-22	5	99.6	Not stereotyped	1	5	4	2	CID	0.83
	DLBCL phase	7-4-1	3-22	5	99.6	Not stereotyped	0	0	0	0	No ID	0.00
3502	CLL phase	4-39	3-22	4	100	Not stereotyped	2	0	2	0	CID	0.27
	DLBCL phase	4-39	3-22	4	100	Not stereotyped	3	3	5	1	CID	0.41
7598	CLL phase	4-39	6-13	5	99.6	8	2	6	7	1	CID	1.10
	DLBCL phase	4-39	6-13	5	99.6	8	1	1	2	0	CID	0.14
3392	CLL phase	4-39	6-13	5	100	8	4	8	10	2	CID	0.82
	DLBCL phase	4-39	6-13	5	100	8	2	0	1	1	CID	0.14
5076	CLL phase	3-49	3-3	2	100	Not stereotyped	1	13	13	1	CID	1.90
	DLBCL phase	3-49	3-3	2	100	Not stereotyped	0	0	0	0	No ID	0.00
4831	CLL phase	3-9	3-3	6	100	7	2	9	9	2	CID	0.76
	DLBCL phase	3-9	3-3	6	100	7	0	0	0	0	No ID	0.00
7840	CLL phase	1-8	Na	6	98.6	Not stereotyped	3	0	2	1	CID	0.42
	DLBCL phase	1-8	Na	6	98.6	Not stereotyped	1	2	2	1	CID	0.21
7838	CLL phase	6-1	4-23	4	98.3	Not stereotyped	2	0	2	0	CID	0.27
	DLBCL phase	6-1	4-23	4	98.3	Not stereotyped	0	9	5	4	UID	0.61
8342	CLL phase	3-21	Na	6	98.9	2	3	6	8	1	CID	0.63
	DLBCL phase	3-21	Na	6	98.9	2	0	5	5	0	UID	0.35
Clonally unrelated RS												
7977	DLBCL phase	1-2	3-22	4	99.3	Not stereotyped	1	5	5	1	CID	0.42
9357	DLBCL phase	4-4	5-24	5	92.0	Not stereotyped	7	16	19	4	CID	1.61
7837	DLBCL phase	3-23	1-26	4	96.2	Not stereotyped	3	5	6	2	CID	0.56
8012	DLBCL phase	1-2	3-22	3	100	Not stereotyped	4	13	14	3	CID	1.18
7836	DLBCL phase	1-2	3-3	4	91.4	Not stereotyped	4	15	15	4	CID	1.32

^aRS, Richter syndrome; CLL, chronic lymphocytic leukemia; DLBCL, diffuse large B-cell lymphoma; *IGHV*, immunoglobulin heavy V genes; *IGHD*, immunoglobulin heavy D genes; *IGHJ*, immunoglobulin heavy J genes; HCDR3, immunoglobulin heavy chain complementarity-determining region 3; CM, confirmed mutations; UCM, unconfirmed mutations; R, replacement; S, silent; ID, intraclonal diversification; CID, confirmed intraclonal diversification; UID, unconfirmed intraclonal diversification; MF, mutation frequency; Na, not assessable.

Herein we utilized analysis of intraclonal diversification (ID) of *IGHV* genes¹⁰⁻¹⁷ as a tool to track the origin of the RS clone and to gain insights into the role of ongoing interactions with antigen in RS clone sustainment. We report that most clonally related RS originate from a tumor cell already present at the time of CLL diagnosis.

Material and methods

Patients

The study included 11 clonally related RS provided with paired CLL samples (Table 1). For comparative purposes, five clonally unrelated RS and 20 *de novo* DLBCL harboring a nongerminal center (GC) phenotype were also analyzed (Table 1 and Supporting Information Table S1). All RS were

classified as DLBCL showing a non-GC phenotype¹⁸ and lacked EBV infection. All samples were derived from lymph node specimens. By immunohistochemistry, the fraction of tumor cells was >80% in all specimens. Samples were centralized at the Oncology Institute of Southern Switzerland-IOSI, Bellinzona, CH, within the framework of an international RS project. Patients provided informed consent in accordance with local institutional review board (IRB) requirements and Declaration of Helsinki. The study was approved by the IOSI IRB.

Analysis of *IGHV-D-J* rearrangements and cloning

IGHV-D-J rearrangements were amplified from genomic DNA using high-fidelity PfuTurbo DNA polymerase (Stratagene, La

Jolla, CA, USA). The number of *IGHV* clones sequenced per case was 50 for RS, 25–50 for CLL and 25–50 for *de novo* DLBCL. Taq error rate was determined by subcloning the germline *IGHV1-69* gene and β -globin exon 1 that had been PCR amplified from the fibroblast cell line MRC5. Fifty clones were sequenced for both *IGHV1-69* and β -globin, obtaining an error rate of 0.19 per 1000 nucleotides in the germline *IGHV1-69* gene and of 0.21 per 1000 nucleotides in β -globin exon 1. The occurrence of stereotyped immunoglobulin heavy chain complementarity-determining region 3 (HCDR3) was analyzed as reported.⁶ The following nomenclature was adopted for mutation analysis of *IGHV* subclones. Mutations occurring in all clones from a given sample were defined as ubiquitous. Mutations occurring in more than one clone from different PCR amplimers of a given sample were defined as confirmed mutations (CM). Mutations occurring in one clone from a given sample were defined as unconfirmed mutations (UCM). Cases were scored positive for ID only in the presence of CM. The mutation frequency was calculated according to the following formula: $\sum(\text{CM} + \text{UCM})/\text{number of cloned sequences} \times \text{sequence length}$.¹⁷

Evolutionary history of subcloned sequences

Phylogenetic analysis of subclones was performed with MEGA4. The evolutionary history of subcloned sequences was inferred using the maximum parsimony method. The maximum parsimony trees were obtained using the close-neighbor-interchange algorithm with search level 2 in which the initial trees were obtained with the random addition of sequences (10 replicates). Branch lengths were calculated using the average pathway method. All positions containing gaps and missing data were eliminated from the dataset.

Analysis of tumor genetic lesions of TP53 and B-cell receptor (BCR) pathway genes

Mutation analysis of *TP53* (exons 4–8; GenBank accession no. NM_000546.4), *CARD11* (exons 5–9; GenBank accession no. NM_032415.3), *CD79A* (ITAM domain on exons 4–5; GenBank accession no. NM_001783.3) and *CD79B* (ITAM domain on exons 5–6; GenBank accession no. NM_000626.2) was performed by DNA direct sequencing on an ABI Prism 3100 automated DNA sequence analyzer (Applied Biosystems, Foster City, CA). FISH analysis for *TP53* deletion (probe LSIp53, Abbott, Rome, Italy) was performed on the bioptic tissue utilized for RS diagnosis by staining nuclei extracted from fresh or frozen cell suspensions. Nuclei were counterstained with 4',6'-diamidino-2-phenylindole (DAPI) and antifade, and signals were visualized using an Olympus BX51 microscope (Olympus Italia, Milan, Italy). For each probe, at least 400 interphase cells with well-delineated fluorescent spots were examined.

Immunohistochemistry

Cell of origin (COO) of DLBCL developing in the context of RS transformation was assigned according to the Choi algo-

rithm¹⁸ after staining for GCET1 (clone RAM341, Abcam, Cambridge, UK), CD10 (clone 56C6, Ventana Medical Systems, Tucson, AZ), BCL6 (polyclonal, Santa Cruz Biotechnology, Santa Cruz, CA), IRF4/MUM1 (clone MRQ-43, Ventana Medical Systems, Tucson, AZ) and FOXP1 (Clone JC12, Abcam, Cambridge, UK).

Statistical analysis

Continuous variables from paired samples were compared by Wilcoxon test. Continuous variables from independent samples were compared by Mann-Whitney test. All statistical tests were two-sided. Statistical significance was defined as $p < 0.05$. Statistical analysis was performed with the Statistical Package for the Social Sciences (SPSS) software v.18.0 (Chicago, IL).

Results

Clonally related RS originate from a cell of the initial CLL clone

Phylogenetic analysis of *IGHV* clones from clonally related CLL/RS pairs revealed that most (10/11, 90.9%) clonally related RS stem from the predominant clone observed at CLL diagnosis (Figs. 1a–1l). One single RS case had a transformation pattern compatible with sequential evolution from a secondary CLL subclone (Fig. 1m). All secondary CLL subclones previously documented in the CLL phase disappeared once RS transformation had occurred and were substituted by the dominant DLBCL clone with its own descendants (Fig. 1).

Paired analysis of clonally related CLL/RS samples documented that the main RS clone maintained the same *IGHV* mutation profile observed in the main CLL clone (Figs. 1a–1l). In only one case (Fig. 1m), the RS clone experienced clonal evolution by acquiring five additional *IGHV* mutations that were absent at the time of CLL diagnosis.

Overall, the mutation frequency significantly decreased during transformation from CLL (median: 0.76×10^{-3} ; range $0.27\text{--}1.90 \times 10^{-3}$) to RS (median: 0.13×10^{-3} ; range: $0\text{--}0.60 \times 10^{-3}$) ($p = 0.013$, Wilcoxon test) (Table 1). In fact, RS transformation was accompanied by the selection of a clone that did not require further ID (4/11, 36.4%) or that reduced the levels of ID (5/11, 45.4%) (Table 1).

The role of ID is different in clonally unrelated RS versus clonally related RS

A distinct and less common subset of RS is represented by clonally unrelated cases that are characterized by the development of a secondary DLBCL in the context of CLL.^{1–4} Despite morphologic and phenotypic similarities,⁴ the ID status of *IGHV* genes of clonally unrelated RS differed from that of both clonally related RS and *de novo* DLBCL with a non-GC phenotype. In fact, all clonally unrelated RS harbored ID and were characterized by a significantly higher mutation frequency (median: 1.18×10^{-3} ; range: $0.42\text{--}1.61 \times 10^{-3}$) compared to clonally related RS ($p = 0.002$, Mann-Whitney test) and *de novo* DLBCL with a non-GC phenotype (0.08×10^{-3}

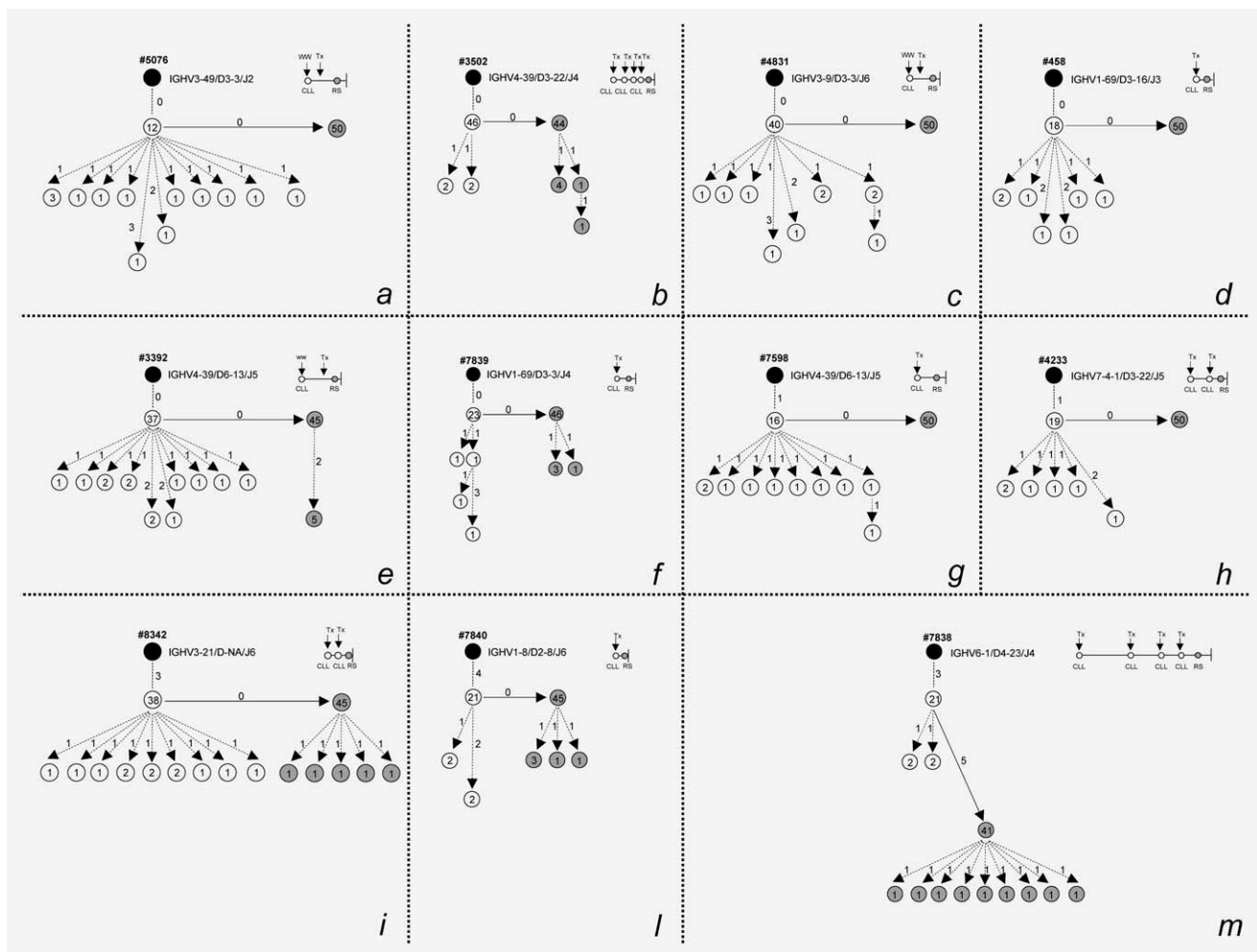


Figure 1. Clonal evolution trees of clonally related RS. The germline sequence is depicted as ●, CLL clones as ○, transformed DLBCL clones as ●. Numbers inside circles correspond to the number of clones with the same sequence; numbers flanking the arrows represent the additional mutations in the clone; numbers flanking the arrows that link the black circle representing the germline sequence to the white circle representing the initial CLL clone show the common mutations shared between CLL and RS paired samples. The graphic timeline representation of the clinical evolution of the disease is also represented for each single case. Vertical arrows represent the management of CLL, including watch and wait (WW) and therapy (Tx). The right-hand vertical bar represents the death of the patient.

range: $0-4.38 \times 10^{-3}$; $p = 0.024$; Mann-Whitney test) (Table 1 and Supporting Information Table S1).

Discussion

The results of this study indicate that (i) clonally related RS stem from a cell that was already present within the initial CLL clone and gains selective advantage over other CLL subclones and (ii) clonally unrelated RS differ from clonally related cases for their need to maintain ID.

Based on other cancer models, at least two distinct scenarios might be envisaged to explain the evolutionary history of clonally related RS as documented in this study.^{19,20} In a first scenario, the genetic lesions initiating RS transformation are acquired by a cell belonging to the initial CLL clone, rather than being progressively accumulated by later CLL subclones. In a second scenario, considering that all patients analyzed in this

study received one or more CLL treatments before developing RS, the selective pressure of CLL treatment might have advantaged the growth of a preexisting drug resistant clone. Notably, most (9/11, 81.1%) clonally related RS harbored a genetic lesion disrupting *TP53* that was already present, though at subclonal levels, in 5/11 (45.5%) samples of the paired CLL phase.

A fraction of clonally related RS switched off ID at transformation. This observation might suggest an antigen selection pressures that advantages clones retaining their BCR in the original conformation. Alternatively, the switch off of ID at transformation might result from the acquisition by RS cells of new genetic lesions subsidizing for BCR activation by antigen. Screening of BCR pathway genes known to be mutated in *de novo* DLBCL (*CD79A*, *CD79B* and *CARD11*)^{21,22} revealed the acquisition of a mutation within the coiled-coil domain of *CARD11* in 1/4 RS that switched off ID at transformation. The

current survey was limited to BCR genes known to be molecularly altered in *de novo* DLBCL and cannot exclude that additional BCR components might be altered in other RS cases characterized by ID switch off.

At variance with clonally related cases, clonally unrelated RS maintained ID in all cases, suggesting an ongoing interaction with antigen. This observation further reinforces the notion that clonally unrelated and clonally related RS are distinct disorders both biologically and clinically and that the diagnosis of RS should be restricted to clonally related cases.⁴

CLL transformation to clonally related RS and follicular lymphoma transformation to DLBCL are the two examples of evolution from indolent to aggressive B-cell malignancy that have been characterized for the molecular history of their clonal progression. In both CLL and FL,⁷ the transformation from an indolent to an aggressive B-cell malignancy arises from the expansion of a cell already present during the initial phases of lymphomagenesis, suggesting that B-cell neoplasms exploit similar pathways of transition from indolent to aggressive histology.

References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW. World Health Organization classification of tumours, pathology and genetics of tumours of haematopoietic and lymphoid tissues. Lyon: IARC; 2008.
2. Tsimberidou AM, Keating MJ. Richter syndrome: biology, incidence, and therapeutic strategies. *Cancer* 2005;103:216–28.
3. Rossi D, Gaidano G. Richter syndrome: molecular insights and clinical perspectives. *Hematol Oncol* 2009;27:1–10.
4. Rossi D, Spina V, Deambrogi C, Rasi S, Laurenti L, Stamatopoulos K, Arcaini L, Lucioni M, Rocque GB, Xu-Monette ZY, Visco C, Chang J, et al. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival post-transformation. *Blood* 2011;117:3391–401.
5. Rossi D, Cerri M, Capello D, Deambrogi C, Rossi FM, Zucchetto A, De Paoli L, Cresta S, Rasi S, Spina V, Franceschetti S, Lunghi M, et al. Biological and clinical risk factors of chronic lymphocytic leukaemia transformation to Richter syndrome. *Br J Haematol* 2008;142:202–15.
6. Rossi D, Spina V, Cerri M, Rasi S, Deambrogi C, De Paoli L, Laurenti L, Maffei R, Forconi F, Bertoni F, Zucca E, Agostinelli C, et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clin Cancer Res* 2009;15:4415–22.
7. Carloti E, Wrench D, Matthews J, Iqbal S, Davies A, Norton A, Hart J, Lai R, Montoto S, Gribben JG, Lister TA, Fitzgibbon J. Transformation of follicular lymphoma to diffuse large B-cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone. *Blood* 2009;113:3553–7.
8. Matolcsy A, Schattner EJ, Knowles DM, Casali P. Clonal evolution of B cells in transformation from low- to high-grade lymphoma. *Eur J Immunol* 1999;29:1253–64.
9. Lévy M, Copie-Bergman C, Gameiro C, Chaumette MT, Delfau-Larue MH, Haioun C, Charachon A, Hemery F, Gaulard P, Leroy K, Delchier JC. Prognostic value of translocation t(11;18) in tumoral response of low-grade gastric lymphoma of mucosa-associated lymphoid tissue type to oral chemotherapy. *J Clin Oncol* 2005;23:5061–6.
10. Chapman CJ, Zhou JX, Gregory C, Rickinson AB, Stevenson FK. VH and VL gene analysis in sporadic Burkitt's lymphoma shows somatic hypermutation, intracлонаl heterogeneity, and a role for antigen selection. *Blood* 1996;88:3562–8.
11. Ottensmeier CH, Thompson AR, Zhu D, Wilkins BS, Sweetenham JW, Stevenson FK. Analysis of VH genes in follicular and diffuse lymphoma shows ongoing somatic mutation and multiple isotype transcripts in early disease with changes during disease progression. *Blood* 1998;91:4292–9.
12. Lossos IS, Alizadeh AA, Eisen MB, Chan WC, Brown PO, Botstein D, Staudt LM, Levy R. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc Natl Acad Sci USA* 2000;97:10209–13.
13. Gurrieri C, McGuire P, Zan H, Yan XJ, Cerutti A, Albesiano E, Allen SL, Vinciguerra V, Rai KR, Ferrarini M, Casali P, Chiorazzi N. Chronic lymphocytic leukemia B cells can undergo somatic hypermutation and intracлонаl immunoglobulin V(H)DJ(H) gene diversification. *J Exp Med* 2002;196:629–39.
14. Zhu D, Orchard J, Oscier DG, Wright DH, Stevenson FK. V(H) gene analysis of splenic marginal zone lymphomas reveals diversity in mutational status and initiation of somatic mutation in vivo. *Blood* 2002;100:2659–61.
15. Volkheimer AD, Weinberg JB, Beasley BE, Whitesides JF, Gockerman JP, Moore JO, Kelsoe G, Goodman BK, Levesque MC. Progressive immunoglobulin gene mutations in chronic lymphocytic leukemia: evidence for antigen-driven intracлонаl diversification. *Blood* 2007;109:1559–67.
16. Campbell PJ, Pleasance ED, Stephens PJ, Dicks E, Rance R, Goodhead I, Follows GA, Green AR, Futreal PA, Stratton MR. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc Natl Acad Sci USA* 2008;105:13081–6.
17. Sutton LA, Kostareli E, Hadzidimitriou A, Darzentas N, Tsaftaris A, Anagnostopoulos A, Rosenquist R, Stamatopoulos K. Extensive intracлонаl diversification in a subgroup of chronic lymphocytic leukemia patients with stereotyped IGHV4–34 receptors: implications for ongoing interactions with antigen. *Blood* 2009;114:4460–8.
18. Choi WW, Weisenburger DD, Greiner TC, Piris MA, Banham AH, Delabie J, Brazier RM, Geng H, Iqbal J, Lenz G, Vose JM, Hans CP, et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin Cancer Res* 2009;15:5494–502.
19. Greaves MF, Wiemels J. Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer* 2003;3:639–49.
20. Roche-Lestienne C, Soenen-Cornu V, Grardel-Duflos N, Lai JL, Philippe N, Facon T, Fenaux P, Preudhomme C. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood* 2002;100:1014–8.
21. Lenz G, Davis RE, Ngo VN, Lam L, George TC, Wright GW, Dave SS, Zhao H, Xu W, Rosenwald A, Ott G, Muller-Hermelink HK, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* 2008;319:1676–9.
22. Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, Kohlhammer H, Lamy L, Zhao H, Yang Y, Xu W, Shaffer AL, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* 2010;463:88–92.

blood

2011 118: 4930-4934
Prepublished online August 31, 2011;
doi:10.1182/blood-2011-06-359166

Alteration of *BIRC3* and multiple other NF- κ B pathway genes in splenic marginal zone lymphoma

Davide Rossi, Silvia Deaglio, David Dominguez-Sola, Silvia Rasi, Tiziana Vaisitti, Claudio Agostinelli, Valeria Spina, Alessio Bruscaggin, Sara Monti, Michaela Cerri, Stefania Cresta, Marco Fangazio, Luca Arcaini, Marco Lucioni, Roberto Marasca, Catherine Thieblemont, Daniela Capello, Fabio Facchetti, Ivo Kwee, Stefano A. Pileri, Robin Foà, Francesco Bertoni, Riccardo Dalla-Favera, Laura Pasqualucci and Gianluca Gaidano

Updated information and services can be found at:
<http://bloodjournal.hematologylibrary.org/content/118/18/4930.full.html>

Articles on similar topics can be found in the following Blood collections
[Brief Reports](#) (1703 articles)
[Lymphoid Neoplasia](#) (1598 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Copyright 2011 by The American Society of Hematology; all rights reserved.



Brief report

Alteration of *BIRC3* and multiple other NF- κ B pathway genes in splenic marginal zone lymphoma

Davide Rossi,¹ Silvia Deaglio,² David Dominguez-Sola,³ Silvia Rasi,¹ Tiziana Vaisitti,² Claudio Agostinelli,⁴ Valeria Spina,¹ Alessio Bruscaggini,¹ Sara Monti,¹ Michaela Cerri,¹ Stefania Cresta,¹ Marco Fangazio,¹ Luca Arcaini,⁵ Marco Lucioni,⁶ Roberto Marasca,⁷ Catherine Thieblemont,⁸ Daniela Capello,¹ Fabio Facchetti,⁹ Ivo Kwee,¹⁰ Stefano A. Pileri,⁴ Robin Foà,¹¹ Francesco Bertoni,¹⁰ Riccardo Dalla-Favera,^{3,12,13} Laura Pasqualucci,^{3,12,14} and Gianluca Gaidano¹

¹Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; ²Department of Genetics, Biology and Biochemistry, University of Turin, Turin, Italy; ³Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY; ⁴Haematopathology, Department L. & A. Seragnoli, University of Bologna, Bologna, Italy; Divisions of ⁵Hematology and ⁶Pathology, University of Pavia Medical School, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy; ⁷Division of Hematology, Department of Oncology and Hematology, University of Modena and Reggio Emilia, Modena, Italy; ⁸Service d'Hémo-Oncologie, Hôpital Saint-Louis, Université Paris Diderot, Paris, France; ⁹Divisions of Pathology, Spedali Civili di Brescia, Brescia, Italy; ¹⁰Laboratory of Experimental Oncology and Lymphoma Unit, Oncology Institute of Southern Switzerland (IOSI), Bellinzona, Switzerland; ¹¹Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy; Departments of ¹²Pathology & Cell Biology and ¹³Genetics & Development, Columbia University, New York, NY; and ¹⁴Institute of Hematology, University of Perugia, Perugia, Italy

Splenic marginal zone lymphoma (SMZL) is one of the few B-cell lymphoma types that remain orphan of molecular lesions in cancer-related genes. Detection of active NF- κ B signaling in 14 (58%) of 24 SMZLs prompted the investigation of NF- κ B molecular alterations in 101 SMZLs. Mutations and copy number abnormalities of NF- κ B genes occurred in 36 (36%) of 101 SMZLs and targeted both canon-

ical (*TNFAIP3* and *IKKB*) and noncanonical (*BIRC3*, *TRAF3*, *MAP3K14*) NF- κ B pathways. Most alterations were mutually exclusive, documenting the existence of multiple independent mechanisms affecting NF- κ B in SMZL. *BIRC3* inactivation in SMZL recurred because of somatic mutations that disrupted the same RING domain that in extranodal marginal zone lymphoma is removed by the t(11;18)

translocation, which points to *BIRC3* disruption as a common mechanism across marginal zone B-cell lymphomagenesis. Genetic lesions of NF- κ B provide a molecular basis for the pathogenesis of more than 30% of SMZLs and offer a suitable target for NF- κ B therapeutic approaches in this lymphoma. (*Blood*. 2011;118(18):4930-4934)

Introduction

Knowledge of the pathogenesis of splenic marginal zone lymphoma (SMZL) is scarce.^{1,2} On the basis of clinical and epidemiologic data, hepatitis C virus infection might play a role in a fraction of SMZLs.^{1,2} The contribution of antigen stimulation is suggested by the highly restricted immunoglobulin gene repertoire and by the stereotyped B-cell receptor in ~10% of SMZLs.^{3,4} Cancer genes involved in SMZL pathogenesis are currently unknown, except for *TP53* disruption in ~10%-15% of cases.^{3,5}

Indirect evidence points to nuclear factor- κ B (NF- κ B) pathway genes as attractive candidates in SMZL. Abnormal marginal zone B-cell expansions are observed frequently in animal models with constitutive NF- κ B activation.⁶⁻⁸ In addition, up-regulated expression of NF- κ B target genes in SMZL suggests the occurrence of NF- κ B activation,⁹ which in other B-cell malignancies is sustained by cancer-specific genetic lesions.^{10,11}

This rationale prompted the assessment of NF- κ B genetic alterations in SMZL. We document that the canonical and noncanonical NF- κ B pathways are affected by multiple genetic lesions in >30% of SMZLs.

Methods

Patient samples

The study was based on 101 SMZLs diagnosed according to World Health Organization classification and SMZL Working Party criteria (supplemental Table 1, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).^{1,2} Matched germline DNA was extracted from saliva. Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki. The study was approved by the ethics committee of the Amedeo Avogadro University.

Molecular studies

Mutation analysis of 20 NF- κ B pathway genes (supplemental Table 2; supplemental Figure 1) and *TP53* was performed by DNA Sanger sequencing. Copy number abnormalities (CNAs) of NF- κ B genes and *TP53* were analyzed by FISH (probes in supplemental Table 2) and single-nucleotide polymorphism array (GeneChip Human Mapping 250K NspI, Affymetrix; Gene Expression Omnibus accession number GSE24881).⁵ Immunoglobulin heavy chain gene rearrangements and stereotyped VH CDR3 were analyzed as reported previously.⁴

Submitted June 7, 2011; accepted August 21, 2011. Prepublished online as *Blood* First Edition paper, August 31, 2011; DOI 10.1182/blood-2011-06-359166.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2011 by The American Society of Hematology

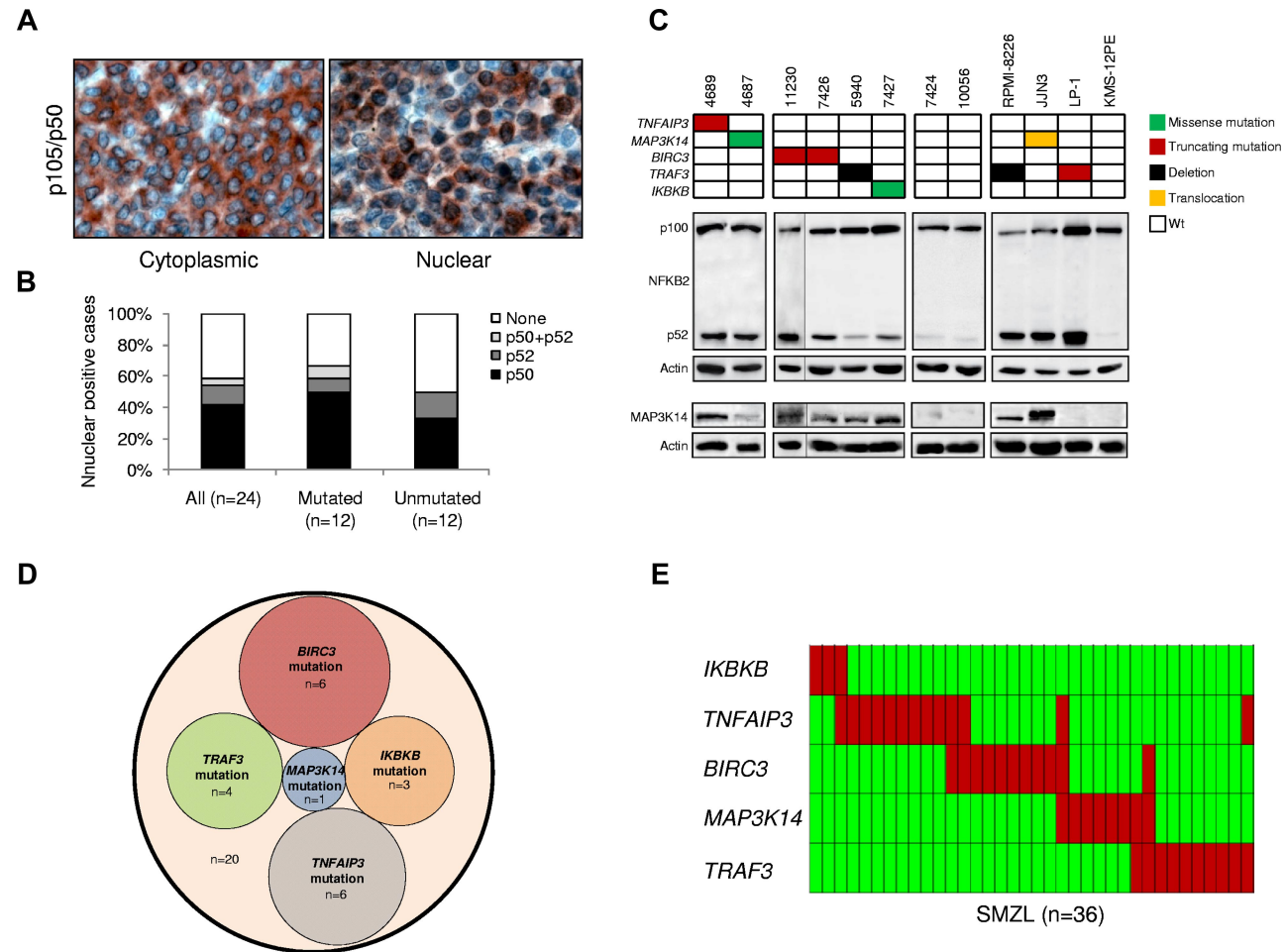


Figure 1. The NF- κ B pathway is activated and frequently targeted by genetic lesions in SMZLs. (A) Immunohistochemical staining of SMZL biopsy samples with anti-NFKB1 (p50) antibody. Nuclear localization of NF- κ B denotes active signaling as opposed to inactive cytoplasmic localization. Original magnification $\times 400$. (B) Prevalence of constitutive NF- κ B activation scored by immunohistochemistry in SMZLs overall and according to mutation status of NF- κ B genes. (C) Western blot analysis showing NFKB2 processing and MAP3K14 expression in purified primary tumor cells from 8 SMZL cases carrying wild-type (Wt) or aberrant NF- κ B genes. Case 11230 was run in a different gel using the same conditions, as indicated by the black dividing line. The RPMI-8226 and LP-1 (both multiple myeloma) and JLN3 (plasma cell leukemia) cell lines were used as positive controls for NFKB2 processing. The KMS-12PE cell line (multiple myeloma) was used as negative control for NFKB2 processing. The RPMI-8226 and JLN3 cell lines were used as positive controls for MAP3K14 expression. The LP-1 and KMS-12PE cell lines were used as negative controls for MAP3K14 expression. Actin was used as loading control. (D) Venn diagram illustrating the absence of overlap between NF- κ B gene somatic mutations. (E) Genetic lesions (including mutations and CNAs) of NF- κ B genes are largely mutually exclusive. In the heat map, rows correspond to identical genes, and columns represent individual patients color-coded on the basis of gene status (green, wild-type; red, mutations of *IKBKB*, mutations or gain of *MAP3K14*, mutations and/or deletion of *TRAF3*, *TNFAIP3*, and *BIRC3*).

Immunohistochemistry

Presence of active NF- κ B was analyzed by immunohistochemistry on formalin-fixed, paraffin-embedded primary SMZL biopsy samples ($n = 24$) with anti-NFKB1 (p50; Cell Signaling, #3035) and anti-NFKB2 (p52; Cell Signaling, #3017) primary antibodies.¹⁰ Cases with $> 20\%$ cells showing positive nuclear staining (of 200 cells counted in representative fields) were scored as NF- κ B active.¹⁰

Western blot studies

B cells from SMZL spleens were purified by negative selection with Dynal magnetic beads (Invitrogen). Proteins were resolved by SDS-PAGE and analyzed by Western blot. Antibodies were anti-NFKB2 (Cell Signaling, #4882), anti-MAP3K14 (Cell Signaling, #4994), and anti-actin (Santa Cruz Biotechnology, #sc-1615). Image acquisition and densitometric analyses were performed with ImageQuant LAS4000 and TL software (GE Healthcare).

Results and discussion

SMZLs were characterized for active NF- κ B signaling by immunohistochemical assays that detected nuclear NFKB1 (p50) and

NFKB2 (p52).¹⁰ Nuclear localization of NF- κ B was observed in tumor cells of 14 (58%) of 24 SMZLs (Figure 1A-B). Both canonical (11/24, 46%) and noncanonical (4/24, 17%) NF- κ B pathways were activated. One patient showed activation of both pathways. Engagement of the noncanonical NF- κ B pathway was confirmed by biochemical detection of NFKB2 processing (Figure 1C).

To investigate the molecular basis of NF- κ B activation, NF- κ B pathway genes were screened for mutations and CNAs in 101 SMZLs. Mutations that targeted multiple NF- κ B genes were detected in 20 (20%) of 101 SMZLs (supplemental Table 3). Mutations affected key regulators of both canonical (*IKBKB*, *TNFAIP3*) and noncanonical (*BIRC3*, *TRAF3*, *MAP3K14*) NF- κ B pathways¹² and were mutually exclusive, which points to multiple molecular mechanisms targeting NF- κ B in SMZLs (Figure 1D). By combining mutations and CNAs identified by single-nucleotide polymorphism array and validated by FISH (supplemental Table 4; Figure 1E), NF- κ B was targeted in 36 (36%) of 101 SMZLs. The mutually exclusive distribution of NF- κ B lesions was confirmed in most cases (29/36, 80%) and when mutations and CNAs were

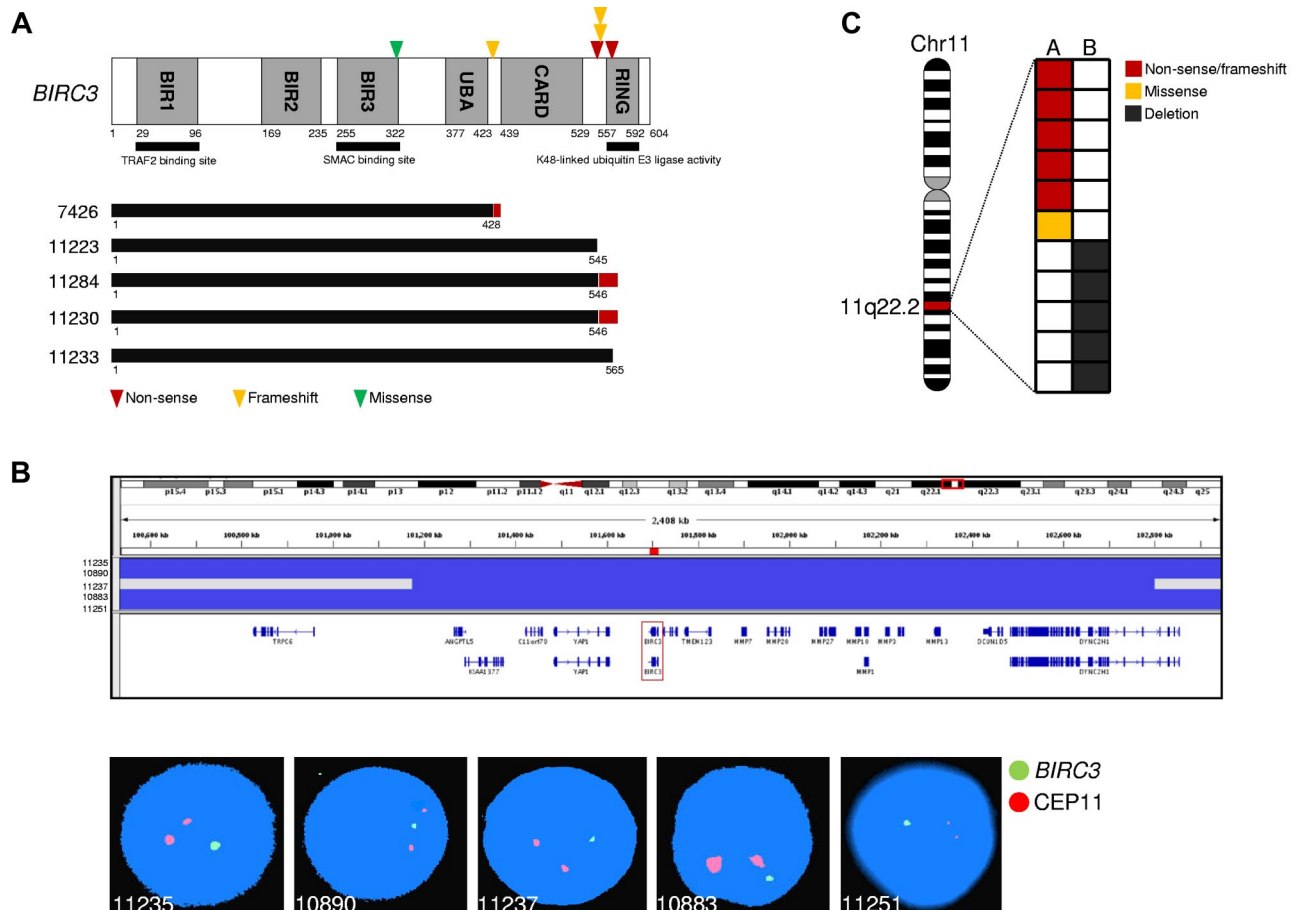


Figure 2. *BIRC3* mutations in SMZL. (A) Schematic representation of the *BIRC3* protein with its key functional domains. Color-coded symbols indicate the type and position of the mutations on *BIRC3*. The predicted *BIRC3* polypeptides in SMZL cases harboring mutations that truncate the *BIRC3* RING domain are aligned below (black bars represent normal coding regions; red bars represent the position of frameshifts and the length of translation before a stop codon is encountered). (B) Graphic representation of segmentation data from SMZL carrying *BIRC3* deletions, visualized with Integrative Genomics Viewer (IGV) software (<http://www.broadinstitute.org/igv>). Each track represents 1 sample, where gray denotes a normal (diploid) copy number and blue indicates region of a copy number loss. Individual genes in the region are aligned in the bottom panel. Dual-color FISH validation of *BIRC3* deletions (RP11-177O8-*BIRC3*-specific probe in green and chromosome 11 centromeric probe in orange). (C) Allelic (A or B) distribution of *BIRC3* genetic lesions in individual SMZL samples. Chr11 indicates chromosome 11.

considered together (Figure 1E). These data document that NF- κ B gene lesions are recurrent in SMZL, a disease that until now was free from the molecular lesions that affect cancer genes.¹

The TRAF3/MAP3K14-TRAF2/*BIRC3* regulatory complex of the noncanonical NF- κ B pathway¹²⁻¹⁴ was targeted in 25 (25%) of 101 SMZLs. *BIRC3* was affected in 11 (11%) of 101 SMZLs by inactivating mutations (3 frameshift and 2 nonsense), missense mutations (n = 1), and gene deletions (n = 5; supplemental Tables 3-4; Figure 2). All inactivating mutations were somatically acquired, were predicted to generate aberrant transcripts carrying premature stop codons, and caused elimination or truncation of the C-terminal RING domain, the E3 ubiquitin ligase activity of which is essential for MAP3K14 proteasomal degradation by *BIRC3* (Figure 2A).^{8,12,14} The sole *BIRC3* missense mutation targeted a conserved Zn binding site within the BIR3 domain that is also occasionally mutated in solid tumors (COSMIC [Catalogue Of Somatic Mutations In Cancer] v52; <http://www.sanger.ac.uk/genetics/CGP>) and was categorized as probably damaging by PolyPhen-2 (Polymorphism Phenotyping v2; supplemental Table 3; Figure 2A). The identification of *BIRC3*-inactivating mutations in SMZL points to *BIRC3* disruption as a common mechanism across marginal zone B-cell lymphomagenesis. In fact, disruption of the *BIRC3* RING domain, which in SMZL is produced by inactivating mutations, in extranodal marginal zone lymphoma is

caused by a t(11;18)(q21;q21) translocation that leads to the *BIRC3*/MALT1 fusion protein, which lacks the RING domain of *BIRC3*.¹⁵

BIRC3 mutations in SMZLs were all monoallelic, in agreement with independent evidence suggesting that *BIRC3* genetic lesions may exert a dominant negative effect. First, *BIRC3* experimental mutants with a disrupted RING domain abrogate the ability of wild-type *BIRC3* to efficiently target proteins for proteasomal degradation.^{16,17} Second, monoallelic disruption of the RING domain of *BIRC3* contributes to *BIRC3*/MALT1 fusion protein oncogenicity in extranodal marginal zone lymphoma.¹⁷ Consistently, all tested SMZL primary cases that harbored monoallelic *BIRC3* disruption displayed constitutive NF- κ B activation, including MAP3K14 accumulation, and active NFKB2 processing (Figure 1C) and/or NF- κ B localization by immunohistochemistry.

TRAF3 was affected in 10 (10%) of 101 SMZLs by inactivating mutations (2 frameshift and 1 nonsense), missense mutations (n = 1), and deletions (n = 7). Inactivating mutations were predicted to generate aberrant transcripts carrying premature stop codons and causing elimination or truncation of the C-terminal MATH domain, which provides the docking site for MAP3K14 and is required for MAP3K14 recruitment to *BIRC3* degradation (supplemental Tables 3-4; supplemental Figure 2).^{6,12-14,18,19} The sole missense mutation of *TRAF3* targeted a TRAF domain that

interacts with TRAF2,¹⁹ was categorized as probably damaging by PolyPhen-2, and was coupled with deletion of the second *TRAF3* allele (supplemental Table 3; supplemental Figure 2). Genetic lesions of *TRAF3* are also recurrent in multiple myeloma, where they contribute to NF- κ B signaling activation.^{20,21} SMZL primary cells harboring monoallelic *TRAF3* inactivation showed NF- κ B localization, MAP3K14 accumulation, and active NFKB2 processing (Figure 1C).

MAP3K14 (also known as *NIK*) was affected by gain ($n = 7$) or missense ($n = 1$) mutations in 8 (8%) of 101 SMZLs (supplemental Tables 3-4; supplemental Figure 3). The missense mutation was somatically acquired and mapped near the kinase domain and far from the TRAF3-binding domain, which suggests a potential role in MAP3K14 downstream signaling activation rather than a role in MAP3K14 protein stabilization (supplemental Table 3; Figure 1C; supplemental Figure 3).

BIRC3 and TRAF3, along with TRAF2, cooperate in the same protein complex that negatively regulates MAP3K14, the central activator of noncanonical NF- κ B signaling.¹²⁻¹⁴ In SMZLs, genetic lesions activate noncanonical NF- κ B either by disrupting BIRC3 and TRAF3 domains required for functional integrity of the MAP3K14 regulatory complex or directly through MAP3K14 gain or mutation (Figure 1C).^{12-14,18-20} Notably, animal models that harbor *BIRC3* inactivation, conditional deletion of *TRAF3*, or *MAP3K14* overexpression are all characterized by B-cell enhanced proliferation and survival, which leads to abnormal marginal zone B-cell expansions.⁶⁻⁸

In addition to noncanonical signaling, the canonical NF- κ B pathway was affected in SMZLs (15/101, 15%) by genetic lesions of *TNFAIP3* and *IKBKB*, which encode for negative and positive regulators of NF- κ B signaling, respectively.¹² *TNFAIP3* was affected in 13 (13%) of 101 SMZLs by inactivating mutations (6 frameshift mutations) and deletions ($n = 9$), including a focal loss that points to *TNFAIP3* as the specific target of deletions (supplemental Tables 3-4; supplemental Figure 4). *TNFAIP3* lesions in SMZLs displayed a preferentially monoallelic distribution, as also observed in other indolent B-cell malignancies in which monoallelic *TNFAIP3* inactivation is associated with low expression of the residual wild-type allele, which suggests a haploinsufficient role in tumor suppression.^{22,23}

IKBKB was mutated in 3 (3%) of 101 SMZLs carrying a recurrent, somatically acquired missense mutation (K171E) that targeted the IKBKB kinase domain near the activation loop (supplemental Table 3; supplemental Figure 5). This mutation leads to an amino acid substitution within a site that is conserved both intraspecies and interspecies and is predicted as possibly being damaging according to PolyPhen-2 (supplemental Table 3; supplemental Figure 5). The potential pathogenetic relevance of this mutation is supported by the notion that mutations within the activation loop constitutively activate *IKBKB* tyrosine kinase activity and NF- κ B signaling in vitro²⁴ and contribute to B-cell lymphomagenesis in mouse models.²⁵

The preferential monoallelic distribution of NF- κ B gene abnormalities in SMZL is consistent with the preferential monoallelic

distribution of these lesions in other indolent B-cell malignancies.^{22,23} The occurrence of NF- κ B molecular lesions in SMZLs was observed independent of hepatitis C virus infection ($P = .402$), *IGHV* gene mutation status ($P = .065$), *IGHV1-2* gene usage ($P = .761$), stereotyped VH CDR3 ($P = .969$), 7q deletion ($P = .621$), or *TP53* disruption ($P = .638$), which suggests that genetic deregulation of NF- κ B is not restricted to specific SMZL clinico-molecular subgroups.¹⁻⁴

The identification of NF- κ B mutations in SMZLs elucidates the molecular basis of a fraction of these lymphomas and expands the number of genetic mechanisms that activate NF- κ B in B-cell neoplasia. In addition to its pathogenetic relevance, the identification of NF- κ B as a pathway repeatedly involved in SMZL provides the rationale for targeted anti-NF- κ B therapeutic approaches in this lymphoma.

Acknowledgments

The authors are grateful to Professor Mariano Rocchi and Dr Giulia Daniele, University of Bari, Bari, Italy, for the generous gift of several FISH probes used in this study.

This study was supported by AIRC, Special Program in Clinical Molecular Oncology 5 $\times$ 1000, No. 10007, Milan, Italy; Progetto FIRB-Programma "Futuro in Ricerca" 2008, PRIN 2008 and PRIN 2009, MIUR, Rome, Italy; Progetto Giovani Ricercatori 2008, Ministero della Salute, Rome, Italy; Novara-Associazione Italiana contro le Leucemie, Linfomi e Mielomi (AIL) Onlus, Novara, Italy; Oncosuisse grant OCS-02034-02-2007; Swiss National Science Foundation grant 205321-112430; Fondazione per la Ricerca e la Cura sui Linfomi, Lugano, Switzerland; and Computational Life Science/Ticino in rete. V.S. has been supported by the Associazione Franca Capurro per Novara Onlus, Novara, Italy.

Authorship

Contribution: D.R. and G.G. designed the study, analyzed and interpreted data, performed statistical analysis, and drafted the manuscript; S.D. and T.V. performed and interpreted biochemical assays; D.D.-S. and C.A. performed and interpreted immunohistochemistry assays; S.R., V.S., A.B., M.C., S.C., and D.C. performed mutation analysis and interpreted data; S.M. performed FISH analysis and interpreted data; M.F., L.A., M.L., R.M., C.T., and F.F. provided well-characterized biologic samples; I.K. and F.B. provided single-nucleotide polymorphism array data; and D.D.-S., S.A.P., R.F., R.D.-F., and L.P. contributed to data interpretation and drafting of the manuscript.

Conflict-of-interest statement: The authors declare no competing financial interests.

Correspondence: Davide Rossi, MD, PhD, Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy; e-mail: rossidav@med.unipmn.it.

References

- Isaacson PG, Piris MA, Berger F, et al. Splenic marginal zone lymphoma. In: Swerdlow SH, Campo E, Harris NL, et al, eds. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC; 2008:185-187.
- Matutes E, Oscier D, Montalban C, et al. Splenic marginal zone lymphoma proposals for a revision of diagnostic, staging and therapeutic criteria. *Leukemia*. 2008;22(3):487-495.
- Salido M, Baró C, Oscier D, et al. Cytogenetic aberrations and their prognostic value in a series of 330 splenic marginal zone B-cell lymphomas: a multicenter study of the Splenic B-Cell Lymphoma Group. *Blood*. 2010;116(9):1479-1488.
- Zibellini S, Capello D, Forconi F, et al. Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. *Haematologica*. 2010;95(10):1792-1796.
- Rinaldi A, Mian M, Chigrinova E, et al. Genome-wide DNA profiling of marginal zone lymphomas identifies subtype-specific lesions with an impact

- on the clinical outcome. *Blood*. 2011;117(5):1595-1604.
6. Xie P, Stunz LL, Larison KD, Yang B, Bishop GA. Tumor necrosis factor receptor-associated factor 3 is a critical regulator of B cell homeostasis in secondary lymphoid organs. *Immunity*. 2007;27(2):253-267.
 7. Sasaki Y, Calado DP, Derudder E, et al. NIK overexpression amplifies, whereas ablation of its TRAF3-binding domain replaces BAFF:BAFF-R-mediated survival signals in B cells. *Proc Natl Acad Sci U S A*. 2008;105(31):10883-10888.
 8. Conze DB, Zhao Y, Ashwell JD. Non-canonical NF-kappaB activation and abnormal B cell accumulation in mice expressing ubiquitin protein ligase-inactive c-IAP2. *PLoS Biol*. 2010;8(10):e1000518.
 9. Ruiz-Ballesteros E, Mollejo M, Rodriguez A, et al. Splenic marginal zone lymphoma: proposal of new diagnostic and prognostic markers identified after tissue and cDNA microarray analysis. *Blood*. 2005;106(5):1831-1838.
 10. Compagno M, Lim WK, Grunn A, et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature*. 2009;459(7247):717-721.
 11. Staudt LM. Oncogenic activation of NF-kappaB. *Cold Spring Harb Perspect Biol*. 2010;2(6):a000109.
 12. Vallabhapurapu S, Karin M. Regulation and function of NF-kappaB transcription factors in the immune system. *Annu Rev Immunol*. 2009;27:693-733.
 13. Vallabhapurapu S, Matsuzawa A, Zhang W, et al. Nonredundant and complementary functions of TRAF2 and TRAF3 in a ubiquitination cascade that activates NIK-dependent alternative NF-kappaB signaling. *Nat Immunol*. 2008;9(12):1364-1370.
 14. Zarnegar BJ, Wang Y, Mahoney DJ, et al. Noncanonical NF-kappaB activation requires coordinated assembly of a regulatory complex of the adaptors cIAP1, cIAP2, TRAF2 and TRAF3 and the kinase NIK. *Nat Immunol*. 2008;9(12):1371-1378.
 15. Dierlamm J, Baens M, Wlodarska I, et al. The apoptosis inhibitor gene API2 and a novel 18q gene, MLT, are recurrently rearranged in the t(11;18)(q21;q21) associated with mucosa-associated lymphoid tissue lymphomas. *Blood*. 1999;93(11):3601-3609.
 16. Li X, Yang Y, Ashwell JD. TNF-RII and c-IAP1 mediate ubiquitination and degradation of TRAF2. *Nature*. 2002;416(6878):345-347.
 17. Hu S, Du MQ, Park SM, et al. cIAP2 is a ubiquitin protein ligase for BCL10 and is dysregulated in mucosa-associated lymphoid tissue lymphomas. *J Clin Invest*. 2006;116(1):174-181.
 18. He JQ, Saha SK, Kang JR, Zarnegar B, Cheng G. Specificity of TRAF3 in its negative regulation of the noncanonical NF-kappa B pathway. *J Biol Chem*. 2007;282(6):3688-3694.
 19. He L, Grammer AC, Wu X, Lipsky PE. TRAF3 forms heterotrimers with TRAF2 and modulates its ability to mediate NF-kappaB activation. *J Biol Chem*. 2004;279(53):55855-55865.
 20. Annunziata CM, Davis RE, Demchenko Y, et al. Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma. *Cancer Cell*. 2007;12(2):115-130.
 21. Keats JJ, Fonseca R, Chesi M, et al. Promiscuous mutations activate the noncanonical NF-kappaB pathway in multiple myeloma. *Cancer Cell*. 2007;12(2):131-144.
 22. Novak U, Rinaldi A, Kwee I, et al. The NF-kB negative regulator TNFAIP3 (A20) is inactivated by somatic mutations and genomic deletions in marginal zone lymphomas. *Blood*. 2009;113(20):4918-4921.
 23. Braggio E, Keats JJ, Leleu X, et al. Identification of copy number abnormalities and inactivating mutations in two negative regulators of nuclear factor-kappaB signaling pathways in Waldenstrom's macroglobulinemia. *Cancer Res*. 2009;69(8):3579-3588.
 24. Mercurio F, Zhu H, Murray BW, et al. IKK-1 and IKK-2: cytokine-activated IkappaB kinases essential for NF-kappaB activation. *Science*. 1997;278(5339):860-866.
 25. Calado DP, Zhang B, Srinivasan L, et al. Constitutive canonical NF-kappaB activation cooperates with disruption of BLIMP1 in the pathogenesis of activated B cell-like diffuse large cell lymphoma. *Cancer Cell*. 2010;18(6):580-589.

blood

2011 118: 6904-6908
Prepublished online October 28, 2011;
doi:10.1182/blood-2011-08-373159

Mutations of the *SF3B1* splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness

Davide Rossi, Alessio Bruscaggin, Valeria Spina, Silvia Rasi, Hossein Khiabanian, Monica Messina, Marco Fangazio, Tiziana Vaisitti, Sara Monti, Sabina Chiaretti, Anna Guarini, Ilaria Del Giudice, Michaela Cerri, Stefania Cresta, Clara Deambrogi, Ernesto Gargiulo, Valter Gattei, Francesco Forconi, Francesco Bertoni, Silvia Deaglio, Raul Rabadan, Laura Pasqualucci, Robin Foà, Riccardo Dalla-Favera and Gianluca Gaidano

Updated information and services can be found at:
<http://bloodjournal.hematologylibrary.org/content/118/26/6904.full.html>

Articles on similar topics can be found in the following Blood collections

[Brief Reports](#) (1703 articles)
[Clinical Trials and Observations](#) (3800 articles)
[Free Research Articles](#) (2175 articles)
[Lymphoid Neoplasia](#) (1598 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>



Brief report

Mutations of the *SF3B1* splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness

*Davide Rossi,¹ *Alessio Bruscaggin,¹ Valeria Spina,¹ Silvia Rasi,¹ Hossein Khiabanian,² Monica Messina,³ Marco Fangazio,¹ Tiziana Vaisitti,⁴ Sara Monti,¹ Sabina Chiaretti,⁵ Anna Guarini,⁵ Ilaria Del Giudice,⁵ Michaela Cerri,¹ Stefania Cresta,¹ Clara Deambrogi,¹ Ernesto Gargiulo,¹ Valter Gattei,⁶ Francesco Forconi,⁷ Francesco Bertoni,⁸ Silvia Deaglio,⁴ Raul Rabadan,² Laura Pasqualucci,^{3,9,10} †Robin Foà,^{5,11} †Riccardo Dalla-Favera,^{3,9,12,13} and †Gianluca Gaidano¹

¹Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; ²Department of Biomedical Informatics and Center for Computational Biology and Bioinformatics and ³Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY; ⁴Department of Genetics, Biology and Biochemistry and Human Genetics Foundation, University of Turin, Turin, Italy; ⁵Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy; ⁶Clinical and Experimental Onco-Hematology, Centro di Riferimento Oncologico, Istituto di Ricovero e Cura a Carattere Scientifico, Aviano, Italy; ⁷Division of Hematology, University of Siena, Siena, Italy; ⁸Oncology Institute of Southern Switzerland–Istituto Oncologico della Svizzera Italiana, Bellinzona, Switzerland; ⁹Department of Pathology & Cell Biology, Columbia University, New York, NY; ¹⁰Institute of Hematology, University of Perugia, Perugia, Italy; ¹¹Fondazione Eleonora Lorillard Spencer Cenci, Sapienza University, Rome, Italy; and Departments of ¹²Genetics and Development and ¹³Microbiology and Immunology, Columbia University, New York, NY

The genetic lesions identified in chronic lymphocytic leukemia (CLL) do not entirely recapitulate the disease pathogenesis and the development of serious complications, such as chemorefractoriness. While investigating the coding genome of fludarabine-refractory CLL, we observed that mutations of *SF3B1*, encoding a splicing factor and representing a critical component of the cell spliceosome, were re-

current in 10 of 59 (17%) fludarabine-refractory cases, with a frequency significantly greater than that observed in a consecutive CLL cohort sampled at diagnosis (17/301, 5%; $P = .002$). Mutations were somatically acquired, were generally represented by missense nucleotide changes, clustered in selected HEAT repeats of the *SF3B1* protein, recurrently targeted 3 hotspots (codons 662, 666, and

700), and were predictive of a poor prognosis. In fludarabine-refractory CLL, *SF3B1* mutations and *TP53* disruption distributed in a mutually exclusive fashion ($P = .046$). The identification of *SF3B1* mutations points to splicing regulation as a novel pathogenetic mechanism of potential clinical relevance in CLL. (*Blood*. 2011; 118(26):6904-6908)

Introduction

The clinical course of chronic lymphocytic leukemia (CLL) ranges from a very indolent disorder with a normal lifespan for the patient to a rapidly progressive disease that leads to death. Occasionally, CLL undergoes a transformation to Richter syndrome (RS).¹⁻³ The variable clinical course of CLL is driven, at least in part, by the disease's immunogenetic and molecular heterogeneity.⁴

Despite recent advances, the genetic lesions identified to date do not fully recapitulate the molecular pathogenesis of CLL and do not entirely explain the development of severe complications, such as chemorefractoriness, which still represent unmet clinical needs.⁵ In approximately 40% of cases, refractoriness to fludarabine is attributable to the disruption of *TP53*, but in a sizeable fraction of patients, the molecular basis of this aggressive phenotype remains unclear.⁶

Recently, 2 independent studies of the CLL coding genome investigated at disease presentation have revealed a restricted number of mutated genes, including *NOTCH1*.^{7,8} These studies have provided a proof of concept that, similar to other malignancies, genome-wide mutational analysis might identify novel lesions

of biologic and clinical relevance in CLL. On these grounds, we have embarked on the investigation of the coding genome of fludarabine-refractory CLL to identify genetic lesions associated with chemorefractoriness. The initial phases of this analysis have revealed recurrent mutations of *SF3B1*, a critical component of the cell spliceosome, pointing to the potential involvement of splicing regulation in CLL pathogenesis and chemorefractoriness.

Methods

Patients

The study population comprised 3 cohorts representative of different disease phases: (1) fludarabine-refractory CLL ($n = 59$), including cases ($n = 11$) subjected to whole-exome sequencing (supplemental Table 1, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article); (2) a consecutive series of newly diagnosed and previously untreated patients with CLL ($n = 301$; supplemental Table 2); and (3) clonally related RS ($n = 33$; all diffuse large B-cell lymphomas;

Submitted August 11, 2011; accepted October 18, 2011. Prepublished online as *Blood* First Edition paper, October 28, 2011; DOI 10.1182/blood-2011-08-373159.

*D.R. and A.B. contributed equally to this study.

†R.F., R.D.-F., and G.G. contributed equally to this study.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2011 by The American Society of Hematology

supplemental Table 3). Diagnosis of CLL and of fludarabine-refractoriness were based on International Workshop on CLL–National Cancer Institute (IWCLL-NCI) criteria²; RS was based on histologic criteria.^{1,9}

Peripheral blood tumor samples were obtained as follows: (1) for patients with fludarabine-refractory CLL, immediately before starting treatment to which the patient did not respond because of stable/progressive disease; and (2) for newly diagnosed and previously untreated CLL, at disease presentation. All studies on RS were performed on RS diagnostic biopsies. Germline DNA samples from the same patients were obtained from saliva or from purified granulocytes and confirmed to be tumor-free by PCR of tumor-specific *IGHV-D-J* rearrangements. Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki. The study was approved by the ethical committee of the Ospedale Maggiore della Carità di Novara associated with the Amedeo Avogadro University of Eastern Piedmont (protocol code 59/CE; study number CE 8/11).

Mutation analysis of *SF3B1*

Mutation analysis of *SF3B1* (exons 1-25, including splice sites; RefSeq NM_012433.2) was performed on PCR amplicons obtained from genomic DNA by a combination of Sanger sequencing (ABI PRISM 3100 Genetic Analyzer; Applied Biosystems) and targeted next generation sequencing (Genome Sequencer Junior, 454 Life Sciences; Roche; mean coverage approximately 200×). Additional details of sequencing strategies are in supplemental Methods.

FISH karyotype; mutation analysis of *IGHV*, *TP53*, and *NOTCH1*; copy number analysis; and gene expression profile analysis

FISH analysis was performed with the use of probes LSI13 and LSID13S319, CEP12, LSIp53, and LSIATM (Abbott).³ *IGHV* sequences were aligned to ImMunoGeneTics directory and considered mutated if their identity to corresponding germline genes was < 98%.³ *TP53* and *NOTCH1* mutations were analyzed by Sanger sequencing.^{3,7} Genome-wide DNA profiles were obtained with the Affymetrix Genome-Wide Human SNP Array 6.0. Gene expression profile analysis was performed with the use of Affymetrix HG-U133Plus2 arrays. Further details are reported in supplemental Methods.

Statistical analysis

Overall survival was measured from date of diagnosis to date of death (event) or last follow-up (censoring). Treatment-free survival was measured from date of diagnosis to date of progression to symptomatic disease requiring treatment according to IWCLL-NCI guidelines (event),² death, or last follow-up (censoring). Further details are reported in supplemental Methods.

Results and discussion

After the initial observation of recurrent *SF3B1* mutations in 3 of 11 fludarabine-refractory CLL analyzed by whole-exome sequencing, we performed targeted resequencing of the *SF3B1* coding sequence and splice sites in 48 additional cases of progressive and fludarabine-refractory CLL (total number of cases analyzed: 59; supplemental Table 1). *SF3B1* was altered in 10 of 59 (17%) fludarabine-refractory CLL by missense mutations ($n = 9$) or in-frame deletions ($n = 1$) clustering in the HEAT3, HEAT4, and HEAT5 repeats of the *SF3B1* protein (Figures 1 and 2A; supplemental Table 4). Two sites that are highly conserved interspecies (codon 662 and codon 700) were recurrently mutated in 3 and 5 cases, respectively (Figure 1). *SF3B1* mutations were monoallelic and were predicted to be functionally significant according to the PolyPhen-2 algorithm (supplemental Table 4).¹⁰ These data document that mutations of *SF3B1*, a splicing factor that is a critical component of the spliceosome, recurrently associate with fludarabine-refractory CLL.

The biologic characteristics of fludarabine-refractory CLL harboring *SF3B1* mutations are summarized in supplemental Table 1. Mutations occurred irrespective of the *IGHV* mutation status, CD38 expression, and ZAP70 expression. At the time of fludarabine-refractoriness, *SF3B1* mutations were enriched in cases harboring a normal FISH karyotype ($P = .008$; supplemental Table 1). In addition, *SF3B1* mutations distributed in a mutually exclusive fashion compared with *TP53* disruption tested by deletion and/or mutation (mutual information $I = 0.0609$; $P = .046$; Figure 2B). By combining *SF3B1* mutations with other genetic lesions enriched in chemorefractory cases (*TP53* disruption, *NOTCH1* mutations, *ATM* deletion),^{7,11-13} fludarabine-refractory CLL appeared to be characterized by multiple molecular alterations that, to some extent, are mutually exclusive (Figure 2B).

To investigate whether *SF3B1* mutations are restricted to chemorefractory cases, we then compared the prevalence of mutations observed at the time of fludarabine-refractoriness to the prevalence of mutations observed in other disease phases. In a consecutive series evaluated at CLL diagnosis, *SF3B1* mutations were rare (17/301; 5%; Figure 2A; supplemental Table 4) and occurred irrespective of other molecular and immunogenetic features (supplemental Table 2; supplemental Table 5; supplemental Figure 1). Remarkably, 5 of 17 (29%) CLL mutated at diagnosis were primary fludarabine-refractory patients. In these 5 cases, *TP53* disruption and *NOTCH1* mutations occurred in 1 case each. None of the 12 remaining cases harbored *TP53* disruption or *NOTCH1* mutations. By univariate analysis, *SF3B1* mutations showed a crude association with short treatment-free survival ($P < .001$) and overall survival ($P = .011$; Figure 2C). By multivariate analysis, the increased risk of death predicted by *SF3B1* mutations was independent (hazard ratio 3.02; 95% confidence interval 1.24-7.35; $P = .015$) of confounding clinical and biologic variables (supplemental Table 6). Confirmation within the frame of prospective clinical trials will be helpful to fully assess the generalization of *SF3B1* mutations as a CLL prognostic marker.

In CLL investigated at diagnosis, the hotspot distribution and molecular spectrum of *SF3B1* mutations, as well as their mutual relationship with other genetic lesions, were similar to those observed in fludarabine-refractory CLL (Figures 1 and 2B; supplemental Table 4). *SF3B1* mutations were only found in 2 of 33 (6%) clonally related RS (Figures 1 and 2A; supplemental Table 4). Across the different disease phases investigated, mutations were confirmed to be somatically acquired in all cases ($n = 18$) for which germline DNA was available (supplemental Table 4). Among the 3 *SF3B1* mutated cases for which serial samples were analyzed, *SF3B1* mutations were acquired in 2 cases. One fludarabine-refractory CLL (case 7915 in supplemental Table 4) acquired the c.2044A > G p.K666E mutation at the time of refractoriness, and one RS (case 7509 in supplemental Table 4) acquired the c.2146A > G p.K700E mutation at the time of transformation. In the remaining case (case 8343 in supplemental Table 4), the *SF3B1* mutation was present in all disease phases.

Although the relative expression of *SF3B1* in CLL was greater compared with normal B-cell subsets (Figure 2D), extensive investigation by single-nucleotide polymorphism array analysis ruled out focal copy number abnormalities of *SF3B1* in this leukemia ($n = 0/323$). *SF3B1* mutations were consistently absent among mature B-cell neoplasms ($n = 136$) other than CLL (supplemental Table 7). These data document that *SF3B1* mutations: (1) are specific for CLL among mature B-cell neoplasms; (2) occur at a low rate at CLL presentation, whereas they are enriched in fludarabine-refractory cases; and (3) play a minor role in RS transformation, corroborating the notion

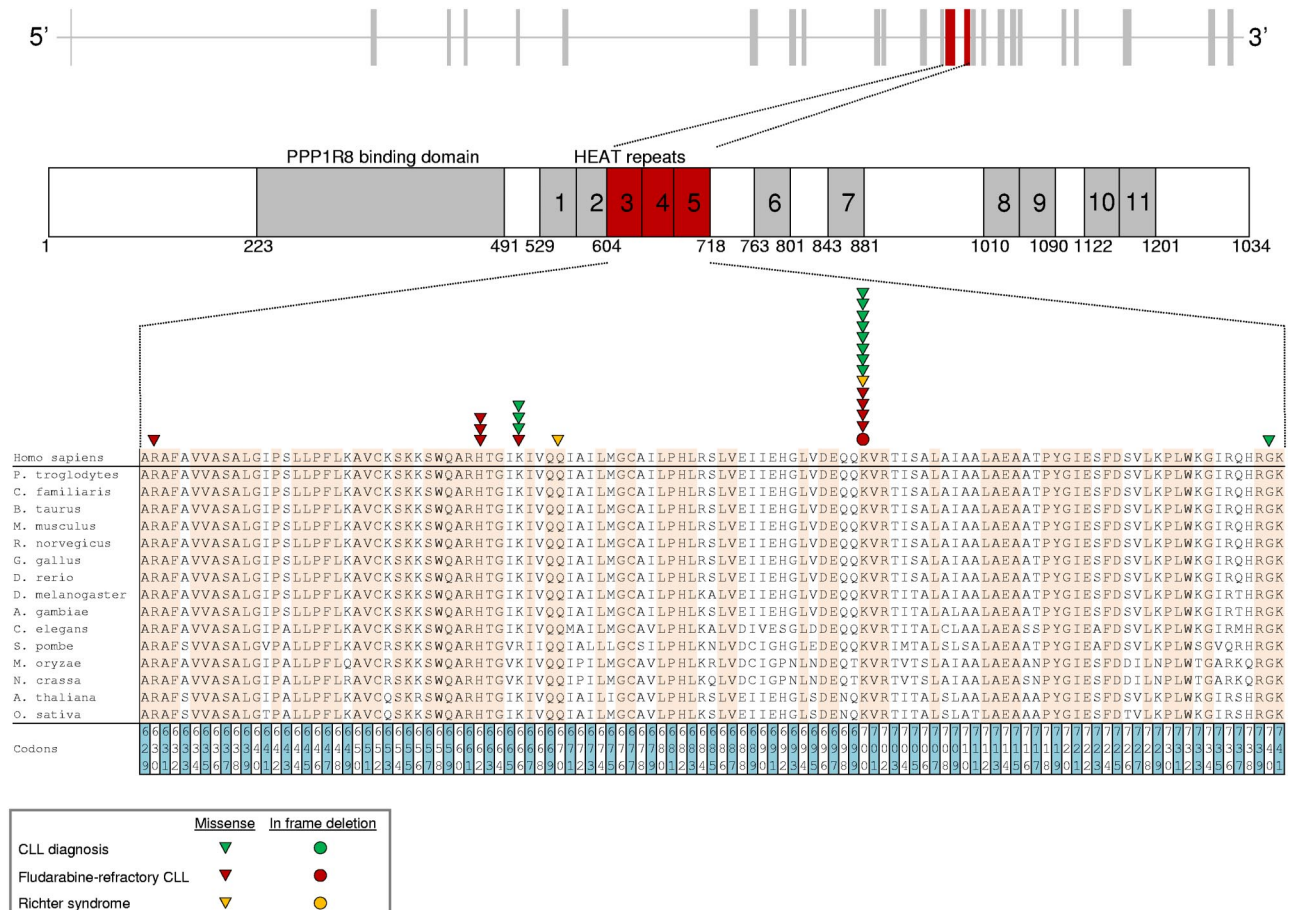


Figure 1. *SF3B1* mutations in CLL and RS. Schematic diagram of the human *SF3B1* gene (top) and protein (bottom) with its functional domains (PPP1R8 binding domain and HEAT repeats), and multiple alignment of the HEAT3, HEAT4, and HEAT5 amino acid sequences of the human *SF3B1* protein with orthologous *SF3B1* proteins (n = 15). Amino acids conserved among species are highlighted. Color-coded shapes indicate the position of the mutations found in CLL at diagnosis, in fludarabine-refractory CLL, and in RS.

that CLL histologic shift is molecularly distinct from chemorefractory progression without RS transformation.³

Our identification of *SF3B1* mutations in CLL, and the recent discovery of *SF3B1* mutations in myelodysplasia, points to the involvement of splicing regulation as a novel pathogenetic mechanism in hematologic malignancies.^{14,15} *SF3B1* is a critical component of both major (U2-like) and minor (U12-like) spliceosomes,¹⁶⁻¹⁸ which enact the precise excision of introns from pre-mRNA.¹⁹⁻²¹ The precise biologic role of *SF3B1* mutations in CLL is currently elusive and will require dedicated studies. The pathogenicity of *SF3B1* mutations in CLL is strongly supported by the clustering of these mutations in evolutionarily conserved hotspots localized within HEAT domains, which are tandemly arranged curlicue-like structures serving as flexible scaffolding on which other components can assemble.^{22,23} In addition, the observation that *SF3B1* regulates the alternative splicing program of genes controlling cell-cycle progression and apoptosis points to a potential contribution of *SF3B1* mutations in modulating tumor cell proliferation and survival.^{20,24,25} In addition to pathogenetic implications, *SF3B1* mutations might also provide a therapeutic target for *SF3B1* inhibitors,^{24,25} which are currently under preclinical development as anticancer drugs.

Acknowledgments

This study was supported by grants from AIRC, Special Program Molecular Clinical Oncology, 5 x 1000, No. 10007, Milan, Italy (to

G.G. and R.F.); Progetto FIRB-Programma “Futuro in Ricerca” 2008 (to D.R.), PRIN 2008 (to G.G.), and PRIN 2009 (to D.R.), MIUR, Rome, Italy; Progetto Giovani Ricercatori 2008 (to D.R.) and Ricerca Sanitaria Finalizzata 2008 (to G.G.), Ministero della Salute, Rome, Italy; Compagnia di San Paolo, Turin, Italy (to R.F.); NIH Grant PO1-CA092625 and a Specialized Center of Research grant from the Leukemia & Lymphoma Society (both to R.D.-F.); and Helmut Horten Foundation and San Salvatore Foundation (to F.B.). S.M. and S. Cresta are being supported by fellowships from Novara-AIL Onlus, Novara.

Authorship

Contribution: D.R., L.P., R.F., R.D.-F., and G.G. designed the study, interpreted data, and wrote the manuscript; A.B., V.S., S.R., M.M., S.M., M.C., S. Cresta, and E.G. performed and interpreted mutational analysis; C.D. performed and interpreted FISH studies; H.K. and R.R. performed and interpreted bioinformatics studies; T.V. and S.D. contributed to molecular data analysis and interpretation; M.F., S. Chiaretti, A.G., I.D.G., and F.F. provided well-characterized clinical samples; V.G. performed immunophenotypic studies; and F.B. and L.P. interpreted SNP array data.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Riccardo Dalla-Favera, MD, Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer

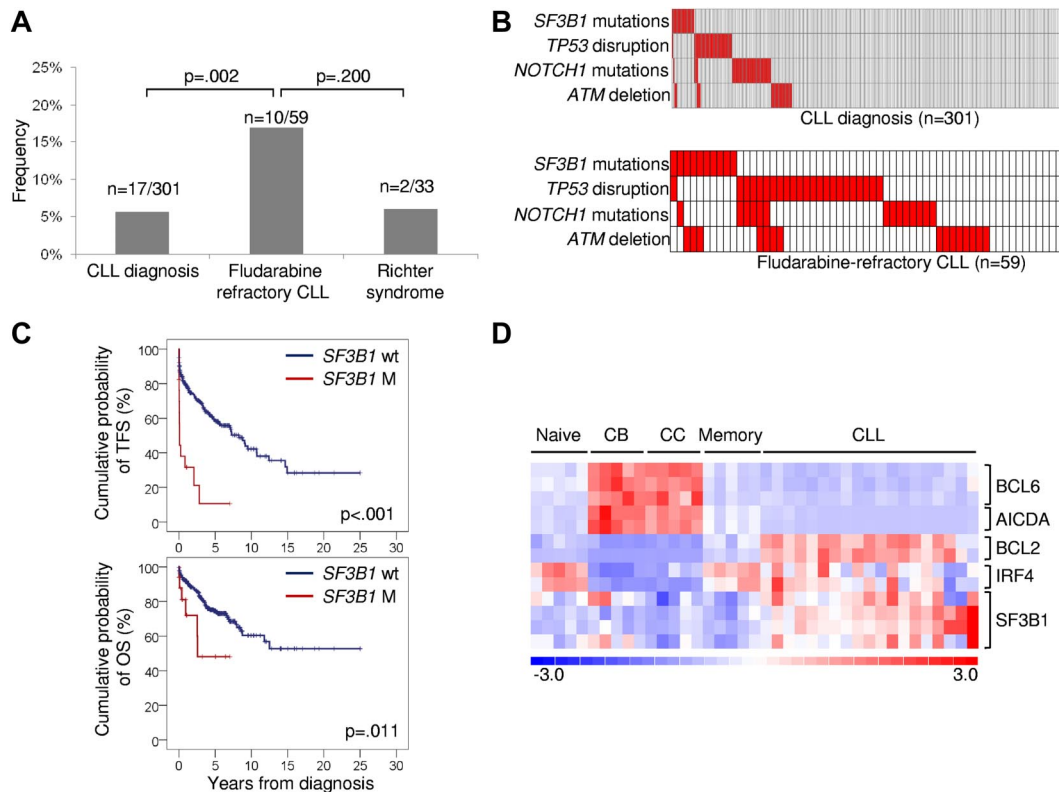


Figure 2. Prevalence, mutual relationship with other genetic lesions, and clinical impact of *SF3B1* mutations in CLL. (A) Prevalence of *SF3B1* mutations in CLL at diagnosis, in fludarabine-refractory CLL, and in RS; numbers on top indicate the actual number of mutated samples over the total number analyzed. (B) Mutual relationship of *SF3B1* mutations with other genetic lesions in CLL at diagnosis and in fludarabine-refractory CLL. In the heat map, rows correspond to identical genes, and columns represent individual patients color-coded based on the gene status (white: wild type; red: mutations of *SF3B1*, mutations of *NOTCH1*, mutations and/or deletion of *TP53*, deletion of *ATM*). (C) Kaplan-Meier estimates of treatment-free survival (TFS) and overall survival (OS) from diagnosis in the consecutive series of newly diagnosed and previously untreated CLL ($n = 301$). *SF3B1* wild-type (*SF3B1* wt) are represented by the blue line. *SF3B1* mutated cases (*SF3B1* M) are represented by the red line. (D) Gene expression levels of *BCL6*, *AICDA*, *BCL2*, *IRF4*, and *SF3B1* in normal B-cell subpopulations (Naive; centroblasts, CB; centrocytes, CC; memory) and CLL samples. Relative levels of gene expression are depicted with a color scale: red represents the greatest level of expression and blue represents the lowest level.

Center, Columbia University, St Nicholas Ave, Rm 508, New York, NY 10032; e-mail: rd10@columbia.edu; and Gianluca Gaidano, MD, PhD, Division of Hematology, Department of Clinical and

Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy; e-mail: gaidano@med.unipmn.it.

References

- Müller-Hermelink HK, Montserrat E, Catovsky D, Campo E, Harris NL, Stein H. Chronic lymphocytic leukemia/small lymphocytic lymphoma. In: Swerdlow SH, Campo E, Harris NL, et al, eds. *World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC; 2008:180-182.
- Hallek M, Cheson BD, Catovsky D, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute—Working Group 1996 guidelines. *Blood*. 2008; 111(12):5446-5456.
- Rossi D, Spina V, Deambrogi C, et al. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood*. 2011;117(12):3391-3401.
- Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med*. 2005;352(8):804-815.
- Kay NE, O'Brien SM, Pettitt AR, Stilgenbauer S. The role of prognostic factors in assessing 'high-risk' subgroups of patients with chronic lymphocytic leukemia. *Leukemia*. 2007;21(9):1885-1891.
- Stilgenbauer S, Zenz T. Understanding and managing ultra high-risk chronic lymphocytic leukemia. *Hematology Am Soc Hematol Educ Program*. 2010;2010:481-488.
- Fabbri G, Rasi S, Rossi D, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of *NOTCH1* mutational activation. *J Exp Med*. 2011;208(7):1389-1401.
- Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475(7354):101-105.
- Stein H, Warnke RA, Chan WC, Jaffe ES. Diffuse large B-cell lymphoma. In: Swerdlow SH, Campo E, Harris NL, et al, eds. *World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC; 2008:233-237.
- Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.
- Döhner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000;343(26):1910-1916.
- Rossi D, Cerri M, Deambrogi C, et al. The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. *Clin Cancer Res*. 2009;15(3):995-1004.
- Zenz T, Häbe S, Denzel T, et al. Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. *Blood*. 2009;114(13):2589-2597.
- Visconte V, Makishima H, Jankowska A, et al. *SF3B1*, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts [published online ahead of print September 2, 2011]. *Leukemia*. doi:10.1038/leu.2011.232.
- Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*. 2011;365(15):1384-1395.
- Luke MM, Della Seta F, Di Como CJ, Sugimoto H, Kobayashi R, Arndt KT. The SAP, a new family of proteins, associate and function positively with

- the SIT4 phosphatase. *Mol Cell Biol.* 1996;16(6):2744-2755.
17. Wang C, Chua K, Seghezzi W, Lees E, Gozani O, Reed R. Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Genes Dev.* 1998;12(10):1409-1414.
18. Das BK, Xia L, Palandjian L, Gozani O, Chyung Y, Reed R. Characterization of a protein complex containing spliceosomal proteins SAPs 49, 130, 145, and 155. *Mol Cell Biol.* 1999;19(10):6796-6802.
19. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell.* 2009;136(4):701-718.
20. David CJ, Manley JL. Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.* 2010;24(21):2343-2364.
21. Ward AJ, Cooper TA. The pathobiology of splicing. *J Pathol.* 2010;220(2):152-163.
22. Andrade MA, Bork P. HEAT repeats in the Huntington's disease protein. *Nat Genet.* 1995;11(2):115-116.
23. Andrade MA, Perez-Iratxeta C, Ponting CP. Protein repeats: structures, functions, and evolution. *J Struct Biol.* 2001;134(2-3):117-131.
24. Kaida D, Motoyoshi H, Tashiro E, et al. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol.* 2007;3(9):576-583.
25. Corrionero A, Miñana B, Valcárcel J. Reduced fidelity of branch point recognition and alternative splicing induced by the antitumor drug spliceostatin A. *Genes Dev.* 2011;25(5):445-459.

NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL

Ilaria Del Giudice,^{1*} Davide Rossi,^{2*} Sabina Chiaretti,^{1*} Marilisa Marinelli,¹ Simona Tavolaro,¹ Sara Gabrielli,¹ Luca Laurenti,³ Roberto Marasca,⁴ Silvia Rasi,² Marco Fangazio,² Anna Guarini,¹ Gianluca Gaidano,^{2§} and Robin Foà^{1§}

¹Division of Hematology, Department of Cellular Biotechnologies and Haematology, University "Sapienza", Rome; ²Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara; ³Institute of Hematology, Catholic University of the Sacred Heart, Rome; and ⁴Division of Hematology, Department of Oncology and Hematology, University of Modena and Reggio Emilia, Modena, Italy

ABSTRACT

Trisomy 12, the third most frequent chromosomal aberration in chronic lymphocytic leukemia (CLL), confers an intermediate prognosis. In our cohort of 104 untreated patients carrying +12, *NOTCH1* mutations occurred in 24% of cases and were associated to unmutated *IGHV* genes ($P=0.003$) and +12 as a sole cytogenetic abnormality ($P=0.008$). *NOTCH1* mutations in +12 CLL associated with an approximately 2.4 fold increase in the risk of death, a significant shortening of survival ($P<0.01$) and proved to be an independent predictor of survival in multivariate analysis. Analogous to +12 CLL with *TP53* disruption or del(11q), *NOTCH1* mutations in +12 CLL conferred a significantly worse survival compared to that of +12 CLL with del(13q) or +12 only. The overrepresentation of cell cycle/proliferation related genes of +12 CLL with *NOTCH1* mutations suggests the biological contribution of *NOTCH1* mutations to determine a poor outcome.

NOTCH1 mutations refine the intermediate prognosis of +12 CLL.

Key words: chronic lymphocytic leukemia, *NOTCH1* mutations, trisomy 12, prognosis, gene expression profile.

Citation: Del Giudice I, Rossi D, Chiaretti S, Marinelli M, Tavolaro S, Gabrielli S, Laurenti L, Marasca R, Rasi S, Fangazio M, Guarini A, Gaidano G, and Foà R. *NOTCH1* mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. *Haematologica* 2012;97(3):437-441. doi:10.3324/haematol.2011.060129

©2012 Ferrata Storti Foundation. This is an open-access paper.

Introduction

Trisomy 12 represents the third most frequent chromosomal aberration in chronic lymphocytic leukemia (CLL) (15-20% of cases) and often (approx. 60% of cases) occurs as the sole cytogenetic lesion.^{1,2} Within the hierarchical model of genetic subgroups commonly used in clinical practice, +12 as single aberration confers an intermediate prognostic risk, with a median time to progression of 33 months and a median overall survival (OS) of 114 months.¹

The *NOTCH1* gene has been shown to have an essential biological role in hematopoiesis.³ Following the pivotal study that identified *NOTCH1* mutations in CLL and provided initial evidence on the unfavorable clinical outcome associated with *NOTCH1* alterations,⁴ two independent studies of the CLL coding genome have recently identified activating muta-

tions of the *NOTCH1* gene in approximately 10% of CLL at diagnosis.^{5,6} The prevalence of *NOTCH1* mutations increases with disease aggressiveness.⁵ At diagnosis, *NOTCH1* mutations show an adverse impact on outcome, confirmed in at least four series,^{4,7} and act independently of other clinico-biological features, including *TP53* disruption.⁷ Among CLL cytogenetic subgroups, *NOTCH1* mutations are distributed in a mutually exclusive fashion with *TP53* disruption and are enriched in CLL carrying +12, where they recur in approximately 25% of patients.⁷

Based on the emerging association between *NOTCH1* alterations and +12, we investigated *NOTCH1* mutations in a series of untreated +12 CLL. We observed that in these patients, *NOTCH1* mutations: i) cluster within cases with no additional cytogenetic abnormalities; ii) induce a particular transcriptional profile; and iii) refine outcome prediction.

The online version of this article has a Supplementary Appendix.

*IDG, DR and SC contributed equally to this study. §GG and RF contributed equally to this study.

Acknowledgments: this work was supported by AIRC (to RF) and AIRC, Special Program Molecular Clinical Oncology, 5 x 1000, n. 10007, Milan, Italy (to GG and to RF), Compagnia di San Paolo, Torino, Italy, Fondazione per le Biotechnologie, Torino, Italy, Fondazione Buzzati-Traverso, Pavia, Italy, Progetto FIRB-Programa "Futuro in Ricerca" 2008 (to DR), PRIN 2008 (to GG) and 2009 (to DR), MIUR, Rome, Italy, Progetto Giovani Ricercatori 2008 (to DR), Ministero della Salute, Rome, Italy, and Novara-AIL Onlus (to GG), Novara, Italy.

Manuscript received on December 12, 2011. Manuscript accepted on December 28, 2011.

Correspondence: Robin Foà, MD, Professor of Hematology, Division of Hematology, Department of Cellular Biotechnologies and Haematology University, "Sapienza" Via Benevento 6, 00161 Rome, Italy. Phone: international +39.06.85795753. Fax: +39.06.85795792.

E-mail: rfoa@bce.uniroma1.it/delgiudice@bce.uniroma1.it

Design and Methods

Patients

This multicenter study evaluated 104 patients carrying +12: 54 were males and 50 females, with a median age of 65 years (interquartile range 56-72). All cases satisfied the IWCLL diagnostic criteria for CLL⁸ and were selected on the basis of: i) untreated disease; ii) availability of biological material; and iii) presence of +12, independent of additional chromosomal abnormalities.

Patients gave their informed consent to blood collection and biological analyses, in agreement with the Declaration of Helsinki. The study was approved by the Ethical Committee of Policlinico Umberto I, "La Sapienza" University of Rome (n. 2182/16.06.2011) and of the Ospedale Maggiore della Carità, Novara, northern Italy, associated with the Amedeo Avogadro University of Eastern Piedmont (protocol code 59/CE; study n. CE 8/11).

Lymphocyte morphology, immunophenotype, FISH analysis, *IGHV* and *TP53* sequencing were performed as previously described.⁹

Mutation analysis of NOTCH1

The *NOTCH1* (exon 34; RefSeq NM_017617.2) mutation hotspot previously identified in CLL⁷ was analyzed by direct sequencing of genomic DNA extracted from blood mononuclear cells. Purified amplicons were subjected to conventional DNA Sanger sequencing using the ABI PRISM 3100 Genetic Analyzer

(Applied Biosystems, CA, USA). The presence of *NOTCH1* c.7544_7545delCT alleles was also investigated by ARMS PCR. Further details are reported in the *Online Supplementary Design and Methods*.

Statistical analysis

Overall survival (OS) was measured from the date of initial presentation to the date of death (event) or last follow up (censoring). Survival analysis was performed by the Kaplan-Meier method. Further details are reported in the *Online Supplementary Design and Methods*.

Gene expression profile analysis

For oligonucleotide array analysis, the HGU133 Plus 2.0 gene chips (Affymetrix, Santa Clara, CA, USA) were used. Sample preparation and microarray processing were performed as previously described.¹⁰ Further details are reported in the *Online Supplementary Design and Methods*.

Results and Discussion

Frequency and distribution of NOTCH1 mutations in +12 CLL

NOTCH1 mutations occurred in 25 of the 104 untreated

Table 1A. Characteristics of CLL patients harboring +12 according to the *NOTCH1* mutation status.

Characteristics ^a	All (n=104)		<i>NOTCH1</i> wild type (n=79)		<i>NOTCH1</i> mutated (n=25)		P
	n ^b	%	n ^b	%	n	%	
Age >65 years	52	50	40	51	12	48	0.819
Male	54	52	43	54	11	44	0.363
Binet B-C	37	36	26	33	11	44	0.313
<i>IGHV</i> identity ≥98%	60	58	39	50	21	84	0.003
13q14 deletion	19	18	19	24	0	0	0.006
Trisomy 12 only	73	70	51	65	23	92	0.008
11q22-q23 deletion	8	8	6	8	2	8	1.000
17p13 deletion	5	5	4	5	1	4	1.000
<i>TP53</i> mutations	4	4	2	3	2	8	0.243
<i>TP53</i> disruption	8	8	5	6	3	12	0.395

^a*IGHV*: immunoglobulin heavy variable gene; ^b*IGHV*: rearrangement was not available for one *NOTCH1* wild-type patient.

Table 1B. Univariate and multivariate analysis for overall survival in CLL patients harboring +12^a.

Characteristics	Events	Total	OS (years)			Univariate analysis				P	Multivariate analysis			
			Median	LCI	UCI	HR	LCI	UCI	P		HR	LCI	UCI	P
Age ≤65 years	15	51	15.6	10.4	20.7	-	-	-	-	-	-	-	-	-
Age >65 years	21	52	6.7	3.6	9.9	2.92	1.44	5.92	0.003	-	6.03	2.60	13.95	<0.001
Binet A	19	69	15.6	9.7	21.4	-	-	-	-	-	-	-	-	-
Binet B-C	17	36	7.8	5.5	10.0	2.47	1.26	4.82	0.008	-	2.40	1.17	4.89	0.016
<i>IGHV</i> identity <98%	12	43	15.6	9.2	21.9	-	-	-	-	-	-	-	-	-
<i>IGHV</i> identity ≥98%	23	59	11.1	7.6	14.6	2.35	1.07	5.12	0.032	-	2.14	0.90	5.11	0.085
<i>NOTCH1</i> wild-type	21	78	13.0	8.9	17.1	-	-	-	-	-	-	-	-	-
<i>NOTCH1</i> mutations	15	25	7.8	1.2	14.3	2.46	1.26	4.81	0.008	-	3.04	1.39	6.64	0.005
<i>TP53</i> wild-type	30	95	13.0	10.5	15.5	-	-	-	-	-	-	-	-	-
<i>TP53</i> disruption	6	8	5.6	3.0	8.3	3.18	1.29	7.80	0.011	-	3.46	1.38	8.70	0.008

^aOS: overall survival; HR: hazard ratio; LCI: 95% lower confidence interval; UCI: 95% upper confidence interval; *IGHV*: immunoglobulin heavy variable gene.

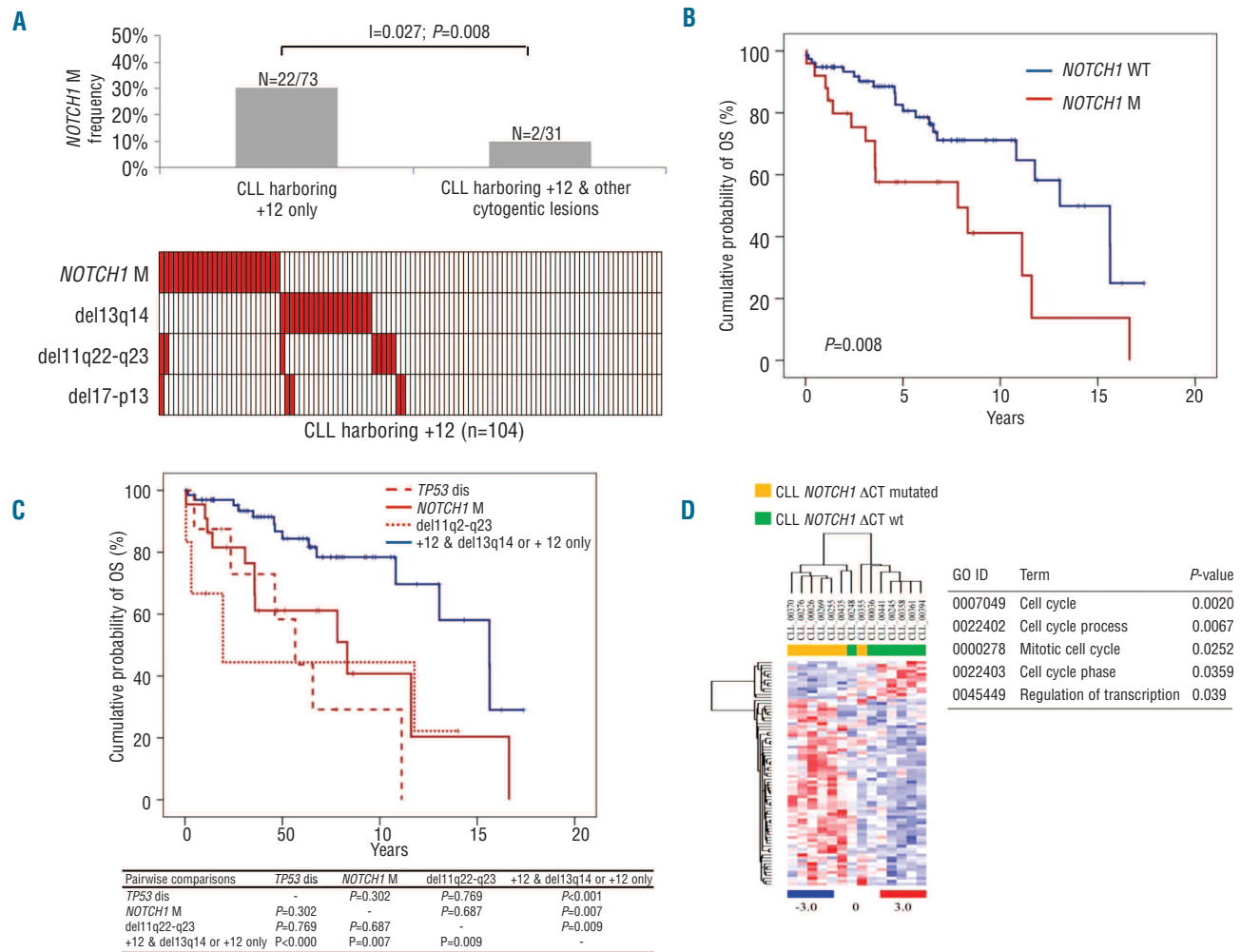


Figure 1. (A) Distribution of *NOTCH1* mutations among genetic subgroups of +12 CLL. (B) Overall survival according to *NOTCH1* mutation status in +12 CLL. (C) Hierarchical stratification of overall survival according to genetic lesions in +12 CLL. (D) Comparison between *NOTCH1* mutated and *NOTCH1* wt CLL samples. Relative levels of gene expression are depicted with a color scale: red represents the highest level of expression and blue represents the lowest level. The table reports the functional annotation analysis, performed using the DAVID database, of differentially expressed genes between *NOTCH1* mutated and *NOTCH1* wt CLL. The biological processes reported are ordered according to their *P* value.

CLL with +12 investigated (24%) (Table 1A), were represented in all cases by frameshift deletions, including the c.7544_7545delCT in 22 of 25 (88%) cases, and preferentially associated with use of unmutated *IGHV* genes (84%, $P=0.003$). *NOTCH1* mutations occurred independent of gender, thus suggesting that *NOTCH1* mutations might be an important marker of unfavorable prognosis in both male and female CLL patients.

Trisomy 12 occurred as an isolated chromosomal abnormality in 73 of 104 (70%) cases, while it was associated to other cytogenetic abnormalities in 31 of 104 (30%) cases (Figure 1A). Mutual information analysis revealed a clustering of *NOTCH1* mutations among CLL harboring +12 as a sole abnormality (22 of 73, 30%) compared to patients harboring +12 in addition to other cytogenetic lesions (2 of 31, 6%) ($I=0.027$; $P=0.008$) (Figure 1A). Consistently, +12 CLL harboring *NOTCH1* mutations carried deletion 13q14 only exceptionally (0 of 19). Consistent with pivotal observations,^{5,7} also in +12 CLL, *NOTCH1* mutations distributed in

a mutually exclusive fashion with deletions of 17p13 and 11q22-q23 (Figure 1A).

This extended cohort corroborates the high prevalence of *NOTCH1* mutations in +12 CLL, where the overall frequency of *NOTCH1* mutations in +12 patients consistently ranges from 24.5 to 28.6%.^{7,11} Furthermore, +12 patients harboring *NOTCH1* mutations prevalently belong to aggressive cases, i.e. cases with an unmutated *IGHV* gene status, in line with recent findings,^{7,11} and expression of CD38 (*NOTCH1* mutated/CD38 positive, $n=20$ of 25 (80.0%) vs. *NOTCH1* wild-type/CD38 positive, $n=39$ of 77 (50.6%) ($P=0.010$)).

At variance, no difference emerged in the distribution of ZAP-70 positivity (*NOTCH1* mutated/ZAP-70 positive, $n=15$ of 24 (62.5%) vs. *NOTCH1* wild-type/ZAP-70 positive, $n=34$ of 77 (44.2%) ($P=0.116$)).

Finally, the analysis of this specific cohort of patients showed an enrichment of *NOTCH1* mutations in CLL harboring +12 as sole cytogenetic abnormality.

NOTCH1 mutations and overall survival in +12 CLL

After a median follow up of seven years, 36 of 103 evaluable patients had died, for a median OS of 11.6 years (95% CI: 10.4-12.7). By univariate analysis, the crude impact of *NOTCH1* mutations on survival in +12 CLL was an approximately 2.4 fold increase in the hazard of death (HR: 2.46; 95% CI: 1.26-4.81) and a significant shortening of OS ($P < 0.01$) (Table 1B; Figure 1B). Other variables associated with shorter OS were age, Binet stage, *IGHV* mutation status and *TP53* disruption.

Multivariate analysis selected *NOTCH1* mutations as an independent risk factor of OS (HR: 3.04; 95% CI: 1.39-6.64; $P < 0.01$), after adjusting for age ($>$ vs. ≤ 65 years), Binet stage (B-C vs. A), *IGHV* mutation status and *TP53* disruption by mutation and/or deletion. *NOTCH1* mutations, 13q14 deletion, 11q22-q23 deletion and *TP53* status were used to build a hierarchical model of genetic subgroups to predict OS in CLL with +12. The outcome of +12 CLL with *NOTCH1* mutations was poor, similar to cases with +12 and *TP53* disruption or 11q22-q23 deletion and significantly worse than patients with +12 as an isolated abnormality or plus 13q14 deletion (Figure 1C).

NOTCH1 mutations represent, therefore, an independent adverse prognostic factor of OS among +12 CLL, providing a new genetic prognostic stratification of patients with this intermediate risk marker.

Gene expression profiling of +12 CLL with NOTCH1 mutations

To understand whether *NOTCH1* mutations induced a distinctive transcriptional profile in +12 CLL, we compared 7 *NOTCH1* mutated vs. 7 *NOTCH1* wild-type cases in a cohort of patients carrying +12 (Online Supplementary Table S1).

This analysis showed that *NOTCH1* mutated cases formed a tight clustering (Figure 1D), with 2 patients incorrectly placed. Of these, one later developed a *TP53* mutation and a myelodysplastic syndrome (CLL_00248, Online Supplementary Table S1).

Sixty-five differentially expressed genes (Online Supplementary Table S2) were selected, the majority being upmodulated in *NOTCH1* mutated samples. DAVID functional annotation analysis highlighted an overrepresentation of cell cycle related genes, indicating that *NOTCH1*

mutations induce a proliferative advantage that might explain the clinically aggressive behavior (Figure 1D). We also observed significantly higher levels of IgM expression in *NOTCH1* mutated cases. It is known that IgM expression is higher in cells with increased ability to respond to external stimuli,^{12,13} indicating that the *NOTCH1* mutated clone might survive and expand also thanks to these interactions. Intriguingly, approximately 30% of the upregulated transcripts were located on chromosome 12. This might be due to the fact that +12 was present as a single alteration in all *NOTCH1* mutated cases, whereas in the *NOTCH1* wild-type subgroup the scenario was more complex, with 3 cases displaying only +12 and 4 cases one or more additional chromosomal aberrations.

Among the transcripts previously reported to be associated with a +12 signature, we confirmed the upregulation of *ANAPC5*, *GLIPR1*, *TIMELESS* and *SLC2A6*.^{14,15}

In summary, this study highlights that *NOTCH1* mutations in +12 CLL: i) preferentially cluster with cases harboring +12 as sole genetic abnormality; ii) account for 24% of cases and are more frequently detected in cases with unfavorable biological markers; iii) associate with a particular gene expression profile; iv) predict a poor outcome, stratifying +12 CLL in two distinct subgroups; and v) are not associated with *TP53* disruption or 11q deletion, thus making them even more useful in a genetic hierarchical prognostic model, given their mutual exclusivity. This observation sheds light on the heterogeneous clinical course of +12 CLL patients and allows us to refine the intermediate prognostic risk of this chromosomal lesion. The transcriptional profile characterized by an overrepresentation of genes involved in cell cycle and proliferation suggests the potential biological contribution of *NOTCH1* mutations in determining the aggressive behavior of the disease.

Authorship and Disclosures

The information provided by the authors about contributions from persons listed as authors and in acknowledgments is available with the full text of this paper at www.haematologica.org.

Financial and other disclosures provided by the authors using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are also available at www.haematologica.org.

References

1. Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000;343(26):1910-6.
2. Matutes E, Oscier D, Garcia-Marco J, Ellis J, Copplestone A, Gillingham R, et al. Trisomy 12 defines a group of CLL with atypical morphology: correlation between cytogenetic, clinical and laboratory features in 544 patients. *Br J Haematol*. 1996;92(2):382-8.
3. Milner LA, Bigas A. Notch as a mediator of cell fate determination in hematopoiesis: evidence and speculation. *Blood*. 1999;93(8):2431-48.
4. Sportoletti P, Baldoni S, Cavalli L, Del Papa B, Bonifacio E, Ciurnelli R, et al. *NOTCH1* PEST domain mutation is an adverse prognostic factor in B-CLL. *Br J Haematol*. 2010;151(4):404-6.
5. Fabbri G, Rasi S, Rossi D, Trifonov V, Khiabanian H, Ma J, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of *NOTCH1* mutational activation. *J Exp Med*. 2011;208(7):1389-401.
6. Puente XS, Pinyol M, Quesada V, Conde L, Ordóñez GR, Villamor N, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475(7354):101-5.
7. Rossi D, Rasi S, Fabbri G, Spina V, Fangazio M, Forconi F, Marasca R, et al. Mutations of *NOTCH1* are an independent predictor of survival in chronic lymphocytic leukemia. *Blood*. 2012;119(2):521-9.
8. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner H, et al. International Workshop on Chronic Lymphocytic Leukemia. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*. 2008;112(13):5259-69.
9. Del Giudice I, Mauro FR, De Propriis MS, Santangelo S, Marinelli M, Peragine N, et al. White blood cell count at diagnosis and immunoglobulin variable region gene mutations are independent predictors of treatment-free survival in young patients with stage A chronic lymphocytic

- leukemia. *Haematologica*. 2011;96(4):626-30.
10. Chiaretti S, Tavoraro S, Marinelli M, Messina M, Del Giudice I, Mauro FR, et al. Evaluation of TP53 mutations with the AmpliChip p53 research test in chronic lymphocytic leukemia: correlation with clinical outcome and gene expression profiling. *Genes Chromosomes Cancer*. 2011;50(4):263-74.
 11. Balatti V, Bottoni A, Palamarchuk A, Alder H, Rassenti LZ, Kipps TJ, et al. NOTCH1 mutations in CLL associated with trisomy 12. *Blood*. 2012;119(2):329-31.
 12. Mockridge CI, Potter KN, Wheatley I, Neville LA, Packham G, Stevenson FK. Reversible anergy of sIgM-mediated signaling in the two subsets of CLL defined by VH-gene mutational status. *Blood*. 2007;109(10):4424-31.
 13. Guarini A, Chiaretti S, Tavoraro S, Maggio R, Peragine N, Citarella F, et al. BCR ligation induced by IgM stimulation results in gene expression and functional changes only in IgV H unmutated chronic lymphocytic leukemia (CLL) cells. *Blood*. 2008;112(3):782-92.
 14. Haslinger C, Schweifer N, Stilgenbauer S, Döhner H, Lichter P, Kraut N, et al. Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. *J Clin Oncol*. 2004;22(19):3937-49.
 15. Porpaczy E, Bilban M, Heinze G, Gruber M, Vanura K, Schwarzingner I, et al. Gene expression signature of chronic lymphocytic leukaemia with Trisomy 12. *Eur J Clin Invest*. 2009;39(7):568-75.

blood

2012 119: 521-529
Prepublished online November 10, 2011;
doi:10.1182/blood-2011-09-379966

Mutations of *NOTCH1* are an independent predictor of survival in chronic lymphocytic leukemia

Davide Rossi, Silvia Rasi, Giulia Fabbri, Valeria Spina, Marco Fangazio, Francesco Forconi, Roberto Marasca, Luca Laurenti, Alessio Bruscaggin, Michaela Cerri, Sara Monti, Stefania Cresta, Rosella Famà, Lorenzo De Paoli, Pietro Bulian, Valter Gattei, Anna Guarini, Silvia Deaglio, Daniela Capello, Raul Rabadan, Laura Pasqualucci, Riccardo Dalla-Favera, Robin Foà and Gianluca Gaidano

Updated information and services can be found at:

<http://bloodjournal.hematologylibrary.org/content/119/2/521.full.html>

Articles on similar topics can be found in the following Blood collections

[Clinical Trials and Observations](#) (3800 articles)

[Lymphoid Neoplasia](#) (1598 articles)

Information about reproducing this article in parts or in its entirety may be found online at:

http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:

<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:

<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Copyright 2011 by The American Society of Hematology; all rights reserved.



Mutations of *NOTCH1* are an independent predictor of survival in chronic lymphocytic leukemia

*Davide Rossi,¹ *Silvia Rasi,¹ Giulia Fabbri,² Valeria Spina,¹ Marco Fangazio,¹ Francesco Forconi,³ Roberto Marasca,⁴ Luca Laurenti,⁵ Alessio Bruscaggini,¹ Michaela Cerri,¹ Sara Monti,¹ Stefania Cresta,¹ Rosella Famà,¹ Lorenzo De Paoli,¹ Pietro Bulian,⁶ Valter Gattei,⁶ Anna Guarini,⁷ Silvia Deaglio,⁸ Daniela Capello,⁹ Raul Rabadan,¹⁰ Laura Pasqualucci,^{2,11} Riccardo Dalla-Favera,^{2,11,12} †Robin Foà,^{7,13} and †Gianluca Gaidano¹

¹Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; ²Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY; ³Division of Hematology, University of Siena, Siena, Italy; ⁴Division of Hematology, Department of Oncology and Hematology, University of Modena and Reggio Emilia, Modena, Italy; ⁵Institute of Hematology, Catholic University of the Sacred Heart, Rome, Italy; ⁶Clinical and Experimental Onco-Hematology, Centro di Riferimento Oncologico, Istituto di Ricovero e Cura a Carattere Scientifico, Aviano, Italy; ⁷Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy; ⁸Department of Genetics, Biology, and Biochemistry and Human Genetics Foundation, University of Turin, Turin, Italy; ⁹Division of Clinical Biochemistry and Clinical Molecular Biology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; ¹⁰Department of Biomedical Informatics and Center for Computational Biology and Bioinformatics, Columbia University, New York, NY; Departments of ¹¹Pathology and Cell Biology and ¹²Genetics and Development, Columbia University, New York, NY; and ¹³Fondazione Eleonora Lorillard Spencer Cenci, Sapienza University, Rome, Italy

Analysis of the chronic lymphocytic leukemia (CLL) coding genome has recently disclosed that the *NOTCH1* proto-oncogene is recurrently mutated at CLL presentation. Here, we assessed the prognostic role of *NOTCH1* mutations in CLL. Two series of newly diagnosed CLL were used as training (n = 309) and validation (n = 230) cohorts. *NOTCH1* mutations occurred in 11.0% and 11.3% CLL of the training and validation series, respectively. In the training series, *NOTCH1* mu-

tations led to a 3.77-fold increase in the hazard of death and to shorter overall survival (OS; *P* < .001). Multivariate analysis selected *NOTCH1* mutations as an independent predictor of OS after controlling for confounding clinical and biologic variables. The independent prognostic value of *NOTCH1* mutations was externally confirmed in the validation series. The poor prognosis conferred by *NOTCH1* mutations was attributable, at least in part, to shorter treatment-free survival

and higher risk of Richter transformation. Although *NOTCH1* mutated patients were devoid of *TP53* disruption in more than 90% cases in both training and validation series, the OS predicted by *NOTCH1* mutations was similar to that of *TP53* mutated/deleted CLL. *NOTCH1* mutations are an independent predictor of CLL OS, tend to be mutually exclusive with *TP53* abnormalities, and identify cases with a dismal prognosis. (*Blood*. 2012;119(2): 521-529)

Introduction

Chronic lymphocytic leukemia (CLL) is the most common leukemia in adults.¹⁻⁴ The clinical course of CLL ranges from very indolent, with a nearly normal life expectancy,⁵⁻⁹ to rapidly progressive leading to death and occasionally undergoing transformation to aggressive lymphoma, known as Richter syndrome (RS).¹⁰⁻¹⁸

At presentation, several clinical and biologic features may help to predict, at least in part, the clinical course of CLL.¹⁹⁻²¹ Of the biologic prognosticators that have been developed, current guidelines for clinical practice recommend screening only for *TP53* disruption by mutation, deletion, or both of the locus, that identifies a fraction of high-risk CLL destined to experience a very short survival.^{2,21-28} High-risk CLL, however, cannot be fully recapitulated by *TP53* disruption, and other lesions of cancer genes may be implicated in this aggressive phenotype.²⁹

Recently, two independent investigations of the CLL coding genome have revealed that activating mutations of the *NOTCH1* proto-oncogene are recurrently associated with CLL.^{30,31} Based on

current knowledge, *NOTCH1* mutations occur in ~10% CLL at diagnosis and their frequency increases in advanced disease phases, as exemplified by the case of RS.^{30,31} The relevance of *NOTCH1* mutations in CLL is reinforced by knowledge of activation of the *NOTCH1* pathway in this leukemia,³² and by the possibility of targeting *NOTCH1* with drugs currently under development in other clinical contexts.³³ Although not designed to fully assess clinical correlates, the pivotal studies that have identified *NOTCH1* mutations in CLL have provided initial evidence suggesting that *NOTCH1* alterations might be associated with an unfavorable clinical outcome.^{30,31,34} However, several aspects of the clinical implications of *NOTCH1* mutations in CLL still remain to be elucidated, including: (1) their distribution among well established CLL genetic subgroups, including those defined by FISH abnormalities and *TP53* status; and (2) their independent prognostic role, given the tight association between *NOTCH1* mutations and unmutated immunoglobulin heavy variable (*IGHV*) genes, one of the most widely accepted prognosticators in CLL.

Submitted September 26, 2011; accepted November 8, 2011. Prepublished online as *Blood* First Edition paper, November 10, 2011; DOI 10.1182/blood-2011-09-379966.

*D.R. and S.R. contributed equally to this study.

†R.F. and G.G. contributed equally to this study.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2012 by The American Society of Hematology

By using a training-validation approach, we hereby report that *NOTCH1* mutations: (1) cluster with CLL harboring trisomy 12, suggesting that aberrant NOTCH signaling plays an important role in this genetic subgroup; (2) tend to be mutually exclusive with *TP53* disruption in the same patient; and (3) are an independent predictor of CLL overall survival (OS) because they identify a subset of high-risk patients with dismal prognosis similar to that associated with *TP53* abnormalities.

Patients and methods

Patients

The study used a training-validation design. The training cohort was a consecutive series of 309 previously untreated CLL who presented for initial evaluation at a single center. The training series was provided with prospectively collected biologic samples drawn at presentation and with a prospectively maintained clinical database updated in May 2010. Median follow-up of alive patients was 6 years. No patient was lost at follow-up. The validation cohort was represented by a retrospective series of 230 previously untreated CLL from 3 institutions participating to the same national CLL network. Inclusion criteria for the validation series were availability of: (1) biologic samples collected at presentation, and (2) clinical follow-up. Median follow-up of alive patients for the validation series was 7 years.

For sample size definition, we assumed a prevalence of *NOTCH1* mutations at presentation of at least 10% and a 5-year OS of 80% for the entire population. Based on these assumptions, the sample size would allow detection of at least 15% and 19% difference in 5-year OS for the training series and the validation series, respectively (power = 80%; $\alpha = .05$).

CLL diagnosis was based on International Working Group on CLL–National Cancer Institute criteria.^{1,2} RS diagnosis was histologically proven and was represented by diffuse large B-cell lymphoma (clonally related to the CLL phase).^{1,35}

The Reporting Recommendations for Tumor Marker Prognostic Studies criteria were followed throughout this study.³⁶ Patients provided informed consent in accordance with local institutional review board requirements and Declaration of Helsinki. The study was approved by the Ethical Committee of the Ospedale Maggiore della Carità di Novara associated with the Amedeo Avogadro University of Eastern Piedmont (protocol code 59/CE; study CE 8/11).

Molecular studies

NOTCH1, *TP53*, and *IGHV* mutations were analyzed by DNA Sanger sequencing.^{17,26,31,37,38} *NOTCH1* c.7544_7545delCT mutation also was investigated by amplification refractory mutation system (ARMS) PCR. Probes (Abbott) used for FISH analysis were LSI13 and LSI13S319, CEP12, LSIp53, and LSIATM.^{18,23} Molecular studies were performed on tumor samples collected from peripheral blood: (1) at CLL presentation for both the training ($n = 309$) and validation ($n = 230$) CLL series; and (2) at the time of first progression requiring treatment for progressive CLL that were treated with fludarabine-based regimens ($n = 113$). Details of molecular methods are in supplemental Methods (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

Flow cytometry

CD38 and ZAP-70 expression was analyzed by flow cytometry on PBMCs collected at presentation. Cut-off points of 30% and 20% were used to define positivity for CD38 and ZAP-70, respectively. Details are reported in supplemental Methods.

Serum β 2-microglobulin quantification

β 2-microglobulin levels were quantified by nephelometry (Siemens Healthcare Diagnostics; reference range, 1.8–2.5 mg/L) on serum samples collected at presentation.

Statistical analysis

OS was measured from date of initial presentation to date of death (event) or last follow-up (censoring). Treatment-free survival (TFS) was measured from date of initial presentation to date of progressive and symptomatic disease requiring treatment according to International Working Group on CLL–National Cancer Institute guidelines (event), death, or last follow-up (censoring).² Time to RS transformation was measured from date of initial presentation to date of the biopsy documenting occurrence of RS transformation (event), death, or last follow-up (censoring).¹⁷ OS from first line treatment was assessed among cases treated with fludarabine-based regimens ($n = 113$) and was measured from date of treatment start to date of death (event) or last follow-up (censoring). Survival was estimated by Kaplan–Meier method.³⁹ The crude association between exposure variables and outcome was estimated by univariate Cox regression analysis.⁴⁰ The independence of *NOTCH1* mutations as a predictor of CLL OS was estimated after controlling for confounding variables by multivariate Cox regression analysis.^{40–42}

Covariates included in the multivariate analysis along with *NOTCH1* mutation status were selected according to the following criteria: (1) wide acceptance as clinical or biologic prognosticators in CLL, (2) availability of the information in both the training and validation series, and (3) limitation of the number of predictors to no more than $\sim 1/10$ uncensored events to avoid overfitting.^{41,42} Based on these criteria, the following variables were included in multivariate analysis: *NOTCH1* mutations (present vs absent), age (continuous variable), sex (male vs female), Rai stage (III–IV vs 0–II), *IGHV* identity $\geq 98\%$ (present vs absent), trisomy 12 (present vs absent), 11q22–q23 deletion (present vs absent), and *TP53* disruption by mutation, deletion, or both (present vs absent). None of the covariates violated the proportional hazard assumption as documented by plotting the smoothed Schoenfeld residuals and by performing a correlation test between time and residuals.^{41–43} The assumption of effect additivity of predictors was not violated, as documented by a global test of additivity including interactions between *NOTCH1* mutations and other covariates.^{41,42} None of the covariates showed collinearity.^{41,42} Age was treated as a continuous variable and did not violate the linearity assumption as assessed by plotting the smoothed martingale residuals.^{41,42,44}

The prediction accuracy of the multivariate model was verified by assessing model discrimination and calibration (see supplemental Methods for details).^{41,42,45} The stability and predictive performance of *NOTCH1* mutations as an independent predictor of CLL OS was validated both internally in the training series and externally in an independent validation series. Internal validation was performed using a bootstrapping resampling procedure (see supplemental Methods for details).^{41,42,46} The more general validity of *NOTCH1* mutations as an independent predictor of CLL OS was tested using an external validation approach. In this step, Cox regression was applied to an independent validation cohort that included *NOTCH1* mutations and the confounding variables also tested in the training series.

Recursive-partitioning analysis for censored survival data was performed to hierarchically classify CLL patients into risk categories based on *NOTCH1* and *TP53* status.⁴⁷ Categorical variables were compared by χ^2 test and Fisher exact test when appropriate. Continuous variables were compared by Mann–Whitney test. All statistical tests were 2-sided. Statistical significance was defined as P value $< .05$. The analysis was performed with the Statistical Package for the Social Sciences Version 18.0 software (SPSS) and with R statistical package 2.13.0 (<http://www.r-project.org>).

Results

Frequency and distribution of *NOTCH1* mutations in the training series

The training series ($n = 309$) was representative of the main clinical and biologic characteristics of CLL (Table 1). *NOTCH1* mutations ($n = 34$, all heterozygous) occurred in 34/309 (11.0%) patients, being mostly represented (26/34, 76.5%) by a recurrent

Table 1. Characteristics of the CLL training series according to NOTCH1 mutation status

Characteristic	All (n = 309)		NOTCH1 wild type (n = 275)		NOTCH1 mutated (n = 34)		P
	n*	%	n*	%	n*	%	
Age, y (range)	69 (60-76)		69 (60-76)		69 (60-76)		.615
Male	170	55.0	152	55.3	18	52.9	.797
Rai stage III-IV	34	11.0	25	9.1	9	26.5	.006
IGHV identity $\geq$ 98%	103	33.3	77	28.0	26	76.5	< .001
13q14 deletion	160	51.8	151	54.9	9	26.5	.002
Normal FISH	92	29.8	81	29.5	11	32.4	.727
Trisomy 12	61	19.7	46	16.7	15	44.1	< .001
11q22-q23 deletion	25	8.1	24	8.7	1	2.9	.334
17p13 deletion	25	8.1	24	8.7	1	2.9	.334
TP53 mutations	23	7.4	21	7.6	2	5.9	1.000
TP53 disruption	33	10.7	30	10.9	3	8.8	1.000

*Median and 25th-75th percentiles are reported for continuous variables.

2-bp frameshift deletion (c.7544_7545delCT). The remaining NOTCH1 mutations (8/34, 23.5%) were frameshift deletions other than c.7544_7545delCT (n = 7) and frameshift insertions (n = 1; supplemental Table 1). All mutations were predicted to disrupt the NOTCH1 PEST domain.

The clinical and biologic features of NOTCH1 mutated CLL are summarized in Table 1. CLL with NOTCH1 mutations preferentially carried unmutated IGHV genes (76.5%; $P < .001$). Other characteristics at presentation associated with NOTCH1 mutations were advanced Rai stage and trisomy 12 (Table 1). Consistent with the mutually exclusive distribution of trisomy 12 and 13q14 deletion,²³ NOTCH1 mutated CLL were less frequently deleted on 13q14 (Table 1).

NOTCH1 mutations are an independent prognosticator of OS in the training series

After a median follow-up of 6 years, 135/309 patients from the training series had received treatment, 19/309 had developed RS and 78/309 had died, accounting for a median TFS of 7.1 year (95% confidence interval [CI], 4.5-9.8), a 5-year risk of RS of 7.9%

(95% CI, 4.4-11.4), and a median OS of 13.0 years (95% CI, 10.2-15.9).

By univariate analysis, the crude impact of NOTCH1 mutations on survival was an ~ 3.8 -fold increase in the hazard ratio (HR, 3.77; 95% CI, 2.14-6.66) and a significant OS shortening ($P < .001$; Table 2; Figure 1A) that occurred irrespective of the NOTCH1 mutation type (c.7544_7545delCT, $P < .001$; other mutations, $P = .009$; supplemental Figure 1). Other variables associated with shorter OS were age, Rai stage, IGHV mutation status, trisomy 12, 11q22-q23 deletion, and TP53 disruption (Table 2; supplemental Figure 2).

The adjusted impact of NOTCH1 mutations on OS was estimated after controlling for confounding variables by multivariate Cox regression analysis. Along with NOTCH1 mutations, other variables included in the analysis were: (1) those known a priori to be widely accepted clinical (age and Rai stage)³⁻¹⁴ and genetic (IGHV mutation status, 11q22-q23 deletion, and TP53 disruption)^{20-28,37,38} risk factors affecting CLL OS; (2) trisomy 12, given its double association with NOTCH1 mutations and OS in this series (Tables 1-2); and (3) sex.

Table 2. Univariate and multivariate analysis for overall survival in the CLL training series

Characteristics	Event	Total	OS, y			Univariate analysis				Multivariate analysis*††				Internal bootstrapping validation			
			Median	LCI	UCI	HR	LCI	UCI	P	HR	LCI	UCI	P	HR	LCI	UCI	Bootstrap selection, %
Age§						1.06	1.03	1.08	< .001	1.07	1.04	1.10	< .001	1.07	1.05	1.10	100
Female	29	139	13.9	10.8	17.0												
Male	49	170	12.2	8.0	16.5	1.53	.096	2.43	.069	1.96	1.20	3.20	.007	2.17	1.28	3.68	91.1
NOTCH1 germline	62	275	13.9	10.5	17.3												
NOTCH1 mutations	16	34	3.5	0	7.3	3.77	2.14	6.66	< .001	3.99	2.05	7.76	< .001	4.55	2.23	9.31	98.0
Rai stage 0-II	55	275	15.6	13.1	18.7												
Rai stage III-IV	23	34	6.0	2.5	9.6	3.99	2.44	6.50	< .001	2.33	1.25	3.99	.007	2.57	1.38	4.82	90.0
IGHV identity < 98%	42	206	NR														
IGHV identity $\geq$ 98%	36	103	11.7	6.3	17.2	2.23	1.42	3.52	.001	1.44	0.83	2.50	.191	1.53	0.86	2.71	50.4
No trisomy 12	52	248	NR														
Trisomy 12	26	61	10.8	5.5	16.0	1.93	1.21	3.01	.006	1.50	0.91	2.64	.108	1.62	0.96	2.74	55.4
No 11q22-q23 deletion	65	284	15.6	13.3	17.9												
11q22-q23 deletion	13	25	6.8	0	14.1	2.25	1.23	4.11	.008	1.57	0.76	3.27	.220	1.77	0.81	3.90	47.4
TP53 germline	61	276	13.9	10.5	17.3												
TP53 disruption	17	33	4.6	1.8	7.3	3.74	2.16	6.49	< .001	3.27	1.80	5.95	< .001	3.38	1.78	6.43	96.5

LCI indicates 95% lower confidence interval; UCI, 95% upper confidence interval; and NR, not reached.

*Shrinkage coefficient, 0.91.

†Discrimination: c-index of the original model, 0.792; bias-corrected c-index, 0.015; and optimism, 0.777.

‡Calibration: calibration slope of the original model, 1.000; bias-corrected calibration slope, 0.910; and optimism, 0.090.

§Age was treated as a continuous variable.

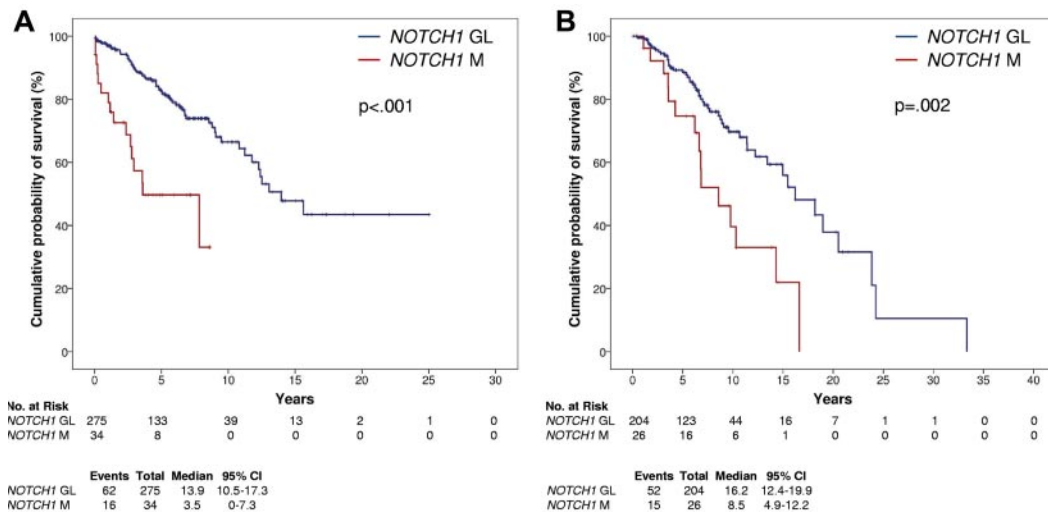


Figure 1. Kaplan-Meier estimates of OS according to *NOTCH1* mutation status. Overall survival according to *NOTCH1* mutation status in the CLL training series (n = 309; A) and in the CLL validation series (n = 230; B). *NOTCH1* germ line cases (*NOTCH1* GL) are represented by the blue line. *NOTCH1* mutated cases (*NOTCH1* M) are represented by the red line.

Multivariate analysis selected *NOTCH1* mutations as an independent risk factor of OS (HR, 4.22; 95% CI, 2.15-8.28; $P < .001$; Table 2). The inclusion of *NOTCH1* mutations in addition to age, sex, Rai stage, *IGHV* mutation status, trisomy 12, 11q22-q23 deletion, and *TP53* disruption significantly improved the fit (-2LL of the model without *NOTCH1* mutations, 692 vs -2LL of the model with *NOTCH1* mutations, 677; likelihood ratio statistics, 14.5; $P < .001$) and the predictive accuracy (c-index of the model without *NOTCH1* mutations, 0.768 vs c-index of the model with *NOTCH1* mutations, 0.792; $P < .001$) of the model. The *IGHV* mutation status that was significant in the model without *NOTCH1* mutations (supplemental Table 2) was no longer retained as an independent prognosticator of OS after inclusion of *NOTCH1* mutations. By bivariate analysis, the *IGHV* mutation status maintained its prognostic relevance in CLL devoid of *NOTCH1* mutations (supplemental Figure 3).

The stability and predictive performance of *NOTCH1* mutations as an independent prognostic factor of CLL OS was internally validated in the training series using a bootstrapping resampling procedure. The first step of the internal validation showed that *NOTCH1* mutations were selected at high frequency (> 98.3%) as an independent prognosticator of CLL OS in each of the 1000 bootstrap samples that were generated (Table 2). This step validated *NOTCH1* mutations as one of the most important variables affecting OS in the training series. The second step of the internal validation demonstrated that the hazard ratios produced from the original series were very close to those produced from the 1000 bootstrap samples (Table 2).

An exploratory analysis applied only to the training series demonstrated that *NOTCH1* mutations maintained their independent prognostic role also after adjusting for β_2 -microglobulin levels, ZAP-70 expression, and CD38 expression (supplemental Table 3).

***NOTCH1* mutations predict an increased risk of CLL progression, RS transformation, and short survival after treatment**

At CLL diagnosis, *NOTCH1* mutations identified CLL patients with rapidly progressive disease and patients at risk of RS development. In fact, patients from the training series carrying

NOTCH1 mutations displayed a shorter time to progression requiring treatment compared with patients without *NOTCH1* mutations ($P < .001$; Figure 2A). In addition, *NOTCH1* mutated patients from the training series displayed a higher cumulative probability of RS compared with patients without *NOTCH1* mutations ($P = .026$; Figure 2B), that occurred irrespective of the *NOTCH1* mutation type.

In CLL patients treated with fludarabine-based regimens at first progression requiring treatment, *NOTCH1* mutations occurred in 23/113 (20.4%) cases (supplemental Table 4) and associated with an OS from treatment similar to that marked by *TP53* disruption ($P = .600$) and significantly shorter compared with that of patients lacking both *NOTCH1* and *TP53* lesions ($P = .041$; Figure 3).

***NOTCH1* mutations are an independent prognosticator of OS in the validation series**

The prognostic value of *NOTCH1* mutations as a risk factor was externally validated in an independent CLL series (n = 230; supplemental Table 5). *NOTCH1* mutations (n = 26, all heterozygous) occurred in 26/230 (11.3%) patients and affected in all cases the PEST domain, with a mutational spectrum similar to that of the training series (c.7544_7545delCT, 21/26 [80.7%]; other mutations, 5/26 [19.3%]; supplemental Table 1). Survival analysis in the validation series confirmed that *NOTCH1* mutations represent an adverse prognostic factor in CLL. By univariate analysis, *NOTCH1* mutated patients were confirmed to display a significantly shorter OS compared with *NOTCH1* germ line patients ($P = .002$; Figure 1B). Similarly to the training series, also in the validation series *NOTCH1* mutations predicted poor OS irrespective of mutation type (supplemental Figure 1). By multivariate analysis, *NOTCH1* mutations were selected as an independent risk factor of OS also in the validation series (HR, 2.15; 95% CI, 1.13-4.11; $P = .019$), after controlling for the same covariates applied to the training series.

CLL with *NOTCH1* mutations have a poor prognosis similar to CLL with *TP53* disruption

NOTCH1 mutations and *TP53* disruption tended to distribute in a mutually exclusive fashion in both the training and validation

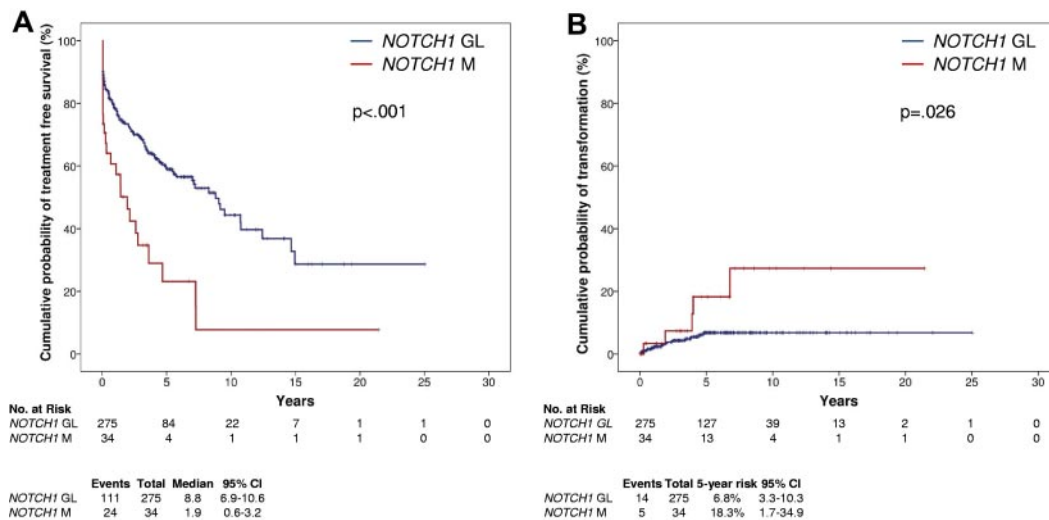


Figure 2. Kaplan-Meier estimates of TFS and cumulative probability of transformation to RS according to *NOTCH1* mutation status. TFS according to *NOTCH1* mutation status (A) and cumulative probability of transformation to RS (B) in the CLL training series ($n = 309$). *NOTCH1* germ line cases (*NOTCH1* GL) are represented by the blue line. *NOTCH1* mutated cases (*NOTCH1* M) are represented by the red line.

series (Figure 4A-B). Patients harboring *NOTCH1* mutations were devoid of *TP53* disruption in 31/34 (91.2%) and in 25/26 (96.2%) cases of the training and validation series, respectively.

Because *TP53* disruption identifies patients with the shortest survival in CLL,²²⁻²⁸ the outcome of *NOTCH1* mutated cases was compared with that of patients with *TP53* disruption. Cases harboring both *NOTCH1* and *TP53* lesions were compiled to cases harboring only *TP53* disruption, because of the low number (3 in the training series and 1 in the validation series) of double mutated cases, and based on a recursive partitioning analysis for risk of death. In both the training and validation series, CLL harboring *NOTCH1* mutations displayed an OS similar to that of CLL harboring *TP53* disruption (Figure 4C-D). The results were superimposable also when the few double mutated cases were analyzed as a separate subgroup (supplemental Figure 4).

ARMS is a useful tool for *NOTCH1* mutation screening

A PCR-based test was designed to detect the c.7544_7545delCT mutation that accounts for ~80% of *NOTCH1* mutations in CLL. ARMS was calibrated to detect a mutation present in greater than or equal to 10% alleles, approximating the sensitivity of DNA Sanger sequencing, and was applied in blind to the training series. Under these conditions, ARMS showed a 100% sensitivity and specificity in detecting *NOTCH1* c.7544_7545delCT ($\kappa = 1$). In fact, all 26 CLL from the training series harboring c.7544_7545delCT by DNA Sanger sequencing scored positive by ARMS, whereas all 283 CLL lacking c.7544_7545delCT by DNA Sanger sequencing scored negative by ARMS (Figure 5). By survival analysis, CLL that scored positive by ARMS for c.7544_7545delCT showed a shorter OS compared with negative patients ($P < .001$; Figure 5). These results confirm that ARMS is a useful tool for *NOTCH1* c.7544_7545delCT mutation screening.

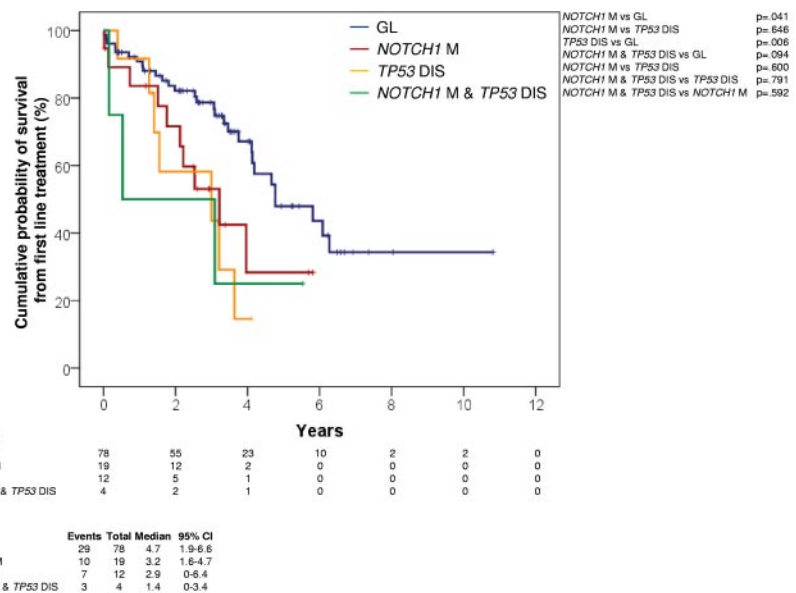


Figure 3. Kaplan-Meier estimates of OS from first-line treatment. OS from first line treatment according to *NOTCH1* mutation status in CLL treated with fludarabine-based regimens series ($n = 113$). Cases with germ line *NOTCH1* and *TP53* genes (GL) are represented by the blue curve. Cases harboring *NOTCH1* mutations without *TP53* disruption (*NOTCH1* M) are represented by the red curve. Cases harboring *TP53* disruption without *NOTCH1* mutations (*TP53* DIS) are represented by the yellow curve. Cases harboring both *NOTCH1* mutations and *TP53* disruption (*NOTCH1* M and *TP53* DIS) are represented by the green curve.

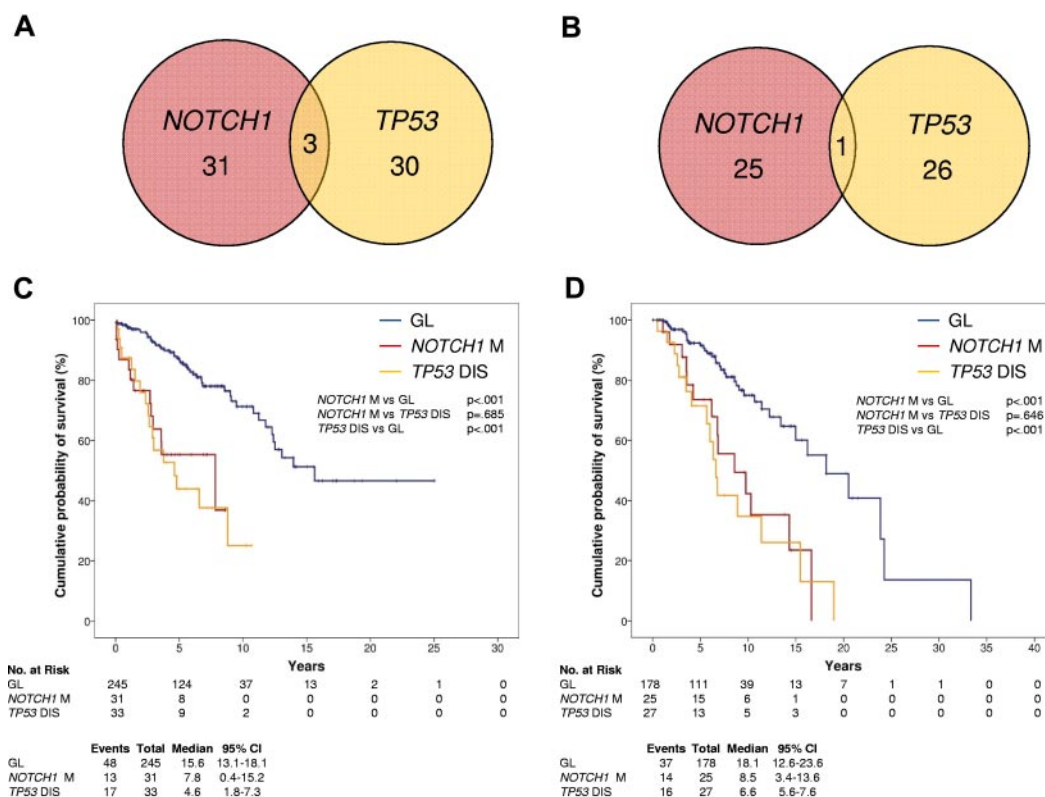


Figure 4. Comparative analysis of *NOTCH1* mutations and *TP53* disruption. (A-B) Venn diagram illustrating the overlap between *NOTCH1* mutations and *TP53* disruption by mutations, deletion, or both in the CLL training series (A) and in the CLL validation series (B). Numbers within the red and yellow circles indicate the number of cases harboring *NOTCH1* mutations without *TP53* disruption (red circle) and cases with *TP53* disruption without *NOTCH1* mutations (yellow circle). Numbers within the overlaps between circles indicate the number of cases harboring both *NOTCH1* mutations and *TP53* disruption. (C-D) Kaplan-Meier estimates of OS according to *NOTCH1* mutation and *TP53* disruption in the CLL training series (n = 309; C) and in the CLL validation series (n = 230; D). Cases with germ line *NOTCH1* and *TP53* genes (GL) are represented by the blue line. Cases harboring *NOTCH1* mutations without *TP53* disruption (*NOTCH1* M) are represented by the red line. Cases harboring *TP53* disruption (*TP53* DIS) are represented by the yellow line. Cases harboring both *NOTCH1* and *TP53* disruption were compiled to cases harboring only *TP53* disruption, because of the low number (3 in the training series and 1 in the validation series) of double-mutated cases, and based on a recursive partitioning analysis for risk of death.

Timing of *NOTCH1* mutations and relationship with *TP53* disruption in high-risk CLL

Paired sequential samples were tested in selected cases of high-risk CLL. *NOTCH1* mutations were acquired at chemorefractoriness in 1/4 cases and at RS transformation in 4/11 (supplemental Table 6).

The relationship between *NOTCH1* mutations and *TP53* disruption was investigated in high-risk CLL, including fludarabine-refractory CLL (n = 59) and RS (n = 26; Figure 6). At the time of fludarabine-refractoriness, *NOTCH1* mutations and *TP53* disruption overlapped in 11.8% (7/59) of cases (Figure 6B). Consistent with the whole CLL series, *NOTCH1* mutations and *TP53* disruption distributed in a mutually exclusive manner also in CLL that subsequently transformed to RS (Figure 6A). However, on RS transformation, *NOTCH1* mutations and *TP53* disruption occurred simultaneously in a fraction of patients as documented by: (1) acquisition of *TP53* disruption in 2 RS that already harbored *NOTCH1* mutations in the CLL phase; and (2) acquisition of both *NOTCH1* mutations and *TP53* disruption in 3 RS devoid of these alterations in the CLL phase (Figure 6A).

Discussion

The current study on 539 CLL documents that *NOTCH1* mutations: (1) represent one of the most frequent cancer gene mutations

known to be involved at CLL presentation; (2) among CLL genetic subgroups, cluster with cases harboring trisomy 12 and tend to be mutually exclusive with *TP53* disruption; (3) identify a high-risk subgroup of patients showing poor survival similar to that associated with *TP53* abnormalities; and (4) exert a prognostic role independent of widely accepted clinical and genetic risk factors, and in series from different institutions, as documented by the training-validation approach chosen for the design of this study.

Of the biologic predictors of CLL identified to date,¹⁹⁻²¹ *TP53* disruption is the sole risk factor consistently associated with high-risk patients.²²⁻²⁹ The genetics of high-risk CLL, however, is not fully recapitulated by *TP53* disruption, because 40% to 50% high-risk CLL are devoid of *TP53* abnormalities.²⁹ Conceivably, other genetic lesions may drive CLL aggressiveness. This study is consistent with a role of *NOTCH1* mutations in contributing to CLL clinical aggressiveness, because these genetic alterations identify patients whose survival is similar to that associated with *TP53* disruption.

The role of *NOTCH1* mutations in determining CLL aggressiveness is independent of the effect exerted by *TP53* disruption. In fact, at presentation, *NOTCH1* mutations in both the training and validation series tend to distribute in a mutually exclusive manner with *TP53* disruption. Consistently, the impact of *NOTCH1* mutations on CLL survival is independent of *TP53* disruption by multivariate analysis. The scenario observed in CLL differs from that of RS, in which mutations of *NOTCH1* associate with *TP53*

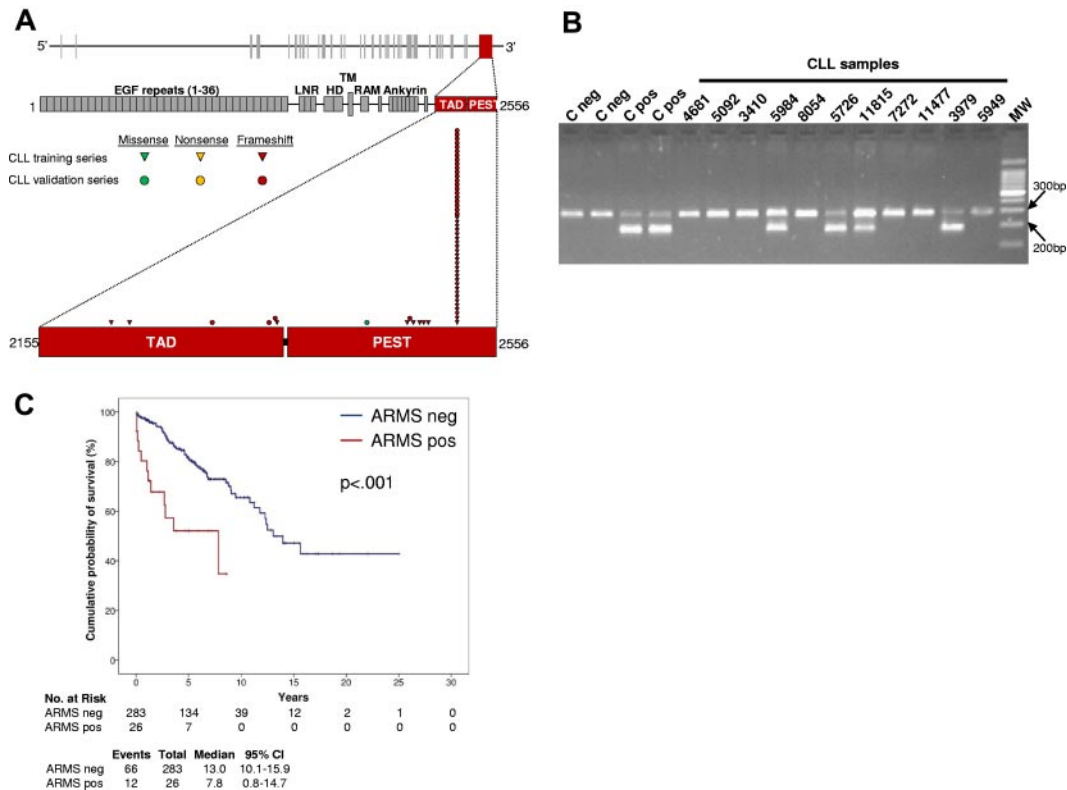


Figure 5. ARMS to detect *NOTCH1* c.7544_7545delCT mutation. (A) Schematic diagram of the *NOTCH1* gene (top) and protein (bottom), with its conserved functional domains (EGF-like repeats: LNR, LIN-12/NOTCH repeats; HD, heterodimerization; TM, transmembrane; Ankyrin repeats: TAD, transactivation domain; PEST, proline, glutamic acid, serine, threonine sequence). The TAD domain and the PEST sequence, both coded by exon 34, are magnified. Color-coded shapes indicate the position of the mutations found in the CLL training series (n = 34) and in the CLL validation series (n = 26). (B) Representative results of the ARMS assay showing 4 CLL samples that scored positive for the c.7544_7545delCT mutation (codes 5984, 5726, 11815, 3979) and 7 CLL samples that scored negative for the c.7544_7545delCT mutation (codes 4681, 5092, 3410, 8054, 7272, 11477, 5949). Negative samples show a normal band of 284 bp. Positive samples show an additional mutant band of 183 bp. Negative (C neg) and positive (C pos) controls also are included. Molecular weight (MW) is the 100-bp DNA ladder. Camera: Gel Doc 1000, BioRad; image acquisition software: Quantity One 4.5.0, BioRad. (C) Kaplan-Meier estimates of overall survival according to the results of the ARMS assay in the CLL training series (n = 309). Cases that scored negative by ARMS for the *NOTCH1* c.7544_7545delCT mutation (ARMS neg) are represented by the blue line. Cases that scored positive by ARMS for the *NOTCH1* c.7544_7545delCT mutation (ARMS pos) are represented by the red line.

disruption in 50% of the patients (Fabbri et al³¹; this study). A likely explanation is that the concomitant occurrence of *NOTCH1* mutations and *TP53* disruption in the same CLL clone causes further clinical aggressiveness and, potentially, histologic transformation to aggressive lymphoma.

The pivotal studies on *NOTCH1* mutations in CLL have provided initial evidence that *NOTCH1* alterations might be associated with an unfavorable clinical outcome.^{30,31,34} However,

these studies were based on small CLL series,³⁴ used TFS as surrogate clinical end point,³⁴ and lacked a formal demonstration that the clinical effect of *NOTCH1* mutations is reproducible and independent of confounders.^{30,31} Our results add to the current knowledge on the clinical aspects of *NOTCH1* mutations in CLL by demonstrating the robustness and reproducibility of these genetic alterations as a risk factor, thus providing a new tool for the early identification of high-risk patients.

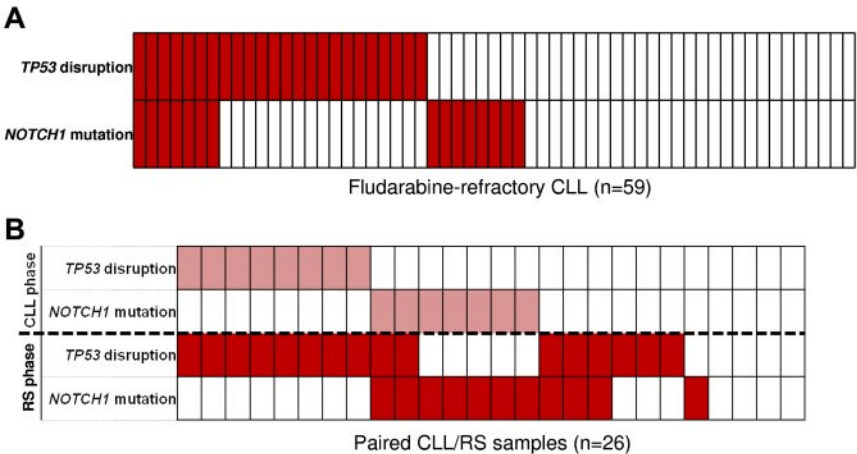


Figure 6. Timing of *NOTCH1* mutations and relationship with *TP53* disruption in high-risk CLL. (A) *NOTCH1* mutations and *TP53* disruption in fludarabine-refractory CLL. In the heatmap, rows correspond to the *NOTCH1* and *TP53* genes, and columns represent individual patients color-coded based on the gene status (white, wild type; red, mutations of *NOTCH1* and disruption of *TP53*). (B) *NOTCH1* mutations and *TP53* disruption in sequential CLL/RS samples. In the heatmap, rows correspond to the *NOTCH1* and *TP53* genes. Columns represent individual patients color-coded based on the gene status (white, wild type; pink, mutations of *NOTCH1* and disruption of *TP53* in the CLL phase; red, mutations of *NOTCH1* and disruption of *TP53* at RS transformation).

Different mechanisms might explain, at least in part, the poor prognosis associated with *NOTCH1* mutations in CLL. First, *NOTCH1* mutations lead to the acquisition of a progressive clinical phenotype that mandates treatment shortly after initial presentation, as documented by a median TFS of ~ 2 years for mutated cases versus ~ 9 years for *NOTCH1* germ line patients. Second, our actuarial analysis indicates that *NOTCH1* mutated patients display a higher risk of developing RS, a condition that is frequently lethal and recurrently harbors *NOTCH1* mutations that, importantly, are present already at the time of CLL presentation in a significant fraction of RS patients.^{15-18,31} A potential association of *NOTCH1* mutations with chemorefractoriness may further explain the poor outcome associated with *NOTCH1* alterations. Although the relationship between *NOTCH1* mutations and response to treatment needs to be formally tested within clinical trials, indirect evidence for this hypothesis comes from the observation that *NOTCH1* mutations are enriched among chemorefractory CLL patients³¹ and that *NOTCH1* activation in vitro confers resistance to apoptosis through NF- κ B pathway activation.³²

The external validation approach exploited in the current study documents that *NOTCH1* mutations are an independent prognostic factor retaining its predictive value in CLL followed at different institutions. This observation suggests that the prognostic value of *NOTCH1* mutations, though detected retrospectively, is independent of a potential bias because of patient referral or patient management at a single center.⁴² The general validity of *NOTCH1* mutations as a prognosticator in CLL is further supported by the consistent association of *NOTCH1* mutations with poor outcome in all series tested to date by independent investigators, although previous studies were not designed for a comprehensive survival analysis, or included a limited number of patients.^{30,31,34} Confirmation within the frame of prospective clinical trials will be helpful to fully assess the generalization of *NOTCH1* mutations as a prognostic marker in CLL.

Consistent with the mutational spectrum of *NOTCH1* in CLL, all mutations disrupted the C-terminal PEST domain that in normal conditions is required to limit the intensity and duration of *NOTCH1* activation.^{33,48,49} Removal of the PEST domain results in *NOTCH1* impaired degradation and accumulation of an active *NOTCH1* isoform sustaining deregulated signaling.³⁰ A practically important feature of the *NOTCH1* mutational spectrum is that one single recurrent mutation, c.7544_7545delCT, accounts for $\sim 80\%$ of all *NOTCH1* mutations detectable in CLL. The high recurrence of c.7544_7545delCT in CLL has prompted the design of a simple PCR-based strategy for its rapid detection. This assay

allows the reliable detection of all cases harboring the c.7544_7545delCT mutation, translates into prognostically meaningful results, and might provide a potentially helpful approach for a first-level screening of *NOTCH1* alterations avoiding the need of DNA sequencing procedures. In addition to prognostic implications, *NOTCH1* mutations might also provide a therapeutic target for *NOTCH1* inhibitors that are currently under development in other clinical contexts^{33,49} and that prompt future studies of molecular therapy for *NOTCH1*-mutated CLL patients.

Acknowledgments

This study was supported by Associazione Italiana Ricerca sul Cancro, Special Program Molecular Clinical Oncology, 5 $\times$ 1000, No. 10007, Milan, Italy (G.G. and R.F.); Progetto Fondo per gli Investimenti della Ricerca di Base-Programma "Futuro in Ricerca" 2008 (D.R.), PRIN 2008 (G.G. and R.M.), and PRIN 2009 (D.R.), Ministero dell'Istruzione, dell'Università e della Ricerca, Rome, Italy; Progetto Giovani Ricercatori 2008 (D.R.) and Ricerca Sanitaria Finalizzata 2008 (G.G.), Ministero della Salute, Rome, Italy; Novara-AIL Onlus, Novara, Italy (G.G.); Compagnia di San Paolo, Turin, Italy (R.F.); National Institutes of Health grant P01-CA092625 and a Specialized Center of Research grant from the Leukemia & Lymphoma Society (both R.D.-F.). S.M. and S.C. are being supported by fellowships from Novara-AIL Onlus, Novara. L.P. is on leave from the University of Perugia Medical School.

Authorship

Contribution: D.R., R.R., L.P., R.D.-F., R. Foà, and G.G. designed the study, interpreted data, and wrote the manuscript; D.R., M.F., and P.B. performed statistical analysis; S.R., V.S., A.B., M.C., S.C., and R. Famà performed molecular analysis; S.M. performed FISH analysis; L.D.P., F.F., L.L., and R.M. provided well characterized biological samples and clinical data; G.F., V.G., A.G., and S.D. contributed to interpretation of biological data; and D.C. was in charge of the ARMS assay.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Davide Rossi, Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy; e-mail: rossidav@med.unipmn.it.

References

- Müller-Hermelink HK, Montserrat E, Catovsky D, Campo E, Harris NL, Stein H. Chronic lymphocytic leukemia/small lymphocytic lymphoma. In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW, eds. *World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC; 2008:180-182.
- Hallek M, Cheson BD, Catovsky D, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*. 2008;111(12):5446-5456.
- Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med*. 2005;352(8):804-815.
- Dighiero G, Hamblin TJ. Chronic lymphocytic leukaemia. *Lancet*. 2008;371(9617):1017-1029.
- Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. *Blood*. 1975;46(2):219-234.
- Montserrat E, Viñolas N, Reverter JC, Rozman C. Natural history of chronic lymphocytic leukemia: on the progression and progression and prognosis of early clinical stages. *Nouv Rev Fr Hematol*. 1988;30(5-6):359-361.
- French Cooperative Group on Chronic Lymphocytic Leukaemia. Natural history of stage A chronic lymphocytic leukaemia untreated patients. *Br J Haematol*. 1990;76(1):45-57.
- Dighiero G, Maloum K, Desablens B, et al. Chlorambucil in indolent chronic lymphocytic leukemia. French Cooperative Group on Chronic Lymphocytic Leukemia. *N Engl J Med*. 1998;338(21):1506-1514.
- Del Giudice I, Mauro FR, De Propriis MS, et al. White blood cell count at diagnosis and immunoglobulin variable region gene mutations are independent predictors of treatment-free survival in young patients with stage A chronic lymphocytic leukemia. *Haematologica*. 2011;96(4):626-630.
- Wierda WG, O'Brien S, Wang X, et al. Prognostic nomogram and index for overall survival in previously untreated patients with chronic lymphocytic leukemia. *Blood*. 2007;109(11):4679-4685.
- Abrisqueta P, Pereira A, Rozman C, et al. Improving survival in patients with chronic lymphocytic leukemia (1980-2008): the Hospital Clinic of Barcelona experience. *Blood*. 2009;114(19):2044-2050.
- Shanafelt TD, Jenkins G, Call TG, et al. Validation

- of a new prognostic index for patients with chronic lymphocytic leukemia. *Cancer*. 2009; 115(2):363-372.
13. Shanafelt TD, Rabe KG, Kay NE, et al. Age at diagnosis and the utility of prognostic testing in patients with chronic lymphocytic leukemia. *Cancer*. 2010;116(20):4777-4787.
 14. Bulian P, Tarnani M, Rossi D, et al. Multicentre validation of a prognostic index for overall survival in chronic lymphocytic leukaemia. *Hematol Oncol*. 2011;29(2):91-99.
 15. Tsimberidou AM, Keating MJ. Richter syndrome: biology, incidence, and therapeutic strategies. *Cancer*. 2005;103(2):216-228.
 16. Rossi D, Gaidano G. Richter syndrome: molecular insights and clinical perspectives. *Hematol Oncol*. 2009;27(1):1-10.
 17. Rossi D, Spina V, Cerri M, et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clin Cancer Res*. 2009;15(13):4415-4422.
 18. Rossi D, Spina V, Deambrogi C, et al. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood*. 2011;117(12):3391-3401.
 19. Cramer P, Hallek M. Prognostic factors in chronic lymphocytic leukemia-what do we need to know? *Nat Rev Clin Oncol*. 2011;8(1):38-47.
 20. Kay NE, O'Brien SM, Pettitt AR, Stilgenbauer S. The role of prognostic factors in assessing 'high-risk' subgroups of patients with chronic lymphocytic leukemia. *Leukemia*. 2007;21(9):1885-1891.
 21. Binet JL, Caligaris-Cappio F, Catovsky D, et al. Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. *Blood*. 2006;107(3):859-861.
 22. Döhner H, Fischer K, Bentz M, et al. p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. *Blood*. 1995;85(6):1580-1589.
 23. Döhner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000;343(26):1910-1916.
 24. Dreger P, Corradini P, Kimby E, et al. Indications for allogeneic stem cell transplantation in chronic lymphocytic leukemia: the EBMT transplant consensus. *Leukemia*. 2007;21(1):12-17.
 25. Zenz T, Kröber A, Scherer K, et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood*. 2008;112(8):3322-3329.
 26. Rossi D, Cerri M, Deambrogi C, et al. The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. *Clin Cancer Res*. 2009;15(3):995-1004.
 27. Malcikova J, Smardova J, Rocnova L, et al. Monoallelic and biallelic inactivation of TP53 gene in chronic lymphocytic leukemia: selection, impact on survival, and response to DNA damage. *Blood*. 2009;114(26):5307-5314.
 28. Dicker F, Herholz H, Schnittger S, et al. The detection of TP53 mutations in chronic lymphocytic leukemia independently predicts rapid disease progression and is highly correlated with a complex aberrant karyotype. *Leukemia*. 2009;23(1):117-124.
 29. Stilgenbauer S, Zenz T. Understanding and managing ultra high-risk chronic lymphocytic leukemia. *Hematology Am Soc Hematol Educ Program*. 2010;2010:481-488.
 30. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475(7354):101-105.
 31. Fabbri G, Rasi S, Rossi D, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med*. 2011;208(7):1389-1401.
 32. Rosati E, Sabatini R, Rampino G, et al. Constitutively activated Notch signaling is involved in survival and apoptosis resistance of B-CLL cells. *Blood*. 2009;113(4):856-865.
 33. Paganin M, Ferrando A. Molecular pathogenesis and targeted therapies for NOTCH1-induced T-cell acute lymphoblastic leukemia. *Blood Rev*. 2011;25(2):83-90.
 34. Sportoletti P, Baldoni S, Cavalli L, et al. NOTCH1 PEST domain mutation is an adverse prognostic factor in B-CLL. *Br J Haematol*. 2010;151(4):404-406.
 35. Stein H, Warnke RA, Chan WC, Jaffe ES. Diffuse large B-cell lymphoma. In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW, eds. *World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC; 2008:233-237.
 36. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol*. 2005;23(36):9067-9072.
 37. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*. 1999;94(6):1848-1854.
 38. Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 1999;94(6):1840-1847.
 39. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Am Stat Assoc*. 1958;53:457-481.
 40. Cox DR. Regression models and life tables. *J R Stat Assoc*. 1972;34:187-220.
 41. Harrell FE Jr, Lee K, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361-387.
 42. Harrell FE Jr. *Regression Modeling Strategies*. New York, NY: Springer-Verlag; 2001:465-522.
 43. Schoenfeld D. Partial residuals for the proportional hazard regression model. *Biometrika*. 1982;69(1):239-241.
 44. Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals for survival models. *Biometrika*. 1990;77(1):216-218.
 45. Van Houwelingen JC, le Cessie S. Predictive value of statistical models. *Stat Med*. 1990;9(1):1303-1325.
 46. Chen CH, George SL. The bootstrap and identification of prognostic factors via Cox's proportional hazards regression model. *Stat Med*. 1995;14(1):39-46.
 47. Ciampi A, Lawless J, McKinney S, et al: Regression and recursive partition strategies in the analysis of medical and survival data. *J Clin Epidemiol*. 1988;41(8):737-748.
 48. Weng AP, Ferrando AA, Lee W, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science*. 2004;306(5694):269-271.
 49. Aster JC, Blacklow SC, Pear WS. Notch signaling in T-cell lymphoblastic leukaemia/lymphoma and other haematological malignancies. *J Pathol*. 2011;223(2):262-273.

blood

2012 119: 2854-2862
Prepublished online February 3, 2012;
doi:10.1182/blood-2011-12-395673

Disruption of *BIRC3* associates with fludarabine chemorefractoriness in *TP53* wild-type chronic lymphocytic leukemia

Davide Rossi, Marco Fangazio, Silvia Rasi, Tiziana Vaisitti, Sara Monti, Stefania Cresta, Sabina Chiaretti, Ilaria Del Giudice, Giulia Fabbri, Alessio Bruscaggin, Valeria Spina, Clara Deambrogi, Marilisa Marinelli, Rosella Famà, Mariangela Greco, Giulia Daniele, Francesco Forconi, Valter Gattei, Francesco Bertoni, Silvia Deaglio, Laura Pasqualucci, Anna Guarini, Riccardo Dalla-Favera, Robin Foà and Gianluca Gaidano

Updated information and services can be found at:
<http://bloodjournal.hematologylibrary.org/content/119/12/2854.full.html>

Articles on similar topics can be found in the following Blood collections
[Lymphoid Neoplasia](#) (1598 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>



Disruption of *BIRC3* associates with fludarabine chemorefractoriness in *TP53* wild-type chronic lymphocytic leukemia

*Davide Rossi,¹ *Marco Fangazio,¹ *Silvia Rasi,¹ Tiziana Vaisitti,² Sara Monti,¹ Stefania Cresta,¹ Sabina Chiaretti,³ Ilaria Del Giudice,³ Giulia Fabbri,⁴ Alessio Bruscaggini,¹ Valeria Spina,¹ Clara Deambrogi,¹ Marilisa Marinelli,³ Rosella Famà,¹ Mariangela Greco,¹ Giulia Daniele,⁵ Francesco Forconi,^{6,7} Valter Gattei,⁸ Francesco Bertoni,^{9,10} Silvia Deaglio,² Laura Pasqualucci,^{4,11,12} Anna Guarini,³ Riccardo Dalla-Favera,^{4,12,13} †Robin Foà,³ and †Gianluca Gaidano¹

¹Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; ²Department of Genetics, Biology and Biochemistry and Human Genetics Foundation, University of Turin, Turin, Italy; ³Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy; ⁴Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY; ⁵Hematology Unit, National Cancer Center of Bari and Department of Biology, University of Bari, Bari, Italy; ⁶Cancer Sciences Unit, Cancer Research UK Clinical Centre, University of Southampton, Southampton, United Kingdom; ⁷Division of Hematology, University of Siena, Siena, Italy; ⁸Clinical and Experimental Onco-Hematology, Centro di Riferimento Oncologico Istituto di Ricovero e Cura a Carattere Scientifico, Aviano, Italy; ⁹Institute of Oncology Research and ¹⁰Oncology Institute of Southern Switzerland, Bellinzona, Switzerland; ¹¹Institute of Hematology, University of Perugia, Perugia, Italy; and Departments of ¹²Pathology and Cell Biology and ¹³Genetics and Development, Columbia University, New York, NY

The genetic lesions identified to date do not fully recapitulate the molecular pathogenesis of chronic lymphocytic leukemia (CLL) and do not entirely explain the development of severe complications such as chemorefractoriness. In the present study, *BIRC3*, a negative regulator of noncanonical NF- κ B signaling, was investigated in different CLL clinical phases. *BIRC3* lesions were absent in monoclonal B-cell lymphocytosis (0 of 63) and were rare in CLL at diagnosis (13 of 306, 4%). Conversely, *BIRC3* disruption selectively affected 12 of

49 (24%) fludarabine-refractory CLL cases by inactivating mutations and/or gene deletions that distributed in a mutually exclusive fashion with *TP53* abnormalities. In contrast to fludarabine-refractory CLL, progressive but fludarabine-sensitive patients were consistently devoid of *BIRC3* abnormalities, suggesting that *BIRC3* genetic lesions associate specifically with a chemorefractory phenotype. By actuarial analysis in newly diagnosed CLL (n = 306), *BIRC3* disruption identified patients with a poor outcome similar to

that associated with *TP53* abnormalities and exerted a prognostic role that was independent of widely accepted clinical and genetic risk factors. Consistent with the role of *BIRC3* as a negative regulator of NF- κ B, biochemical studies revealed the presence of constitutive noncanonical NF- κ B activation in fludarabine-refractory CLL patients harboring molecular lesions of *BIRC3*. These data identify *BIRC3* disruption as a recurrent genetic lesion of high-risk CLL devoid of *TP53* abnormalities. (*Blood*. 2012;119(12):2854-2862)

Introduction

The clinical course of chronic lymphocytic leukemia (CLL) ranges from a very indolent disorder with a normal lifespan to a rapidly progressive disease leading to death.¹⁻³ Occasionally, CLL undergoes histological transformation to Richter syndrome (RS).⁴⁻⁶ The variable clinical course of CLL is driven, at least in part, by the immunogenetic and molecular heterogeneity of the disease.^{7,8}

The genetic lesions identified to date in CLL do not fully recapitulate the molecular pathogenesis of the disease and do not entirely explain the development of severe complications such as chemorefractoriness that still represent a major unmet clinical need.⁹⁻¹⁶ In clinical trials, fludarabine refractoriness is due to *TP53* disruption in approximately 40% of CLL patients failing treatment, but in a sizeable fraction of patients, the molecular basis of this aggressive clinical phenotype remains unclear.¹⁷⁻¹⁹ Recently, whole-exome sequencing studies have revealed novel genetic alterations such as *NOTCH1* and *SF3B1* mutations that might potentially

contribute to high-risk CLL, although the precise clinical impact of these molecular lesions is still under scrutiny.²⁰⁻²³

In CLL, activation of the NF- κ B pathway is regarded as a mechanism of resistance to disease eradication.²⁴⁻²⁸ Specific interactions between protective microenvironmental niches and CLL cells activate NF- κ B signaling, which in turn provides prosurvival signals to the leukemic clone through the up-regulation of several antiapoptotic genes. NF- κ B activation is correlated with CLL outcome and enhanced resistance to fludarabine.^{25,27,28} In addition to microenvironmental interactions, NF- κ B signaling may also be activated in B-cell neoplasia through an array of molecular lesions affecting genes at different levels of the pathway.^{29,30} The Baculoviral IAP Repeat Containing 3 (*BIRC3*) gene, along with TRAF2 and TRAF3, cooperates in the same protein complex that negatively regulates the MAP3K14 serine-threonine kinase, the central activator of noncanonical NF- κ B signaling.³¹⁻³⁴

Submitted November 30, 2011; accepted January 29, 2012. Prepublished online as *Blood* First Edition paper, February 3, 2012; DOI 10.1182/blood-2011-12-395673.

*D.R., M.F., and S.R. contributed equally to this study.

†R.F. and G.G. contributed equally to this study.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2012 by The American Society of Hematology

Table 1. Clinical and genetic characteristics of the consecutive series of newly diagnosed and previously untreated CLL patients

	All (N = 306)		BIRC3 disrupted (n = 13)		BIRC3 WT (n = 293)		P
	n	%	n	%	n	%	
Age > 65 y	200	65	10	77	190	65	.553
Male	169	55	8	61	161	55	.779
CLL phenotypic score > 3	306	100	13	100	293	100	1.000
Rai stage III-IV	33	11	4	31	29	10	.040
IGHV identity ≥ 98%	98	32	10	77	88	30	.001
TP53 disruption	31	10	0	0	31	11	.376
NOTCH1 mutations	33	11	2	15	31	11	.641
SF3B1 mutations	13	4	2	15	16	5	.173
11q22-q23 deletion	18	6	9	69	9	3	< .001
Trisomy 12	59	19	4	31	55	19	.285
13q14 deletion	160	52	9	69	151	293	.263
Normal FISH	91	30	0	0	91	31	.012

WT indicates wild-type.

After our initial observation of recurrent mutations of *BIRC3* in splenic marginal zone lymphoma,³⁵ we performed targeted resequencing studies of the *BIRC3* coding sequence and splicing sites across the spectrum of B-cell neoplasia. In the present study, we report that *BIRC3*-disrupting mutations and deletions recurrently and selectively associate with *TP53* wild-type, fludarabine-refractory CLL and predict a dismal clinical outcome in consecutive CLL series.

Methods

Patients

The study population comprised 4 clinical CLL panels representative of different disease phases, including: (1) a consecutive series of newly diagnosed and previously untreated CLL (n = 306; Table 1; median survival from diagnosis, 12.5 years); (2) a cohort of fludarabine-refractory CLL (n = 49; Table 2; median survival from fludarabine-based treatment, 1.8 years); (3) a cohort of fludarabine-sensitive CLL (n = 68; supplemental Table 1 [available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article]; median survival from fludarabine-based treatment, 5.1 years); and (4) a cohort of clonally related RS (n = 33; all diffuse large B-cell lymphomas; supplemental Table 2; median survival from RS transformation, 1.1 year). A consecutive series of clinical

monoclonal B-cell lymphocytosis (MBL) with a CLL-like phenotype was also investigated (supplemental Table 3). Diagnosis of MBL, CLL, and fludarabine refractoriness was based on the International Workshop on Chronic Lymphocytic Leukemia-National Cancer Institute criteria.² Diagnosis of RS was based on histological criteria.¹

In addition to CLL, RS, and MBL, we also investigated a cohort of 194 lymphoid tumors representative of the main categories of mature B-cell neoplasms (follicular lymphoma, n = 20; diffuse large B-cell lymphoma, n = 30; Burkitt lymphoma, n = 38; extranodal marginal zone lymphoma, n = 65; hairy-cell leukemia, n = 19; and multiple myeloma, n = 22). The number of the mature B-cell neoplasms included in each panel allows a 90% probability of identifying genes that are mutated in at least 10% of patients.

CLL and clinical MBL samples were extracted from fresh or frozen PBMCs isolated by Ficoll-Paque gradient centrifugation. PBMCs were obtained at the following times: (1) for clinical MBL and for the consecutive series of newly diagnosed and previously untreated CLL, at disease presentation; (2) for the fludarabine-refractory CLL cohort, immediately before starting treatment to which the patient failed to respond because of stable/progressive disease; and (3) for the fludarabine-sensitive CLL cohort, at the time of progression immediately before starting the treatment to which the patient responded. All RS studies were performed on RS diagnostic biopsies. Samples of the other lymphoid neoplasms were obtained at diagnosis from the involved site (lymph nodes or extranodal sites in the case of lymphoma; CD138⁺ cells purified from BM aspirate in

Table 2. Clinical and genetic characteristics of the fludarabine-refractory CLL cohort

	All (N = 49)		BIRC3 disrupted (n = 12)		BIRC3 WT (n = 37)		P
	N	%	N	%	N	%	
Age > 65 y	36	73	8	67	28	75	.708
Male	31	63	7	58	24	65	.738
CLL phenotypic score > 3	49	100	12	100	37	100	1.000
Rai stage III-IV	28	57	8	67	20	54	.443
Treatment regimen at refractoriness							.499
FCR	17	35	5	42	12	32	
FR	2	4	0	0	2	5	
FC	14	29	2	17	12	32	
F	16	33	5	42	11	30	
IGHV identity ≥ 98%	39	80	10	82	29	78	1.000
TP53 disruption	17	35	0	0	17	37	.004
NOTCH1 mutations	12	24	3	25	9	24	1.000
SF3B1 mutations	12	24	2	17	7	19	1.000
11q22-q23 deletion	13	26	7	53	6	16	.008
Trisomy 12	14	29	3	25	11	30	1.000
13q14 deletion	28	57	7	58	21	57	1.000
Normal FISH	6	12	1	8	5	13	1.000

WT indicates wild-type; FCR, fludarabine, cyclophosphamide, rituximab; FR, fludarabine, rituximab; FC, fludarabine, cyclophosphamide; and F, fludarabine.

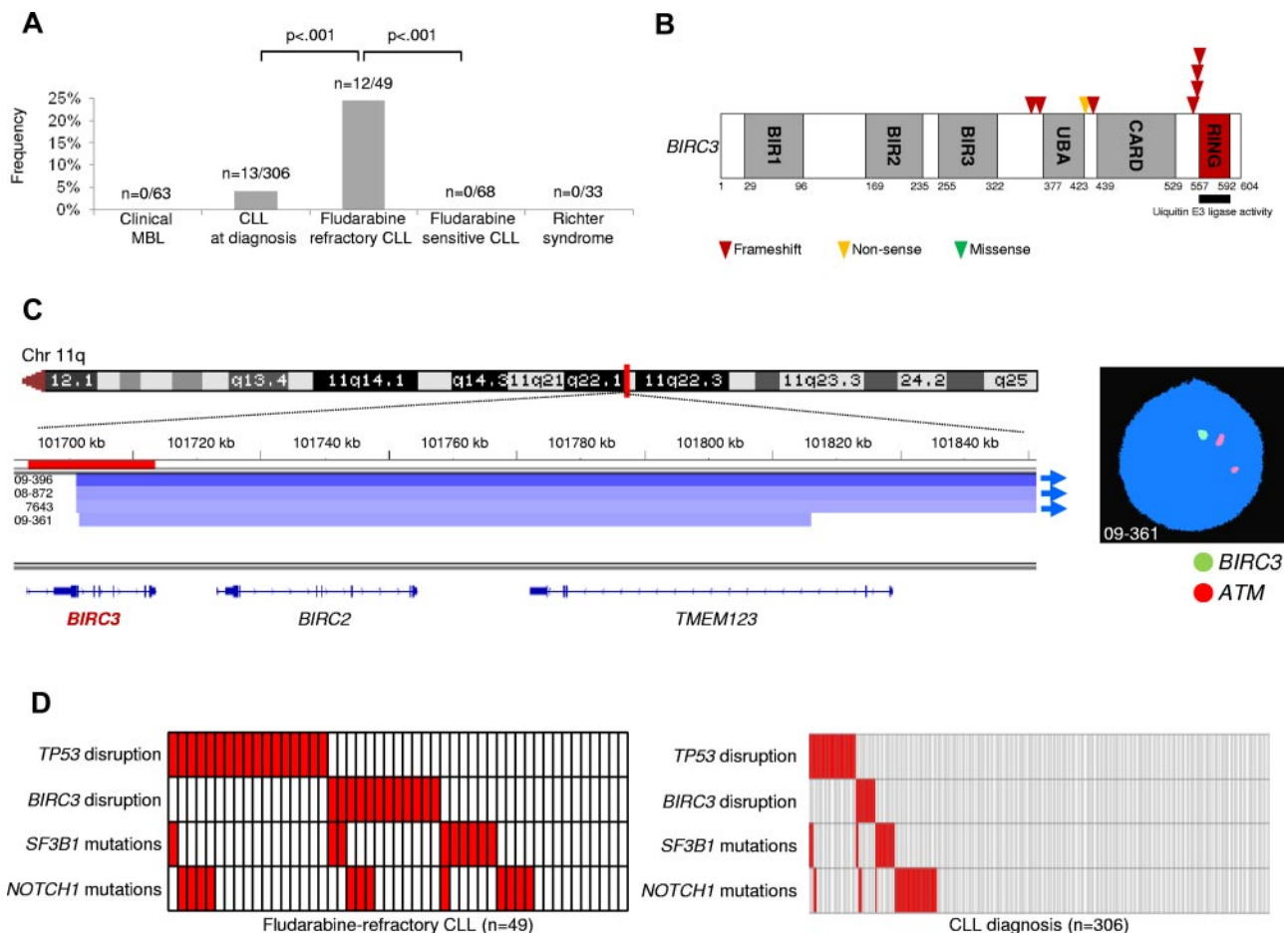


Figure 1. *BIRC3* disruption in CLL. (A) prevalence of *BIRC3* disruption in clinical MBL, in CLL at diagnosis, in fludarabine-refractory CLL, in fludarabine-sensitive CLL, and in RS. Numbers on top indicate the actual number of mutated samples over the total number analyzed. (B) Schematic diagram of the *BIRC3* protein with its key functional domains. Color-coded symbols indicate the type and position of the mutations in *BIRC3*. (C) Graphic representation of segmentation data from 4 CLL patients carrying *BIRC3* deletion. Deletions start from a centromeric break that truncates *BIRC3* and removes its terminal exons, including exon 9, which encodes the RING domain. Sample 09-361 harbors a focal loss of 411 kb on 11q22 involving *BIRC3* and its homolog, *BIRC2*. Each track represents one sample; white denotes a normal (diploid) copy number and blue indicates region of a copy number loss (Integrative Genomics Viewer software; <http://www.broadinstitute.org/igv>; assembly NCBI36/hg18). Individual genes in the region are aligned in the bottom panel. Dual-color FISH validates the occurrence of 11q22 deletion involving *BIRC3* but sparing *ATM* in sample 09-361 (RP11-17708-*BIRC3* specific probe in green and LSIATM probe in orange). (D) Mutual relationship of the *BIRC3* disruption with other genetic lesions in CLL at diagnosis and in fludarabine-refractory CLL. In the heat map, rows correspond to identical genes and columns represent individual patients color-coded based on the gene status (white indicates wild-type and red, mutations and/or deletion of *TP53*, mutations and/or deletion of *BIRC3*, mutations of *SF3B1*, and mutations of *NOTCH1*).

the case of multiple myeloma; and peripheral blood purified B cells in the case of hairy-cell leukemia). In all samples, the fraction of tumor cells corresponded to > 70% as assessed by flow cytometry and/or immunohistochemistry. Matched normal DNAs from the same patients were obtained from saliva or from purified granulocytes and confirmed to be tumor-free by PCR of tumor-specific *IGHV-D-J* rearrangements.³⁶

Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki. The study was approved by the Ethical Committee of the Ospedale Maggiore della Carità di Novara affiliated with the Amedeo Avogadro University of Eastern Piedmont (protocol code 59/CE; study number CE 8/11).

Mutation analysis of *BIRC3* and *TP53*

The *BIRC3* (exons 2-9, including splicing sites; RefSeq NM_001165.3), *TRAF2* (exons 2-11, including splicing sites; RefSeq NM_021138.3), *TRAF3* (exons 3-12, including splicing sites; RefSeq NM_145725.2), *MAP3K14* (exons 1-17, including splicing sites; RefSeq NM_003954.2), *TP53* (exons 4-8, including splicing sites; RefSeq NM_000546.4), *NOTCH1* (exons 26, 27 and 34; RefSeq NM_017617.2), and *SF3B1* (exons 1-25, including splice sites; RefSeq NM_012433.2) genes were analyzed by PCR amplification and DNA direct sequencing of high-molecular-weight genomic DNA.^{11,20,22,23,35} Purified amplicons were

subjected to conventional DNA Sanger sequences using the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Sequences were compared with the corresponding germline RefSeq sequences using the Mutation Surveyor Version 2.41 software package (SoftGenetics) after both automated and manual curation. All variants were sequenced from both strands on independent PCR products. Synonymous mutations, previously reported polymorphisms (dbSNP135, Ensembl Database, UCSC Genome Browser, or the 1000 Genome Project) and changes present in the matched normal DNA were removed from the analysis. In patients displaying more than one mutational event within the *BIRC3* gene, the allelic distribution of the mutations was determined by cloning and sequencing full-length PCR products obtained from cDNA ($n = 20$ clones). Molecular studies were performed blindly with respect to clinical data. All PCR primers and conditions are available on request.

IGHV mutation status

The Ig heavy variable gene (*IGHV*) mutational status was investigated as described previously.⁷ Sequences were aligned to the ImMunoGeneTics sequence directory and considered mutated if identity to the corresponding germline gene was < 98%.^{37,38}

Interphase FISH

Probes used for FISH analysis were: LSI13 and LSI13S319 (13q14 deletion), CEP12 (trisomy 12), LSIp53 (17p13/*TP53* deletion), and LSI-ATM (11q2-q23/*ATM* deletion; all from Abbott Laboratories) and the RP11-177O8 (*BIRC3*), RP11-769N4 (*TRAF2*), RP11-676M2 (*TRAF3*), and RP11-666C2 (*MAP3K14*) BAC clones. For each probe, at least 400 interphase cells with well-delineated fluorescent spots were examined. The presence of 13q14 deletion, trisomy 12, 11q22-q23 deletion, 17p13 deletion, *BIRC3* deletion, *TRAF3* deletion, *MAP3K14* gain, and *TRAF2* abnormalities was scored when the percentage of nuclei with the abnormality was above our internal cutoff (5%, 5%, 7%, 10%, 10%, 10%, 10%, and 10%, respectively), defined as the means plus 3 SDs of the frequency of normal control cells exhibiting the abnormality.^{6,35}

Copy-number analysis

Copy-number data were derived from a previously reported dataset.³⁹ Genome-wide DNA profiles were obtained from high-molecular-weight genomic DNA of CLL patients using the Genome-Wide Human SNP Array Version 6.0 (Affymetrix) following the manufacturer's instructions. The bioinformatics pipeline used for the identification of copy-number alterations is described in detail in Rinaldi et al.³⁹

Western blot studies

CLL cells were purified by negative selection using anti-CD3, anti-CD14, and anti-CD16 mAbs and Dynal magnetic beads (Invitrogen). Proteins were resolved by SDS-PAGE and analyzed by Western blot. Abs were anti-NFKB2 (#4882; Cell Signaling Technology), anti-BIRC3 (#3130; Cell Signaling Technology), and anti-actin (#sc-1615; Santa Cruz Biotechnology). Image acquisition and densitometric analyses were performed using ImageQuant LAS4000 and TL Version 7.0 software (GE Healthcare). Band intensities were calculated by standardizing over actin and using the JJN3 cell line as an internal loading control in different gels.³⁵

Statistical analysis

Overall survival (OS) was measured from date of initial presentation to date of death (event) or last follow-up (censoring).² Survival analysis was performed by the Kaplan-Meier method.⁴⁰ The crude association between exposure variables and outcome was estimated by univariate Cox regression analysis.⁴¹ The independence of *BIRC3* disruption as a predictor of OS for CLL was estimated after controlling for confounding variables by multivariate Cox regression analysis.⁴¹ The following variables were included in multivariate analysis: *BIRC3* disruption by mutation and/or deletion (present vs absent), age (> 65 years vs ≤ 65 years), Rai stage (III-IV vs 0-II), *IGHV* identity ≥ 98% (present vs absent), 11q22-q23 deletion (present vs absent), *TP53* disruption by mutation and/or deletion (present vs absent), *NOTCH1* mutations (present vs absent), and *SF3B1* mutations (present vs absent). None of the covariates violated the proportional hazard assumption as documented by plotting the smoothed Schoenfeld residuals and by performing a correlation test between time and residuals.^{42,43} The assumption of effect additivity of predictors was not violated, as documented by a global test of additivity including interactions between *BIRC3* disruption and other covariates. None of the covariates showed collinearity.⁴²

The prediction accuracy of the multivariate survival model was verified by assessing model discrimination and calibration.⁴² Discrimination indicates how well the model separates patients who experienced the event of interest from patients who did not, and was measured by calculating the c-index (where 1 is perfect discrimination and 0.5 is equivalent to chance).⁴² Calibration describes how well the survival probability from the model corresponds to the survival probability from the observed data and was assessed graphically by plotting the predicted probability of survival against the actual probability of survival.⁴² The heuristic shrinkage estimator, which quantifies model overfitting, was calculated using the formula (model likelihood ratio χ^2) – (number of degree of freedom in the model)/(model likelihood ratio χ^2). According to this formula, the more the shrinkage approximates 1, the lower the overfitting of the model.⁴² A

bootstrap resampling (with 1000 resamples) was used to calculate the bias-corrected c-index and to construct a bias-corrected calibration curve and slope.⁴² This approach provides a bias-corrected estimate of prediction accuracy of the model to protect against overfitting.

The stability and predictive performance of *BIRC3* disruption as an independent predictor of OS from CLL was validated internally using a bootstrapping resampling procedure.^{42,44} In the first step, 1000 bootstrap samples were generated randomly with replacement from the original CLL population. Cox regression was applied to each bootstrap sample with the same covariates as the original modeling. The percentage of bootstrap samples for which each covariate was selected as significant in the model was then calculated. The percentage of selection reflects the prognostic importance of a covariate, because it is expected that an important covariate will be selected for the majority of bootstrap samples. In the second step, 1000 additional bootstrap samples were generated randomly with replacement from the original CLL population. Cox regression was applied to each bootstrap sample with the same covariates as the original modeling. For each covariate, the mean SD and confidence intervals were computed for the 1000 bootstrap replications.^{42,44}

Categorical variables were compared by χ^2 test and Fisher exact test when appropriate. Continuous variables were compared by the Mann-Whitney test. All statistical tests were 2-sided. Statistical significance was defined as $P < .05$. The analysis was performed with the Statistical Package for the Social Sciences (SPSS) Version 18.0 software and with R statistical package Version 2.13.0 (<http://www.r-project.org>).

Results

BIRC3 mutations across the spectrum of B-cell neoplasia

After the initial observation of recurrent *BIRC3* mutations in splenic marginal zone lymphoma,³⁵ we performed targeted resequencing of the *BIRC3* coding sequence and splicing sites across the spectrum of B-cell neoplasias. This initial screening revealed that *BIRC3* mutations are selectively restricted to CLL (2 of 20; both cases were primary fludarabine refractory), whereas they are absent in diffuse large B-cell lymphoma (0 of 30), Burkitt lymphoma (0 of 38), follicular lymphoma (0 of 20), extranodal marginal zone lymphoma (0 of 65), hairy-cell leukemia (0 of 19), and multiple myeloma (0 of 22). These results prompted the investigation of the prevalence and clinical impact of *BIRC3* genetic lesions in different clinical phases of CLL.

BIRC3 is frequently targeted by inactivating lesions in fludarabine-refractory CLL

BIRC3 lesions were absent in MBL with a CLL-like phenotype (0 of 63) and were rare in CLL at diagnosis (13 of 306, 4%; $n = 3$ monoallelic truncating mutations, $n = 8$ monoallelic deletions, and $n = 2$ biallelic disruptions by truncating mutation coupling with a deletion; Figure 1A-B and supplemental Tables 3 and 4). Conversely, *BIRC3* was affected in 12 of 49 (24%) fludarabine-refractory CLL patients by inactivating mutations (7 frameshift and 1 nonsense) and/or gene deletions ($n = 7$; nuclei harboring *BIRC3* deletion: 20%-91%; Figure 1A-B and supplemental Tables 3 and 4). *BIRC3* lesions were similarly represented among patients who received fludarabine-cyclophosphamide-rituximab (FCR) or other fludarabine-based regimens (Table 2). All inactivating mutations were somatically acquired, were predicted to generate aberrant transcripts carrying premature stop codons, and caused elimination or truncation of the C-terminal RING domain, the E3 ubiquitin ligase activity of which is essential for proteasomal degradation of MAP3K14 (Figure 1B).³¹⁻³⁴

BIRC3 lesions were monoallelic in all but 2 CLL patients (patients 5977 and 6550) who carried inactivation of both *BIRC3* alleles (Table 3 and supplemental Table 4). In patient 5977, a

Table 3. *BIRC3* mutations in CLL

Sample ID	Nucleotide change†	Amino acid change§	Predicted functional consequence
5977*†	c.1101_1132del32	p.G367fs*6	Truncated protein
6550*†	c.1270G > T	p.E424*	Truncated protein
6550*†	c.1183_1352del4894	p.V395fs*78	Truncated protein
3878*†	c.1279_1280insA	p.I427fs*11	Truncated protein
5610*†	c.1638_1639insA	p.Q547fs*12	Truncated protein
5889*†	c.1663_1666delAGAA	p.R555fs*12	Truncated protein
12632*	c.4388delA	p.E553fs*22	Truncated protein
12534*	c.4441_4444delAGAA	p.E553fs*14	Truncated protein

*For these patients, paired normal DNA was available and confirmed the somatic origin of the mutation.

†These patients were treated at diagnosis with fludarabine-based regimens to which they failed to respond; therefore, they are included both in the fludarabine-refractory CLL cohort and in the consecutive series of newly diagnosed CLL patients.

‡Numbering according to GenBank accession number NM_001165.3.

§Numbering according to GenBank accession number NP_001156.1.

frameshift mutation of one allele was coupled with a deletion of the second allele. Patient 6550 harbored 3 distinct genetic events on *BIRC3*, including: (1) a monoallelic deletion of the entire *BIRC3* locus in 84% of the nuclei; (2) a large intragenic deletion of 4894 bp that removed the last 142 bp of exon 6, the entire intron 6-7, and the first 285 bp of exon 7 and caused a frameshift mutation (c.1183_1352del, p.V395fs*78) in the gene-coding sequence; and (3) a nonsense truncating mutation (c.1270G > T, p.E424*). In patient 6550, sequencing analysis of *BIRC3* transcripts after cDNA amplification and subcloning demonstrated that the c.1183_1352del intragenic deletion and the c.1270G > T nonsense mutation were located on separate alleles and that the c.1183_1352del intragenic deletion was represented in the majority of the clones, whereas the c.1270G > T nonsense mutation was restricted to 15% of the clones. These data suggest a tumor mosaicism for *BIRC3* in patient 6550.

SNP array analysis of a large CLL dataset (n = 339, including 158 newly diagnosed CLL investigated in this study) identified a focal loss of 411 kb involving *BIRC3* and its homolog *BIRC2* in one CLL (case 09-361; Figure 1C and supplemental Table 4). This deletion truncated *BIRC3* and removed its terminal exons, including exon 9 that encodes the RING domain. Three additional CLL harbored large 11q22 deletions starting from a centromeric break that truncated *BIRC3* and removed its terminal exons, including exon 9 (Figure 1C and supplemental Table 4).

Overall, these data document that *BIRC3* disruption by mutation and/or deletion occurs frequently in fludarabine-refractory CLL, and point to removal of the C-terminal RING domain as a common mechanism altering *BIRC3* function in this context.

Genetic lesions of *BIRC3* cause the formation of truncated *BIRC3* proteins and associate with activated NF- κ B signaling

Among cases carrying truncating gene mutations, Western blot analysis using Abs directed against the N-terminus of the *BIRC3* protein revealed the expression of an aberrant band of lower molecular weight corresponding in size to the predicted truncated *BIRC3* (Figure 2). A full-length *BIRC3* protein, corresponding to the wild-type allele, was also present. Patient 6550, who carried 3 genetic events, expressed very low levels of both the wild-type and the truncated *BIRC3* proteins, with a pattern that is consistent with the occurrence of 2 clonal disrupting deletions and a subclonal truncating mutation in this patient. Patients carrying a monoallelic *BIRC3* deletion displayed reduced levels of *BIRC3* protein that were consistent with a residual expression of the wild-type allele (Figure 2).

Western blot analysis of NF κ B2 processing from p100 to p52 revealed a constitutive noncanonical NF- κ B activation in CLL harboring *BIRC3* disruption by mutations or deletion (Figure 2;

mean p52/p100 ratio, 0.565). In contrast, noncanonical NF- κ B activation was not observed in patients devoid of *BIRC3* abnormalities, as documented by the predominant expression of p100 over p52 (mean p52/p100 ratio, 0.138) in these patients.

BIRC3-inactivating lesions are specific for fludarabine-refractory CLL and are mutually exclusive with *TP53* disruption

To investigate whether *BIRC3* genetic lesions were restricted to chemorefractory cases in progressive CLL patients, we analyzed *BIRC3* for mutations and deletions in progressive but fludarabine-sensitive CLL and in RS (supplemental Tables 1-3).

Fludarabine-sensitive CLL patients were consistently devoid of *BIRC3* disruption in all cases (n = 68), suggesting that *BIRC3* genetic lesions specifically associate with a chemorefractory phenotype among progressive CLL cases requiring treatment (Figure 1A). *BIRC3* genetic lesions were absent in clonally related RS (0 of 33; Figure 1A and supplemental Table 2), thus strengthening the notion that RS is molecularly distinct from chemorefractory progression without transformation.^{6,22}

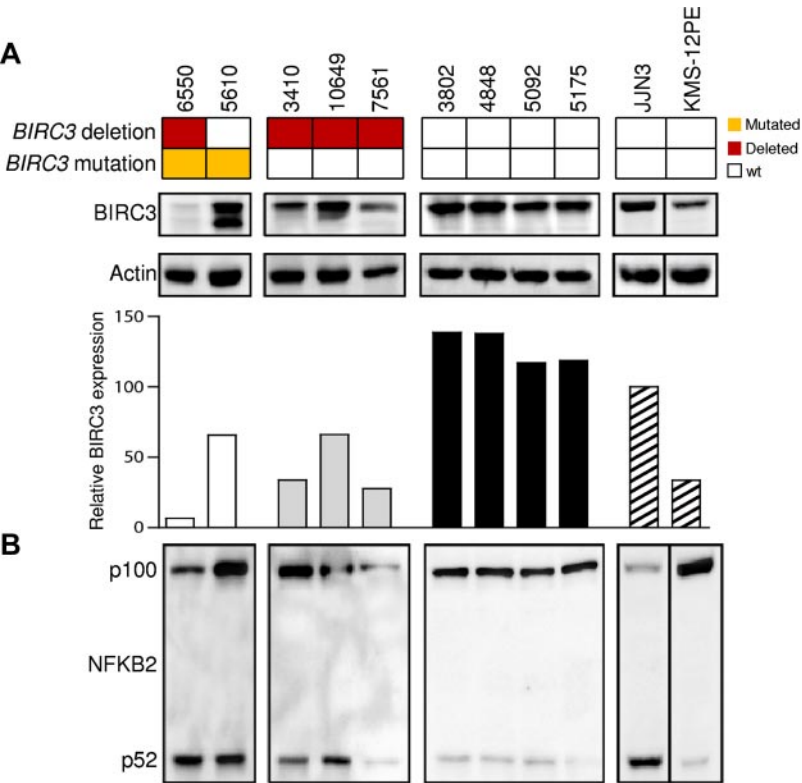
We then tested the relationship between disruption of *BIRC3* and *TP53* in fludarabine-refractory CLL. In the current series of fludarabine-refractory patients, *TP53* was disrupted by mutations (n = 10) and/or deletions (n = 14) in 17 of 49 (35%) patients. *BIRC3* lesions distributed in a mutually exclusive fashion with *TP53* disruption (Figure 1D). *BIRC3* lesions segregated with *TP53* wild-type, fludarabine-refractory CLL patients (12 of 32, 37%), whereas they were consistently absent among *TP53*-disrupted patients (0 of 17; *P* = .004). These data show that *BIRC3* lesions selectively associate with fludarabine-refractory but *TP53* wild-type CLL across the clinical spectrum of the disease.

By combining *BIRC3* disruption with other genetic lesions enriched in chemorefractory patients (ie, *TP53* disruption, *NOTCH1* mutations, and *SF3B1* mutations), fludarabine-refractory CLL appeared to be characterized by multiple molecular alterations that to some extent are mutually exclusive and account for approximately 80% of cases.

BIRC3 disruption identifies CLL patients whose survival is similar to that associated with *TP53* abnormalities

In a consecutive CLL series evaluated at diagnosis, *BIRC3* disruption (13 of 306, 4%) was associated with unfavorable clinical and genetic features (Table 1) and predicted poor outcome because of primary chemorefractoriness among patients requiring treatment. As shown by univariate analysis, the crude impact of *BIRC3* disruption on survival was an approximately 5-fold increase in the hazard of death (hazard ratio = 5.09; 95% confidence interval,

Figure 2. The NF- κ B pathway is activated in CLL patients harboring *BIRC3* disruption. Western blot analysis showing *BIRC3* expression and NFKB2 processing in purified primary tumor cells from 9 CLL patients carrying wild-type or disrupted *BIRC3*. The JJN3 (plasma cell leukemia) cell line was used as a positive control for NFKB2 processing and *BIRC3* expression. The KMS-12PE cell line (a multiple myeloma cell line) was used as a negative control for NFKB2 processing. Actin was used as a loading control. The histogram shows the relative fold change of *BIRC3* expression compared with the *BIRC3* expression level observed in the JJN3 cell line (arbitrarily set as 100). The aberrant *BIRC3* bands in patients 6550 and 5610 correspond in size to the predicted *BIRC3*-truncated protein.



2.52-10.28; $P < .001$) and a shortening of OS (median OS in *BIRC3*-disrupted patients, 3.1 years vs not reached in *BIRC3* wild-type patients; $P < .001$; Figure 3A and Table 4), which occurred irrespective of the type of *BIRC3* lesion (for mutations, $P = .001$; for deletions, $P = .001$; supplemental Figure 1).

As shown by multivariate analysis, the increased risk of death predicted by *BIRC3* disruption (hazard ratio = 3.04; 95% confidence interval, 1.21-7.61; $P = .017$) was independent of confound-

ing clinical (ie, age and Rai stage) and genetic (ie, *IGHV* mutation status, 11q22-q23 deletion, *TP53* abnormalities, *NOTCH1* mutations, and *SF3B1* mutations) variables at diagnosis (Table 4).

Analogous to CLL analyzed at the time of chemorefractoriness, *BIRC3* disruption and *TP53* abnormalities were also distributed in a mutually exclusive fashion at CLL presentation (mutual information = 0.004; Figure 1D). Because *TP53* abnormalities identify patients with the shortest survival in CLL, the outcome of

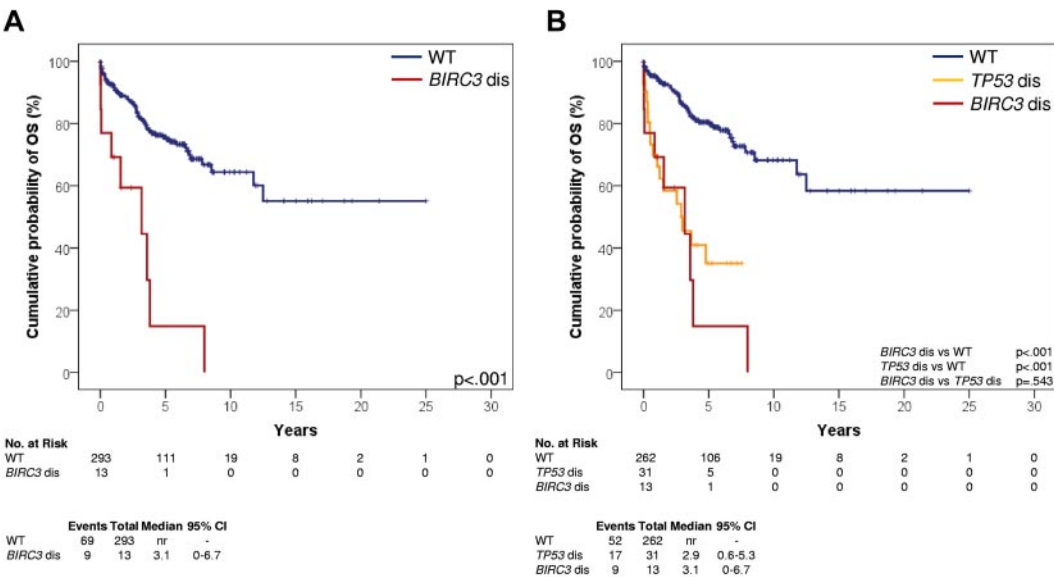


Figure 3. Kaplan-Meier estimates of OS according to *BIRC3* disruption. (A) Kaplan-Meier estimates of OS according to *BIRC3* disruption in newly diagnosed CLL patients ($n = 306$). *BIRC3* wild-type (WT) patients are represented by the blue line. Patients with *BIRC3* disruption by mutations and/or deletions (*BIRC3* dis) are represented by the red line. (B) Kaplan-Meier estimates of OS according to *BIRC3* and *TP53* disruption in newly diagnosed CLL patients ($n = 306$). *BIRC3* and *TP53* wild-type patients are represented by the blue line. Patients harboring *BIRC3* disruption but who are wild-type on the *TP53* gene (*BIRC3* dis) are represented by the red line. Patients harboring *TP53* disruption but who are wild-type on the *BIRC3* gene (*TP53* dis) are represented by the yellow line.

Table 4. Univariate and multivariate analysis for OS in newly diagnosed and previously untreated CLL patients

Characteristic	Events	Total	OS, y			Univariate analysis				Multivariate analysis*†‡				Internal bootstrapping validation			
			Median	LCI	UCI	HR	LCI	UCI	P	HR	LCI	UCI	P	Bootstrap parameters, mean			Bootstrap selection, %
<i>BIRC3</i> WT	69	293	NR														
<i>BIRC3</i> disruption	9	13	3.1	0	6.7	5.09	2.52	10.28	< .001	3.04	1.21	7.61	.017	3.90	1.32	13.17	85%
Age ≤ 65 y	13	106	NR														
Age > 65 y	65	200	8.5	4.6	12.4	3.57	1.95	6.52	< .001	3.88	2.08	7.22	< .001	4.41	2.21	8.87	100%
Rai stage 0-II	56	273	NR														
Rai stage III-IV	22	33	2.7	0.6	4.9	5.44	3.31	8.97	< .001	3.10	1.72	5.59	< .001	3.44	1.83	6.47	97%
<i>IGHV</i> identity < 98%	42	208	NR														
<i>IGHV</i> identity ≥ 98%	36	98	11.7	3.0	20.4	2.03	1.29	3.18	.002	1.24	0.72	2.12	.426	1.31	0.75	2.28	33%
No 11q22-q23 deletion	67	288	NR														
11q22-q23 deletion	11	18	3.1	0.8	5.4	4.08	2.14	7.77	< .001	2.18	0.85	5.62	.105	2.51	0.85	5.64	57%
<i>TP53</i> WT	61	275	NR														
<i>TP53</i> disruption	17	31	2.9	0.6	5.3	3.85	2.23	6.65	< .001	4.97	2.75	8.97	< .001	5.30	2.84	9.92	99%
<i>NOTCH1</i> WT	63	273	NR														
<i>NOTCH1</i> mutated	15	33	3.3	1.2	14.3	2.15	1.22	3.79	.007	2.08	1.04	4.16	.038	2.31	1.13	4.76	74%
<i>SF3B1</i> WT	71	288	NR														
<i>SF3B1</i> mutated	7	18	2.5	1.7	3.3	3.29	1.50	7.21	.003	2.49	1.08	5.72	.032	2.94	1.20	7.35	71%

WT indicates wild-type; HR, hazard ratio; LCI, 95% lower confidence interval; UCI, 95% upper confidence interval; and NR, not reached.

*Shrinkage coefficient, 0.91.

†Discrimination: c-index of the original model, 0.788; bias-corrected c-index, 0.776; optimism, 0.012.

‡Calibration: calibration slope of the original model, 1.000; bias-corrected calibration slope, 0.894; optimism, 0.106.

BIRC3-disrupted cases was compared with that of patients with *TP53* abnormalities. As shown by survival analysis, CLL patients harboring *BIRC3* disruption displayed an OS (median, 3.1 year) similar to that of CLL patients harboring *TP53* abnormalities (median, 2.9 years; $P = .543$; Figure 3B).

***BIRC3* is the sole noncanonical NF-κB pathway gene mutated in fludarabine-refractory CLL**

To determine the occurrence of genetic lesions of other members of the *BIRC3* protein complex,³³ we performed an extensive mutation and FISH analysis of the *TRAF2*, *TRAF3*, and *MAP3K14* genes in fludarabine-refractory CLL patients ($n = 49$). This analysis did not reveal alterations of *TRAF2*, *TRAF3*, or *MAP3K14*, documenting that *BIRC3* is the sole gene targeted by molecular lesions across all members of the *BIRC3* complex.

Discussion

The results of the present study document that *BIRC3* genetic lesions: (1) recurrently and selectively associate with fludarabine-refractory but *TP53* wild-type CLL; (2) at diagnosis, identify patients with a poor outcome similar to that associated with *TP53* abnormalities; (3) exert a prognostic role independent of widely accepted clinical and genetic risk factors; and (4) activate NF-κB signaling.

Fludarabine refractoriness in CLL may be explained by *TP53* disruption in approximately 40% of patients, whereas approximately 60% of high-risk CLL patients are devoid of *TP53* abnormalities.^{8,17} This observation prompted the search for other markers of fludarabine refractoriness. Although next-generation sequencing studies have allowed the identification of several previously unrecognized mutated genes in chemorefractory CLL, including *NOTCH1* and *SF3B1*,²⁰⁻²² at present, these novel mutations are not numerically sufficient to fully recapitulate the genetics of fludarabine-refractory CLL patients who are wild-type on the *TP53* gene.^{20,22}

The candidate gene approach used in this study revealed that genetic lesions of *BIRC3* contribute to clinical aggressiveness and chemorefractoriness in *TP53* wild-type CLL. *BIRC3* disruption selectively occurs in approximately 40% of fludarabine-refractory CLL patients harboring a wild-type *TP53*, whereas it is consistently absent in progressive CLL patients who require treatment but prove to be sensitive to fludarabine-based regimens. At diagnosis, *BIRC3* disruption is rare, but identifies a subgroup of high-risk patients displaying poor survival similar to that associated with *TP53* abnormalities. Confirmation within the frame of prospective clinical trials will be helpful to fully assess the generalization of *BIRC3* disruption as a CLL prognostic marker and as a tool for the early identification of chemorefractory cases.

BIRC3 genetic lesions were absent among clinical MBL cases with a CLL-like phenotype. This observation is consistent with the low frequency in clinical MBL of genetic lesions that are otherwise associated with high-risk CLL, namely *NOTCH1* mutations and *TP53* or *ATM* disruption, and corroborates the notion that clinical MBL is characterized by an indolent biologic phenotype.⁴⁵

BIRC3, along with *TRAF2* and *TRAF3*, cooperates in the same protein complex that negatively regulates *MAP3K14*, the central activator of noncanonical NF-κB signaling.³¹⁻³⁴ All inactivating mutations of *BIRC3* detected in CLL are predicted to cause the elimination or truncation of the C-terminal RING domain, the E3 ubiquitin ligase activity of which is essential for *MAP3K14* proteasomal degradation by *BIRC3*.³¹⁻³⁴ Consistently, fludarabine-refractory CLL patients harboring *BIRC3* disruption by either mutation and/or deletion display constitutive NF-κB activation.

NF-κB activation is known to provide prosurvival signals to leukemic cells through the up-regulation of several antiapoptotic genes, including members of the *BCL-2* family, and to be correlated with both survival and enhanced fludarabine resistance of CLL cells.²⁴⁻²⁸ In CLL, NF-κB activation is generally viewed as a consequence of specific interactions between leukemic cells and the microenvironment.²⁴⁻²⁸ The present study expands the mechanisms that may sustain NF-κB activation in CLL. In fact, our data suggest that, at least in a fraction of high-risk CLL patients,

leukemic cells might become independent from microenvironmental interactions and may gain active NF- κ B signaling through the acquisition of functionally active mutations targeting *BIRC3*. Genome-wide studies have pointed to the potential role of other NF- κ B gene alterations, namely mutations of *MYD88*, in CLL development.^{20,21} On these bases, the poor prognosis and refractory phenotype associated with *BIRC3* abnormalities might be a consequence of the constitutive NF- κ B activation specifically observed among fludarabine-refractory CLL patients harboring *BIRC3* disruption. NF- κ B inhibitors are under development in CLL, and preclinical findings suggest that these compounds might be active against chemoresistant CLL clones.⁴⁶⁻⁴⁹ On these bases, our data might provide a molecular rationale for targeting NF- κ B in poor-risk refractory CLL patients.

Identification of *BIRC3* involvement in CLL may be important for elucidating the molecular genetics of 11q22-q23 deletion. In fact, although *ATM* has been regarded as the relevant gene of this chromosomal abnormality,^{9,12} biallelic inactivation of *ATM* does not exceed 30% in patients with 11q22-q23 deletion.¹³⁻¹⁶ On these bases, a second tumor suppressor in the 11q22-q23 region has been postulated along with *ATM*.⁵⁰ *BIRC3*, which maps on 11q22.2 approximately 6 Mb centromeric to the *ATM* gene, might represent an attractive candidate, because: (1) the *BIRC3* locus is included in focal deletions that affect the 11q22.2 band but spare *ATM*; (2) a fraction of 11q22-q23 deletions truncate *BIRC3* and, similarly to *BIRC3*-disrupting mutations, remove its terminal exons encoding the RING domain; (3) analogous to *BIRC3* mutations, *BIRC3* deletions also activate NF- κ B signaling; and (4) the poor prognosis marked by *BIRC3* disruption is independent of *ATM* deletion in multivariate analysis.

The identification of *BIRC3* disruption elucidates the molecular basis of a fraction of high-risk CLL cases and expands knowledge of the molecular mechanisms activating NF- κ B in this leukemia. In addition to its pathogenetic relevance, *BIRC3* disruption might represent a molecular marker for the early identification of chemorefractoriness among CLL patients who are wild-type on *TP53*, and may provide the rationale for targeted anti-NF- κ B therapeutic approaches in this unfavorable clinical setting.

Acknowledgments

This study was supported by the Italian Association for Cancer Research Special Program in Molecular Clinical Oncology (5 × 1000, no. 10007, Milan, Italy, to G.G. and R.F.); Futuro in Ricerca 2008 (to D.R. and S.D.); Programma di ricerca di Rilevante Interesse Nazionale (PRIN) 2008 (to G.G. and R.M.) and PRIN 2009 (to D.R.); Ministero dell'Istruzione, dell'Università della Ricerca, Rome, Italy; Progetto Giovani Ricercatori 2008 (to D.R. and S.D.); Ricerca Sanitaria Finalizzata 2008 (to G.G.); Ministero della Salute, Rome, Italy; Novara-Associazione Italiana contro le Leucemie, linfomi e i mieloma (AIL) Onlus, Novara, Italy (to G.G. and D.R.); Compagnia di San Paolo, Turin, Italy (to R.F.); and the Helmut Horten Foundation and the San Salvatore Foundation (to F.B.). S.M. and S.C. are supported by fellowships from Novara-AIL Onlus, Novara. L.P. is on leave from the University of Perugia Medical School, Perugia, Italy.

Authorship

Contribution: D.R., R. Foà, and G.G. designed the study, interpreted the data, and wrote the manuscript; D.R. and M.F. performed the statistical analysis; M.F., S.R., S. Cresta, A.B., V.S., R. Famà, and M.G. performed the molecular analysis; S.M., C.D., and G.D. performed the FISH analysis; T.V. and S.D. performed and interpreted the biochemical assays; F.B. provided the SNP array data; S. Chiaretti, I.D.G., M.M., and F.F. provided well-characterized biological samples and clinical data; and G.F., V.G., L.P., A.G., and R.D.-F. contributed to data interpretation.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Davide Rossi, MD, PhD, Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy; e-mail: rossidav@med.unipmn.it.

References

- Müller-Hermelink HK, Montserrat E, Catovsky D, Campo E, Harris NL, Stein H. Chronic lymphocytic leukemia/small lymphocytic lymphoma. Swerdlow S, Campo E, Harris NL, eds; International Agency for Research on Cancer. In: *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissue*. Geneva, Switzerland: World Health Organization; 2008:180-182.
- Hallek M, Cheson BD, Catovsky D, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*. 2008; 111(12):5446-5456.
- Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med*. 2005;352(8):804-815.
- Rossi D, Gaidano G. Richter syndrome: molecular insights and clinical perspectives. *Hematol Oncol*. 2009;27(1):1-10.
- Rossi D, Spina V, Cerri M, et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clin Cancer Res*. 2009;15(13):4415-4422.
- Rossi D, Spina V, Deambrogi C, et al. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood*. 2011;117(12):3391-3401.
- Cramer P, Hallek M. Prognostic factors in chronic lymphocytic leukemia-what do we need to know? *Nat Rev Clin Oncol*. 2011;8(1):38-47.
- Stilgenbauer S, Zenz T. Understanding and managing ultra high-risk chronic lymphocytic leukemia. *Hematology Am Soc Hematol Educ Program*. 2010;2010:481-488.
- Döhner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000;343(26):1910-1916.
- Zenz T, Kröber A, Scherer K, et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood*. 2008;112(8):3322-3329.
- Rossi D, Cerri M, Deambrogi C, et al. The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. *Clin Cancer Res*. 2009;15(3):995-1004.
- Stilgenbauer S, Liebis P, James MR, et al. Molecular cytogenetic delineation of a novel critical genomic region in chromosome bands 11q22.3-923.1 in lymphoproliferative disorders. *Proc Natl Acad Sci U S A*. 1996;93(21):11837-41.
- Schaffner C, Stilgenbauer S, Rappold GA, Döhner H, Lichter P. Somatic ATM mutations indicate a pathogenic role of ATM in B-cell chronic lymphocytic leukemia. *Blood*. 1999;94(2):748-753.
- Austen B, Powell JE, Alvi A, et al. Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. *Blood*. 2005; 106(9):3175-3182.
- Guarini A, Marinelli M, Tavoraro S, et al. ATM gene alterations in chronic lymphocytic leukemia patients induce a distinct gene expression profile and predict disease progression. *Haematologica*. 2012;97(1):47-55.
- Skowronska A, Austen B, Powell JE, et al. ATM germline heterozygosity does not play a role in CLL initiation but influences rapid disease progression through loss of the remaining ATM allele. *Haematologica*. 2012;97(1):142-146.
- Zenz T, Häbe S, Denzel T, et al. Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. *Blood*. 2009;114(13):2589-2597.

18. Zenz T, Eichhorst B, Busch R, et al. TP53 mutation and survival in chronic lymphocytic leukemia. *J Clin Oncol*. 2010;28(29):4473-4479.
19. Gonzalez D, Martinez P, Wade R, et al. Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. *J Clin Oncol*. 2011;29(16):2223-2229.
20. Fabbri G, Rasi S, Rossi D, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med*. 2011;208(7):1389-1401.
21. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475(7354):101-105.
22. Rossi D, Bruscaggini A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood*. 2011;118(26):6904-6908.
23. Rossi D, Rasi S, Fabbri G, et al. Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. *Blood*. 2012;119(2):521-529.
24. Kern C, Cornuel JF, Billard C, et al. Involvement of BAFF and APRIL in the resistance to apoptosis of B-CLL through an autocrine pathway. *Blood*. 2004;103(2):679-688.
25. Endo T, Nishio M,ENZLER T, et al. BAFF and APRIL support chronic lymphocytic leukemia B-cell survival through activation of the canonical NF-kappaB pathway. *Blood*. 2007;109(2):703-710.
26. Hewamana S, Alghazal S, Lin TT, et al. The NF-kappaB subunit Rel A is associated with in vitro survival and clinical disease progression in chronic lymphocytic leukemia and represents a promising therapeutic target. *Blood*. 2008;111(9):4681-4689.
27. Buggins AG, Pepper C, Patten PE, et al. Interaction with vascular endothelium enhances survival in primary chronic lymphocytic leukemia cells via NF-kappaB activation and de novo gene transcription. *Cancer Res*. 2010;70(19):7523-7533.
28. Herishanu Y, Pérez-Galán P, Liu D, et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood*. 2011;117(2):563-574.
29. Compagno M, Lim WK, Grunn A, et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature*. 2009;459(7247):717-721.
30. Staudt LM. Oncogenic activation of NF-kappaB. *Cold Spring Harb Perspect Biol*. 2010;2(6):a000109.
31. Vallabhapurapu S, Karin M. Regulation and function of NF-kappaB transcription factors in the immune system. *Annu Rev Immunol*. 2009;27:693-733.
32. Li X, Yang Y, Ashwell JD. TNF-R1I and c-IAP1 mediate ubiquitination and degradation of TRAF2. *Nature*. 2002;416(6878):345-347.
33. Zarnegar BJ, Wang Y, Mahoney DJ, et al. Noncanonical NF-kappaB activation requires coordinated assembly of a regulatory complex of the adaptors cIAP1, cIAP2, TRAF2 and TRAF3 and the kinase NIK. *Nat Immunol*. 2008;9(12):1371-1378.
34. Conze DB, Zhao Y, Ashwell JD. Non-canonical NF-kB activation and abnormal B cell accumulation in mice expressing ubiquitin protein ligase-inactive c-IAP2. *PLoS Biol*. 2010;8(10):e1000518.
35. Rossi D, Deaglio S, Dominguez-Sola D, et al. Alteration of BIRC3 and multiple other NF-(kappa)B pathway genes in splenic marginal zone lymphoma. *Blood*. 2011;118(18):4930-4934.
36. Rasi S, Bruscaggini A, Rinaldi A, et al. Saliva is a reliable and practical source of germline DNA for genome-wide studies in chronic lymphocytic leukemia. *Leuk Res*. 2011;35(10):1419-1422.
37. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*. 1999;94(6):1848-1854.
38. Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 1999;94(6):1840-1847.
39. Rinaldi A, Mian M, Kwee I, et al. Genome-wide DNA profiling better defines the prognosis of chronic lymphocytic leukaemia. *Br J Haematol*. 2011;154(5):590-599.
40. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Am Stat Assoc*. 1958;53:457-481.
41. Cox DR. Regression models and life tables. *J R Stat Assoc*. 1972;34:187-220.
42. Harrell FE Jr. *Regression Modeling Strategies*. New York, NY: Springer-Verlag; 2001:465-522.
43. Schoenfeld D. Partial residuals for the proportional hazard regression model. *Biometrika*. 1982;69(1):239-241.
44. Chen CH, George SL. The bootstrap and identification of prognostic factors via Cox's proportional hazards regression model. *Stat Med*. 1995;14(1):39-46.
45. Rasi S, Monti S, Spina V, Foa R, Gaidano G, Rossi D. Analysis of NOTCH1 mutations in monoclonal B cell lymphocytosis. *Haematologica*. 2012;97(1):153-154.
46. Hewamana S, Lin TT, Jenkins C, et al. The novel nuclear factor-kappaB inhibitor LC-1 is equipotent in poor prognostic subsets of chronic lymphocytic leukemia and shows strong synergy with fludarabine. *Clin Cancer Res*. 2008;14(24):8102-8111.
47. Hertlein E, Wagner AJ, Jones J, et al. 17-DMAG targets the nuclear factor-kappaB family of proteins to induce apoptosis in chronic lymphocytic leukemia: clinical implications of HSP90 inhibition. *Blood*. 2010;116(1):45-53.
48. Hertlein E, Byrd JC. Signalling to drug resistance in CLL. *Best Pract Res Clin Haematol*. 2010;23(1):121-131.
49. Herman SE, Gordon AL, Hertlein E, et al. Bruton tyrosine kinase represents a promising therapeutic target for treatment of chronic lymphocytic leukemia and is effectively targeted by PCI-32765. *Blood*. 2011;117(23):6287-6296.
50. Kalla C, Scheuermann MO, Kube I, et al. Analysis of 11q22-q23 deletion target genes in B-cell chronic lymphocytic leukaemia: evidence for a pathogenic role of NPAT, CUL5, and PPP2R1B. *Eur J Cancer*. 2007;43(8):1328-1335.

Letter to the Editor

Analysis of *SF3B1* mutations in monoclonal B-cell lymphocytosis

To the Editor

Clinical monoclonal B-cell lymphocytosis (cMBL) represents an asymptomatic monoclonal B-cell expansion with less than $5.0 \times 10^9/L$ circulating cells that is occasionally identified during a diagnostic workup of an otherwise asymptomatic lymphocytosis [1–5]. Most cMBLs are characterized by a chronic lymphocytic leukaemia (CLL) phenotype and might represent the initial phases of CLL [2–5]. At the molecular level, cMBL with a CLL phenotype carry CLL-related abnormalities in ~50% cases, which are mainly represented by the 13q14 deletion [2–4]. Consistent with the indolent phenotype of this condition, genetic lesions predicting poor prognosis in CLL, such as disruption of *TP53* and *ATM* as well as mutations of *NOTCH1*, are rare in cMBL with a CLL phenotype [2–4,6].

Recent investigations of the CLL coding genome have revealed recurrent mutations of the *SF3B1* splicing gene [7–9]. In CLL, *SF3B1* mutations occur at a low rate (~5%) at presentation, where they identify poor survival patients, while they are recurrent (~17%) in more advanced phases of the disease, thus suggesting a relationship between *SF3B1* mutations and CLL aggressiveness [7–9].

Here, we investigated the occurrence of *SF3B1* mutations in 63 consecutive cMBL with a CLL phenotype presenting at our clinic for the initial evaluation of an asymptomatic lymphocytosis. Peripheral blood mononuclear cell samples drawn at cMBL presentation were analysed for *SF3B1*, *NOTCH1*, *TP53* and *IGHV* mutations by DNA Sanger sequencing, and for 13q14 deletion, +12, 11q22–q23 deletion and 17p13 deletion by FISH [3,6,8]. An *SF3B1* mutation (K700E) that is known to be highly recurrent in CLL was also independently investigated by amplification refractory mutation system (ARMS) PCR [7–9]. Patients provided informed consent in accordance with local Institutional Review Board requirements and the Declaration of Helsinki. The study was approved by the Ethical Committee of the Ospedale Maggiore della Carità di Novara associated with the Amedeo Avogadro University of Eastern Piedmont (Protocol Code 59/CE; Study Number CE 8/11).

The clinical profile of the cMBL cohort was representative of this condition [2–4]. Median age was 68 years (range: 40–88 years). The male : female ratio was 33:30. The median absolute lymphocyte count and median B-cell count were $5.0 \times 10^9/L$ (range: $1.7\text{--}7.6 \times 10^9/L$) and $2.9 \times 10^9/L$ (range: $0.2\text{--}4.9 \times 10^9/L$), respectively. The median value of Hb was 14.1 g/dL (range: 11.0–16.9 g/dL) and of platelets $220 \times 10^9/L$ (range: 148–420). The biological profile

of cMBL was consistent with the indolent nature of this condition. Expression on CD38 and ZAP70 occurred in 8/60 (13.3%) and in 9/53 (17.0%) cases, respectively. An *IGHV* identity $\geq 98\%$ was present in 9/59 (15.3%) cMBLs and stereotyped VH CDR3 in 9/59 (15.3%). Deletion of 13q14 was observed in 29/63 (46.0%) cases, trisomy 12 in 0/63 and 11q22–q23 deletion in 2/63 (3.2%). One of 63 (1.5%) cMBLs harboured a biallelic *TP53* disruption by the deletion of one allele and mutation of the second allele. *NOTCH1* mutations were present in 2/63 (3.2%) cMBL [6].

One single *SF3B1* mutation occurred in 1/63 (1.5%) cMBL and was represented by a missense mutation of somatic origin leading to the K666N substitution (Figure 1). The cMBL case harbouring this *SF3B1* mutation presented with an absolute lymphocyte count of $4.7 \times 10^9/L$ and a B-cell count of $2.3 \times 10^9/L$, was CD38 and ZAP70 negative, expressed a mutated *IGHV3-48* gene (*IGHV* identity 95.15%), lacked a stereotyped VH CDR3, showed a normal FISH karyotype and lacked mutations of *TP53* and *NOTCH1*. It progressed to CLL (absolute lymphocyte count of $16.3 \times 10^9/L$ and a B-cell count of $10.2 \times 10^9/L$, Rai stage 0) after 44 months of follow-up. Analysis of a sequential sample collected at cMBL progression to CLL confirmed the *SF3B1* K666N substitution [8].

The sensitivity of DNA Sanger sequencing does not allow the identification of a mutation whose allelic representation is $<10\%$. Because the peripheral blood monoclonal B-cell content in cMBL is lower than in CLL, we specifically designed a high sensitivity ARMS PCR assay [10] for detecting the *SF3B1* K700E mutation, that accounts for ~60% *SF3B1* substitutions in CLL [7–9]. Despite the high sensitivity (1%) of this assay, ARMS did not identify additional K700E mutations in cMBL.

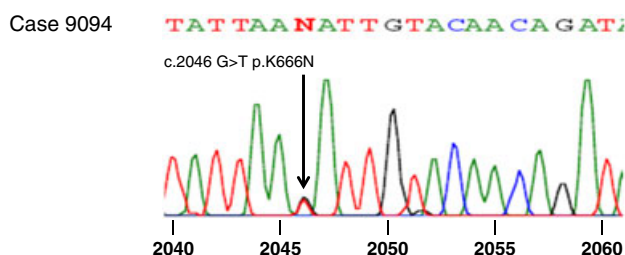


Figure 1. *SF3B1* mutation in a case of clinical monoclonal B-cell lymphocytosis (cMBL). Sequencing traces of the cMBL tumour sample 9094 harbouring the *SF3B1* c.2046G > T p.K666N (RefSeq NM_012433.2) mutation; the arrow points to the position of the nucleotide change

The low prevalence of *SF3B1* mutations observed in cMBL is consistent with the rarity in this indolent condition of other genetic lesions otherwise associated with high-risk CLL.

Conflict of interest

The authors have no conflict of interest to disclose.

Acknowledgements

This study was supported by AIRC, Special Program Molecular Clinical Oncology, 5 × 1000, No. 10007, Milan, Italy (to G.G. and to R.F.); Progetto FIRB-Programma 'Futuro in Ricerca' 2008 (to D.R.), PRIN 2008 (to G.G.) and 2009 (to D.R.), MIUR, Rome, Italy; Progetto Giovani Ricercatori 2008 (to D.R.) and Ricerca Sanitaria Finalizzata 2008 (to G.G.) and 2009 (to D.C.), Ministero della Salute, Rome, Italy; Ricerca Sanitaria Finalizzata 2009 (to D.C.), Regione Piemonte, Turin, Italy; Novara-AIL Onlus, Novara, Italy (to G.G. and D.R.); Compagnia di San Paolo, Turin, Italy (to R.F.). S.M. is being supported by fellowships from Novara-AIL Onlus, Novara. S.R. is being supported by a fellowship from Società Italiana di Ematologia Sperimentale.

M.G., A.B., V.S., S.R., C.C., S.C. and M.F. performed molecular studies and interpreted data; D.C. contributed to study design and data interpretation; S.M. performed and interpreted FISH analysis; G.G., R.F. and D.R. designed the study and wrote the manuscript.

References

1. Hallek M, Cheson BD, Catovsky D, *et al.* Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* 2008; **111**: 5446–5456.
2. Rawstron AC, Bennett FL, O'Connor SJ, *et al.* Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med* 2008; **359**: 575–583.
3. Rossi D, Sozzi E, Puma A, *et al.* The prognosis of clinical monoclonal B cell lymphocytosis differs from prognosis of Rai 0 chronic lymphocytic leukaemia and is recapitulated by biological risk factors. *Br J Haematol* 2009; **146**: 64–75.
4. Shanafelt TD, Kay NE, Rabe KG, *et al.* Brief report: natural history of individuals with clinically recognized monoclonal B-cell lymphocytosis compared with patients with Rai 0 chronic lymphocytic leukemia. *J Clin Oncol* 2009; **27**: 3959–3963.
5. Shanafelt TD, Ghia P, Lanasa MC, Landgren O, Rawstron AC. Monoclonal B-cell lymphocytosis (MBL): biology, natural history and clinical management. *Leukemia* 2010; **24**: 512–520.
6. Rasi S, Monti S, Spina V, Foà R, Gaidano G, Rossi D. Analysis of NOTCH1 mutations in monoclonal B cell lymphocytosis. *Haematologica* 2012; **97**: 153–154.
7. Quesada V, Conde L, Villamor N, *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* 2011; **44**: 47–52.
8. Rossi D, Bruscaggin A, Spina V, *et al.* Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* 2011; **118**: 6904–6908.
9. Wang L, Lawrence MS, Wan Y, *et al.* SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med* 2011; **365**: 2497–2506.
10. Baxter EJ, Scott LM, Campbell PJ, *et al.* Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* 2005; **365**: 1054–1061.

Mariangela Greco¹
 Daniela Capello¹
 Alessio Bruscaggin¹
 Valeria Spina¹
 Silvia Rasi¹
 Sara Monti¹
 Carmela Ciardullo¹
 Stefania Cresta¹
 Marco Fangazio¹
 Gianluca Gaidano¹
 Robin Foà²
 Davide Rossi^{1*}

¹Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, 28100 Novara, Italy

²Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy

*rossidav@med.unipmn.it

Different impact of *NOTCH1* and *SF3B1* mutations on the risk of chronic lymphocytic leukemia transformation to Richter syndrome

Richter syndrome (RS) represents the development of an aggressive lymphoma, most commonly diffuse large B-cell lymphoma (DLBCL), in the context of chronic lymphocytic leukaemia (CLL). At least two types of RS exist: (i) transformation of CLL into a clonally related DLBCL, that accounts for ~80% of cases; and (ii) development of a DLBCL unrelated to the CLL clone. Clonally related RS and clonally unrelated RS are distinct disorders (Rossi *et al*, 2011a). Clinically, transformation into a clonally related RS is frequently lethal with an expected survival of a few months, while CLL patients developing a clonally unrelated RS display a survival probability in the range of *de novo* DLBCL (Rossi *et al*, 2011a). Biologically, clonally related RS frequently acquire genetic lesions of *TP53*, *MYC* and *NOTCH1*, which are otherwise absent or exceptional in clonally unrelated RS (Rossi *et al*, 2011a).

NOTCH1 and *SF3B1* mutations are predictors of poor outcome in CLL (Puente *et al*, 2011; Rossi *et al*, 2011b, 2012; Wang *et al*, 2011; Quesada *et al*, 2012). Here, we investigated a large multi-institutional series of 605 CLL patients to verify whether *NOTCH1* and *SF3B1* mutations at diagnosis may help in the identification of CLL at high risk of RS transformation.

The study was approved by the Ethical Committee of the Ospedale Maggiore della Carità di Novara associated with the Amedeo Avogadro University of Eastern Piedmont (Protocol Code 59/CE; Study Number CE 8/11). Patients provided informed consent in accordance with local Institutional Review Board requirements and the Declaration of Helsinki. CLL was diagnosed according to National Cancer Institutes-International Workshop on CLL criteria (Hallek *et al*, 2008). RS diagnosis was histologically-proven (Müller-Hermelink *et al*, 2008). *NOTCH1*, *SF3B1*, *TP53* and *IGHV* mutations were analysed by Sanger sequencing (Rossi *et al*, 2012). Probes (Abbott) used for fluorescent *in situ* hybridization analysis were: LSI13 and LSID13S319, CEP12, LSIp53 and LSIATM (Rossi *et al*, 2011a). CD38 expression was analysed by flow cytometry, utilizing a cut-off of 30% to define positivity (Rossi *et al*, 2009). VH CDR3 subset analysis and nomenclature were according to Murray *et al* (2008). Clonal relationship between CLL/DLBCL paired samples was established by comparing *IGHV-D-J* sequences. Time-to-event analysis utilized the Kaplan–Meier method. The cumulative incidence of RS development was calculated accounting

for death as a competing risk, and was compared across groups with the Gray's test. The association between exposure variables and outcome was estimated by univariate and multivariate Cox regression analysis. Further details are reported in Supplementary Methods.

Features of the CLL cohort are reported in Table SI. At CLL presentation, *NOTCH1* mutations occurred in 12.2% patients, being mostly represented (82.4%) by a recurrent two bp frameshift deletion (c.7544_7545delCT). *SF3B1* mutations occurred in 5.8% cases and were mainly represented (50.0%) by the K700E missense substitution.

After a median follow-up of 6.4 years, 30/605 (4.9%) CLL had transformed into a clonally related RS, while 10/605 (1.6%) had developed a clonally unrelated DLBCL. *NOTCH1* mutations at disease presentation recurred in 12/30 (40.0%) CLL that subsequently transformed into a clonally related RS, whereas they were consistently absent in the 10 CLL that developed a clonally unrelated RS. *SF3B1* mutations were absent in all CLL that subsequently transformed to a RS, both clonally related or unrelated.

The close relationship between *NOTCH1* mutations and clonally related RS was reinforced by the following observations: (i) *NOTCH1* mutations were the second most frequent genetic lesion (40.0%) at transformation after *TP53* disruption (60.0%); and (ii) *NOTCH1* mutations observed in the RS phase were already present at CLL presentation in all patients, either at the clonal ($n = 5$) or subclonal ($n = 7$) level, suggesting a preferential selection of *NOTCH1* mutated clones during disease progression and transformation. Notably, none of the clonally related RS acquired *SF3B1* mutations at transformation.

NOTCH1 mutations were associated with a ~5.8-fold increase in the crude hazard of transformation into a clonally related RS (HR: 5.81; $P < 0.001$) (Table I). Consistently, CLL harbouring *NOTCH1* mutations had a significantly higher cumulative probability of transforming into a clonally related RS (45.0% at 15 years) compared to CLL without *NOTCH1* mutations (4.6% at 15 years) ($P < 0.001$) (Fig. 1A). The association between *NOTCH1* mutations and transformation into a clonally related RS also remained significant after correcting for death as a competing risk ($P < 0.001$). Conversely, *NOTCH1* mutations had no impact on the development of a clonally unrelated DLBCL ($P = 0.494$).

Table I. Univariate and multivariate analysis for CLL transformation into clonally related RS*

Characteristics	Univariate analysis				Multivariate analysis			
	HR	LCI	UCI	P	HR	LCI	UCI	P
Age ≤ 65 years	—	—	—					
Age > 65 years	0.91	0.44	1.90	0.811				
Rai stage 0–II	—	—	—					
Rai stage III–IV	2.14	0.82	5.61	0.119				
IGHV identity <98%	—	—	—					
IGHV identity ≥ 98%	6.20	2.61	14.7	<0.001				
No stereotyped VH CDR3	—	—	—					
Stereotyped VH CDR3	3.76	1.85	7.78	<0.001				
No IGHV4-39 usage	—	—	—					
IGHV4-39 usage	4.63	2.13	10.04	<0.001	3.14	1.01	9.74	0.047
No + 12/del11q22-q23/TP53 disruption	—	—	—					
+12/del11q22-q23/TP53 disruption	4.32	2.05	9.09	<0.001	3.26	1.59	8.66	0.002
CD38 expression <30%	—	—	—					
CD38 expression ≥ 30%	4.63	2.12	10.04	<0.001	2.02	0.83	4.91	0.117
NOTCH1 wild type	—	—	—					
NOTCH1 mutated	5.81	2.78	12.12	<0.001	3.71	1.59	8.66	0.002

*HR, hazard ratio; LCI, 95% lower confidence interval; UCI, 95% upper confidence interval; IGHV, immunoglobulin heavy variable gene; VH CDR3, immunoglobulin heavy chain complementarity-determining region 3.

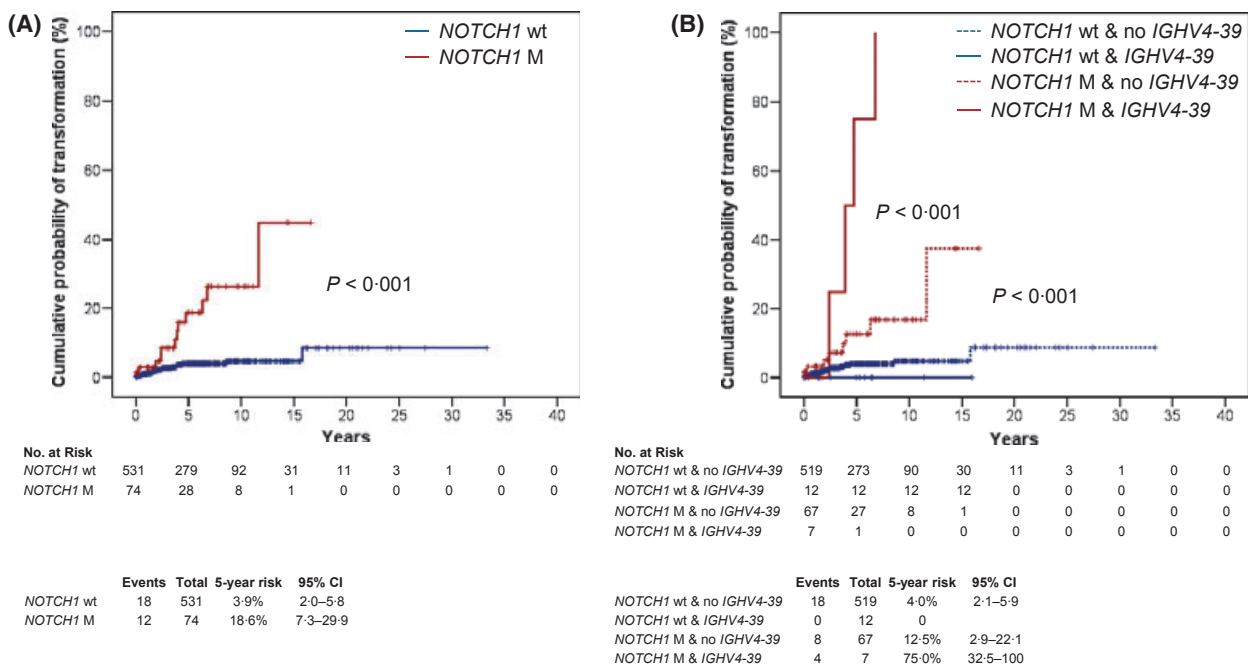


Fig 1. Kaplan-Meier estimates of the cumulative probability of transformation into a clonally related RS. Panel A: cumulative probability of transformation into a clonally related RS according to *NOTCH1* mutation status in the CLL cohort ($n = 605$). *NOTCH1* wild type cases (*NOTCH1* wt) are represented by the blue line. *NOTCH1* mutated cases (*NOTCH1* M) are represented by the red line. Panel B: cumulative probability of transformation into a clonally related RS according to the *NOTCH1* mutation type and *IGHV4-39* usage in the CLL cohort ($n = 605$). *NOTCH1* wild type cases utilizing *IGHV* genes other than *IGHV4-39* (*NOTCH1* wt & no *IGHV4-39*) are represented by the blue dashed line. *NOTCH1* wild type cases utilizing the *IGHV4-39* gene (*NOTCH1* wt & *IGHV4-39*) are represented by the blue solid line. *NOTCH1* mutated cases utilizing *IGHV* genes other than *IGHV4-39* (*NOTCH1* M & no *IGHV4-39*) are represented by the red dashed line. *NOTCH1* mutated cases utilizing the *IGHV4-39* gene (*NOTCH1* M & *IGHV4-39*) are represented by the red solid line.

NOTCH1 mutations were significantly enriched among cases harbouring other factors known to affect RS risk, including CD38 expression (27.6%, $P < 0.001$) and *IGHV4-39* gene

usage (36.8%, $P = 0.005$), especially if rearranged in a stereotyped fashion in subset 8 (66.6%; $P < 0.001$) (Rossi *et al*, 2009).

By multivariate analysis, *NOTCH1* mutations were selected as an independent risk factor of transformation into a clonally related RS (HR: 3.71; $P = 0.002$), after adjusting for trisomy 12 and/or 11q22-q23 deletion and/or *TP53* disruption, *IGHV4-39* usage, and CD38 expression (Table I), which, at CLL diagnosis, are well established risk factors for subsequent RS development (Rossi *et al*, 2009).

The observation of a high frequency of *NOTCH1* mutations in *IGHV4-39* positive CLL prompted investigation of the interaction between *IGHV4-39* usage and *NOTCH1* mutations in the model for clonally RS prediction. Accordingly, bivariate analysis segregated three risk categories (Fig. 1B): (i) CLL harbouring both *NOTCH1* mutations and *IGHV4-39*, which constituted the disease category with the highest risk of transformation; (ii) CLL harbouring *NOTCH1* mutations but lacking *IGHV4-39*, which showed an intermediate risk of transformation; and (iii) CLL lacking *NOTCH1* mutations, independent of *IGHV4-39* usage, which showed a very low risk of transformation.

SF3B1 mutations, a new biomarker of CLL progression (Rossi *et al*, 2011b; Wang *et al*, 2011; Quesada *et al*, 2012), showed no impact on RS development ($P = 0.616$) (Table I). This observation further reinforces the notion that predictors of CLL transformation are different, at least in part, from predictors of other CLL clinical outcomes (Rossi *et al*, 2009).

This study documents that *NOTCH1* mutations are an independent prognosticator of CLL transformation into clonally related RS, whereas they have no impact on the development of clonally unrelated RS. Importantly, up to 40–50% of *NOTCH1* mutated patients are ultimately projected to develop and die of clonally related RS. On these bases, the poor prognosis associated with *NOTCH1* mutations in CLL may be explained, at least in part, by a substantial risk of developing RS.

Acknowledgements

This study was supported by AIRC, Special Program Molecular Clinical Oncology, 5 × 1000, No. 10007, Milan, Italy (to G.G. and to R.Foà.); Progetto FIRB-Programma “Futuro in Ricerca” 2008 (to D.R.), PRIN 2008 (to G.G.) and 2009 (to D.R.), MIUR, Rome, Italy; Progetto Giovani Ricercatori 2008 (to D.R.) and Ricerca Sanitaria Finalizzata 2008 (to G.G. and P.M.), Ministero della Salute, Rome, Italy; Novara-AIL Onlus, Novara, Italy (to G.G.); Compagnia di San Paolo, Turin, Italy (to R.Foà.). S.M. is supported by fellowships from Novara-AIL Onlus, Novara.

Contribution

D.R. and G.G. designed the study, interpreted data and wrote the manuscript; S.R., V.S., M.G., C.C., R.Famà, S.C. and A.B. performed molecular studies and interpreted data; S.M. performed FISH analysis and interpreted data; D.R. and M.F.

performed statistical analysis L.L., M.M., F.F. and R.M. provided well characterized biological samples and clinical data; P.M., L.M.L., and R.Foà. contributed to data analysis and interpretation.

Conflict of interest

The authors have no conflict of interest to disclose.

Davide Rossi¹
Silvia Rasi¹
Valeria Spina¹
Marco Fangazio¹
Sara Monti¹
Mariangela Greco¹
Carmela Ciardullo¹
Rosella Famà¹
Stefania Cresta¹
Alessio Bruscaggin¹
Luca Laurenti²
Maurizio Martini³
Pellegrino Musto⁴
Francesco Forconi^{5,6}
Roberto Marasca⁷
Luigi M. Larocca³
Robin Foà^{8,9}
Gianluca Gaidano¹

¹Division of Haematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, ²Institute of Haematology, Catholic University of the Sacred Heart, Rome, ³Pathology, Catholic University of the Sacred Heart, Rome, ⁴Onco-Haematology Department, Istituto di Ricovero e Cura a carattere Scientifico (IRCCS), Centro di Riferimento Oncologico della Basilicata, Rionero in Vulture, Italy, ⁵Cancer Sciences Unit, CRUK Clinical Centre, University of Southampton, Southampton, UK, ⁶Division of Haematology, University of Siena, Siena, ⁷Division of Haematology, Department of Oncology and Haematology, University of Modena and Reggio Emilia, Modena, ⁸Division of Haematology, Department of Cellular Biotechnologies and Haematology, Sapienza University, Rome, and ⁹Fondazione Eleonora Lorillard Spencer Cenci, Sapienza University, Rome, Italy
 E-mail: rossidav@med.unipmn.it

Keywords: Richter syndrome, chronic lymphocytic leukaemia, *NOTCH1*, *SF3B1*, mutations

First published online 10 May 2012
 doi: 10.1111/j.1365-2141.2012.09155.x

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Characteristics of the CLL cohort according to the *NOTCH1* mutation status

Data S1. Materials and methods

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied

by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

- Hallek, M., Cheson, B.D., Catovsky, D., Caligaris-Cappio, F., Dighiero, G., Döhner, H., Hillmen, P., Keating, M.J., Montserrat, E., Rai, K.R. & Kipps, T.J. (2008) Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*, **111**, 5446–5456.
- Müller-Hermelink, H.K., Montserrat, E., Catovsky, D., Campo, E., Harris, N.L. & Stein, H. (2008) Chronic lymphocytic leukemia/small lymphocytic lymphoma. In: *World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues* (ed. by S.H. Swerdlow, E. Campo, N.L. Harris, E.S. Jaffe, S.A. Pileri, H. Stein, J. Thiele & J.W. Vardiman), pp. 180–182. IARC, Lyon, France.
- Murray, F., Darzentas, N., Hadzidimitriou, A., Tobin, G., Boudjogra, M., Scielzo, C., Laoutaris, N., Karlsson, K., Baran-Marzszak, F., Tsaftaris, A., Moreno, C., Anagnostopoulos, A., Caligaris-Cappio, F., Vaur, D., Ouzounis, C., Belessi, C., Ghia, P., Davi, F., Rosenquist, R. & Stamatopoulos, K. (2008) Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood*, **111**, 1524–1533.
- Puente, X.S., Pinyol, M., Quesada, V., Conde, L., Ordóñez, G.R., Villamor, N., Escaramis, G., Jares, P., Beà, S., González-Díaz, M., Bassaganyas, L., Baumann, T., Juan, M., López-Guerra, M., Colomer, D., Tubío, J.M., López, C., Navarro, A., Tornador, C., Aymerich, M., Rozman, M., Hernández, J.M., Puente, D.A., Freije, J.M., Velasco, G., Gutiérrez-Fernández, A., Costa, D., Carrió, A., Guijarro, S., Enjuanes, A., Hernández, L., Yagüe, J., Nicolás, P., Romeo-Casabona, C.M., Himmelbauer, H., Castillo, E., Dohm, J. C., de Sanjosé, S., Piris, M.A., de Alava, E., San Miguel, J., Royo, R., Gelpi, J.L., Torrents, D., Orozco, M., Pisano, D.G., Valencia, A., Guigó, R., Bayés, M., Heath, S., Gut, M., Klatt, P., Marshall, J., Raine, K., Stebbings, L.A., Futreal, P.A., Stratton, M.R., Campbell, P.J., Gut, I., López-Guillermo, A., Estivill, X., Montserrat, E., López-Otín, C. & Campo, E. (2011) Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*, **475**, 101–105.
- Quesada, V., Conde, L., Villamor, N., Ordóñez, G. R., Jares, P., Bassaganyas, L., Ramsay, A.J., Beà, S., Pinyol, M., Martínez-Trillos, A., López-Guerra, M., Colomer, D., Navarro, A., Baumann, T., Aymerich, M., Rozman, M., Delgado, J., Giné, E., Hernández, J.M., González-Díaz, M., Puente, D.A., Velasco, G., Freije, J.M., Tubío, J.M., Royo, R., Gelpi, J.L., Orozco, M., Pisano, D.G., Zamora, J., Vázquez, M., Valencia, A., Himmelbauer, H., Bayés, M., Heath, S., Gut, M., Gut, I., Estivill, X., López-Guillermo, A., Puente, X.S., Campo, E. & López-Otín, C. (2012) Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature Genetics*, **44**, 47–52.
- Rossi, D., Spina, V., Cerri, M., Rasi, S., Deambroggi, C., De Paoli, L., Laurenti, L., Maffei, R., Forconi, F., Bertoni, F., Zucca, E., Agostinelli, C., Cabras, A., Lucioni, M., Martini, M., Magni, M., Deaglio, S., Ladetto, M., Nomdedeu, J.F., Besson, C., Ramponi, A., Canzonieri, V., Paulli, M., Marasca, R., Larocca, L.M., Carbone, A., Pileri, S.A., Gattei, V. & Gaidano, G. (2009) Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clinical Cancer Research*, **15**, 4415–4422.
- Rossi, D., Spina, V., Deambroggi, C., Rasi, S., Laurenti, L., Stamatopoulos, K., Arcaini, L., Lucioni, M., Rocque, G.B., Xu-Monette, Z.Y., Visco, C., Chang, J., Chigrinova, E., Forconi, F., Marasca, R., Besson, C., Papadaki, T., Paulli, M., Larocca, L.M., Pileri, S.A., Gattei, V., Bertoni, F., Foà, R., Young, K.H. & Gaidano, G. (2011a) The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood*, **117**, 3391–3401.
- Rossi, D., Bruscaggini, A., Spina, V., Rasi, S., Khiabanian, H., Messina, M., Fangazio, M., Vaisitti, T., Monti, S., Chiaretti, S., Guarini, A., Del Giudice, I., Cerri, M., Cresta, S., Deambroggi, C., Gargiulo, E., Gattei, V., Forconi, F., Bertoni, F., Deaglio, S., Rabadan, R., Pasqualucci, L., Foà, R., Dalla-Favera, R. & Gaidano, G. (2011b) Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood*, **118**, 6904–6908.
- Rossi, D., Rasi, S., Fabbri, G., Spina, V., Fangazio, M., Forconi, F., Marasca, R., Laurenti, L., Bruscaggini, A., Cerri, M., Monti, S., Cresta, S., Famà, R., De Paoli, L., Bulian, P., Gattei, V., Guarini, A., Deaglio, S., Capello, D., Rabadan, R., Pasqualucci, L., Dalla-Favera, R., Foà, R. & Gaidano, G. (2012) Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. *Blood*, **119**, 521–529.
- Wang, L., Lawrence, M.S., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., De Luca, D.S., Zhang, L., Zhang, W., Vartanov, A.R., Fernandes, S.M., Goldstein, N.R., Folco, E.G., Cibulskis, K., Tesar, B., Sievers, Q.L., Shefler, E., Gabriel, S., Hacohen, N., Reed, R., Meyerson, M., Golub, T.R., Lander, E. S., Neuberger, D., Brown, J.R., Getz, G. & Wu, C. J. (2011) SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *The New England Journal of Medicine*, **365**, 2497–2506.

The coding genome of splenic marginal zone lymphoma: activation of *NOTCH2* and other pathways regulating marginal zone development

Davide Rossi,¹ Vladimir Trifonov,² Marco Fangazio,^{1,3} Alessio Bruscaggini,¹ Silvia Rasi,¹ Valeria Spina,¹ Sara Monti,¹ Tiziana Vaisitti,⁴ Francesca Arruga,⁴ Rosella Famà,¹ Carmela Ciardullo,¹ Mariangela Greco,¹ Stefania Cresta,¹ Daniela Piranda,¹ Antony Holmes,³ Giulia Fabbri,³ Monica Messina,³ Andrea Rinaldi,⁵ Jiguang Wang,² Claudio Agostinelli,⁶ Pier Paolo Piccaluga,⁶ Marco Lucioni,⁷ Fabrizio Tabbò,⁸ Roberto Serra,⁹ Silvia Franceschetti,¹ Clara Deambrogi,¹ Giulia Daniele,¹⁰ Valter Gattei,¹¹ Roberto Marasca,¹² Fabio Facchetti,¹³ Luca Arcaini,¹⁴ Giorgio Inghirami,⁸ Francesco Bertoni,⁵ Stefano A. Pileri,⁶ Silvia Deaglio,⁴ Robin Foà,¹⁵ Riccardo Dalla-Favera,^{3,16,17} Laura Pasqualucci,^{3,16,18} Raul Rabadan,² and Gianluca Gaidano¹

¹Division of Hematology and ⁹Laboratory of Medical Informatics, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, 28100 Novara, Italy

²Department of Biomedical Informatics and Center for Computational Biology and Bioinformatics, ³Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, ¹⁶Department of Pathology and Cell Biology, and ¹⁷Department of Genetics and Development, Columbia University, New York, NY 10032

⁴Department of Genetics, Biology and Biochemistry and Human Genetics Foundation, and ⁸Department of Pathology, Center for Experimental Research and Medical Studies (CeRMS), University of Turin, 10126 Turin, Italy

⁵Institute of Oncology Research and Oncology Institute of Southern Switzerland, CH-6500 Bellinzona, Switzerland

⁶Haematopathology, Department L. and A. Seragnoli, University of Bologna, 40138 Bologna, Italy

⁷Division of Pathology and ¹⁴Division of Hematology, Fondazione IRCCS Policlinico San Matteo, University of Pavia, 27100 Pavia, Italy

¹⁰Hematology Unit, National Cancer Center of Bari and Department of Biology, University of Bari, 70126 Bari, Italy

¹¹Clinical and Experimental Onco-Hematology, CRO, IRCCS, 33081 Aviano, Italy

¹²Division of Hematology, University of Modena and Reggio Emilia, 41124 Modena, Italy

¹³Division of Pathology, Spedali Civili, University of Brescia, 26123 Brescia, Italy

¹⁵Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, 00618 Rome, Italy

¹⁸Institute of Hematology, University of Perugia, 06132 Perugia, Italy.

Splenic marginal zone lymphoma (SMZL) is a B cell malignancy of unknown pathogenesis, and thus an orphan of targeted therapies. By integrating whole-exome sequencing and copy-number analysis, we show that the SMZL exome carries at least 30 nonsilent gene alterations. Mutations in *NOTCH2*, a gene required for marginal-zone (MZ) B cell development, represent the most frequent lesion in SMZL, accounting for ~20% of cases. All *NOTCH2* mutations are predicted to cause impaired degradation of the NOTCH2 protein by eliminating the C-terminal PEST domain, which is required for proteasomal recruitment. Among indolent B cell lymphoproliferative disorders, *NOTCH2* mutations are restricted to SMZL, thus representing a potential diagnostic marker for this lymphoma type. In addition to *NOTCH2*, other modulators or members of the NOTCH pathway are recurrently targeted by genetic lesions in SMZL; these include *NOTCH1*, *SPEN*, and *DTX1*. We also noted mutations in other signaling pathways normally involved in MZ B cell development, suggesting that deregulation of MZ B cell development pathways plays a role in the pathogenesis of ~60% SMZL. These findings have direct implications for the treatment of SMZL patients, given the availability of drugs that can target NOTCH, NF- κ B, and other pathways deregulated in this disease.

CORRESPONDENCE

Riccardo Dalla-Favera:
rd10@columbia.edu
OR

Gianluca Gaidano:
gaidano@med.unipmn.it

Abbreviations used: DLBCL, diffuse large B cell lymphoma; HCV, hepatitis C virus; IGHV, immunoglobulin heavy chain variable; MZ, marginal zone; OS, overall survival; PFS, progression-free survival; SMZL, splenic marginal zone lymphoma; SNP, single nucleotide polymorphism.

D. Rossi, V. Trifonov, and M. Fangazio contributed equally to this paper.

R. Dalla-Favera, L. Pasqualucci, R. Rabadan, and G. Gaidano contributed equally to this paper.

© 2012 Rossi et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

Supplemental Material can be found at:
<http://jem.rupress.org/content/suppl/2012/08/09/jem.20120904.DC1.html>

Splenic marginal zone (MZ) lymphoma (SMZL) is a neoplasm of mature B cells that affects elderly patients and involves the spleen, bone marrow, and peripheral blood, while typically sparing peripheral lymph nodes and other sites populated by MZ B cells, namely mucosa-associated lymphoid tissue sites (Matutes et al., 2008; Swerdlow et al., 2008; Traverse-Glehen et al., 2011). Within the spleen, tumor cells are represented by small lymphocytes that occupy the MZ surrounding germinal centers and infiltrate the red pulp (Swerdlow et al., 2008). Despite an indolent behavior in the majority of patients, ~25% of SMZL patients experience a progressive course leading to early death (Matutes et al., 2008; Salido et al., 2010).

The pathogenesis of SMZL is currently elusive. Clinical and epidemiological data point to an association with hepatitis C virus (HCV); however, the infection rate in SMZL patients does not exceed ~15%, even in geographic areas in which the virus is endemic (Hermine et al., 2002; Suarez et al., 2006). The contribution of antigen stimulation to SMZL pathogenesis is suggested by the highly restricted immunoglobulin gene repertoire, including selective usage of the immunoglobulin heavy chain variable (*IGHV*) 1-2*04 allele in ~20–30% of cases, but the relevant antigen has not been identified (Warsame et al., 2011; Bikos et al., 2012).

Although the molecular pathogenesis of many lymphoma entities has been elucidated in detail (Swerdlow et al., 2008), relatively little is known about the genetic lesions associated with SMZL. Deletions of 7q31–q32 and gains of 3q are recurrent in ~20–30% and ~10–20% of cases, respectively, but the genes targeted by these lesions are unknown (Salido et al., 2010; Watkins et al., 2010; Rinaldi et al., 2011; Robledo et al., 2011). Cancer genes known to harbor genetic lesions in SMZL are not specific for this lymphoma type, and are limited to *TP53*, which is disrupted in ~15% of cases, most likely as a secondary genetic alteration (Gruszka-Westwood et al., 2001; Salido et al., 2010; Rinaldi et al., 2011), and to genes of the NF- κ B pathway, which are mutated in ~30% of cases (Rossi et al., 2011; Yan et al., 2012). Scant knowledge of somatic gene lesions associated with SMZL limits the present understanding of its pathogenesis and hampers the possibility of diagnosis and classification based on genetics, one of the mainstay criteria adopted by the World Health Organization Classification of Tumours of Hematopoietic and Lymphoid Tissues for the diagnosis of B cell lymphoma (Swerdlow et al., 2008).

By integrating whole-exome sequencing and genome-wide high-density single nucleotide polymorphism (SNP) array data, this study identifies alterations of key regulators of MZ development in ~60% of SMZL patients, with *NOTCH2* representing the most frequently mutated gene in this lymphoma type.

RESULTS

Identification of recurrent targets of genetic alterations in SMZL

To discover somatic, nonsilent mutations and copy number aberrations (CNAs) that are clonally represented in the SMZL coding genome, and presumably contributed to the initial expansion of the tumor clone, we performed massively parallel

sequencing and high-density SNP array analysis of paired tumor and normal DNA from eight untreated individuals diagnosed with SMZL (Table S1).

After enrichment of protein-coding genes by a hybridization-capture method, next-generation sequencing was performed using the Illumina HiSeq2000 instrument (Table S2). The whole-exome sequencing approach allowed us to map an average of ~102.5 million reads per sample at a mean depth of 110.6 $\times$ (range, 59.0–137.0 $\times$ per sample), with an average of 83.3% of the target sequence being covered by at least 30 reads (range, 72.0–86.0%; Table S2). Bioinformatic analysis followed by Sanger resequencing validation confirmed the presence of 203 somatic, nonsilent mutations (mean, 25.3/case; range, 12–44/case), affecting 191 distinct genes (validation rate, 92.6%; Fig. 1, A–D, and Table S3). The relative expression of the variant allele was determined by transcriptome analysis, performed in 6 of the 8 cases (Table S3).

By using the Affymetrix SNP6.0 platform, 41 somatic CNAs (30 deletions and 11 gains) were identified in the 8 discovery SMZL cases (mean, 5.1/case; range, 0–19/case; Fig. 1 E and Table S4). Alterations known to be associated with SMZL (3q gain, 7q31–q32 deletion, 17p deletion) and previously detected by FISH were correctly identified using the SNP array approach. Of the 41 CNAs, 8 (7 losses and 1 gain) were defined as focal, i.e., spanning ≤ 3 genes, with none being recurrently observed. When combining point mutations and CNAs, the overall load of tumor-acquired lesions was heterogeneous across the 8 SMZL cases investigated, ranging from 13–60 lesions/case (mean load, 30.5 lesions/case; Fig. 1 F).

Genes identified through the whole-exome sequencing and high-resolution SNP array approaches were prioritized for further assessment of their mutation frequency according to the fulfillment of one or more of the following criteria: (a) mutation recurrence in the discovery panel; (b) involvement by both point mutations and focal copy number changes or truncating deletions; and (c) involvement in cellular pathways known a priori to be potentially relevant for SMZL biology. The analysis was also extended to selected genes that were either involved in the same pathways as the genes found to be altered in the discovery genomes ($n = 7$, including *NOTCH1*, *FBXW7*, *SPEN*, *PTEN*, *CTNNB1*, *CD79A*, and *CD79B*) and/or those previously implicated in SMZL pathogenesis ($n = 7$, including *IKBKB*, *TNFAIP3*, *TRAF3*, *MAP3K14*, *TP53*, *CARD11*, and *MYD88*; Rossi et al., 2011; Yan et al., 2012), even if they were not identified in the discovery phase. Based on these criteria, 61 genes (47 revealed by the genomic approach in the discovery panel, and 14 involved in the same pathways as genes mutated in the discovery genome) were analyzed by Sanger-based resequencing of their coding exons and consensus splice sites in an independent screening panel of 32 SMZL cases (Table S5); for selected genes, an additional extension panel of 77 SMZL cases was also analyzed (Table S6; total number of cases, 109).

Of the 61 genes investigated, 21 were found to be mutated in at least two SMZL cases, and 18 were altered in $\geq 5\%$

of cases. Overall, the genes that are recurrently mutated in SMZL point to the involvement of specific programs implicated in normal MZ development (NOTCH, NF- κ B, B cell receptor, and Toll-like receptor signaling), as well as in chromatin remodeling and transcriptional regulation.

Alterations of *NOTCH2*

A molecular feature of SMZL revealed by this study is the presence of recurrent lesions in genes encoding for components of the NOTCH signaling pathway, a master regulator of normal MZ development (Kuroda et al., 2003; Saito et al., 2003;

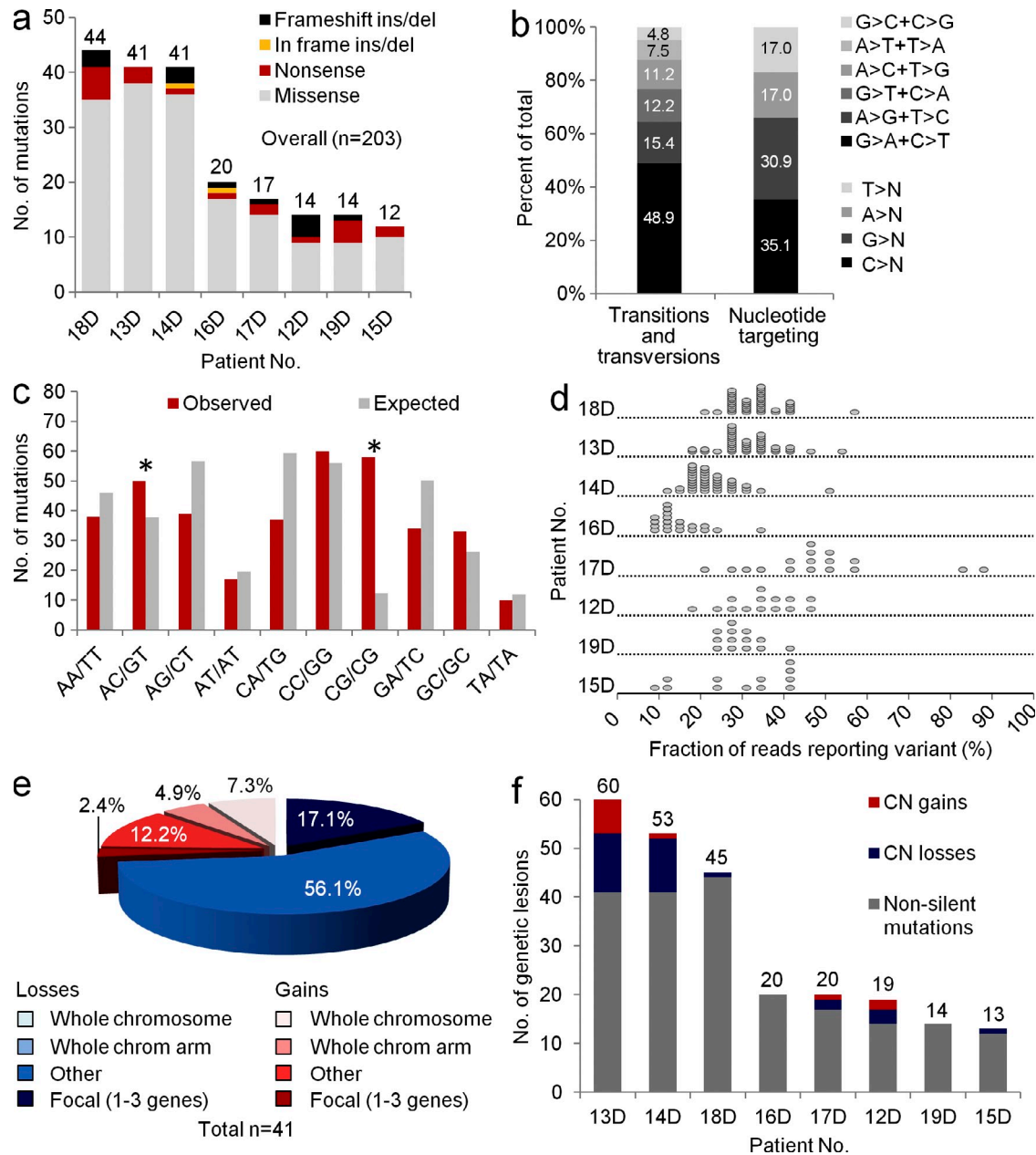


Figure 1. SMZL coding genome complexity. (a) Number and type of nonsilent mutations identified in the 8 discovery genomes. (b) The pattern of nucleotide substitutions in the discovery genomes revealed a predominance of transitions over transversions (121:67, ratio of 1.8) and a preferential targeting of G and C nucleotides (66.0% affecting G/C compared with 34.0% affecting A/T nucleotides). (c) Mutation frequency at specific dinucleotides (red bars). A significant bias toward alterations at 5'-CpG-3' dinucleotides, which accounted for 15.4% of all missense and nonsense changes, was documented. The expected frequencies (gray bars) correspond to the dinucleotide sequence composition of the Consensus CDS. Asterisks denote statistically significant differences in overrepresented changes, as assessed by a Poisson distribution after correction for multiple hypotheses. (d) Fraction of sequencing reads reporting individual somatic nonsilent variants (gray circles) in the discovery genomes. The large majority of the nonsilent mutations (82.1%) were present in at least 20% of the reads. (e) Overall number and frequency of somatically acquired CNAs. Losses of whole chromosomal arms were not observed and are thus not reported in the figure. (f) Combined load of somatically acquired genetic lesions in the discovery genomes, including nonsilent mutations and CNAs.

Moran et al., 2007; Santos et al., 2007; Pillai and Cariappa, 2009; Hampel et al., 2011). Targeted resequencing of *NOTCH2*, a gene required for MZ B cell differentiation (Saito et al., 2003; Moran et al., 2007; Pillai and Cariappa, 2009; Hampel et al., 2011), showed recurrent mutations in 25/117 (21.3%) SMZLs, including 2/8 (25.0%) discovery cases and 23/109 (21.1%) cases from the screening/extension panels. These data establish *NOTCH2* as the most frequently mutated gene in SMZL (Fig. 2).

NOTCH2 mutations were represented in all instances by truncating events (14 frameshift indels and 11 nonsense mutations), and clustered within a hotspot region in exon 34, including a recurrent p.R2400* nonsense mutation in 6/25 (24.0%) cases (Fig. 3 A, Tables S3, and Table S7). *NOTCH2* mutations were consistently absent in the heterodimerization domain or in other portions of the gene that are targeted by inactivating mutations in different cancer types (Wang et al., 2011b). Based on their distribution, all mutations were predicted to cause impaired degradation of the *NOTCH2* protein through the elimination or truncation of the C-terminal PEST domain (Fig. 3 A). Analysis of paired normal DNA confirmed the somatic origin of the mutations in all cases for which material was available ($n = 13$). *NOTCH2* mutations were present in ~40% of the reads by transcriptome sequencing, indicating allelic balance in the expression of the WT and mutated

allele (Table S3). Western blot analysis of *NOTCH2* expression consistently revealed the presence of an aberrant band of lower molecular weight, corresponding in size to the predicted truncated *NOTCH2* protein, in all mutated cases studied (Fig. 4). Copy number gains of *NOTCH2* were never observed in 110 SMZL cases analyzed by SNP array and/or FISH analysis.

In 3 patients for which multiple samples were available from different involved organs, the same *NOTCH2* mutation was detectable at all sites, namely peripheral blood, bone marrow, and spleen. This observation suggests that the mutation had been acquired before dissemination of the lymphoma clone to multiple anatomical sites.

To verify the specificity of *NOTCH2* mutations for SMZL, *NOTCH2* was investigated across the clinico-pathological spectrum of mature B cell tumors ($n = 399$). *NOTCH2* mutations were consistently absent in nodal MZ lymphoma (0/18), chronic lymphocytic leukemia (0/100), mantle cell lymphoma (0/20), follicular lymphoma (0/20), hairy cell leukemia (0/20), and multiple myeloma (0/22), whereas they occurred sporadically in extranodal MZ lymphoma (1/65, 1.5%) and, in accordance with a previous study (Lee et al., 2009), were restricted to 5/134 (3.7%) diffuse large B cell lymphomas (DLBCL; Fig. 3 B).

The clinical impact of *NOTCH2* mutations on SMZL overall survival (OS) was assessed in 94 patients with available

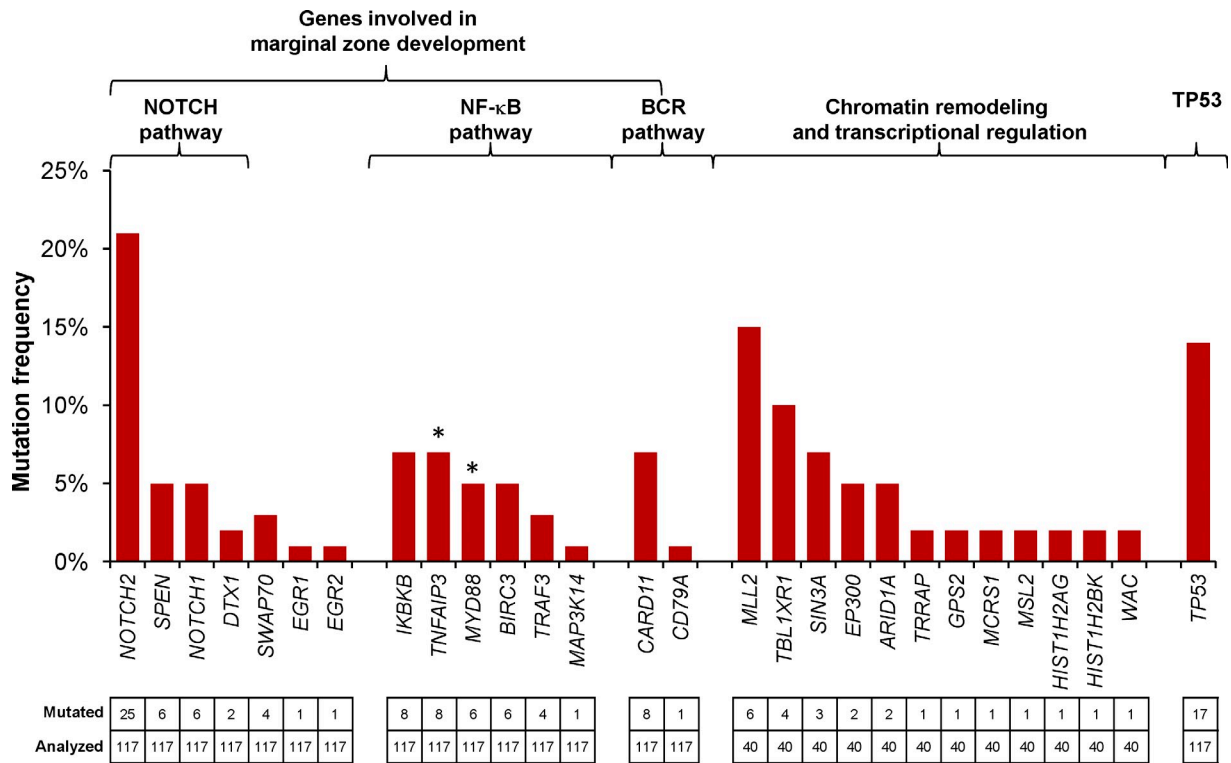


Figure 2. Recurrently targeted pathways in SMZL. Percentage of SMZL cases harboring mutations in selected genes belonging to cellular pathways that are recurrently altered in SMZL. Numbers at the bottom indicate the actual number of mutated cases over the total samples analyzed. Asterisks denote genes that are also implicated in Toll-like receptor responses. BCR, B cell receptor.

follow up. Consistent with the indolent behavior of this lymphoma type, the 5-yr OS of the SMZL cohort was 78.1% (95% CI, 67.8–88.4%). At 5 yr, SMZL patients harboring *NOTCH2* mutations were characterized by a significantly higher OS probability (93.3%; 95% CI, 80.0–100%) compared with patients harboring a WT *NOTCH2* (74.3%; 95% CI, 61.4–87.2%; $P = 0.048$; Fig. 5). Consistent with the improved OS, cases harboring *NOTCH2* mutations displayed a longer progression-free survival (PFS) after first-line treatment compared with *NOTCH2* WT patients (5-yr PFS in *NOTCH2* mutated patients, 83%; 95% CI, 65.4–100%; 5-yr PFS in *NOTCH2* WT patients, 44.1%; 95% CI, 28.9–58.3%; $P = 0.020$). There were no differences in major adverse events between *NOTCH2*-mutated and unmutated patients, as documented by a similar treatment-related mortality in the two groups (*NOTCH2*-mutated patients, 0/18; *NOTCH2* WT patients 1/53, 1.8%; $P = 1.000$).

Overall, these data document that, among B cell neoplasia, *NOTCH2*-activating mutations are predominantly associated with SMZL, and underscore the genetic individuality

of SMZL versus other MZ-derived lymphomas and versus indolent B cell lymphoproliferative disorders clinically mimicking SMZL.

Alterations in other genes of the NOTCH pathway

In addition to *NOTCH2*, other genes involved in NOTCH signaling and known to be relevant for normal MZ differentiation were affected by genomic lesions in SMZL, including *NOTCH1*, *SPEN*, and *DTX1* (Fig. 2; Kuroda et al., 2003; Saito et al., 2003; Santos et al., 2007; Pillai and Cariappa, 2009).

NOTCH1, a paralogue of *NOTCH2* implicated in several phases of lymphoid development, including terminal B cell differentiation into immunoglobulin-secreting cells (Santos et al., 2007; Yuan et al., 2010), was affected in 6/117 (5.1%) SMZLs by a recurrent 2-bp deletion (p.P2515fs*4) that truncates the PEST domain, similar to *NOTCH2* lesions (Fig. 6 A and Table S7). This deletion also represents the most common *NOTCH1* alteration in chronic lymphocytic leukemia and mantle cell lymphoma (Fabbri et al., 2011; Wang et al., 2011a; Kridel et al., 2012;

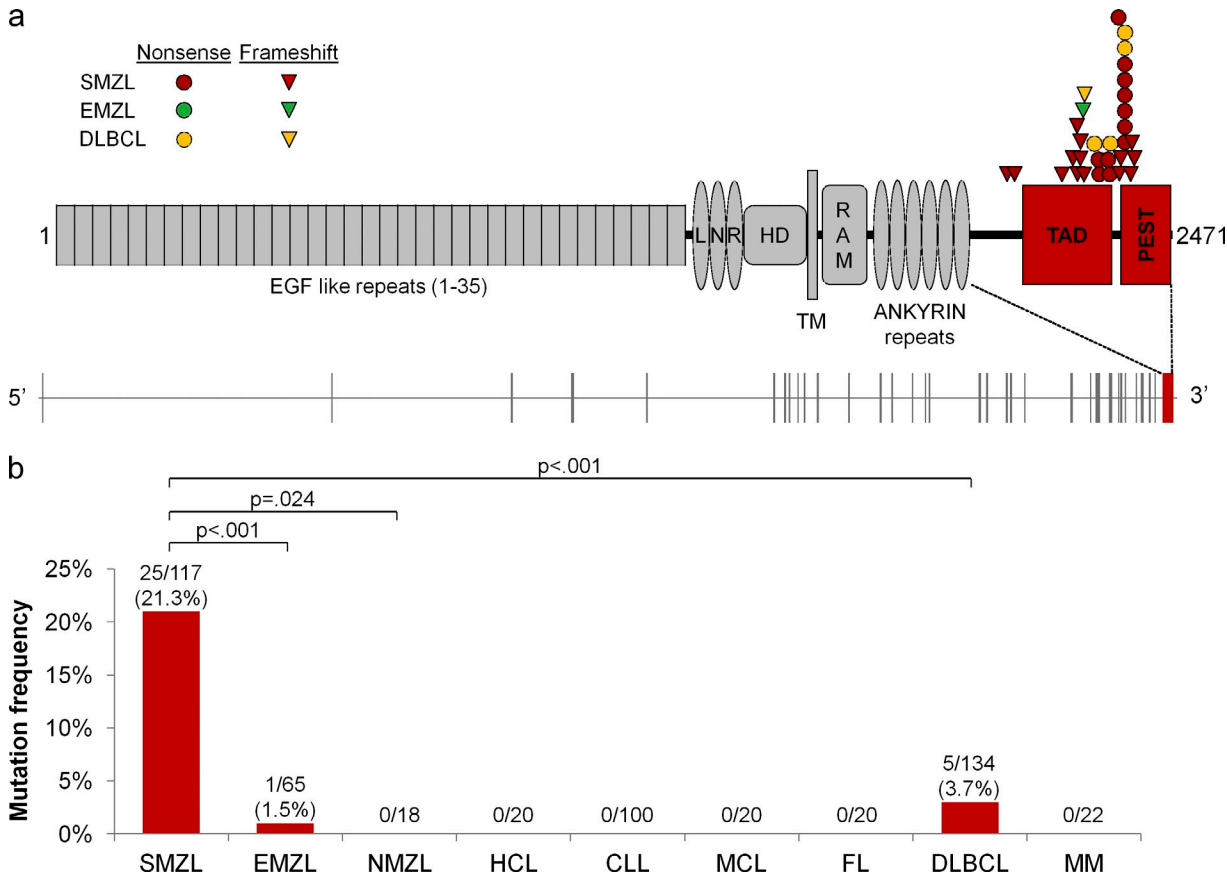


Figure 3. *NOTCH2* is frequently mutated in SMZL. (a) Schematic representation of the human *NOTCH2* gene (bottom) and protein (top), with its key functional domains (EGF, epithelial growth factor; LNR, LIN-12/NOTCH repeats; HD, heterodimerization; TM, transmembrane; RAM, regulation of amino acid metabolism; TAD, transactivation domain). Color-coded symbols indicate the type and position of the mutations. (b) Prevalence of *NOTCH2* mutations among mature B cell tumors (EMZL, extranodal MZ lymphoma; NMZL, nodal MZ lymphoma; HCL, hairy cell leukemia; CLL, chronic lymphocytic leukemia; MCL, mantle cell lymphoma; FL, follicular lymphoma; MM, multiple myeloma). Numbers on the top indicate the actual number of mutated cases over the total samples analyzed.

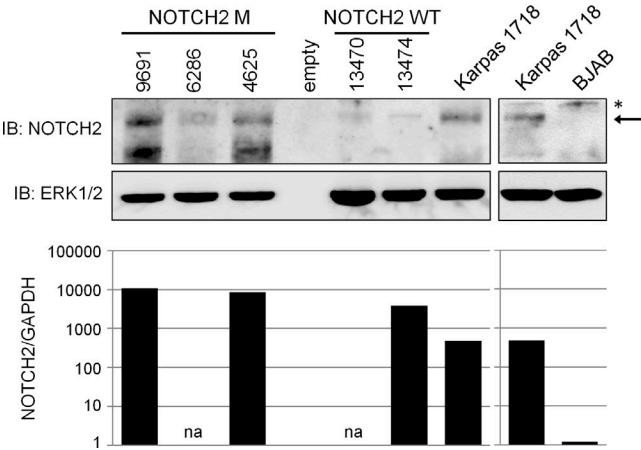


Figure 4. NOTCH2 expression in SMZL. Western blot analysis of NOTCH2 protein expression in purified primary tumor cells from 5 SMZL cases carrying WT or mutated (M) *NOTCH2*, and in the SMZL cell line Karpas 1718, also WT for *NOTCH2* (left); the specificity of the antibody was validated by using the BJAB cell line, which lacks *NOTCH2* mRNA expression (right; asterisk indicates nonspecific band). Arrow indicates the intact NOTCH2 protein; a band of lower molecular weight, consistent with the predicted size of the NOTCH2 protein encoded by the mutant allele, can be detected in all *NOTCH2*-mutated patients, but not in *NOTCH2* WT samples. Where available, the relative *NOTCH2* mRNA levels in the same samples are quantified by qRT-PCR (bottom; na, not available).

Quesada et al., 2012; Rossi et al., 2012), and the most recurrent *NOTCH1* PEST domain mutation in T cell acute lymphoblastic leukemia, where it causes impaired degradation of the NOTCH1 protein (Weng et al., 2004).

SPEN, also known as *MINT*, has been reported to repress NOTCH signaling by physically interacting with and causing inhibition of the transcription factor RBPJ, which

is implicated in the NOTCH signaling cascade (Kuroda et al., 2003; Li et al., 2005; VanderWielen et al., 2011). Consistent with its biochemical function, *SPEN* physiologically acts in the immune system as a negative regulator of B lymphocyte differentiation into MZ B cells by counteracting NOTCH activation (Kuroda et al., 2003). *SPEN* was found to be mutated in 6/117 (5.1%) SMZL cases (Fig. 2). With one exception, *SPEN* mutations were represented by inactivating events, including three frameshift indels and two nonsense substitutions (Fig. 6 A and Table S7), which were documented to be of somatic origin in all cases with available paired normal DNA. *SPEN*-mutated alleles were predicted to encode truncated proteins lacking the C-terminal domain involved in the interaction of SPEN with RBPJ, and to be necessary for NOTCH signaling inhibition (Fig. 6 A; Kuroda et al., 2003). In addition to truncating substitutions, *SPEN* was affected by one somatic missense substitution, categorized as probably damaging by PolyPhen-2 (Fig. 6 A and Table S7), and by four monoallelic deletions (Table S8). In all evaluable cases, *SPEN* mutations were restricted to a single allele and were not accompanied by deletions of the second allele. This pattern of predominant monoallelic inactivation is in agreement with in vitro models, suggesting that *SPEN* mutants may exert a dominant-negative effect (Li et al., 2005).

DTX1 encodes a RING finger ubiquitin ligase that binds the NOTCH family proteins and modulates their signaling, is highly expressed in MZ B cells, and may be important for late steps of B cell differentiation (Izon et al., 2002; Saito et al., 2003). In 2/117 (1.7%) SMZLs, *DTX1* was affected by somatically acquired missense substitutions mapping within two functionally relevant domains: the WWE1 domain, which mediates physical interactions between DTX1 and NOTCH (Zweifel et al., 2005), and the proline-rich domain, which may serve as a docking site for NOTCH inhibitory factors (Matsuno et al., 2002; Fig. 6 A; Tables S3 and S7).

In summary, ~32% (37/117) of SMZL patients carry alterations of genes belonging to the NOTCH pathway, with *NOTCH2* mutations accounting for approximately 2/3 of the events (25/37; 67.5%; Fig. 7).

Alteration of additional pathways involved in MZ development

Besides NOTCH signaling, the development of a normal MZ requires B cell migration to and retention within the spleen marginal sinus until they are activated (Pillai and Cariappa, 2009). *SWAP70* encodes for an F-actin-binding/Rho GTPase-interacting protein that is necessary for

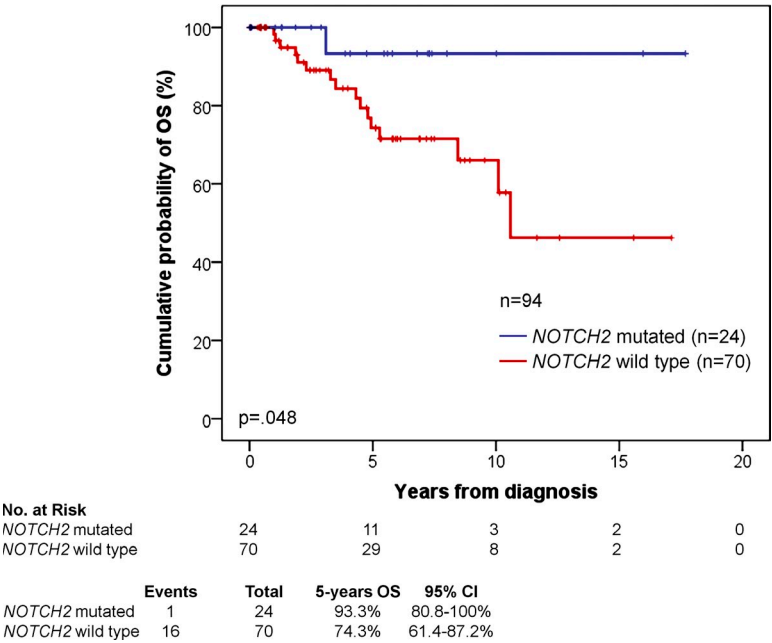


Figure 5. NOTCH2 mutations are associated with better OS. Kaplan-Meier estimates of OS in SMZL patients ($n = 94$), according to *NOTCH2* mutation status.

normal B cell trafficking across spleen compartments (Chopin et al., 2010a, 2011). In 4/117 (3.4%) cases, *SWAP70* was disrupted by truncating mutations that removed the C-terminal domain of the protein required for binding to F-actin

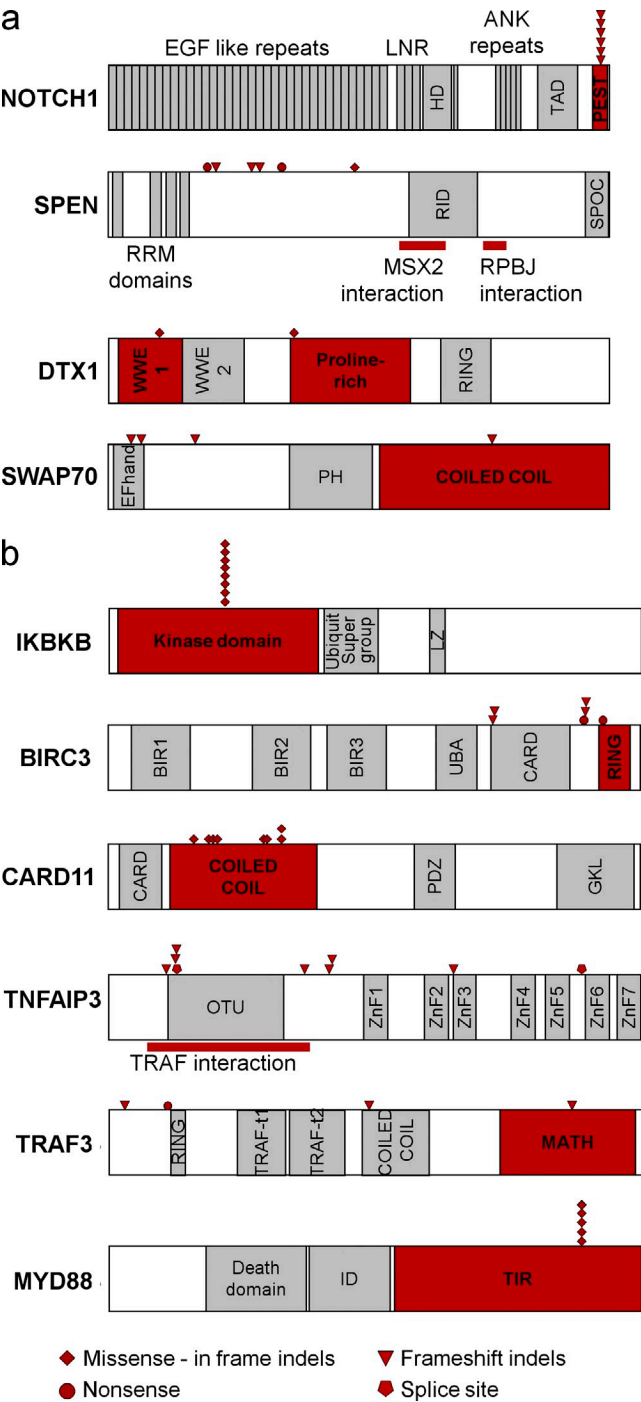


Figure 6. Mutations of genes belonging to the NOTCH, migration/adhesion, NF- κ B, and B cell receptor pathways in SMZL. Schematic diagram of the proteins targeted by mutations in SMZL, with their key functional domains (a, NOTCH pathway; b, NF- κ B pathway). Symbols indicate the type of mutation.

(Fig. 6 A; Tables S3 and S7). *SWAP70* was also affected by a monoallelic deletion in one additional SMZL (Table S4). Moreover, one SMZL (0.9%) displayed a somatic mutation in the *EGR1* B cell transcription factor, which also plays a role in MZ development (Table S3; Gururajan et al., 2008). *EGR2*, a paralogue of *EGR1*, was somatically mutated in one additional case (0.9%; Table S3).

Active NF- κ B signaling is necessary for the generation and/or maintenance of normal MZ B cells (Calado et al., 2010; Chu et al., 2011; Conze et al., 2010; Pappu and Lin, 2006; Moran et al., 2007; Xie et al., 2007; Sasaki et al., 2008; Pillai and Cariappa, 2009). Mutations of both the canonical (17/117, 14.5% cases) and noncanonical (11/117, 9.4% cases) NF- κ B pathways were recurrent in SMZL (Fig. 2), and affected several previously identified genes in this lymphoma type (Rossi et al., 2011; Yan et al., 2012), including *IKBKB* (8/117, 6.8%), *TNFAIP3* (8/117, 6.8%), *BIRC3* (6/117, 5.1%), *TRAF3* (4/117, 3.4%), and *MAP3K14* (1/117, 0.9%; Fig. 6 B; Table S3 and Table S7). These genes were also targeted by CNAs, including deletions of *TNFAIP3* (8/110, 7.2%), *BIRC3* (7/110, 6.3%), and *TRAF3* (6/110, 5.4%), and gains of *MAP3K14* (7/110, 6.3%; Table S8).

In addition to genes specifically attributed to the NF- κ B pathway, recurrent mutations were also found in *CARD11* and *MYD88*, which, among their many functions, act as positive regulators of NF- κ B in signaling from the B cell receptor and Toll-like receptor, respectively (Fig. 2; Lenz et al., 2008; Ngo et al., 2011). *CARD11* mutations (6 missense and 2 in frame deletions) cluster within the coiled-coil domain in 8/117 (6.8%) cases (Fig. 6 B and Table S7). In 6/117 (5.1%) SMZLs, *MYD88* was targeted by a recurrent missense substitution (p.L265P) that has been previously reported in activated B cell-type DLBCL (Fig. 6 B and Table S7; Ngo et al., 2011).

Cases harboring NF- κ B pathway mutations and cases harboring NOTCH pathway mutations did not differ in terms of clinical and biological features at presentation, including age, sex, performance status, levels of hemoglobin, LDH, β -2-microglobulin and albumin, HCV infection, *IGHV* mutation status, usage of the *IGHV1-2*04* allele, stereotyped VH CDR3, 7q31-q32 deletion, 3q gain, or *TP53* disruption ($P > 0.05$ in all instances).

Overall, mutations of positive and negative NF- κ B regulators accounted for 40/117 (34.1%) SMZL cases (Fig. 7), implicating activation of NF- κ B as a major contributor to the pathogenesis of this disease.

Alteration of genes involved in chromatin remodeling and transcriptional regulation

Apart from genes implicated in MZ development, whole-exome sequencing revealed a second set of genes recurrently mutated in SMZL and regulating chromatin remodeling (Fig. 2). The *MLL2* histone methyltransferase controls gene transcription by modifying the lysine-4 position of histone 3 and is recurrently mutated in FL and DLBCL (Morin et al., 2011; Pasqualucci et al., 2011b). Inactivating mutations of

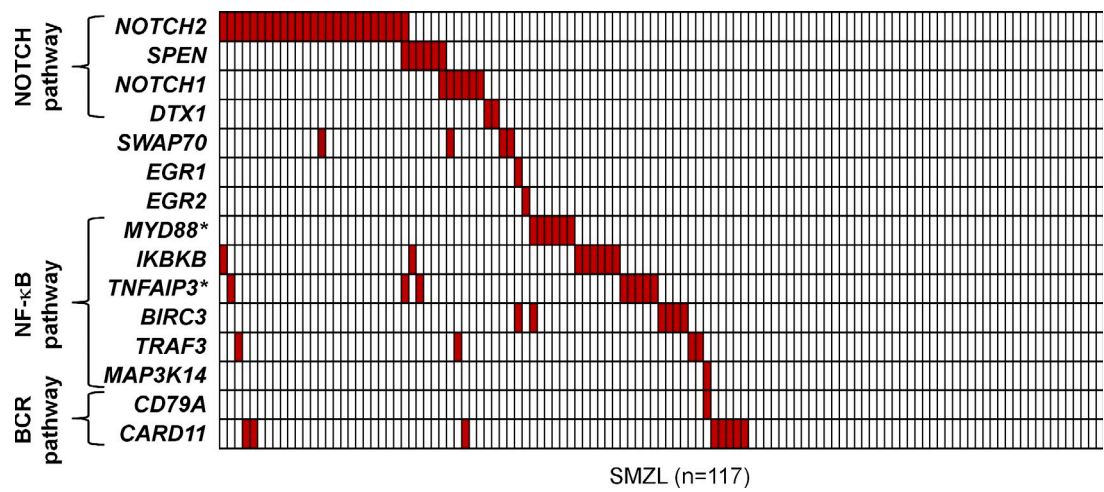


Figure 7. Mutually exclusive involvement of genes implicated in MZ development. In the heatmap, rows correspond to genes and columns represent individual patients. Color coding is based on gene mutation status (white, WT; red, mutated). Asterisks denote genes that also modulate Toll-like receptor responses. BCR, B cell receptor.

MLL2 occurred in 6/40 (15.0%) SMZLs (Fig. 8; Tables S3 and S7), leading to removal or truncation of the C-terminal SET domain required for its enzymatic activity. In addition, two SMZLs showed a p.P692T somatic missense substitution. *ARID1A*, a member of the SWI-SNF chromatin remodeling family that was reported as mutated in several solid tumors (Wiegand et al., 2010), was targeted by genetic lesions in 4/40 (10.0%) cases, including 2 point mutations and 2 deletions (Fig. 8; Tables S3, S7 and S8). *EP300* and *CREBBP* are two highly related acetyltransferases that are also recurrently disrupted in B cell lymphoma (Pasqualucci et al., 2011a) and acute lymphoblastic leukemia (Mullighan et al., 2011). Somatic missense mutations of *EP300* were found in 2/40 (5%) SMZL cases (Fig. 8 and Table S3), whereas one additional patient harbored a small deletion that juxtaposes in-frame the first 16 exons of *CREBBP* to the C-terminal portion of the *ZNF434* gene, thus abrogating the *CREBBP* acetyltransferase domain, as confirmed by transcriptome sequencing analysis (not depicted). Somatic mutations of *SIN3A*, encoding for a core component of the SIN3–HDAC1/2 histone deacetylase complex, occurred in 3/40 (7.5%) SMZLs, and, in two cases, targeted regions that are required for the recruitment of histone modification complexes (Fig. 8; Tables S3 and S7; Grzenda et al., 2009). Other genes involved in chromatin remodeling and found to be somatically mutated in single SMZL cases are represented in Fig. 2.

Mutations were also detected in other transcriptional regulators. *TBL1XR1*, an intrinsic component of the SMRT–N-CoR transcription co-repressor machinery that was recently found to be recurrently disrupted in B cell tumors (Perissi et al., 2008), showed mutations in 4/40 (10.0%) SMZLs (Fig. 2). Mutations (two frameshift deletions and three somatically acquired missense substitutions) targeted the protein WD domains implicated in the exchange of nuclear

receptor co-repressors for co-activators (Fig. 8; Table S3 and Table S7; Oberoi et al., 2011). *TBL1XR1* mutations were monoallelic in three cases, whereas one case harbored a disrupting mutation coupled to a missense substitution (Table S7).

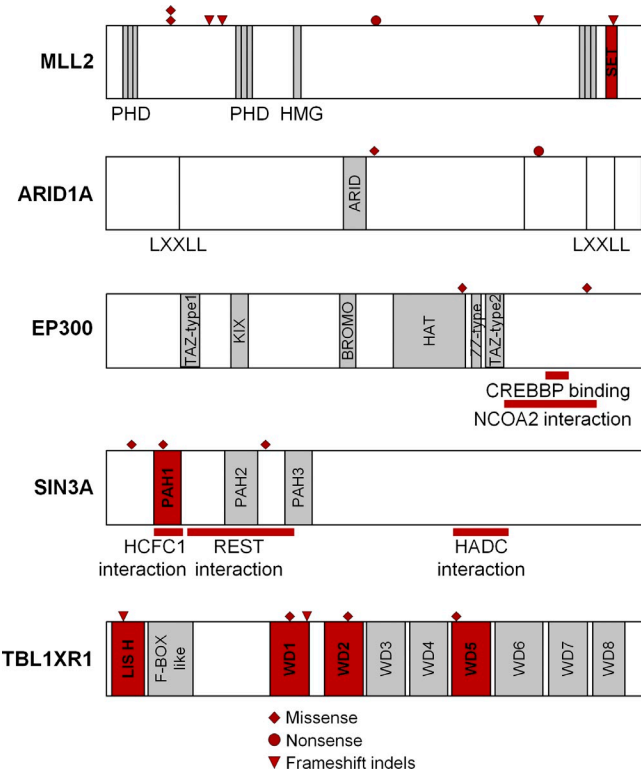


Figure 8. Mutations of genes involved in chromatin remodeling and transcriptional regulation. Schematic diagram of the indicated proteins, with their key functional domains. Symbols indicate the type of mutation.

Three additional SMZLs showed monoallelic *TBL1XR1* deletion (Table S8). *GPS2*, another member of the SMRT-N-CoR transcription co-repressor machinery that physically interacts with *TBL1XR1* (Oberoi et al., 2011), was biallelically inactivated in one SMZL by a nonsense mutation with deletion of the second allele (Table S3 and Table S4). The mutation targeted the *GPS2* domain required for *TBL1XR1* interaction. Monoallelic deletions of *GPS2* were observed in additional eight cases (Table S8).

Overall, mutations of chromatin remodeling and transcriptional regulation genes occurred in 14/40 (35.0%) SMZLs and, as observed in other cancer types (Gui et al., 2011), were frequently concurrent in the same patient (~50% of the mutations; Fig. 9), suggesting that they cooperate to promote tumorigenesis.

Mutations of genes regulating MZ development and NF-κB activation characterize ~60% of SMZLs

Given the biological heterogeneity of SMZLs, we investigated the relationship between mutations of genes controlling MZ development and established genetic, immunogenetic, and biological subgroups of SMZLs, including those defined by deletion of 7q31-q32 (21/110, 19.1%), 3q gain (14/110, 12.7%), expression of unmutated *IGHV* genes (29/110, 26.3%), usage of the *IGHV1-2*04* allele (26/110, 23.6%), usage of stereotyped VH CDR3 (9/110, 8.1%), positive HCV serology (11/86, 12.7%), and *TP53* mutations (17/117, 14.5%). This analysis did not disclose any selective association, suggesting that mutations of genes controlling MZ development play an independent role in lymphomagenesis (Table S9). A preferential, though not selective, association was observed between deletion of 7q31-q32, a cytogenetic abnormality characteristic of SMZL, and cases harboring mutations of MZ development genes (19/64, 29.7% in cases harboring mutations vs. 2/46, 4.3% in cases devoid of mutations; $P = 0.001$; Table S9). In particular, deletion of 7q31-q32 significantly associated with SMZL harboring *NOTCH2* mutations (11/24, 45.8%; $P = 0.001$).

Mutations of genes implicated in MZ development (NOTCH, NF-κB, B cell receptor, and Toll-like receptor pathway, as well as *SWAP70*, *EGR1*, *EGR2*; Izon et al., 2002; Kuroda et al., 2003; Saito et al., 2003; Moran et al., 2007; Santos et al., 2007; Gururajan et al., 2008; Pillai and Cariappa, 2009; Chopin et al., 2010b, 2011; Hampel et al., 2011) show a largely mutually exclusive distribution pattern, although the number of cases ($n = 117$) lacks statistical power for this type of analysis (Fig. 7). Collectively, these lesions account for the majority of SMZL cases (70/117; 59.8%), and suggest that alterations in genes affecting B cell fate decision toward MZ differentiation represent alternative mechanisms converging on the deregulation of a common downstream target, which may include NF-κB.

DISCUSSION

One goal of this study was to provide initial information on the landscape of somatic genetic lesions (type and frequency) that characterize the SMZL coding genome. Given the presence of ~25 nonsilent mutations and ~5 CNAs per case (on average, ~30 alterations/case), SMZL appears to show a degree of genomic complexity that is intermediate between that of aggressive lymphoma, namely DLBCL (~90 nonsilent mutations per case; Morin et al., 2011; Pasqualucci et al., 2011b; Lohr et al., 2012), and that of previously untreated chronic lymphocytic leukemia (~12 nonsilent mutations per case; Fabbri et al., 2011; Quesada et al., 2012), which shares an indolent course with SMZL.

A second key finding of this study was the identification of alterations in genes affecting B cell fate decision toward MZ differentiation as significantly associated with SMZL. Alterations in genes of the NOTCH pathway emerged as highly recurrent (~30%) in SMZL, with *NOTCH2*, a key regulator of MZ development, being the most frequently mutated gene (~20%). All *NOTCH2* mutations observed in SMZL cause disruption of the protein inhibitory PEST domain and are predicted to activate NOTCH2 signaling (Lee et al., 2009). Because of the lack of xenograft and cell line models of SMZL, we are currently unable to assess whether these mutations

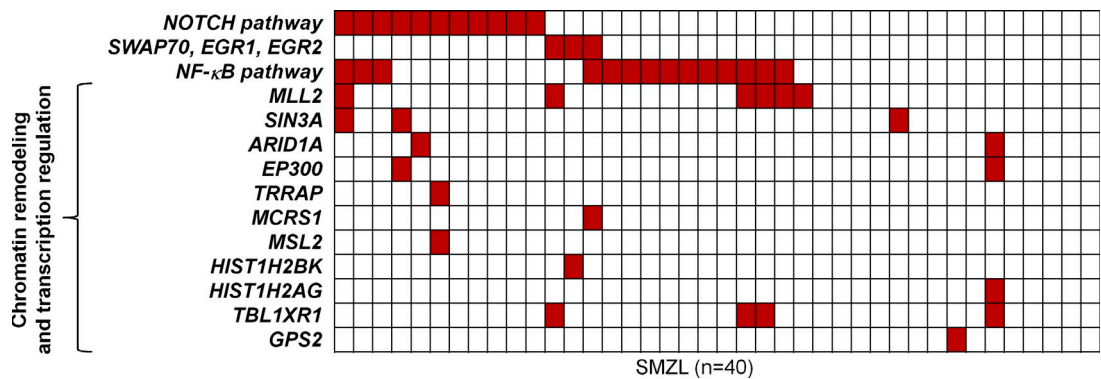


Figure 9. Integrated analysis of mutations affecting chromatin remodeling genes in SMZL. In the heatmaps, rows correspond to mutated genes and columns represent individual patients. Color coding is based on gene mutation status (white, WT; red, mutated). Analysis of chromatin remodeling genes was extended to the screening panel only (in total, 40 cases), and is thus shown separately.

confer ligand independence, or whether the mutated proteins still require activation by ligand binding. Disruption of the PEST domain renders SMZL-associated mutations of *NOTCH2* analogous to those involving *NOTCH1* in chronic lymphocytic leukemia and mantle cell lymphoma, and to a subset of *NOTCH1* mutations in T cell acute lymphoblastic leukemia (Fabbri et al., 2011; Wang et al., 2011a; Kridel et al., 2012; Quesada et al., 2012). Nevertheless, *NOTCH2* mutations appear to be relatively specific for SMZL, being virtually absent in other major types of mature B cell neoplasia and rare in DLBCL (Lee et al., 2009). Consistent with our findings, one single mutation of *NOTCH2* was recently described in one case of SMZL (Trøen et al., 2008). In addition to *NOTCH2*, other modulators or members of the NOTCH pathway were targeted by genetic lesions in SMZL, including *SPEN*, *DTX1*, and *NOTCH1*. The alternative involvement of multiple genes converging on the NOTCH pathway, together with the well-established functional role of some of these alterations, namely *NOTCH1* mutations in T cell acute lymphoblastic leukemia (Weng et al., 2004), strongly support their pathogenic role in SMZL.

As is the case for most cancer-associated genetic lesions, activation of *NOTCH2* may not be sufficient for malignant transformation. In fact, patients affected by the Hajdu-Cheney syndrome, an autosomal-dominant genetic disease that causes severe and progressive bone loss and is associated with germline *NOTCH2* mutations analogous to those identified in SMZL, do not develop lymphomas (Brennan and Pauli, 2001; Simpson et al., 2011). Similarly, transgenic mice engineered to express activating *NOTCH2* mutations in mature B cells display an expansion of the MZ at the expense of the follicular compartment, but do not develop lymphoma (Hampel et al., 2011). It is important to note, however, that lymphoma development may require longer times than those observed so far both in Hajdu-Cheney patients (no diagnosed individual was >50 yr old; Brennan and Pauli, 2001; Simpson et al., 2011) and in mice (1 yr), consistent with the elderly age and indolent course of SMZL (Hampel et al., 2011).

Collectively, the genetic alterations associated with SMZL appear to predominantly involve signaling pathways that regulate MZ development, including: (a) NOTCH, via the alternative mutation of multiple genes (*NOTCH2*, *NOTCH1*, *SPEN*, and *DTX1*); (b) NF- κ B, via mutation of *TNFAIP/A20*, *BIRC3*, *TRAF3* and *CARD11*; and (c) the B cell receptor, via mutation of *CARD11*. Indeed, at least in mice, the commitment of mature B cells to the MZ compartment requires signaling from *NOTCH2* (Kuroda et al., 2003; Saito et al., 2003; Moran et al., 2007; Santos et al., 2007; Pillai and Cariappa, 2009; Hampel et al., 2011), as well as activation of the NF- κ B transcription complex (Xie et al., 2007; Sasaki et al., 2008; Calado et al., 2010; Conze et al., 2010; Chu et al., 2011) and possibly antigen stimulation through the B cell receptor and Toll-like receptor (Pappu and Lin, 2006). The alteration of genes implicated in the physical retention of MZ B cells within the spleen might also be important for SMZL pathogenesis, as suggested by the observation of recurrent mutations in *SWAP70* (Chopin et al.,

2010a, 2011). Thus, the finding that ~60% of SMZL cases display the alternative deregulation of these pathways suggests that one major component of SMZL pathogenesis is the constitutive activation of signals normally deputed to the differentiation and homing of B cells into the MZ.

Finally, the results herein have important implications for the clinical management of SMZL. Because the differential diagnosis of SMZL from other indolent B cell lymphoproliferative disorders clinically mimicking SMZL is often complex (Matutes et al., 2008; Swerdlow et al., 2008), some of the genetic alterations reported here, specifically mutations of *NOTCH2*, may serve as potentially helpful markers. More importantly, the results of this study provide a rationale for the design of novel therapeutic strategies for SMZL, which, to date, remains a disease orphan of specifically targeted drugs (Traverse-Glehen et al., 2011). New regimens may be useful to avoid splenectomy, a well-established therapeutic option in SMZL, which has perioperative and long-term morbidity and mortality, especially in elderly or unfit individuals (Matutes et al., 2008).

The NOTCH pathway, which we show is affected by genetic lesions in up to 30% of SMZLs, may represent an attractive candidate therapeutic target, as some NOTCH inhibitors, such as those preventing its enzymatic conversion to active transcription factor, are already available, and others are under active clinical development (Real et al., 2009). The NF- κ B pathway represents an additional potential target; available proteasome inhibitors already approved for other malignancies, or more specific anti-NF- κ B compounds currently under development, should be tested for their efficacy in SMZL, either alone or in combination with NOTCH inhibitors.

MATERIALS AND METHODS

Patients and tumor biopsies. The study panel comprised a total of 117 SMZL samples obtained from frozen spleen biopsies of newly diagnosed, previously untreated patients, and was distinguished into a discovery panel ($n = 8$ cases), a screening panel ($n = 32$ cases), and an extension panel ($n = 77$ cases). Out of the 109 SMZLs used as screening and extension panel, 61 were already reported (Rossi et al., 2011). In all cases, the SMZL diagnosis was based on spleen histology and was confirmed by centralized pathological revision (S.A. Pileri). Consistent with a SMZL diagnosis, all cases of the discovery and screening panels lacked the t(11;18) and the t(14;18) translocations (Matutes et al., 2008), and all 117 cases lacked the *BRAF* p.V600E mutation (Swerdlow et al., 2008; Tiacci et al., 2012). Matched normal DNA was obtained from saliva or peripheral blood granulocytes in 48 patients ($n = 8$ discovery cases and 40 cases from the screening and extension panel). The clinical and biological characteristics of cases belonging to the SMZL discovery, screening, and extension panels are summarized in Tables S1, S5, and S6, respectively.

For comparative purposes, 399 B cell tumors other than SMZL were also included in the study (18 nodal MZ lymphomas, 65 extranodal MZ lymphomas, 100 chronic lymphocytic leukemias, 20 mantle cell lymphomas, 20 follicular lymphomas, 134 DLBCLs, 20 *BRAF* p.V600E mutation-positive hairy cell leukemias, and 22 multiple myelomas). All of the 399 samples had been obtained at diagnosis from the involved site (lymph nodes or extranodal sites in the case of lymphoma; CD138⁺ cells purified from bone marrow aspirates in the case of multiple myeloma; peripheral blood purified B cells in the case of hairy cell leukemia; and peripheral blood mononuclear cells in the case of chronic lymphocytic leukemia). The number of mature

B cell neoplasms included in each panel was estimated to allow a 90% probability of identifying genes that are mutated in at least 10% of cases.

Patients provided informed consent in accordance with local IRB requirements and The Declaration of Helsinki. The study was approved by the Ethical Committee of the Ospedale Maggiore della Carità di Novara affiliated with the Amedeo Avogadro University of Eastern Piedmont (Protocol Code 59/CE; Study Number CE 8/11) and by the Institutional Review Board of Columbia University.

DNA extraction. High molecular weight (HMW) genomic DNA was extracted from tumor and normal samples according to standard procedures (Rossi et al., 2011). In all tumor cases, the fraction of tumor cells in the tissue biopsy section used for molecular studies was estimated to be >70% by morphology, immunohistochemistry, and/or flow cytometry. DNA was quantified by the Quant-iT PicoGreen reagent (Invitrogen) in the discovery panel, and by the NanoDrop 2000C spectrophotometer (Thermo Fisher Scientific) in the screening and extension panels. All DNA samples were verified for integrity by 1% agarose gel electrophoresis. Tumor cell clonality was established by amplification of the rearranged *IGH* genes, as described in detail in the paragraph “*IGHV-IGHD-IGHJ* rearrangement analysis.” Analysis of patient-specific *IGHV-IGHD-IGHJ* rearrangements were also performed in the paired normal DNA to exclude contamination from tumor cells.

Whole-exome capture and massively parallel sequencing. Purified tumor and germline genomic DNA (3 µg) from the 8 discovery SMZL cases was enriched in protein coding sequences using the in-solution exome capture SureSelect Human All Exon 50Mb kit (Agilent Technologies), according to the manufacturer's protocol. The SureSelect Human All Exon 50 Mb kit encompasses all coding exons annotated by the GENCODE project, including all exons in the Consensus CDS (CCDS, March 2009) database, 10 bp of flanking sequence for each targeted region, and small noncoding RNAs from miRBase (v.13) and Rfam. The captured targets were subjected to massively parallel sequencing using the Illumina HiSeq 2000 analyzer (Illumina) with the paired-end 2 × 100 bp read option, following the manufacturer's instructions. As quality controls for the precapture and post-capture steps, randomly selected PCR-amplified clones of the PCR products were subjected to Sanger sequencing to verify their preferential alignment to human genomic regions and to human coding transcripts ($n = 50/\text{library}$). Exome capture and massively parallel sequencing were performed at the HiSeq Service of Fasteer SA (Plan-les-Ouates, Switzerland).

Sequence mapping and identification of tumor-specific variants. Paired-end reads (~102.5 million per case) obtained by high-throughput sequencing were aligned to the human genome reference hg19/NCBI GRCh37 using the BWA alignment tool version 0.5.9, and provided a mean depth of 111x with at least 83% of the target exome covered at 30x (Table S2). Sequence variants, i.e., differences from the reference sequence, were identified separately for each tumor and normal sample. The frequency of each variant was estimated from the total number of reads covering the position of that variant. Using the SAVI (Statistical Algorithm for Variant Identification) algorithm developed at Columbia University (Tiacchi et al., 2012), an empirical prior was constructed for the variant frequencies. From that prior, we obtained a corresponding high-credibility interval (posterior probability $\geq 1-10^{-5}$) for the frequency of each variant and a high-credibility interval for the corresponding change in frequency between the tumor and the normal samples. 218 nonsilent variants, which are not reported in dbSNP 135 and had high posterior probability ($\geq 1-10^{-5}$) of nonzero presence in the tumor, as well as at least 1% change from the normal with high posterior probability ($\geq 1-10^{-5}$), were kept for validation by Sanger sequencing. A comparison with data obtained by high-density SNP array analysis of the same tumor/normal pairs established the sensitivity of the method at 96.6% for heterozygous SNP calls and 98% for homozygous SNP calls.

Validation of candidate somatic mutations by DNA Sanger sequencing. Candidate nonsilent somatic mutations were subjected to validation by

conventional Sanger-based resequencing of PCR products obtained from both tumor and paired normal HMW genomic DNA using primers specific for the exon encompassing the variant. The sequences surrounding the genomic locations of the candidate tumor-specific nonsilent mutations were obtained from the UCSC Human Genome database, and PCR primers were derived from previously published studies (Parsons et al., 2011) or custom-designed using the Primer 3 online software (<http://frodo.wi.mit.edu/primer3/>). Of the 218 predicted nonsilent variants, 202 (validation rate, 92.6%) were confirmed to be somatic in origin by Sanger sequencing, 6 were also found in the matched normal DNA, thus representing previously nonannotated germline polymorphisms that had not been detected by the high-throughput sequencing analysis, and the remaining 10 variants were absent in both tumor and normal genomic DNA, when tested by Sanger sequencing. Confirmed, somatic nonsynonymous mutations were tested for their functional consequences in silico by using the PolyPhen-2 (Polymorphism Phenotyping) algorithm (<http://genetics.bwh.harvard.edu/pph2/>), which is based on structure and sequence conservation. Genes found to be mutated were verified for their presence in the Catalogue of Somatic Mutations in Cancer database (Forbes et al., 2011) and in the Cancer Gene Census database.

Mutation screening of identified genes. The complete coding sequences and exon-intron junctions of selected genes identified through the whole-exome sequencing and/or the SNP array approach were analyzed in the SMZL screening panel and extension panel by PCR amplification and direct sequencing of whole-genome-amplified DNA obtained using the Repli-g Mini kit (QIAGEN). Sequences for all annotated exons and flanking splice sites were retrieved from the UCSC Human Genome database using the corresponding mRNA accession number as a reference. PCR primers, located ~50 bp upstream or downstream to target exon boundaries, were either derived from previously published studies (Parsons et al., 2011) or designed in the Primer 3 program and filtered using UCSC in silico PCR to exclude pairs yielding more than a single product. All PCR primers and conditions are available upon request. Purified amplicons were subjected to conventional DNA Sanger sequencing using the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems), and compared with the corresponding germline sequences using the Mutation Surveyor Version 3.97 software package (SoftGenetics) after automated and/or manual curation. Of the evaluated sequences, 99% had a Phred score of ≥ 20 and 96% had a score of ≥ 30 . Candidate somatic mutations were confirmed from both strands on independent PCR products obtained from HMW genomic DNA. Synonymous mutations, previously reported germline polymorphisms, and changes in the matched normal DNA (when available) were removed from the analysis. The following databases were used to exclude known germline variants in primary cases for which paired normal DNA was not available: Human dbSNP Database at NCBI (build 136), Ensembl Database, the 1000 Genomes Project, five single-genome projects available at the UCSC Genome Bioinformatics resource.

High-density SNP array analysis. Copy number abnormalities were assessed by high-density SNP array analysis in 81 of the 117 SMZL cases, including the 8 discovery cases (paired tumor and normal DNA), 21 SMZL from the screening panel (both sets using the Genome-Wide Human SNP Array 6.0; Affymetrix), and 52 SMZL cases from the extension panel (using the GeneChip Human Mapping 250K NspI; Affymetrix; GEO accession no. GSE24881). In brief, HMW genomic DNA was restriction enzyme digested, ligated, PCR amplified, purified, labeled, fragmented, and hybridized to the arrays according to the manufacturer's instructions. Identification of segments of abnormal copy numbers were performed using the dChipSNP software, and a karyotype-guided normalization procedure according to a published workflow (Pasqualucci et al., 2011b).

Whole-transcriptome sequencing (RNA-seq). Whole-transcriptome sequencing was performed in 6 SMZL cases belonging to the discovery panel (cases 12D through 17D). RNA was extracted from CD19⁺ B cells

purified from spleen disaggregation using the Allprep DNA/RNA Mini kit (QIAGEN). Polyadenylated RNA was selected after DNaseI treatment and used as a template for double-stranded cDNA synthesis. The 190–210 bp fraction was then isolated and PCR amplified, and libraries were constructed using the Illumina Genome Analyzer paired-end library protocol according to manufacturer's instructions. As quality controls for the precapture and postcapture steps, randomly selected PCR-amplified clones of the PCR products were subjected to Sanger sequencing to verify their preferential alignment to human genomic regions and to human coding transcripts (50 clones/library). Transcriptome sequencing was performed at the HiSeq Service of Fasteris SA (Plan-les-Ouates, Switzerland).

Whole-transcriptome data analysis. Reads (mean, 17 M per case) obtained by high-throughput sequencing were aligned to a human transcriptome reference (113 M), obtained by stitching together the exons belonging to the RefSeq transcripts as reported at the UCSC Genome Browser, using the BWA alignment tool version 0.5.9. Sequence variants, i.e., differences from the reference sequence, were identified separately for each sample, and credibility intervals for those frequencies, as well as for the change in frequency from the corresponding whole-exome DNA sequencing, were obtained as described in "Sequence mapping and identification of tumor-specific variants." Expression for each RefSeq transcript was calculated using non-clonal (reads that map to the same exact position are counted once) reads per kilobase of exon model per million mapped reads. For each transcript, we first filter out clonal reads mapping to the same position, and then calculate the mean nonclonal read depth per base and normalize by the total number of nonclonal reads mapped to the transcriptome.

IGHV-IGHD-IGHJ rearrangement analysis. PCR amplification of *IGHV-IGHD-IGHJ* rearrangements was performed on HMW genomic DNA using *IGHV* leader primers or consensus primers for the *IGHV* FR1, along with appropriate *IGHJ* genes, as previously described (Rossi et al., 2011). PCR products were directly sequenced with the ABI PRISM Big-Dye Terminator v1.1 Ready Reaction Cycle Sequencing kit using the ABI PRISM 3100 Genetic Analyzer (both from Applied Biosystems). Sequences were analyzed using the IMGT databases and the IMGT/V-QUEST tool (version 3.2.17; Centre National de la Recherche Scientifique, LIGM, Université Montpellier 2, Montpellier, France). The following immunogenetic information were recorded for all *IGHV-IGHD-IGHJ* rearrangements: *IGHV* gene and allele usage; percentage of identity to the closest germline *IGHV* allele; VH CDR3 length and composition, including *IGHD*; and *IGHJ* gene usage and *IGHD* gene reading frame. To identify clusters of sequences with common VH CDR3 motifs, sequences were clustered based on the patterns they shared. The VH CDR3 from SMZL was aligned to the VH CDR3 sequences from a database of 28721 samples. Subsets identified as having stereotyped VH CDR3 AA sequences were those characterized by at least 50% amino acid identity and 70% similarity between stereotyped sequences, usage of the same *IGHV* gene, and identical VH CDR3 length (Bikos et al., 2012).

Fluorescence in situ hybridization (FISH). In 36 SMZL cases lacking high resolution SNP array data (11 SMZL cases belonging to the screening panel and 25 SMZL cases from the extension panel) the presence of copy number aberrations in selected candidate genes was assessed by FISH analysis, using the following probes: (a) BAC clones RP11-696E2 (*ARID1A*), RP11-177O8 (*BIRC3*), RP11-292B10 (*CREBBP*), RP11-1078O11 (*EP300*), RP11-1113D20 (*GPS2*), RP11-666C2 (*MAP3K14*), RP11-45L15 (*MLL2*), RP11-383D9 (*PTEN*), RP11-622D13 (*SPEN*), RP11-753H20 (*SWAP70*), RP11-122K3 and RP11-996H19 (*TBL1XR1*), RP11-102P5 and RP11-703G8 (*TNFAIP3*), RP11-676M2 (*TRAF3*), RP11-1077G19 (*WAC*); and (b) the commercial probes LSIBCL6 (3q27), LSID7S522-CEP7 (7q31), LSITP53 (17p13.1; Abbott). Labeled BAC probes were tested against normal control metaphases to verify the specificity of the hybridization. For each probe, at least 200 interphase cells with well-delineated fluorescent spots were examined. Nuclei were counterstained with DAPI and antifade

reagent, and signals were visualized using a BX51 microscope (Olympus). The presence of copy number abnormalities was scored when the percentage of nuclei showing the abnormality was >10%.

Immunoblotting. B cells were purified from the spleen of 3 *NOTCH2* WT and 2 *NOTCH2*-mutated SMZL patients by positive selection, using an anti-CD19 PE antibody, followed by anti-PE microbeads (Miltenyi Biotec). Whole-cell extracts were obtained from CD19-purified SMZL cells or exponentially growing cell lines in RIPA Buffer (140 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.5% sodium deoxycholate 10x, 0.1% sodium-dodecyl-sulfate + 1% Triton-X 100), according to standard protocols, and proteins were quantified using the Bradford assay. Equivalent amounts of lysates were resolved on 6% SDS-PAGE in reducing conditions. The immunoblot analysis of *NOTCH2* expression was performed using a specific anti-*NOTCH2* antibody (Bethyl Laboratories) following the manufacturer's instructions. An anti-rabbit HRP-conjugated antibody (Santa Cruz Biotechnology) was used as a secondary reagent. ERK1/2 (BD) was used as loading control. Image acquisition was performed using the ImageQuant LAS4000 software (GE Healthcare).

RNA extraction and quantitative real-time PCR (qRT-PCR). RNA was extracted using RNeasy Plus Mini kit (QIAGEN) and converted to cDNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems). qRT-PCR was performed using the 7900 HT Fast Real Time PCR System (SDS2.3 software) using commercial primers (TaqMan Gene Expression Assays; Applied Biosystems). The comparative CT method was used to calculate the expression relative to the endogenous control (Serra et al., 2011).

Statistical analysis. OS was measured from date of initial presentation to date of death (event) or last follow-up (censoring). PFS was measured from date of initial presentation to date of disease progression or death as a result of any cause, or last follow-up (censoring). Treatment-related mortality was defined as any deaths reported by investigators as probably/possibly related to SMZL treatment. Survival was analyzed by the Kaplan-Meier method and compared by the log-rank test. Categorical variables were compared by χ^2 test and Fisher's exact test when appropriate. Continuous variables were compared by the *t* test. All statistical tests were two-sided. Statistical significance was defined as *P* value < 0.05. The analysis was performed with the Statistical Package for the Social Sciences software v.19.

Accession codes. The whole-exome sequencing and copy number data reported in this paper have been deposited in dbGaP under accession no. phs000502.v1.p1 (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000502.v1.p1).

Online supplemental material. Table S1 shows the features of the 8 SMZL discovery cases analyzed by whole-exome sequencing. Table S2 reports the results of Illumina sequencing after whole-exome capture. Table S3 reports the validated somatic mutations identified by whole-exome sequencing in the SMZL discovery panel. Table S4 illustrates the segments (regions) of tumor-acquired copy number alterations identified in the SMZL discovery panel. Table S5 and Table S6 list the patient's features in the screening and extension panel, respectively. Table S7 shows the mutations identified in the screening and extension panels by targeted resequencing of candidate genes. Table S8 reports the copy number aberrations encompassing any of the 61 genes that were subjected to targeted resequencing. Table S9 relates the genetic, immunogenetic and biological features of SMZL and the mutations of MZ development genes. Online supplemental material is available at <http://www.jem.org/cgi/content/full/jem.20120904/DC1>.

We would like to thank Vladan Miljkovic and the Genomics Technologies Shared Resource of the Herbert Irving Comprehensive Cancer Center at Columbia University for hybridization of the Affymetrix SNP6.0 arrays; Laurent Farinelli and the Fasteris SA facility at Plan-les-Ouates, Geneva, Switzerland, for assistance with the whole-exome capture and sequencing procedure.

This study was supported by AIRC, Special Program Molecular Clinical Oncology, 5 x 1000, No. 10007, Milan, Italy (to G. Gaidano and to R. Foà); Progetto FIRB-Programma "Futuro in Ricerca" 2008 (to D. Rossi and S. Deaglio), PRIN 2008 (to G. Gaidano), and PRIN 2009 (to D. Rossi), MIUR, Rome, Italy; Progetto Giovani Ricercatori 2008 (to D. Rossi and S. Deaglio) and Ricerca Sanitaria Finalizzata 2008 (to G. Gaidano), Ministero della Salute, Rome, Italy; Novara-AIL Onlus, Novara, Italy (to G. Gaidano and D. Rossi); National Institutes of Health grant P01-CA092625 and a Specialized Center of Research grant from the Leukemia & Lymphoma Society (both to R. Dalla-Favera); Compagnia di San Paolo, Turin, Italy (to R. Foà); Onco-suisse grant OCS-02034-02-2007, Bern, Switzerland (to F. Bertoni); Fondazione Ticinese per la Ricerca sul Cancro, Lugano, Switzerland (to F. Bertoni); SwissLife, Zurich, Switzerland (to F. Bertoni); Nelia and Amedeo Barletta Foundation, Lausanne, Switzerland (to F. Bertoni). S. Monti and S. Cresta are being supported by fellowships from Novara-AIL Onlus, Novara, Italy. S. Rasi is being supported by a fellowship from Società Italiana di Ematologia Sperimentale. G. Fabbri is a Fellow of the American Italian Cancer Foundation.

The authors have no competing financial interests.

Author contributions: D. Rossi, R. Dalla-Favera, L. Pasqualucci, and G. Gaidano designed the study, interpreted data, and wrote the manuscript; V.T., J.W. and R.R. developed bioinformatic tools, performed bioinformatic analysis, and contributed to writing the manuscript; M.F. performed and interpreted molecular studies and contributed to writing the manuscript; A.B., S.R., V.S., R.Fa., C.C., M.G., S.C., D.P. and G.F. performed and interpreted mutational analysis; T.V., F.A., V.G. and S.D. performed and interpreted protein expression studies and contributed to writing the manuscript; S.M., C.D. and G.D. performed FISH analysis; A.H., M.M., A.R. and F.B. contributed to analysis of copy number abnormalities; R.S. performed Circos plot analysis; C.A., P.P.P., M.L., F.T., S.F., R.M., F.F., L.A., G.I., S.A.P. and R.Fo. performed diagnosis, provided well characterized pathological samples, and contributed to study design and data interpretation.

Submitted: 27 April 2012

Accepted: 18 July 2012

REFERENCES

- Bikos, V., N. Darzentas, A. Hadzidimitriou, Z. Davis, S. Hockley, A. Traverse-Glehen, P. Algara, A. Santoro, D. Gonzalez, M. Mollejo, et al. 2012. Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. *Leukemia*. 26:1638–1646. <http://dx.doi.org/10.1038/leu.2012.3>
- Brennan, A.M., and R.M. Pauli. 2001. Hajdu-Cheney syndrome: evolution of phenotype and clinical problems. *Am. J. Med. Genet.* 100:292–310.
- Calado, D.P., B. Zhang, L. Srinivasan, Y. Sasaki, J. Seagal, C. Unitt, S. Rodig, J. Kutok, A. Tarakhovsky, M. Schmidt-Suppran, and K. Rajewsky. 2010. Constitutive canonical NF- κ B activation cooperates with disruption of BLIMP1 in the pathogenesis of activated B cell-like diffuse large cell lymphoma. *Cancer Cell*. 18:580–589. <http://dx.doi.org/10.1016/j.ccr.2010.11.024>
- Chopin, M., L. Quemeneur, T. Ripich, and R. Jessberger. 2010a. SWAP-70 controls formation of the splenic marginal zone through regulating T1B-cell differentiation. *Eur. J. Immunol.* 40:3544–3556. <http://dx.doi.org/10.1002/eji.201040556>
- Chopin, M., L. Quemeneur, T. Ripich, and R. Jessberger. 2010b. SWAP-70 controls formation of the splenic marginal zone through regulating T1B-cell differentiation. *Eur. J. Immunol.* 40:3544–3556. <http://dx.doi.org/10.1002/eji.201040556>
- Chopin, M., C.A. Chacón-Martínez, and R. Jessberger. 2011. Fine tuning of IRF-4 expression by SWAP-70 controls the initiation of plasma cell development. *Eur. J. Immunol.* 41:3063–3074. <http://dx.doi.org/10.1002/eji.201141742>
- Chu, Y., J.C. Vahl, D. Kumar, K. Heger, A. Bertossi, E. Wójtowicz, V. Soberon, D. Schenten, B. Mack, M. Reutelschöfer, et al. 2011. B cells lacking the tumor suppressor TNFAIP3/A20 display impaired differentiation and hyperactivation and cause inflammation and autoimmunity in aged mice. *Blood*. 117:2227–2236. <http://dx.doi.org/10.1182/blood-2010-09-306019>
- Conze, D.B., Y. Zhao, and J.D. Ashwell. 2010. Non-canonical NF- κ B activation and abnormal B cell accumulation in mice expressing ubiquitin proteinligase-inactive c-IAP2. *PLoS Biol.* 8:e1000518. <http://dx.doi.org/10.1371/journal.pbio.1000518>
- Fabbri, G., S. Rasi, D. Rossi, V. Trifonov, H. Khiabani, J. Ma, A. Grunn, M. Fangazio, D. Capello, S. Monti, et al. 2011. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J. Exp. Med.* 208:1389–1401. <http://dx.doi.org/10.1084/jem.20110921>
- Forbes, S.A., N. Bindal, S. Bamford, C. Cole, C.Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, et al. 2011. *Nucleic Acids Res.* 39:D945–D950. <http://dx.doi.org/10.1093/nar/gkq929>
- Gruszka-Westwood, A.M., R.A. Hamoudi, E. Matutes, E. Tuset, and D. Catovsky. 2001. p53 abnormalities in splenic lymphoma with villous lymphocytes. *Blood*. 97:3552–3558. <http://dx.doi.org/10.1182/blood.V97.11.3552>
- Grzenda, A., G. Lomberg, J.S. Zhang, and R. Urrutia. 2009. Sin3: master scaffold and transcriptional corepressor. *Biochim. Biophys. Acta*. 1789:443–450. <http://dx.doi.org/10.1016/j.bbarm.2009.05.007>
- Gui, Y., G. Guo, Y. Huang, X. Hu, A. Tang, S. Gao, R. Wu, C. Chen, X. Li, L. Zhou, et al. 2011. Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. *Nat. Genet.* 43:875–878. <http://dx.doi.org/10.1038/ng.907>
- Gururajan, M., A. Simmons, T. Dasu, B.T. Spear, C. Calulot, D.A. Robertson, D.L. Wiest, J.G. Monroe, and S. Bondada. 2008. Early growth response genes regulate B cell development, proliferation, and immune response. *J. Immunol.* 181:4590–4602.
- Hampel, F., S. Ehrenberg, C. Hojer, A. Draeseke, G. Marschall-Schröter, R. Kühn, B. Mack, O. Gires, C.J. Vahl, M. Schmidt-Suppran, et al. 2011. CD19-independent instruction of murine marginal zone B-cell development by constitutive Notch2 signaling. *Blood*. 118:6321–6331. <http://dx.doi.org/10.1182/blood-2010-12-325944>
- Hermine, O., F. Lefrère, J.P. Bronowicki, X. Mariette, K. Jondeau, V. Eclache-Saudreau, B. Delmas, F. Valensi, P. Cacoub, C. Brechot, et al. 2002. Regression of splenic lymphoma with villous lymphocytes after treatment of hepatitis C virus infection. *N. Engl. J. Med.* 347:89–94. <http://dx.doi.org/10.1056/NEJMoa013376>
- Izon, D.J., J.C. Aster, Y. He, A. Weng, F.G. Karnell, V. Patriub, L. Xu, S. Bakkour, C. Rodriguez, D. Allman, and W.S. Pear. 2002. Deltex1 redirects lymphoid progenitors to the B cell lineage by antagonizing Notch1. *Immunity*. 16:231–243. [http://dx.doi.org/10.1016/S1074-7613\(02\)00271-6](http://dx.doi.org/10.1016/S1074-7613(02)00271-6)
- Kridel, R., B. Meissner, S. Rogic, M. Boyle, A. Telenius, B. Woolcock, J. Gunawardana, C. Jenkins, C. Cochrane, S. Ben-Neriah, et al. 2012. Whole transcriptome sequencing reveals recurrent NOTCH1 mutations in mantle cell lymphoma. *Blood*. 119:1963–1971. <http://dx.doi.org/10.1182/blood-2011-11-391474>
- Kuroda, K., H. Han, S. Tani, K. Tanigaki, T. Tun, T. Furukawa, Y. Taniguchi, H. Kurooka, Y. Hamada, S. Toyokuni, and T. Honjo. 2003. Regulation of marginal zone B cell development by MINT, a suppressor of Notch/RBP-J signaling pathway. *Immunity*. 18:301–312. [http://dx.doi.org/10.1016/S1074-7613\(03\)00029-3](http://dx.doi.org/10.1016/S1074-7613(03)00029-3)
- Lee, S.Y., K. Kumano, K. Nakazaki, M. Sanada, A. Matsumoto, G. Yamamoto, Y. Nannya, R. Suzuki, S. Ota, Y. Ota, et al. 2009. Gain-of-function mutations and copy number increases of Notch2 in diffuse large B-cell lymphoma. *Cancer Sci.* 100:920–926. <http://dx.doi.org/10.1111/j.1349-7006.2009.01130.x>
- Lenz, G., R.E. Davis, V.N. Ngo, L. Lam, T.C. George, G.W. Wright, S.S. Dave, H. Zhao, W. Xu, A. Rosenwald, et al. 2008. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science*. 319:1676–1679. <http://dx.doi.org/10.1126/science.1153629>
- Li, J., J. Li, X. Yang, H. Qin, P. Zhou, Y. Liang, and H. Han. 2005. The C terminus of MINT forms homodimers and abrogates MINT-mediated transcriptional repression. *Biochim. Biophys. Acta*. 1729:50–56. <http://dx.doi.org/10.1016/j.bbaexp.2005.02.001>
- Lohr, J.G., P. Stojanov, M.S. Lawrence, D. Auclair, B. Chapuy, C. Sougnez, P. Cruz-Gordillo, B. Knoechel, Y.W. Asmann, S.L. Slager, et al. 2012. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. USA*. 109:3879–3884. <http://dx.doi.org/10.1073/pnas.1121343109>

- Matsuno, K., M. Ito, K. Hori, F. Miyashita, S. Suzuki, N. Kishi, S. Artavanis-Tsakonas, and H. Okano. 2002. Involvement of a proline-rich motif and RING-H2 finger of Deltex in the regulation of Notch signaling. *Development*. 129:1049–1059.
- Matutes, E., D. Oscier, C. Montalban, F. Berger, E. Callet-Bauchu, A. Dogan, P. Felman, V. Franco, E. Iannitto, M. Mollejo, et al. 2008. Splenic marginal zone lymphoma proposals for a revision of diagnostic, staging and therapeutic criteria. *Leukemia*. 22:487–495. <http://dx.doi.org/10.1038/sj.leu.2405068>
- Moran, S.T., A. Cariappa, H. Liu, B. Muir, D. Sgroi, C. Boboila, and S. Pillai. 2007. Synergism between NF-kappa B1/p50 and Notch2 during the development of marginal zone B lymphocytes. *J. Immunol.* 179:195–200.
- Morin, R.D., M. Mendez-Lago, A.J. Mungall, R. Goya, K.L. Mungall, R.D. Corbett, N.A. Johnson, T.M. Severson, R. Chiu, M. Field, et al. 2011. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 476:298–303. <http://dx.doi.org/10.1038/nature10351>
- Mullighan, C.G., J. Zhang, L.H. Kasper, S. Lerach, D. Payne-Turner, L.A. Phillips, S.L. Heatley, L. Holmfeldt, J.R. Collins-Underwood, J. Ma, et al. 2011. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature*. 471:235–239. <http://dx.doi.org/10.1038/nature09727>
- Ngo, V.N., R.M. Young, R. Schmitz, S. Jhavar, W. Xiao, K.H. Lim, H. Kohlhammer, W. Xu, Y. Yang, H. Zhao, et al. 2011. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 470:115–119. <http://dx.doi.org/10.1038/nature09671>
- Oberoi, J., L. Fairall, P.J. Watson, J.C. Yang, Z. Czimmerer, T. Kampmann, B.T. Goult, J.A. Greenwood, J.T. Gooch, B.C. Kallenberger, et al. 2011. Structural basis for the assembly of the SMRT/NCOR core transcriptional repression machinery. *Nat. Struct. Mol. Biol.* 18:177–184. <http://dx.doi.org/10.1038/nsmb.1983>
- Pappu, B.P., and X. Lin. 2006. Potential role of CARMA1 in CD40-induced splenic B cell proliferation and marginal zone B cell maturation. *Eur. J. Immunol.* 36:3033–3043. <http://dx.doi.org/10.1002/eji.200535663>
- Parsons, D.W., M. Li, X. Zhang, S. Jones, R.J. Leary, J.C. Lin, S.M. Boca, H. Carter, J. Samayoa, C. Bettegowda, et al. 2011. The genetic landscape of the childhood cancer medulloblastoma. *Science*. 331:435–439. <http://dx.doi.org/10.1126/science.1198056>
- Pasqualucci, L., D. Dominguez-Sola, A. Chiarenza, G. Fabbri, A. Grunn, V. Trifonov, L.H. Kasper, S. Lerach, H. Tang, J. Ma, et al. 2011a. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature*. 471:189–195. <http://dx.doi.org/10.1038/nature09730>
- Pasqualucci, L., V. Trifonov, G. Fabbri, J. Ma, D. Rossi, A. Chiarenza, V.A. Wells, A. Grunn, M. Messina, O. Eliot, et al. 2011b. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* 43:830–837. <http://dx.doi.org/10.1038/ng.892>
- Perissi, V., C. Scafoglio, J. Zhang, K.A. Ohgi, D.W. Rose, C.K. Glass, and M.G. Rosenfeld. 2008. TBL1 and TBLR1 phosphorylation on regulated gene promoters overcomes dual CtBP and NCoR/SMRT transcriptional repression checkpoints. *Mol. Cell*. 29:755–766. <http://dx.doi.org/10.1016/j.molcel.2008.01.020>
- Pillai, S., and A. Cariappa. 2009. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat. Rev. Immunol.* 9:767–777. <http://dx.doi.org/10.1038/nri2656>
- Quesada, V., L. Conde, N. Villamor, G.R. Ordóñez, P. Jares, L. Bassaganyas, A.J. Ramsay, S. Beà, M. Pinyol, A. Martínez-Trillos, et al. 2012. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat. Genet.* 44:47–52. <http://dx.doi.org/10.1038/ng.1032>
- Real, P.J., V. Tosello, T. Palomero, M. Castillo, E. Hernando, E. de Stanchina, M.L. Sulis, K. Barnes, C. Sawai, I. Homminga, et al. 2009. Gamma-secretase inhibitors reverse glucocorticoid resistance in T cell acute lymphoblastic leukemia. *Nat. Med.* 15:50–58. <http://dx.doi.org/10.1038/nm.1900>
- Rinaldi, A., M. Mian, E. Chigrinova, L. Arcaini, G. Bhagat, U. Novak, P.M. Rancoita, C.P. De Campos, F. Forconi, R.D. Gascoyne, et al. 2011. Genome-wide DNA profiling of marginal zone lymphomas identifies subtype-specific lesions with an impact on the clinical outcome. *Blood*. 117:1595–1604. <http://dx.doi.org/10.1182/blood-2010-01-264275>
- Robledo, C., J.L. García, R. Benito, T. Flores, M. Mollejo, J.A. Martínez-Climent, E. García, N.C. Gutiérrez, M.A. Piris, and J.M. Hernández. 2011. Molecular characterization of the region 7q22.1 in splenic marginal zone lymphomas. *PLoS ONE*. 6:e24939. <http://dx.doi.org/10.1371/journal.pone.0024939>
- Rossi, D., S. Deaglio, D. Dominguez-Sola, S. Rasi, T. Vaisitti, C. Agostinelli, V. Spina, A. Bruscaggin, S. Monti, M. Cerri, et al. 2011. Alteration of BIRC3 and multiple other NF-κB pathway genes in splenic marginal zone lymphoma. *Blood*. 118:4930–4934. <http://dx.doi.org/10.1182/blood-2011-06-359166>
- Rossi, D., S. Rasi, G. Fabbri, V. Spina, M. Fangazio, F. Forconi, R. Marasca, L. Laurenti, A. Bruscaggin, M. Cerri, et al. 2012. Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. *Blood*. 119:521–529. <http://dx.doi.org/10.1182/blood-2011-09-379966>
- Saito, T., S. Chiba, M. Ichikawa, A. Kunisato, T. Asai, K. Shimizu, T. Yamaguchi, G. Yamamoto, S. Seo, K. Kumano, et al. 2003. Notch2 is preferentially expressed in mature B cells and indispensable for marginal zone B lineage development. *Immunity*. 18:675–685. [http://dx.doi.org/10.1016/S1074-7613\(03\)00111-0](http://dx.doi.org/10.1016/S1074-7613(03)00111-0)
- Salido, M., C. Baró, D. Oscier, K. Stamatopoulos, J. Dierlamm, E. Matutes, A. Traverse-Glehen, F. Berger, P. Felman, C. Thieblemont, et al. 2010. Cytogenetic aberrations and their prognostic value in a series of 330 splenic marginal zone B-cell lymphomas: a multicenter study of the Splenic B-Cell Lymphoma Group. *Blood*. 116:1479–1488. <http://dx.doi.org/10.1182/blood-2010-02-267476>
- Santos, M.A., L.M. Sarmento, M. Rebelo, A.A. Doce, I. Maillard, A. Dumortier, H. Neves, F. Radtke, W.S. Pear, L. Parreira, and J. Demengeot. 2007. Notch1 engagement by Delta-like-1 promotes differentiation of B lymphocytes to antibody-secreting cells. *Proc. Natl. Acad. Sci. USA*. 104:15454–15459. <http://dx.doi.org/10.1073/pnas.0702891104>
- Sasaki, Y., D.P. Calado, E. Derudder, B. Zhang, Y. Shimizu, F. Mackay, S. Nishikawa, K. Rajewsky, and M. Schmidt-Supprian. 2008. NIK overexpression amplifies, whereas ablation of its TRAF3-binding domain replaces BAFF:BAFF-R-mediated survival signals in B cells. *Proc. Natl. Acad. Sci. USA*. 105:10883–10888. <http://dx.doi.org/10.1073/pnas.0805186105>
- Serra, S., A.L. Horenstein, T. Vaisitti, D. Brusa, D. Rossi, L. Laurenti, G. D'Arena, M. Coscia, C. Tripodo, G. Inghirami, S.C. Robson, G. Gaidano, F. Malavasi, and S. Deaglio. 2011. CD73-generated extracellular adenosine in chronic lymphocytic leukemia creates local conditions counteracting drug-induced cell death. *Blood*. 118:6141–6152.
- Simpson, M.A., M.D. Irving, E. Asilmaz, M.J. Gray, D. Dafou, F.V. Elmslie, S. Mansour, S.E. Holder, C.E. Brain, B.K. Burton, et al. 2011. Mutations in NOTCH2 cause Hajdu-Cheney syndrome, a disorder of severe and progressive bone loss. *Nat. Genet.* 43:303–305. <http://dx.doi.org/10.1038/ng.779>
- Suarez, F., O. Lortholary, O. Hermine, and M. Lecuit. 2006. Infection-associated lymphomas derived from marginal zone B cells: a model of antigen-driven lymphoproliferation. *Blood*. 107:3034–3044. <http://dx.doi.org/10.1182/blood-2005-09-3679>
- Swerdlow, S.H., E. Campo, N.L. Harris, E.S. Jaffe, S.A. Pileri, H. Stein, J. Thiele, and J.W. Vardiman. 2008. WHO classification of tumours of haematopoietic and lymphoid tissues, Fourth Edition. Lyon, France.
- Tiacci, E., G. Schiavoni, F. Forconi, A. Santi, L. Trentin, A. Ambrosetti, D. Cecchini, E. Sozzi, P. Francia di Celle, C. Di Bello, et al. 2012. Simple genetic diagnosis of hairy cell leukemia by sensitive detection of the BRAF-V600E mutation. *Blood*. 119:192–195. <http://dx.doi.org/10.1182/blood-2011-08-371179>
- Traverse-Glehen, A., L. Baseggio, G. Salles, P. Felman, and F. Berger. 2011. Splenic marginal zone B-cell lymphoma: a distinct clinico-pathological and molecular entity. Recent advances in ontogeny and classification. *Curr. Opin. Oncol.* 23:441–448. <http://dx.doi.org/10.1097/CCO.0b013e328349ab8d>
- Trøen, G., I. Wlodarska, A. Warsame, S. Hernández Llodrà, C. De Wolf-Peeters, and J. Delabie. 2008. NOTCH2 mutations in marginal zone lymphoma. *Haematologica*. 93:1107–1109.

- VanderWielen, B.D., Z. Yuan, D.R. Friedmann, and R.A. Kovall. 2011. Transcriptional repression in the Notch pathway: thermodynamic characterization of CSL-MINT (Msx2-interacting nuclear target protein) complexes. *J. Biol. Chem.* 286:14892–14902. <http://dx.doi.org/10.1074/jbc.M110.181156>
- Wang, L., M.S. Lawrence, Y. Wan, P. Stojanov, C. Sougnez, K. Stevenson, L. Werner, A. Sivachenko, D.S. DeLuca, L. Zhang, et al. 2011a. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N. Engl. J. Med.* 365:2497–2506. <http://dx.doi.org/10.1056/NEJMoa1109016>
- Wang, N.J., Z. Sanborn, K.L. Arnett, L.J. Bayston, W. Liao, C.M. Proby, I.M. Leigh, E.A. Collisson, P.B. Gordon, L. Jakkula, et al. 2011b. Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proc. Natl. Acad. Sci. USA.* 108:17761–17766. <http://dx.doi.org/10.1073/pnas.1114669108>
- Warsame, A.A., H.C. Aasheim, K. Nustad, G. Trøen, A. Tierens, V. Wang, U. Randen, H.P. Dong, S. Heim, A. Brech, and J. Delabie. 2011. Splenic marginal zone lymphoma with VH1-02 gene rearrangement expresses poly- and self-reactive antibodies with similar reactivity. *Blood.* 118:3331–3339. <http://dx.doi.org/10.1182/blood-2011-03-341651>
- Watkins, A.J., Y. Huang, H. Ye, E. Chanudet, N. Johnson, R. Hamoudi, H. Liu, G. Dong, A. Attygalle, E.D. McPhail, et al. 2010. Splenic marginal zone lymphoma: characterization of 7q deletion and its value in diagnosis. *J. Pathol.* 220:461–474.
- Weng, A.P., A.A. Ferrando, W. Lee, J.P. Morris IV, L.B. Silverman, C. Sanchez-Irizarry, S.C. Blacklow, A.T. Look, and J.C. Aster. 2004. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science.* 306:269–271. <http://dx.doi.org/10.1126/science.1102160>
- Wiegand, K.C., S.P. Shah, O.M. Al-Agha, Y. Zhao, K. Tse, T. Zeng, J. Senz, M.K. McConechy, M.S. Anglesio, S.E. Kalloger, et al. 2010. ARID1A mutations in endometriosis-associated ovarian carcinomas. *N. Engl. J. Med.* 363:1532–1543. <http://dx.doi.org/10.1056/NEJMoa1008433>
- Xie, P., L.L. Stunz, K.D. Larison, B. Yang, and G.A. Bishop. 2007. Tumor necrosis factor receptor-associated factor 3 is a critical regulator of B cell homeostasis in secondary lymphoid organs. *Immunity.* 27:253–267. <http://dx.doi.org/10.1016/j.immuni.2007.07.012>
- Yan, Q., Y. Huang, A.J. Watkins, S. Kocalkowski, N. Zeng, R.A. Hamoudi, P.G. Isaacson, L. de Leval, A. Wotherspoon, and M.Q. Du. 2012. BCR and TLR signaling pathways are recurrently targeted by genetic changes in splenic marginal zone lymphomas. *Haematologica.* 97:595–598. <http://dx.doi.org/10.3324/haematol.2011.054080>
- Yuan, J.S., P.C. Kousis, S. Suliman, I. Visan, and C.J. Guidos. 2010. Functions of notch signaling in the immune system: consensus and controversies. *Annu. Rev. Immunol.* 28:343–365. <http://dx.doi.org/10.1146/annurev.immunol.021908.132719>
- Zweifel, M.E., D.J. Leahy, and D. Barrick. 2005. Structure and Notch receptor binding of the tandem WWE domain of Deltex. *Structure.* 13:1599–1611. <http://dx.doi.org/10.1016/j.str.2005.07.015>

SUPPLEMENTAL MATERIAL

Rossi et al., <http://www.jem.org/cgi/content/full/jem.20120904/DC1>

Table S1, which is presented as an Excel file, shows the features of the 8 SMZL discovery cases analyzed by whole-exome sequencing.

Table S2, which is presented as an Excel file, reports the results of Illumina sequencing after whole-exome capture.

Table S3, which is presented as an Excel file, reports the validated somatic mutations identified by whole-exome sequencing in the SMZL discovery panel.

Table S4, which is presented as an Excel file, illustrates the segments (regions) of tumor-acquired copy number alterations identified in the SMZL discovery panel.

Table S5, which is presented as an Excel file, and Table S6 list the patient's features in the screening and extension panel, respectively.

Table S7, which is presented as an Excel file, shows the mutations identified in the screening and extension panels by targeted resequencing of candidate genes.

Table S8, which is presented as an Excel file, reports the copy number aberrations encompassing any of the 61 genes that were subjected to targeted resequencing.

Table S9, which is presented as an Excel file, relates the genetic, immunogenetic, and biological features of SMZL and the mutations of MZ development genes.

Review Article

The Spectrum of Genetic Defects in Chronic Lymphocytic Leukemia

Davide Rossi,¹ Marco Fangazio¹ and Gianluca Gaidano¹

¹ Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy

Correspondence to: Davide Rossi, M.D., Ph.D., Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy; Tel: +39-0321-660698; Fax: +39-0321-620421; E-mail rossidav@med.unipmn.it

Competing interests: The authors have declared that no competing interests exist.

Published: November 13, 2012

Received: October 26, 2012

Accepted: November 9, 2012

Citation: Mediterr J Hematol Infect Dis 2012, 4(1): e2012076, DOI: 10.4084/MJHID.2012.076

This article is available from: <http://www.mjhid.org/article/view/11209>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. Chronic lymphocytic leukemia (CLL) is the most common leukemia in the Western world and shows a remarkable heterogeneity in the clinical course. Understand the genetic basis of CLL may help in clarifying the molecular bases of this clinical heterogeneity. Recurrent chromosomal aberrations at 13q14, 12q, 11q22-q23 and 17p13, and *TP53* mutations are the first genetic lesions identified as drivers of the disease. While some of these lesions are associated with poor outcome (17p13 deletion, *TP53* mutations and, to a lesser extent, 11q22-q23 deletion) others are linked to a favorable course (13q14 deletion as sole aberration). Recently, next generation sequencing has revealed additional recurrent alterations in CLL targeting the *NOTCH1*, *SF3B1*, and *BIRC3* genes. *NOTCH1*, *SF3B1*, and *BIRC3* lesions provide: *i*) new insights on the mechanisms of leukemogenesis, tumor progression and chemoresistance in this leukemia; *ii*) new biomarkers for the identification of poor risk patients, having individually shown correlations with survival in CLL; and *iii*) new therapeutic targets, especially in the setting of high risk disease. This review will summarize the most important genetic aberrations in CLL and how our improved knowledge of the genome of leukemic cells may translate into improved patients' management.

Introduction. In Western countries, chronic lymphocytic leukemia (CLL) is the most frequent mature B-cell malignancy.^{1,2} The course CLL ranges from very indolent, with a nearly normal life expectancy, to rapidly progressive leading to early death.³⁻⁸ Understand the genetic basis of CLL may help in clarifying the molecular determinants of this clinical heterogeneity and improve patients' prognostication.

Recurrent chromosomal aberrations at 13q14, 12q, 11q22-q23 and 17p13 are the first genetic lesions

identified as drivers of the disease, and has enabled the construction of a hierarchical model of cytogenetic abnormalities that correlates with outcome.⁹ Cytogenetic lesions, however, may not entirely explain the genetic basis of CLL clinical heterogeneity, as documented by the contribution of *TP53* mutation assessment in identifying high risk patients.⁹ The recent major improvements in massive parallel sequencing technologies have provided an opportunity to examine the CLL genome, allowing for the identification of

genomic alterations underlying the disease and for the discovery of new therapeutic targets and clinically predictive biomarkers such as *NOTCH1*, *SF3B1* and *BIRC3*.¹⁰⁻¹⁶

Prevalence of Genetic Lesions at Different CLL Clinical Phases. During its history, CLL may proceed through distinct clinical phases, ranging from a pre-malignant condition known as monoclonal B-cell lymphocytosis (MBL), to overt CLL, and even transformation into an aggressive lymphoma (Richter syndrome).^{1,2}

Similarly to other pre-malignant conditions, also MBL frequently harbor genetic changes that can be found in the overt disease. In MBL, 13q14 deletion occurs at the same prevalence as in overt CLL (~40-50% of cases), even when the number of circulating monoclonal CLL-like cells is extremely small, thus indicating that this lesion occurs early during the natural history of the disease.¹⁷⁻²¹ What distinguishes MBL from CLL is the rate of occurrence of genetic lesions that are considered secondary events and that associate with poor outcome in this leukemia.^{19,21} In clinical MBL, 11q22-q23 deletion, 17p13 deletion and mutations of *BIRC3*, *TP53*, *NOTCH1* and *SF3B1* may be observed in ~1-3% of cases, a prevalence that is significantly lower than that of CLL (Table I).^{17,19,21,22} High risk cytogenetic abnormalities have been occasionally described also in low count MBL, but the biological implications of this observation are currently unknown.^{18,20}

When CLL is overt, three major clinical phases can be envisaged, including: *i*) newly diagnosed CLL; *ii*) progressive CLL; and *iii*) relapsed and fludarabine-refractory CLL (Table I).² *TP53* abnormalities, including mutations and 17p13 deletions, are observed in ~5-10% newly diagnosed CLL, in ~10% progressive CLL requiring first treatment,^{9,23-32} and in ~40-50% relapsed and fludarabine-refractory CLL,³³⁻³⁵ thus representing the most frequent lesions in this high risk clinical condition. Deletion of 11q22-q23 occurs in 10-15% in newly diagnosed CLL,^{9,36} while its prevalence raises to 20-25% at the time of first treatment and 25-30% at fludarabine-refractoriness.^{24,29,33,34} Mutations of *ATM*, which is included in the minimal common region of deletion on 11q22-q23, have been shown to

be present in 12% of newly diagnosed patients and in 15% progressive CLL requiring first treatment.³⁷⁻⁴⁰ By combining mutations and deletions, genetic lesions of *ATM* occur in 25% of diagnostic samples of CLL and in 37% cases requiring first treatment.³⁷⁻⁴⁰ These frequencies make *ATM* alterations the most common genetic lesions predicting poor outcome at CLL presentation and treatment requirement.

Among the novel genetic alterations disclosed by whole genome/exome sequencing, *NOTCH1*, *SF3B1* and *BIRC3* lesions follows the same distribution across CLL clinical phases as *TP53* and *ATM* abnormalities (Table 1). *NOTCH1* mutations recur in ~10% unselected newly diagnosed CLL while their prevalence increases to 15-20% in progressive and relapsed cases.^{10,11,14} *SF3B1* mutations have been identified in ~7% unselected newly diagnosed CLL, while their prevalence rises to 17% in relapsed and fludarabine-refractory patients.^{12,13,16} *BIRC3* lesions occur at low rate (4% of cases) in unselected newly diagnosed CLL, while are enriched among relapsed and fludarabine-refractory CLL (24% of cases).¹⁵ Because of their recent identification and the lack of information from large clinical trials, the precise rate of occurrence of *NOTCH1*, *BIRC3*, and *SF3B1* lesions at the time of first treatment requirement still remains to be clarified.

Within the spectrum of the various aspects of CLL, Richter syndrome (RS) is the most aggressive clinical phenotype because of the combined effect of chemoresistance and rapid disease kinetics. The clinical behavior of RS is strongly related to its genetic background (Table I). The high rate of *TP53* abnormalities, which occur in ~60% cases and represent the most frequent genetic lesion at the time of transformation, accounts for the chemoresistance that is very common in RS.⁴¹ *NOTCH1* mutations are the second most frequent genetic lesion in RS, where they occur in ~30% of cases.¹⁰ Among the other high risk genetic lesions, *ATM* abnormalities, *BIRC3* genetic lesions and *SF3B1* mutations that are otherwise enriched at the time of chemorefractoriness are rare or absent in RS, thus strengthening the notion that RS is molecularly distinct from chemorefractory progression without transformation.^{13,14,41}

Table 1. Prevalence of CLL recurrent lesion stratified according the disease phase

	<i>TP53</i> disruption	del 11q22-q23	<i>NOTCH1</i> mutations	<i>SF3B1</i> mutations	<i>BIRC3</i> disruption
MBL	1-2%	0-3%	3%	1-2%	0
Diagnosis	5-10%	10-15%	8-11%	4-7%	0.05
First treatment	10-11%	20-25%	10-15%	17%	n.a.
Chemorefractoriness	40-50%	25-30%	15-20%	17%	25%
Richter Syndrome	50-60%	10%	30-40%	6%	0
CLL, Chronic lymphocytic leukemia;			MBL, Monoclonal B-cell lymphocytosis		

TP53 Abnormalities. The tumor suppressor gene *TP53* codes for a central regulator of the DNA-damage-response pathway, and its activation leads to cell-cycle arrest, DNA repair, apoptosis, or senescence through both transcription-dependent and transcriptional-independent activities.⁴² Among CLL harboring *TP53* abnormalities, mutations of *TP53* co-occurred with deletion of the corresponding locus in ~70% of cases, consistent with a dual hit mechanism of inactivation.⁴³ The remaining ~30% of cases have 17p13 deletion in the absence of *TP53* mutations (~20%), or *TP53* mutations in the absence of 17p13 deletion (~10%). *TP53* mutations are mainly represented by missense substitutions targeting the DNA-binding domain, while the remaining are truncating lesions. Mutations either directly disrupt the DNA binding domain of *TP53* or cause conformational changes of the TP53 protein, thus leading to severely impaired TP53 function.^{43,44}

The clinical importance of *TP53* abnormalities in CLL is tightly linked to their close association with poor outcome and refractoriness, as documented by a number of observational studies and prospective trials led in both the chemotherapy and immuno-chemotherapy era. Among unselected newly diagnosed CLL, patients harboring 17p13 deletion have an estimated median overall survival (OS) of only 3-5 years.^{9,45} However, it is important to stress that there is a small subgroup of patients with 17p13 deletion (and mostly mutated immunoglobulin genes) who may exhibit stable disease for years without treatment indications.⁴⁵

The outcome of patients with 17p13 deletion and need for treatment is very poor. With the most effective regimen available today for CLL, i.e. FCR (fludarabine-cyclophosphamide-rituximab), patients with 17p13 deletion have a poor response (5% of complete response vs ~50% in non 17p13 deleted CLL), a short progression free survival (PFS) (11.2 months vs 51.8 months) and OS (38.1% at 36 months).²⁹ This is in line with the established importance of the wild-type TP53 protein in mediating the cytotoxicity of DNA-damaging agents including purine analogs.

A number of prospective studies suggest that, in addition to 17p13 deletion, also *TP53* mutations, even in the absence of 17p13 deletion, predict poor outcome in CLL. In the GCLLSG CLL4 trial (fludarabine vs fludarabine-cyclophosphamide) no complete response were observed in *TP53* mutated CLL, and the median PFS (23.3 vs 62.2 months) and OS (29.2 vs 84.6 months) were significantly shorter in the group with *TP53* mutation.³⁰ In the GCLLSG CLL8 trial (fludarabine-cyclophosphamide vs FCR), patients with *TP53* mutations showed the lowest complete response

and overall response rates (6.9% vs. 36.4% and 62.1% vs. 95.3%), translating into shorter PFS (12.4 months vs. 45 months) and OS (39.3 months vs not reached in all other patients).⁴⁴ In the UK LRF CLL4 trial (chlorambucil vs fludarabine vs fludarabine-cyclophosphamide), the complete response rate of *TP53* mutated patients was only 5% with a 5-years PFS of 5% and a 5-years OS of 20%.³¹

Based on these data, 17p13 deletion is the sole cytogenetic abnormality that is recommended to be tested by FISH in CLL patients requiring treatment.² Since CLL with *TP53* mutations experience poor prognosis regardless of the presence of 17p13 deletion, the *TP53* mutation analysis should be integrated into the evaluation of CLL patients before treatment initiation.⁴⁴ CLL patients carrying *TP53* alterations, regardless of whether mutated or deleted, should be redirected to different therapeutic regimens compared to the standard chemo/chemoimmunotherapies.^{2,33,35,44,46}

NOTCH1 Mutations. The *NOTCH1* gene encodes a heterodimeric transmembrane protein that functions as a ligand-activated transcription factor with a high conserved pathway.⁴⁷ When the NOTCH1 receptor interacts with its ligands through the extracellular subunit, two consecutive proteolytic cleavages of the protein are initiated and lead to pathway activation.^{47,48} The S2 cleavage in the heterodimerization domain is performed by ADAM10, and is followed by the S3 cleavage by the γ -secretase complex. Upon activation the cleaved intracellular portion of NOTCH1 (ICN) translocates into the nucleus where it modifies the expression of target genes, including the *MYC* oncogene. As a transcriptional factor, *NOTCH1* plays an important role in a number of cellular functions during embryogenesis and in self-renewing tissues of the adult organism, including maintenance of stem cells, cell fate specification, proliferation, and apoptosis.⁴⁸ One of the mechanisms of the NOTCH1 signal suppression is operated through the PEST [proline (P), glutamic acid (E), serine (S), and threonine (T) rich] domain that directs the activated NOTCH1 towards proteosomal degradation.⁴⁷ A major role of *NOTCH1* in lymphoid cells in the adult organism is the commitment of hematopoietic progenitors to differentiate toward T lineage.⁴⁹ Conversely, in mature B-lymphocytes, *NOTCH1* signaling promotes terminal differentiation to antibody-secreting cells.⁵⁰

NOTCH1 mutations were the first molecular lesion identified through massive parallel next generation sequencing in CLL by two independent groups.^{10,11} *NOTCH1* mutations are significantly more frequent in CLL with unmutated, rather than mutated,

immunoglobulin genes, are significantly enriched in CLL harboring trisomy 12, and identify a distinct clinico-molecular subgroup of CLL with deregulated cell cycle and short survival.^{10-12,14,16,51-53}

NOTCH1 mutations in CLL mainly clusters within a hotspot in exon 34, and are commonly represented by a single 2-bp deletion (c.7544_7545delCT) that accounts for ~80-95% of all *NOTCH1* mutations in this leukemia (Figure 1).^{10-12,14,16,51-53} The predicted functional consequence of *NOTCH1* mutations in CLL is the disruption of the C-terminal PEST domain resulting in activated NOTCH1 protein, impaired degradation and accumulation, and sustaining deregulated signaling.¹¹ Consistent with this notion, a number of cellular pathways are specifically altered in CLL harboring *NOTCH1* mutations.^{11,52}

Beside their pathogenetic role, *NOTCH1* mutations also represent a new biomarker for the identification of poor risk CLL patients. *NOTCH1* mutated patients have a rapidly progressive disease and a significantly shorter survival probability (21-45% at 10 years) compared to *NOTCH1* wild type cases (56-66% at 10 years).^{10,11,14} The poor prognosis associated with *NOTCH1* mutations in CLL may be explained, at least in part, by a substantial risk (~40-50%) of developing Richter syndrome.^{10,11,14}

NOTCH1 is a potential therapeutic target in CLL. Treatment with γ -secretase inhibitors induces apoptosis of CLL cells by inhibiting the enzymatic S3 cleavage necessary for NOTCH1 activation.^{47,54,55} However, the limitations due to toxicity of γ -secretase inhibitors in the clinical setting suggest that alternative strategies may be needed for the therapeutic targeting of NOTCH1.

SF3B1 Mutations. The spliceosome machinery, a complex of five small nuclear ribonucleoproteins (snRNPs), contributes to the formation of mature mRNA through the removal of introns in the precursor messenger RNA (pre-mRNA) of protein-encoding genes, and is involved in both normal and alternative splicing.⁵⁶ Alternative splicing can generate numerous transcript variants from a single gene, contributing to genomic complexity and potentially to cancer.⁵⁷

SF3B1 is a core component of the U2 snRNP that recognizes the 3' splice site at the intron-exon junctions.^{56,58-61} Structurally, the SF3B1 protein has two well-defined regions: *i*) the N-terminal amino acid region which contains several protein-binding motifs and functions as a scaffold to facilitate its interaction with other splicing factors; and *ii*) the C-terminal region which contains 22 non-identical tandem repeats of the HEAT motif that meander around the SF3b complex.^{56,58-61}

Whole genome/exome sequencing technologies

allowed for the identification of *SF3B1* as a recurrently mutated gene in CLL.^{12,13,16} *SF3B1* mutations in CLL cluster in selected HEAT repeats of the SF3B1 protein, target a number of hotspots (codons 662, 666, 700, 742), and are generally represented by missense substitutions (Figure 1).^{12,13,16} Notably, an identical spectrum of *SF3B1* mutations has been identified in other hematopoietic tumors of the myeloid compartment.⁶²

The precise biological consequences of *SF3B1* mutations in CLL are currently unknown. However, the clustering of *SF3B1* mutations within the HEAT domains suggests that they are selected to modify SF3B1 interactions with other proteins of the spliceosome complex, thus resulting in deregulated normal and alternative mRNA splicing.^{12,16}

Consistent with their accumulation in the more advanced phases of the disease, *SF3B1* mutated patients show a significantly shorter overall survival (34-48% at 10 years) compared to wild type cases (60-73% at 10-years).^{12,13,16}

BIRC3 Abnormalities. In CLL, activation of the NF- κ B pathway contributes to the acquisition of a chemorefractory clinical phenotype and correlates with poor outcome.⁶³⁻⁶⁷ The Baculoviral IAP repeat containing 3 (*BIRC3*) gene is one of the components of a protein complex that negatively regulates the MAP3K14 serin-threonine kinase, the downstream activator of non-canonical NF- κ B signaling.⁶³⁻⁶⁶

BIRC3 was found to be recurrently disrupted by mutations, deletions, or a combination of mutations and deletions in CLL patients.¹⁵ *BIRC3* inactivating mutations and a fraction of *BIRC3* deletions cause a truncation of the C-terminal RING domain of the BIRC3 protein, essential for ubiquitination, and the following proteasome degradation, of MAP3K14, and drives constitutive non-canonical NF- κ B activation (Figure 1).¹⁵

The *BIRC3* gene maps to 11q22.2, approximately 6Mb centromeric to the *ATM* locus. The identification of *BIRC3* involvement in CLL might be important for elucidating the molecular genetics of 11q22-q23 deletion, a frequent cytogenetic abnormality predictive of poor outcome. In fact, although *ATM* has been regarded as the relevant gene of this chromosomal abnormality, biallelic inactivation of *ATM* does not exceed ~30% of cases with 11q22-q23 deletion.³⁶⁻³⁹ The presence of an additional tumor suppressor in the 11q22-q23 region has been postulated,⁴⁰ and *BIRC3* implicates a suitable candidate.

From a clinical standpoint, *BIRC3* lesions contribute to clinical aggressiveness and fludarabine refractoriness in CLL.¹⁵ Indeed, *BIRC3* lesions identify a subgroup of CLL displaying poor survival

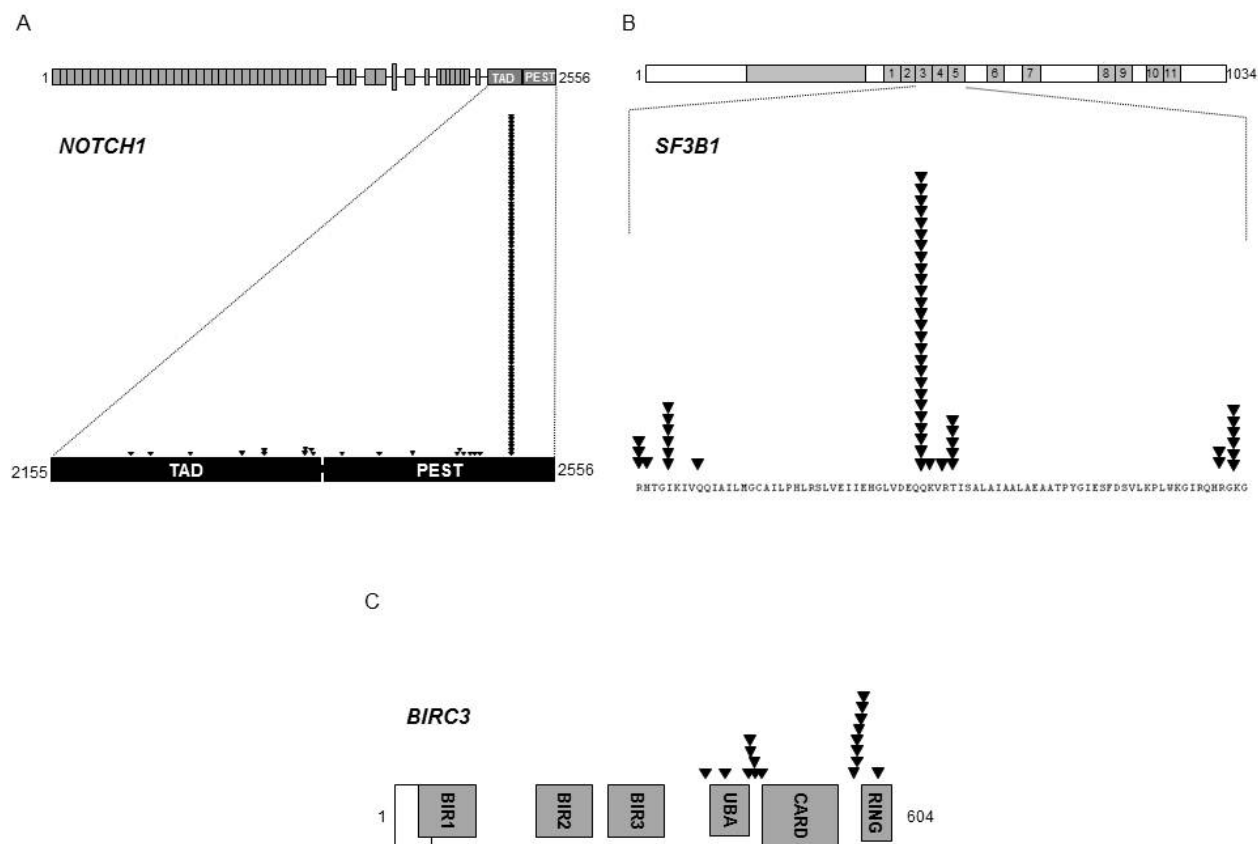


Figure 1. *NOTCH1*, *SF3B1*, and *BIRC3* mutation distribution in CLL. Schematic representation of the human *NOTCH1* (panel A), *SF3B1* (panel B), and *BIRC3* (panel C) proteins, with their key functional domains. Symbols indicate the position of the mutations. Mutations are from the Novara CLL mutation database and from the COSMIC database (v61).

(median 3.1 years) similar to that associated with *TP53* abnormalities.¹⁵

In CLL, fludarabine refractoriness may be explained by *TP53* disruption in ~40% of patients, while ~60% high risk CLL do not present *TP53* abnormalities.³⁴ Intriguingly the distribution of *BIRC3* disruption and *TP53* abnormalities is mutually exclusive and *BIRC3* abnormalities can recapitulate the genetics of ~40% chemorefractory and *TP53* wild type CLL.

On these bases, *BIRC3* disruption may contribute to expand the panel of biomarkers for the early identification of chemorefractory cases.¹⁵ In addition, *BIRC3* abnormalities provide a molecular rationale for targeting NF- κ B in poor risk and chemorefractory CLL. NF- κ B inhibitors are under development in CLL and pre-clinical findings suggest that these compounds might be active against chemoresistant CLL clones.^{67,68}

References:

1. Swerdlow, S.H., Campo, E., Harris, N.L., Jaffe, E.S., Pileri, S.A., Stein, H., Thiele, J., Vardiman, J.W. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, Fourth Edition. (Lyon, France, 2008).
2. Hallek, M., Cheson, B.D., Catovsky, D., Caligaris-Cappio, F., Dighiero, G., Dohner, H., Hillmen, P., Keating, M.J., Montserrat, E., Rai, K.R., Kipps, T.J. and International Workshop on Chronic Lymphocytic Leukemia, Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood 2008 111: 5446-5456. <http://dx.doi.org/10.1182/blood-2007-06-093906> PMID:18216293 PMCID:2972576
3. Damle, R.N., Wasil, T., Fais, F., Ghiotto, F., Valetto, A., Allen, S.L., Buchbinder, A., Budman, D., Dittmar, K., Kolitz, J., Lichtman, S.M., Schulman, P., Vinciguerra, V.P., Rai, K.R., Ferrarini, M. and Chiorazzi, N., Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999 94: 1840-1847. PMID:10477712
4. Hamblin, T.J., Davis, Z., Gardiner, A., Oscier, D.G. and Stevenson, F.K., Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood 1999 94: 1848-1854. PMID:10477713
5. Oscier, D.G., Gardiner, A.C., Mould, S.J., Glide, S., Davis, Z.A., Ibbotson, R.E., Corcoran, M.M., Chapman, R.M., Thomas, P.W., Copplestone, J.A., Orchard, J.A. and Hamblin, T.J., Multivariate analysis of prognostic factors in CLL: clinical stage, IGVH gene mutational status, and loss or mutation of the p53 gene are independent prognostic factors. Blood 2002 100: 1177-1184. PMID:12149195
6. Crespo, M., Bosch, F., Villamor, N., Bellosillo, B., Colomer, D.,

- Rozman, M., Marce, S., Lopez-Guillermo, A., Campo, E. and Montserrat, E., ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. *N Engl J Med* 2003 348: 1764-1775. <http://dx.doi.org/10.1056/NEJMoa023143> PMID:12724482
7. Vasconcelos, Y., Davi, F., Levy, V., Oppezio, P., Magnac, C., Michel, A., Yamamoto, M., Pritsch, O., Merle-Beral, H., Maloum, K., Ajchenbaum-Cymbalista, F. and Dighiero, G., Binet's staging system and VH genes are independent but complementary prognostic indicators in chronic lymphocytic leukemia. *J Clin Oncol* 2003 21: 3928-3932. <http://dx.doi.org/10.1200/JCO.2003.02.134> PMID:14581416
8. Rassenti, L.Z., Huynh, L., Toy, T.L., Chen, L., Keating, M.J., Gribben, J.G., Neuberg, D.S., Flinn, I.W., Rai, K.R., Byrd, J.C., Kay, N.E., Greaves, A., Weiss, A. and Kipps, T.J., ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *N Engl J Med* 2004 351: 893-901. <http://dx.doi.org/10.1056/NEJMoa040857> PMID:15329427
9. D'haer, H., Stilgenbauer, S., Benner, A., Leupolt, E., Krober, A., Bullinger, L., Dohner, K., Bentz, M. and Lichter, P., Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med* 2000 343: 1910-1916. <http://dx.doi.org/10.1056/NEJM200012283432602> PMID:11136261
10. Fabbri, G., Rasi, S., Rossi, D., Trifonov, V., Khiabani, H., Ma, J., Grun, A., Fangazio, M., Capello, D., Monti, S., Cresta, S., Gargiulo, E., Forconi, F., Guarini, A., Arcaini, L., Paulli, M., Laurenti, L., Larocca, L.M., Marasca, R., Gattei, V., Oscier, D., Berton, F., Mullighan, C.G., Foa, R., Pasqualucci, L., Rabadan, R., Dalla-Favera, R. and Gaidano, G., Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med* 2011 208: 1389-1401. <http://dx.doi.org/10.1084/jem.20110921> PMID:21670202 PMID:3135373
11. Puente, X.S., Pinyol, M., Quesada, V., Conde, L., Ordóñez, G.R., Villamor, N., Escarot, G., Jares, P., Bea, S., González-Díaz, M., Bassaganyas, L., Baumann, T., Juan, M., López-Guerra, M., Colomer, D., Tubio, J.M., López, C., Navarro, A., Tarnador, C., Aymerich, M., Rozman, M., Hernández, J.M., Puente, D.A., Freije, J.M., Velasco, G., Gutiérrez-Fernández, A., Costa, D., Carro, A., Gujarró, S., Enjuanes, A., Hernández, L., Yague, J., Nicolás, P., Romeo-Casabona, C.M., Himmelbauer, H., Castillo, E., Dohm, J.C., de Sanjose, S., Piris, M.A., de Alava, E., San Miguel, J., Royo, R., Gelpi, J.L., Torrents, D., Orozco, M., Pisano, D.G., Valencia, A., Guigo, R., Bayes, M., Heath, S., Gut, M., Klatt, P., Marshall, J., Raine, K., Stebbings, L.A., Futreal, P.A., Stratton, M.R., Campbell, P.J., Gut, I., López-Guillermo, A., Estivill, X., Montserrat, E., López-Otin, C. and Campo, E., Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* 2011 475: 101-105. <http://dx.doi.org/10.1038/nature10113> PMID:21642962 PMID:3322590
12. Quesada, V., Conde, L., Villamor, N., Ordóñez, G.R., Jares, P., Bassaganyas, L., Ramsay, A.J., Bea, S., Pinyol, M., Martínez-Trillos, A., López-Guerra, M., Colomer, D., Navarro, A., Baumann, T., Aymerich, M., Rozman, M., Delgado, J., Gine, E., Hernández, J.M., González-Díaz, M., Puente, D.A., Velasco, G., Freije, J.M., Tubio, J.M., Royo, R., Gelpi, J.L., Orozco, M., Pisano, D.G., Zamora, J., Vazquez, M., Valencia, A., Himmelbauer, H., Bayes, M., Heath, S., Gut, M., Gut, I., Estivill, X., López-Guillermo, A., Puente, X.S., Campo, E. and López-Otin, C., Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* 2012 44: 47-52. <http://dx.doi.org/10.1038/ng.1032> PMID:22158541
13. Rossi, D., Bruscaggin, A., Spina, V., Rasi, S., Khiabani, H., Messina, M., Fangazio, M., Vaisitti, T., Monti, S., Chiaretti, S., Guarini, A., Del Giudice, I., Cerri, M., Cresta, S., Deambrogi, C., Gargiulo, E., Gattei, V., Forconi, F., Berton, F., Deaglio, S., Rabadan, R., Pasqualucci, L., Foa, R., Dalla-Favera, R. and Gaidano, G., Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* 2011 118: 6904-6908. <http://dx.doi.org/10.1182/blood-2011-08-373159> PMID:22039264
14. Rossi, D., Rasi, S., Fabbri, G., Spina, V., Fangazio, M., Forconi, F., Marasca, R., Laurenti, L., Bruscaggin, A., Cerri, M., Monti, S., Cresta, S., Fama, R., De Paoli, L., Bulian, P., Gattei, V., Guarini, A., Deaglio, S., Capello, D., Rabadan, R., Pasqualucci, L., Dalla-Favera, R., Foa, R. and Gaidano, G., Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. *Blood* 2012 119: 521-529. <http://dx.doi.org/10.1182/blood-2011-09-379966> PMID:22077063
15. Rossi, D., Fangazio, M., Rasi, S., Vaisitti, T., Monti, S., Cresta, S., Chiaretti, S., Del Giudice, I., Fabbri, G., Bruscaggin, A., Spina, V., Deambrogi, C., Marinelli, M., Fama, R., Greco, M., Daniele, G., Forconi, F., Gattei, V., Berton, F., Deaglio, S., Pasqualucci, L., Guarini, A., Dalla-Favera, R., Foa, R. and Gaidano, G., Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. *Blood* 2012 119: 2854-2862. <http://dx.doi.org/10.1182/blood-2011-12-395673> PMID:22308293
16. Wang, L., Lawrence, M.S., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D.S., Zhang, L., Zhang, W., Vartanov, A.R., Fernandes, S.M., Goldstein, N.R., Folco, E.G., Cibulskis, K., Tesar, B., Sievers, Q.L., Shefler, E., Gabriel, S., Hacohen, N., Reed, R., Meyerson, M., Golub, T.R., Lander, E.S., Neuberg, D., Brown, J.R., Getz, G. and Wu, C.J., SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med* 2011 365: 2497-2506. <http://dx.doi.org/10.1056/NEJMoa109016> PMID:22150006
17. Rawstron, A.C., Bennett, F.L., O'Connor, S.J., Kwok, M., Fenton, J.A., Plummer, M., de Tute, R., Owen, R.G., Richards, S.J., Jack, A.S. and Hillmen, P., Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med* 2008 359: 575-583. <http://dx.doi.org/10.1056/NEJMoa075290> PMID:18687638
18. Nieto, W.G., Almeida, J., Romero, A., Teodosio, C., Lopez, A., Henriques, A.F., Sanchez, M.L., Jara-Acevedo, M., Rasillo, A., Gonzalez, M., Fernandez-Navarro, P., Vega, T., Orfao, A. and Primary Health Care Group of Salamanca for the Study of, M.B.L., Increased frequency (12%) of circulating chronic lymphocytic leukemia-like B-cell clones in healthy subjects using a highly sensitive multicolor flow cytometry approach. *Blood* 2009 114: 33-37. <http://dx.doi.org/10.1182/blood-2009-01-197368> PMID:19420353
19. Rossi, D., Sozzi, E., Puma, A., De Paoli, L., Rasi, S., Spina, V., Gozzetti, A., Tassi, M., Cencini, E., Raspadori, D., Pinto, V., Berton, F., Gattei, V., Lauria, F., Gaidano, G. and Forconi, F., The prognosis of clinical monoclonal B cell lymphocytosis differs from prognosis of Rai 0 chronic lymphocytic leukaemia and is recapitulated by biological risk factors. *Br J Haematol* 2009 146: 64-75. <http://dx.doi.org/10.1111/j.1365-2141.2009.07711.x> PMID:19438485
20. Fazi, C., Scarfo, L., Pecciarini, L., Cottini, F., Dagklis, A., Janus, A., Talarico, A., Scielzo, C., Sala, C., Toniolo, D., Caligaris-Cappio, F. and Ghia, P., General population low-count CLL-like MBL persists over time without clinical progression, although carrying the same cytogenetic abnormalities of CLL. *Blood* 2011 118: 6618-6625. <http://dx.doi.org/10.1182/blood-2011-05-357251> <http://dx.doi.org/10.1182/blood-2011-05-357251> PMID:21876118
21. Kern, W., Bacher, U., Haferlach, C., Dicker, F., Alpermann, T., Schnittger, S. and Haferlach, T., Monoclonal B-cell lymphocytosis is closely related to chronic lymphocytic leukaemia and may be better classified as early-stage CLL. *Br J Haematol* 2012 <http://dx.doi.org/10.1111/j.1365-2141.2011.09010.x> PMID:22224978
22. Greco, M., Capello, D., Bruscaggin, A., Spina, V., Rasi, S., Monti, S., Ciardullo, C., Cresta, S., Fangazio, M., Gaidano, G., Foa, R. and Rossi, D., Analysis of SF3B1 mutations in monoclonal B-cell lymphocytosis. *Hematol Oncol* 2012 <http://dx.doi.org/10.1002/hon.2013> PMID:22461140
23. D'haer, H., Fischer, K., Bentz, M., Hansen, K., Benner, A., Cabot, G., Diehl, D., Schlenk, R., Coy, J., Stilgenbauer, S. and et al., p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. *Blood* 1995 85: 1580-1589.
24. Catovsky, D., Richards, S., Matutes, E., Oscier, D., Dyer, M.J., Bezares, R.F., Pettitt, A.R., Hamblin, T., Milligan, D.W., Child, J.A., Hamilton, M.S., Dearden, C.E., Smith, A.G., Bosanquet, A.G., Davis, Z., Brito-Babapulle, V., Else, M., Wade, R., Hillmen, P., Group, U.K.N.C.R.I.H.O.C.S. and Group, N.C.L.L.W., Assessment of fludarabine plus cyclophosphamide for patients with

- chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. *Lancet* 2007 370: 230-239. [http://dx.doi.org/10.1016/S0140-6736\(07\)61125-8](http://dx.doi.org/10.1016/S0140-6736(07)61125-8) PMID:18689542
25. Zenz, T., Krober, A., Scherer, K., Habe, S., Buhler, A., Benner, A., Denzel, T., Winkler, D., Edelmann, J., Schwaben, C., Dohner, H. and Stilgenbauer, S., Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood* 2008 112: 3322-3329. <http://dx.doi.org/10.1182/blood-2008-04-154070> PMID:18843282
 26. Dicker, F., Herholz, H., Schnittger, S., Nakao, A., Patten, N., Wu, L., Kern, W., Haferlach, T. and Haferlach, C., The detection of TP53 mutations in chronic lymphocytic leukemia independently predicts rapid disease progression and is highly correlated with a complex aberrant karyotype. *Leukemia* 2009 23: 117-124. <http://dx.doi.org/10.1038/leu.2008.274> PMID:19850740
 27. Malcikova, J., Smardova, J., Rocnova, L., Tichy, B., Kuglik, P., Vranova, V., Cejkova, S., Svitakova, M., Skuhrova Francova, H., Brychtova, Y., Doubek, M., Brejcha, M., Klabusay, M., Mayer, J., Pospisilova, S. and Trbusek, M., Monoallelic and biallelic inactivation of TP53 gene in chronic lymphocytic leukemia: selection, impact on survival, and response to DNA damage. *Blood* 2009 114: 5307-5314. <http://dx.doi.org/10.1182/blood-2009-07-234708> PMID:19188171
 28. Rossi, D., Cerri, M., Deambrogi, C., Sozzi, E., Cresta, S., Rasi, S., De Paoli, L., Spina, V., Gattei, V., Capello, D., Forconi, F., Lauria, F. and Gaidano, G., The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. *Clinical Cancer Research* 2009 15: 995-1004. <http://dx.doi.org/10.1158/1078-0432.CCR-08-1630>
 29. Hallek, M., Fischer, K., Fingerle-Rowson, G., Fink, A.M., Busch, R., Mayer, J., Hensel, M., Hopfinger, G., Hess, G., von Grunhagen, U., Bergmann, M., Catalano, J., Zinzani, P.L., Caligaris-Cappio, F., Seymour, J.F., Berrebi, A., Jager, U., Cazin, B., Trneny, M., Westermann, A., Wendtner, C.M., Eichhorst, B.F., Staib, P., Buhler, A., Winkler, D., Zenz, T., Bottcher, S., Ritgen, M., Mendila, M., Kneba, M., Dohner, H., Stilgenbauer, S., International Group of, I. and German Chronic Lymphocytic Leukemia Study, G., Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet* 2010 376: 1164-1174. [http://dx.doi.org/10.1016/S0140-6736\(10\)61381-5](http://dx.doi.org/10.1016/S0140-6736(10)61381-5) PMID:20697090
 30. Zenz, T., Eichhorst, B., Busch, R., Denzel, T., Habe, S., Winkler, D., Buhler, A., Edelmann, J., Bergmann, M., Hopfinger, G., Hensel, M., Hallek, M., Dohner, H. and Stilgenbauer, S., TP53 mutation and survival in chronic lymphocytic leukemia. *J Clin Oncol* 2010 28: 4473-4479. <http://dx.doi.org/10.1200/JCO.2009.27.8762> PMID:21483000
 31. Gonzalez, D., Martinez, P., Wade, R., Hockley, S., Oscier, D., Matutes, E., Dearden, C.E., Richards, S.M., Catovsky, D. and Morgan, G.J., Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. *J Clin Oncol* 2011 29: 2223-2229. <http://dx.doi.org/10.1200/JCO.2010.32.0838> PMID:20870288
 32. Zainuddin, N., Murray, F., Kanduri, M., Gunnarsson, R., Smedby, K.E., Enblad, G., Jurlander, J., Juliusson, G. and Rosenquist, R., TP53 Mutations are infrequent in newly diagnosed chronic lymphocytic leukemia. *Leuk Res* 2011 35: 272-274. <http://dx.doi.org/10.1016/j.leukres.2010.08.023> PMID:19597025
 33. Stilgenbauer, S., Zenz, T., Winkler, D., Buhler, A., Schlenk, R.F., Groner, S., Busch, R., Hensel, M., Duhrsen, U., Finke, J., Dreger, P., Jager, U., Lengfelder, E., Hohloch, K., Soling, U., Schlag, R., Kneba, M., Hallek, M., Dohner, H. and German Chronic Lymphocytic Leukemia Study, G., Subcutaneous alemtuzumab in fludarabine-refractory chronic lymphocytic leukemia: clinical results and prognostic marker analyses from the CLL2H study of the German Chronic Lymphocytic Leukemia Study Group. *J Clin Oncol* 2009 27: 3994-4001. <http://dx.doi.org/10.1200/JCO.2008.21.1128> PMID:19643983
 34. Zenz, T., Habe, S., Denzel, T., Mohr, J., Winkler, D., Buhler, A., Samo, A., Groner, S., Mertens, D., Busch, R., Hallek, M., Dohner, H. and Stilgenbauer, S., Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. *Blood* 2009 114: 2589-2597. <http://dx.doi.org/10.1182/blood-2009-05-224071> PMID:22493413
 35. Pettitt, A.R., Jackson, R., Carruthers, S., Dodd, J., Dodd, S., Oates, M., Johnson, G.G., Schuh, A., Matutes, E., Dearden, C.E., Catovsky, D., Radford, J.A., Bloor, A., Follows, G.A., Devereux, S., Kruger, A., Blundell, J., Agrawal, S., Allsup, D., Proctor, S., Heartin, E., Oscier, D., Hamblin, T.J., Rawstron, A. and Hillmen, P., Alemtuzumab in combination with methylprednisolone is a highly effective induction regimen for patients with chronic lymphocytic leukemia and deletion of TP53: final results of the national cancer research institute CLL206 trial. *J Clin Oncol* 2012 30: 1647-1655. <http://dx.doi.org/10.1200/JCO.2011.35.9695>
 36. Stilgenbauer, S., Liebisch, P., James, M.R., Schroder, M., Schlegelberger, B., Fischer, K., Bentz, M., Lichter, P. and Dohner, H., Molecular cytogenetic delineation of a novel critical genomic region in chromosome bands 11q22.3-923.1 in lymphoproliferative disorders. *Proc Natl Acad Sci USA* 1996 93: 11837-11841. <http://dx.doi.org/10.1073/pnas.93.21.11837> PMID:10397742
 37. Schaffner, C., Stilgenbauer, S., Rappold, G.A., Dohner, H. and Lichter, P., Somatic ATM mutations indicate a pathogenic role of ATM in B-cell chronic lymphocytic leukemia. *Blood* 1999 94: 748-753. PMID:16014569
 38. Austen, B., Powell, J.E., Alvi, A., Edwards, I., Hooper, L., Starczynski, J., Taylor, A.M., Fegan, C., Moss, P. and Stankovic, T., Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. *Blood* 2005 106: 3175-3182. <http://dx.doi.org/10.1182/blood-2004-11-4516> PMID:21993670
 39. Guarini, A., Marinelli, M., Tavorolo, S., Bellacchio, E., Magliozzi, M., Chiaretti, S., De Propriis, M.S., Peragine, N., Santangelo, S., Paoloni, F., Nanni, M., Del Giudice, I., Mauro, F.R., Torrente, I. and Foa, R., ATM gene alterations in chronic lymphocytic leukemia patients induce a distinct gene expression profile and predict disease progression. *Haematologica* 2012 97: 47-55. <http://dx.doi.org/10.3324/haematol.2011.049270> PMID:22952040
 40. Ouillette, P., Li, J., Shaknovich, R., Li, Y., Melnick, A., Shedden, K. and Malek, S.N., Incidence and clinical implications of ATM aberrations in chronic lymphocytic leukemia. *Genes Chromosomes Cancer* 2012 <http://dx.doi.org/10.1002/gcc.21997> PMID:21266718
 41. Rossi, D., Spina, V., Deambrogi, C., Rasi, S., Laurenti, L., Stamatopoulos, K., Arcaini, L., Lucioni, M., Rocque, G.B., Xu-Monette, Z.Y., Visco, C., Chang, J., Chigrinova, E., Forconi, F., Marasca, R., Besson, C., Papadaki, T., Paulli, M., Larocca, L.M., Pileri, S.A., Gattei, V., Berton, F., Foa, R., Young, K.H. and Gaidano, G., The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood* 2011 117: 3391-3401. <http://dx.doi.org/10.1182/blood-2010-09-302174> PMID:22275381
 42. Xu-Monette, Z.Y., Medeiros, L.J., Li, Y., Orlowski, R.Z., Andreeff, M., Bueso-Ramos, C.E., Greiner, T.C., McDonnell, T.J. and Young, K.H., Dysfunction of the TP53 tumor suppressor gene in lymphoid malignancies. *Blood* 2012 119: 3668-3683. <http://dx.doi.org/10.1182/blood-2011-11-366062> PMID:20861914
 43. Zenz, T., Vollmer, D., Trbusek, M., Smardova, J., Benner, A., Soussi, T., Helfrich, H., Heuberger, M., Hoth, P., Fuge, M., Denzel, T., Habe, S., Malcikova, J., Kuglik, P., Truong, S., Patten, N., Wu, L., Oscier, D., Ibbotson, R., Gardiner, A., Tracy, I., Lin, K., Pettitt, A., Pospisilova, S., Mayer, J., Hallek, M., Dohner, H., Stilgenbauer, S. and European Research Initiative on, C.L.L., TP53 mutation profile in chronic lymphocytic leukemia: evidence for a disease specific profile from a comprehensive analysis of 268 mutations. *Leukemia* 2010 24: 2072-2079. <http://dx.doi.org/10.1038/leu.2010.208> PMID:22297721
 44. Pospisilova, S., Gonzalez, D., Malcikova, J., Trbusek, M., Rossi, D., Kater, A.P., Cymbalista, F., Eichhorst, B., Hallek, M., Dohner, H., Hillmen, P., van Oers, M., Gribben, J., Ghia, P., Montserrat, E., Stilgenbauer, S., Zenz, T. and European Research Initiative on, C.L.L., ERIC recommendations on TP53 mutation analysis in chronic lymphocytic leukemia. *Leukemia* 2012 26: 1458-1461. <http://dx.doi.org/10.1038/leu.2012.25> PMID:19414856

45. Tam, C.S., Shanafelt, T.D., Wierda, W.G., Abruzzo, L.V., Van Dyke, D.L., O'Brien, S., Ferrajoli, A., Lerner, S.A., Lynn, A., Kay, N.E. and Keating, M.J., De novo deletion 17p13.1 chronic lymphocytic leukemia shows significant clinical heterogeneity: the M. D. Anderson and Mayo Clinic experience. *Blood* 2009 114: 957-964. <http://dx.doi.org/10.1182/blood-2009-03-210591> PMID:20595516
46. Dreger, P., Dohner, H., Ritgen, M., Bottcher, S., Busch, R., Dietrich, S., Bunjes, D., Cohen, S., Schubert, J., Hegenbart, U., Beelen, D., Zeis, M., Stadler, M., Hasenkamp, J., Uharek, L., Scheid, C., Humpe, A., Zenz, T., Winkler, D., Hallek, M., Kneba, M., Schmitz, N., Stilgenbauer, S. and German, C.L.L.S.G., Allogeneic stem cell transplantation provides durable disease control in poor-risk chronic lymphocytic leukemia: long-term clinical and MRD results of the German CLL Study Group CLL3X trial. *Blood* 2010 116: 2438-2447. <http://dx.doi.org/10.1182/blood-2010-03-275420> PMID:20967796 PMID:2996483
47. Aster, J.C., Blacklow, S.C. and Pear, W.S., Notch signalling in T-cell lymphoblastic leukaemia/lymphoma and other haematological malignancies. *J Pathol* 2011 223: 262-273. <http://dx.doi.org/10.1002/path.2789> PMID:21948802 PMID:3182047
48. Lobry, C., Oh, P. and Aifantis, I., Oncogenic and tumor suppressor functions of Notch in cancer: it's NOTCH what you think. *J Exp Med* 2011 208: 1931-1935. <http://dx.doi.org/10.1084/jem.20111855> PMID:21646301 PMID:3111953
49. Rothenberg, E.V., T cell lineage commitment: identity and renunciation. *J Immunol* 2011 186: 6649-6655. <http://dx.doi.org/10.4049/jimmunol.1003703> PMID:17878313 PMID:2000509
50. Santos, M.A., Sarmento, L.M., Rebelo, M., Doce, A.A., Maillard, I., Dumortier, A., Neves, H., Radtke, F., Pear, W.S., Parreira, L. and Demengeot, J., Notch1 engagement by Delta-like-1 promotes differentiation of B lymphocytes to antibody-secreting cells. *Proc Natl Acad Sci USA* 2007 104: 15454-15459. <http://dx.doi.org/10.1073/pnas.0702891104> PMID:22086416
51. Balatti, V., Bottoni, A., Palamarchuk, A., Alder, H., Rassenti, L.Z., Kipps, T.J., Pekarsky, Y. and Croce, C.M., NOTCH1 mutations in CLL associated with trisomy 12. *Blood* 2012 119: 329-331. <http://dx.doi.org/10.1182/blood-2011-10-386144> PMID:22207691 PMID:3291600
52. Del Giudice, I., Rossi, D., Chiaretti, S., Marinelli, M., Tavoraro, S., Gabrielli, S., Laurenti, L., Marasca, R., Rasi, S., Fangazio, M., Guarini, A., Gaidano, G. and Foa, R., NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. *Haematologica* 2012 97: 437-441. <http://dx.doi.org/10.3324/haematol.2011.060129> PMID:22619094
53. Lopez, C., Delgado, J., Costa, D., Conde, L., Ghita, G., Villamor, N., Navarro, A., Cazorla, M., Gomez, C., Arias, A., Munoz, C., Baumann, T., Rozman, M., Aymerich, M., Colomer, D., Cobo, F., Campo, E., Lopez-Guillermo, A., Montserrat, E. and Carrio, A., Different distribution of NOTCH1 mutations in chronic lymphocytic leukemia with isolated trisomy 12 or associated with other chromosomal alterations. *Genes Chromosomes Cancer* 2012 51: 881-889. <http://dx.doi.org/10.1002/gcc.21972> PMID:3415400
54. Groth, C. and Fortini, M.E., Therapeutic approaches to modulating Notch signaling: current challenges and future prospects. *Semin Cell Dev Biol* 2012 23: 465-472. <http://dx.doi.org/10.1016/j.semcdb.2012.01.016> PMID:20965628 PMID:3033461
55. Paganin, M. and Ferrando, A., Molecular pathogenesis and targeted therapies for NOTCH1-induced T-cell acute lymphoblastic leukemia. *Blood Rev* 2011 25: 83-90. <http://dx.doi.org/10.1016/j.blre.2010.09.004>
56. Will, C.L. and Luhrmann, R., Spliceosome structure and function. *Cold Spring Harb Perspect Biol* 2011 3: <http://dx.doi.org/10.1101/cshperspect.a003707> PMID:23142775 PMID:3493507
57. David, C.J. and Manley, J.L., Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Gene Dev* 2010 24: 2343-2364. <http://dx.doi.org/10.1101/gad.1973010> PMID:8649382 PMID:231265
58. Luke, M.M., Della Seta, F., Di Como, C.J., Sugimoto, H., Kobayashi, R. and Arndt, K.T., The SAP, a new family of proteins, associate and function positively with the SIT4 phosphatase. *Mol Cell Biol* 1996 16: 2744-2755.
59. Wang, C., Chua, K., Seghezzi, W., Lees, E., Gozani, O. and Reed, R., Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Gene Dev* 1998 12: 1409-1414. PMID:10490618 PMID:84676
60. Das, B.K., Xia, L., Palandjian, L., Gozani, O., Chyung, Y. and Reed, R., Characterization of a protein complex containing spliceosomal proteins SAPs 49, 130, 145, and 155. *Mol Cell Biol* 1999 19: 6796-6802. PMID:19239890
61. Wahl, M.C., Will, C.L. and Luhrmann, R., The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009 136: 701-718. <http://dx.doi.org/10.1016/j.cell.2009.02.009> PMID:21995386 PMID:3322589
62. Papaemmanuil, E., Cazzola, M., Boulton, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J.S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., Godfrey, A.L., Rapado, I., Cvejic, A., Rance, R., McGee, C., Ellis, P., Mudie, L.J., Stephens, P.J., McLaren, S., Massie, C.E., Tarpey, P.S., Varela, I., Nik-Zainal, S., Davies, H.R., Shlien, A., Jones, D., Raine, K., Hinton, J., Butler, A.P., Teague, J.W., Baxter, E.J., Score, J., Galli, A., Della Porta, M.G., Travaglino, E., Groves, M., Tauro, S., Munshi, N.C., Anderson, K.C., El-Naggar, A., Fischer, A., Mustonen, V., Warren, A.J., Cross, N.C., Green, A.R., Futreal, P.A., Stratton, M.R., Campbell, P.J. and Chronic Myeloid Disorders Working Group of the International Cancer Genome, C., Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med* 2011 365: 1384-1395. <http://dx.doi.org/10.1056/NEJMoa1103283> PMID:11907583
63. Li, X., Yang, Y. and Ashwell, J.D., TNF-RII and c-IAP1 mediate ubiquitination and degradation of TRAF2. *Nature* 2002 416: 345-347. <http://dx.doi.org/10.1038/416345a> PMID:18997794 PMID:2676931
64. Zarnegar, B.J., Wang, Y., Mahoney, D.J., Dempsey, P.W., Cheung, H.H., He, J., Shiba, T., Yang, X., Yeh, W.C., Mak, T.W., Korneluk, R.G. and Cheng, G., Noncanonical NF-kappaB activation requires coordinated assembly of a regulatory complex of the adaptors cIAP1, cIAP2, TRAF2 and TRAF3 and the kinase NIK. *Nature Immunol* 2008 9: 1371-1378. <http://dx.doi.org/10.1038/ni.1676>
65. Vallabhapurapu, S. and Karin, M., Regulation and function of NF-kappaB transcription factors in the immune system. *Mol Cell Biol* 2009 29: 693-733. <http://dx.doi.org/10.1146/annurev.immunol.021908.132641> PMID:21048983 PMID:2964333
66. Conze, D.B., Zhao, Y. and Ashwell, J.D., Non-canonical NF-kappaB activation and abnormal B cell accumulation in mice expressing ubiquitin protein ligase-inactive c-IAP2. *PLoS Biol* 2010 8: e1000518. <http://dx.doi.org/10.1371/journal.pbio.1000518>
67. Hewamana, S., Lin, T.T., Jenkins, C., Burnett, A.K., Jordan, C.T., Fegan, C., Brennan, P., Rowntree, C. and Pepper, C., The novel nuclear factor-kappaB inhibitor LC-1 is equipotent in poor prognostic subsets of chronic lymphocytic leukemia and shows strong synergy with fludarabine. *Clin Cancer Res* 2008 14: 8102-8111. <http://dx.doi.org/10.1158/1078-0432.CCR-08-1673> PMID:20351313 PMID:2904580
68. Hertlein, E., Wagner, A.J., Jones, J., Lin, T.S., Maddocks, K.J., Towns, W.H., 3rd, Goettl, V.M., Zhang, X., Jarjoura, D., Raymond, C.A., West, D.A., Croce, C.M., Byrd, J.C. and Johnson, A.J., 17-DMAG targets the nuclear factor-kappaB family of proteins to induce apoptosis in chronic lymphocytic leukemia: clinical implications of HSP90 inhibition. *Blood* 2010 116: 45-53. <http://dx.doi.org/10.1182/blood-2010-01-263756>

Genetics of Follicular Lymphoma Transformation

Laura Pasqualucci,^{1,2,3,*} Hossein Khiabanian,⁴ Marco Fangazio,¹ Mansi Vasishtha,¹ Monica Messina,¹ Antony B. Holmes,¹ Peter Ouillette,⁵ Vladimir Trifonov,⁴ Davide Rossi,⁶ Fabrizio Tabbò,⁷ Maurilio Ponzoni,⁸ Amy Chadburn,⁹ Vundavalli V. Murty,^{1,2,3} Govind Bhagat,^{2,3} Gianluca Gaidano,⁶ Giorgio Inghirami,⁷ Sami N. Malek,⁵ Raul Rabadan,⁴ and Riccardo Dalla-Favera^{1,2,3,10,11,*}

¹Institute for Cancer Genetics, Columbia University, New York, NY 10032, USA

²Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY 10032, USA

³Department of Pathology & Cell Biology, Columbia University, New York, NY 10032, USA

⁴Department of Biomedical Informatics and Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10032, USA

⁵Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan, Ann Arbor, MI 48109, USA

⁶Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara 28100, Italy

⁷Department of Molecular Biotechnology and Health Science, Center for Experimental Research and Medical Studies (CeRMS), University of Torino, Torino 10126, Italy

⁸Pathology and Lymphoid Malignancies Units, San Raffaele Scientific Institute, Milan 20132, Italy

⁹Department of Pathology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

¹⁰Department of Genetics & Development, Columbia University, New York, NY 10032, USA

¹¹Department of Microbiology & Immunology, Columbia University, New York, NY 10032, USA

*Correspondence: lp171@columbia.edu (L.P.), rd10@columbia.edu (R.D.-F.)

<http://dx.doi.org/10.1016/j.celrep.2013.12.027>

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

SUMMARY

Follicular lymphoma (FL) is an indolent disease, but 30%–40% of cases undergo histologic transformation to an aggressive malignancy, typically represented by diffuse large B cell lymphoma (DLBCL). The pathogenesis of this process remains largely unknown. Using whole-exome sequencing and copy-number analysis, we show here that the dominant clone of FL and transformed FL (tFL) arise by divergent evolution from a common mutated precursor through the acquisition of distinct genetic events. Mutations in epigenetic modifiers and antiapoptotic genes are introduced early in the common precursor, whereas tFL is specifically associated with alterations deregulating cell-cycle progression and DNA damage responses (*CDKN2A/B*, *MYC*, and *TP53*) as well as aberrant somatic hypermutation. The genomic profile of tFL shares similarities with that of germinal center B cell-type de novo DLBCL but also displays unique combinations of altered genes with diagnostic and therapeutic implications.

INTRODUCTION

Follicular lymphoma (FL) is the second most common type of B cell non-Hodgkin lymphoma, comprising ~25% of all new diagnoses (Swerdlow et al., 2008) (<http://seer.cancer.gov/statistics/>). Although initially indolent and responsive to a variety of treatments, this disease remains largely incurable (Kridel et al.,

2012). One particularly compelling problem in the clinical history of FL is its histologic transformation to a more aggressive malignancy, typically represented by a diffuse large B cell lymphoma (DLBCL) (Montoto and Fitzgibbon, 2011). FL transformation has been reported to occur in 16% to 70% of patients over time, with a consensus rate of 3% per year, and is associated with a mean survival posttransformation of less than 2 years (Montoto and Fitzgibbon, 2011). Thus, there is a strong need for an increased understanding of both the dynamics of tumor clonal evolution and the mechanisms that are responsible for transformation, which may in turn be translated into more effective therapies.

Although the process of transformation to DLBCL was originally described several decades ago, few studies have specifically addressed this question in longitudinal series with documented clonal relationship between the two phases (Lossos and Gascoyne, 2011). Current knowledge of the biology of transformation suggests the involvement of heterogeneous genetic, epigenetic, and microenvironment-dependent factors, most notably mutations of *TP53* (Lo Coco et al., 1993; Sander et al., 1993), genetic and/or epigenetic inactivation of the *CDKN2A/p16* tumor suppressor gene (Pinyol et al., 1998), translocations deregulating the *BCL6* proto-oncogene (Akasaka et al., 2003), alterations involving chromosome 1p36 (Martinez-Clement et al., 2003), and changes in *MYC* expression (Lossos et al., 2002). Additionally, analysis of selected genes in few cases revealed an association between progression to DLBCL and aberrant somatic hypermutation (ASHM) (Rossi et al., 2006), a mechanism of genetic instability resulting from the abnormal functioning of the physiologic somatic hypermutation (SHM) process that operates in germinal center (GC) B cells (Pasqualucci et al., 2001). However, these findings were based on small number of cases and a candidate gene approach as opposed to an

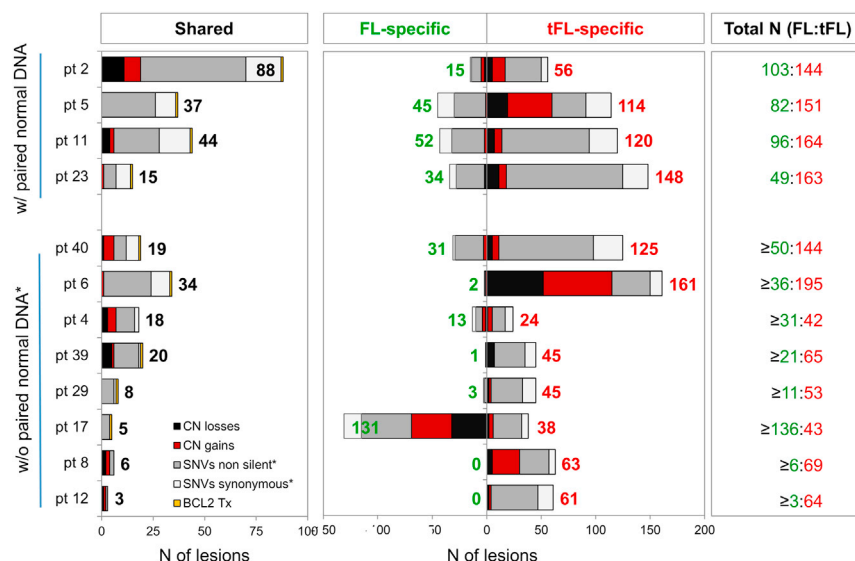


Figure 1. FL and tFL Display Shared and Unique Genomic Aberrations

Overall load of genetic lesions identified by WES and CN analysis in the dominant clone of the 12 discovery cases. Color codes denote distinct types of aberrations (Tx, translocation). *In cases lacking matched normal DNA, shared SNVs are limited to those affecting 52 selected genes with well-established roles in lymphomagenesis (see the [Experimental Procedures](#)); thus, the total number of genetic lesions in these patients (right column) most likely represents an underestimate. FL-specific SNVs that could be due to genomic loss or cnLOH of the same region in the tFL phase were excluded.

analysis allows to discriminate between a linear, sequential model, wherein the tFL dominant clone will maintain all lesions present in the FL dominant clone, along with additional tFL-acquired alter-

ations, and a divergent evolution model, which postulates the existence of lesions that are unique to the dominant clone of the FL or the tFL in addition to the set of shared alterations ([Experimental Procedures](#) and [Figure S1](#)).

Overall, we found 52 clonally represented, shared copy-number aberrations (CNAs; average, 4.3 per sample; range, 0 to 19 per sample) and 234 shared single-nucleotide variants (SNVs), including silent and nonsilent mutations (average, 38.5 per sample in the four patients with matched normal DNA; in the remaining eight pairs, shared SNVs were only considered if they affected 52 genes that have been previously validated as functional targets of somatic mutations in lymphoid malignancies, because of the exceedingly high number of variants that are predicted in the absence of matched normal DNA, most likely reflecting private SNPs not reported in public databases; see the [Experimental Procedures](#)). The presence of shared genetic alterations was documented in all sample pairs analyzed, confirming the original clonal relationship between the FL and tFL sample ([Figure 1](#), left).

RESULTS

Divergent Evolution of FL and tFL from a Common Mutated Precursor

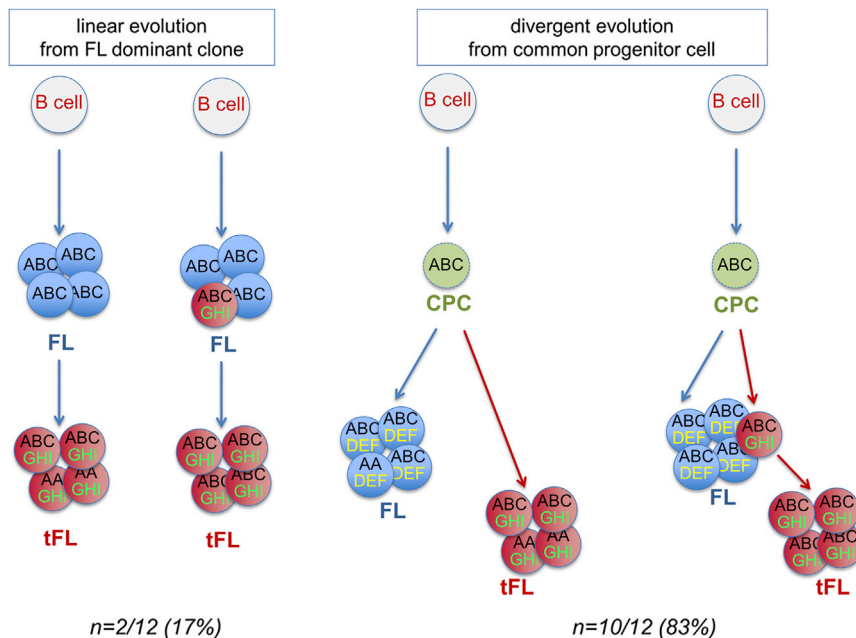
To investigate whether transformation of FL evolves as a linear process (i.e., through the emergence of an aggressive subclone from the initial dominant FL population) or derives from the divergent evolution of an ancestral common precursor cell (CPC) that acquired distinct mutations to become a FL or a transformed FL (tFL), we integrated massively parallel whole-exome sequencing (WES) and genome-wide high-resolution SNP array analysis in a “discovery panel” of sequential FL and tFL biopsies obtained from 12 patients, including four with available matched normal DNA ([Tables S1 and S2](#) and [Figure S1](#)). In all cases, investigation of the rearranged immunoglobulin (Ig) genes by Sanger sequencing and/or SNP array analysis confirmed the clonal relationship between the two phases, whereas the inferred copy-number value at the segment of deletional recombination within the Ig loci was used to quantify the percentage of tumor cells in the biopsy ([Bergsagel and Kuehl, 2013](#)), allowing to normalize the data for clonal representation ([Table S1](#)). Fluorescence in situ hybridization (FISH) analysis was used to assess the presence of chromosomal translocations affecting *BCL2*, *MYC*, and *BCL6*.

We extrapolated the evolutionary history of transformation by defining genomic alterations that are present in the dominant clone of both pre- and posttransformation specimens (“shared lesions”) and contrasting them to those that are present exclusively in the FL or tFL biopsy (“phase-specific lesions”). This

analysis allows to discriminate between a linear, sequential model, wherein the tFL dominant clone will maintain all lesions present in the FL dominant clone, along with additional tFL-acquired alter-

ations, and a divergent evolution model, which postulates the existence of lesions that are unique to the dominant clone of the FL or the tFL in addition to the set of shared alterations ([Experimental Procedures](#) and [Figure S1](#)).

In addition to shared lesions, all tFL cases harbored unique mutations and CNAs that were not present in the major FL clone at diagnosis, indicating acquisition during the transformation process or selection of a minor subclone, the size of which was below the detection threshold of the methodologies used. The number of tFL-specific lesions ($n = 709$ SNVs and 291 CNAs, including 119 losses and 172 gains) was widely heterogeneous across different patients, ranging from 24 to 161 per case (average, 83 per sample) ([Figure 1](#), right; see also [Figure S2A](#) and [Table S3](#)). Importantly, unique, clonally represented events were also detected in 10 of 12 baseline FL biopsies ($n = 327$, including 229 SNVs and 98 CNAs; [Figure 1](#), middle, and [Figure S2A](#)). The presence of FL-specific lesions was not due to CN loss or copy-neutral loss of heterozygosity (cnLOH) affecting the same region in the sequential tFL biopsy, as documented by both SNP array and WES analysis. Thus, these events had been acquired independently by the dominant FL clone, consistent with divergent evolution.



Evidence of nonlinear evolution was also observed at the individual gene level. As an example, both FL and tFL of patient #23 harbored biallelic *MLL2* mutations in the dominant clone, but only one of the two events (S286fs) was shared between the pre- and posttransformation biopsy, consistent with its presence in the common ancestor clone, whereas distinct mutations were detected in the second allele of the FL (R2687*) and tFL (R280_splice) specimen, indicating that they had been acquired independently by the ancestor clone during evolution to these two diseases (Figures S2B–S2D).

Overall, 10 of 12 patients analyzed (83.3%; 95% confidence interval [CI], 55% to 95%) showed a mutation pattern suggestive of divergent evolution, indicating that this is the predominant mode in the history of FL transformation (Figure 2). The remaining two patients (#8 and #12) did not harbor FL-specific events; furthermore, a significant proportion of tFL-specific lesions (13 of 63 in patient #8 and 34 of 61 in patient #12) could be detected at low frequencies (4% to 15%) in the FL specimen, suggesting that the tFL arose from a minor subclone within the dominant FL population, which subsequently acquired additional mutations in a linear evolution pattern (Figure 2 and Table S3).

With one exception (patient #17), the number of events acquired by the tFL dominant clone (including CNAs and SNVs) was significantly higher than that acquired by the FL dominant clone, ranging from >50-fold in case #6 to 2-fold in case #11 ($p < 0.005$) and underscoring the genomic complexity of the tFL genome (Figures 1 and S3A). Notably, at least one of these ten patients did not receive any treatment between the original FL diagnosis and transformation, indicating that the higher mutation load of tFL does not simply reflect the consequence of the mutagenic effect or the selective pressure of chemotherapy. In at least two patients, several chromosomes displayed convoluted intrachromosomal rearrangements due to alternating gains and losses of genomic material, frequently accounting for over ten

Figure 2. FL and tFL Arise through Divergent Evolution in Most Patients

Inferred models of clonal evolution during FL transformation. The original B cell clone is on top; blue and red circles depict the FL and tFL clones, respectively, whereas green, dotted circles represent the postulated common mutated precursor cell (CPC). In the linear model (2 of 12 patients analyzed), the tFL dominant clone originates directly from the FL dominant clone after the acquisition of additional mutations. In the divergent model (10 of 12 patients analyzed), the FL and tFL dominant clones derive from a CPC through the independent acquisition of distinct mutations. In both scenarios, the tFL dominant clone may be present as a minor subclone already at the time of the FL biopsy (not addressed in this study). ABC, mutations shared between FL and tFL; DEF, mutations unique to the FL dominant clone; GHI, mutations unique to the tFL dominant clone.

switches per chromosome (Figure S3B and Table S4). Although the sequencing approach adopted in our study (WES, as opposed to whole-genome sequencing)

prevents from distinguishing true chromothripsis from localized lesions that occurred progressively (Korbel and Campbell, 2013), these data highlight a remarkable genomic instability in tFL cases in comparison to both FL and other lymphoid malignancies (Fabbri et al., 2011; Mullighan et al., 2007; Rossi et al., 2012).

Collectively, these findings support a divergent evolution model in a significant proportion of patients undergoing transformation, whereby FL and tFL arise from a common mutated ancestor through the independent acquisition of distinct genetic lesions.

Recurrent Genetic Lesions

In order to identify lesions potentially relevant for transformation among the large number of candidates that emerged from the analysis of the discovery panel (710 unique genes, including those targeted by nonsilent SNVs, small indels, and/or CNAs if they were within minimal common regions involving a maximum of three loci), we extended the WES and SNP array analysis to 27 additional tFL cases (screening panel; combined, 39 tFL cases). Then, these data were interrogated for the presence of recurrent alterations, a common readout for functionally relevant genes, and by three independent analytical methods for the identification of key targets of functional genomic alterations: (1) MutComFocal, a recently developed computational algorithm that isolates candidate cancer genes from high throughput CN and SNV data (Trifonov et al., 2013a), (2) MutSigCV, a tool that analyzes SNV data in order to identify genes mutated more often than expected by chance (Lawrence et al., 2013), and (3) GISTIC, a computational approach identifying significant targets of somatic CNAs.

Figure 3 illustrates the overall proportion of tFL cases harboring genetic lesions in genes altered at $\geq 10\%$ frequency and recognized as functionally relevant targets by at least one

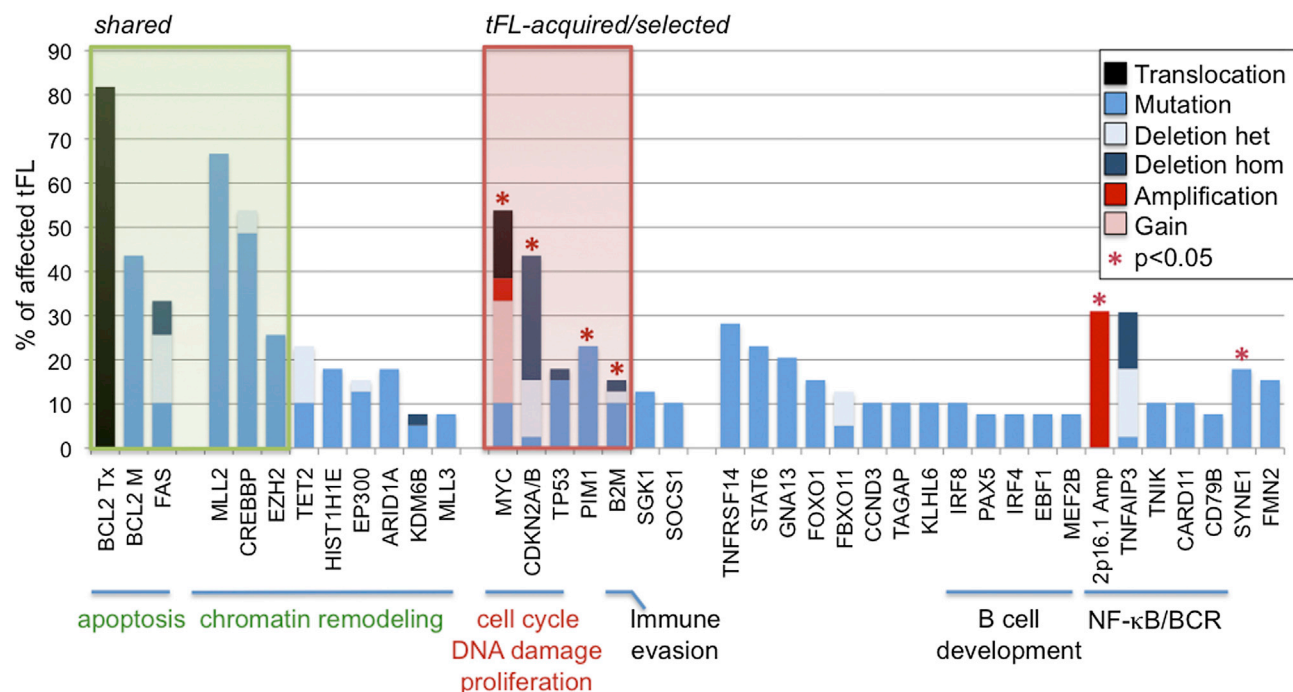


Figure 3. Recurrently Mutated Genes in tFL

Proportion of tFL cases ($n = 39$) carrying genetic lesions (SNVs and CNAs) in genes altered at $\geq 10\%$ frequency and deemed significant by one of three independent algorithms (see the [Supplemental Information](#) and [Table S5](#)); additional genes of functional relevance in the same pathways are also shown. Deletions and gains were only computed if defining MCRs of aberration encompassing a maximum of three genes. Green shade highlights alterations consistently shared between the FL and tFL sample, and red shade identifies alterations enriched in tFL (* $p < 0.05$ when comparing tFL frequency versus unselected FL frequency; note that *TP53* was considered as a tFL-specific target despite its borderline p value because it was invariably unmutated in the diagnostic FL sample of all three informative tFL cases and because of the broad literature data. The remaining genes did not appear to be phase-specific or could not be unequivocally assigned to a given category because of the relatively small numbers (for the full list of SNVs found in $>5\%$ of tFL cases, see [Table S7](#)).

of the three independent approaches, with few additional genes of functional annotation within the same pathway (see [Figures S4–S6](#) and [Table S5](#) for the full list of genes found mutated in $\geq 10\%$ of cases; see also [Table S6](#)). Genes are grouped into biological categories and annotated in order to indicate whether alterations were also found in the diagnostic FL biopsy (available in 24 of the 39 cases) or predominantly acquired or selected at transformation. Collectively, these aberrations point to a number of biological programs and signaling pathways that are either dysregulated early in the putative FL and tFL precursor cell (i.e., “shared” lesions) or selected during transformation (“tFL-acquired” lesions).

Genetic Lesions Shared by FL and tFL

The most commonly affected genes in both FL and tFL were those encoding for histone modification enzymes, including methyltransferases and acetyltransferases (36 of 39 cases [92.3%]). In line with previous findings in unselected FL ([Morin et al., 2010; Morin et al., 2011; Pasqualucci et al., 2011a; Pasqualucci et al., 2011b](#)), the H3K4 trimethyltransferase *MLL2* was mutated in 26 of 39 tFL cases (66.7%) with 36 truncating events and nine missense mutations ([Figures 3, S4, and S6](#) and [Table S7](#)). These lesions were already present at FL diagnosis in all but one patient and were never lost at transformation, consistent with an early acquisition by the CPC. The activity of

the *MLL2*-containing complex was also impaired by mutually exclusive alterations of *KDM6B*, encoding for an H3K4 histone demethylase interacting with *MLL2* ($n = 3$ of 39 cases, including two SNVs and one homozygous deletion) and *MLL3* ($n = 3$ of 39 cases; [Figures S4 and S6](#) and [Table S7](#)). Additionally, one-fourth of tFL cases ($n = 10$ of 39 [25.6%]) harbored *EZH2* gain-of-function mutations that almost invariably replace the hotspot tyrosine residue Y641 within the protein SET domain ($n = 9$ of 10; [Table S7](#)). *EZH2* mutations have been reported in 7% unselected FL cases and 22% de novo GCB-DLBCL ([Morin et al., 2010; Morin et al., 2011](#)), where they increase H3K27 levels through altered substrate specificity.

Another class of chromatin modifiers was represented by the acetyltransferases *CREBBP* ($n = 21$ of 39 patients [53.8%]; 19 point mutations and two focal deletions) and *EP300* (6 of 39 cases [15.4%], of which five mutated and one deleted; [Figures 3, S4, and S6](#)). In both genes, the mutation pattern was highly reminiscent of what has been reported in unselected FL and de novo DLBCL with respect to the inactivating nature of the lesions, the evidence of mutational hotspots (R1446 in five patients, F1484 in two patients, Y1503 in two patients, and Δ S1680 in three patients; [Table S7](#)), and the predominantly monoallelic distribution (19 of 21 affected cases), indicating a haploinsufficient tumor suppressor role ([Pasqualucci et al., 2011a](#)).

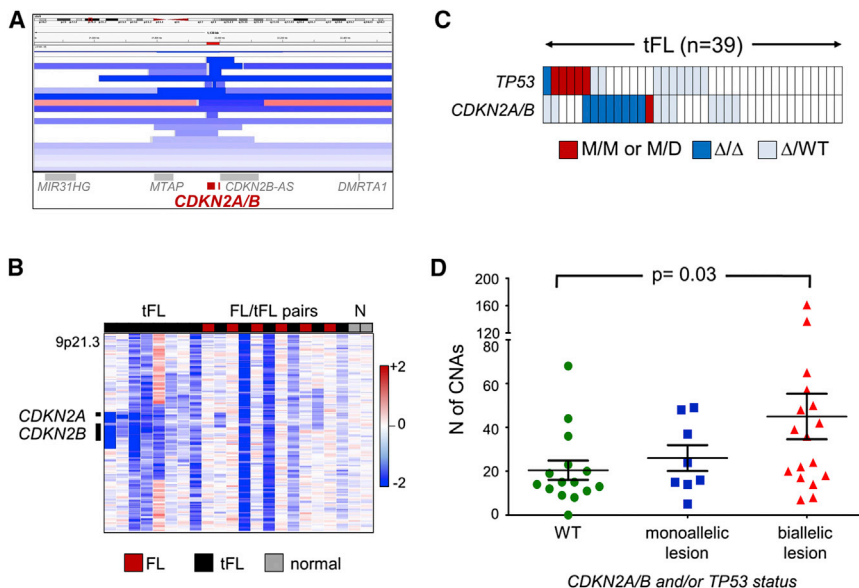


Figure 4. Biallelic Loss of *CDKN2A/B* Is Specifically Acquired during Transformation

(A) Segmentation data from 18 DLBCL samples harboring *CDKN2A/B* deletions, visualized with the Integrative Genomics Viewer (<http://www.broadinstitute.org/igv>). Each track represents one sample, wherein white denotes a normal (diploid) copy number, blue indicates a region of copy-number loss, and red corresponds to a region of copy-number gain. The inferred copy number and the corresponding color intensity may vary across samples because of the presence of nontumor cells infiltrating the biopsy. Individual genes in the region are aligned in the bottom panel, and the area defined by the red bar at the top corresponds to the *CDKN2A/B* locus.

(B) dChipSNP heatmap showing median-smoothed log2 CN ratio in 14 tFL biopsies harboring *CDKN2A/B* deletions, including six with matched FL biopsies and two control DNAs (N). A vertical black bar indicates the location of the *CDKN2A/B* loci.

(C) Relative distribution of biallelic lesions affecting *CDKN2A/B* and *TP53*. Columns correspond to individual patients, and the genomic status of the two genes is color-coded as indicated.

(D) Samples carrying biallelic alterations of *CDKN2A/B* and/or *TP53* are characterized by a significantly higher number of CN aberrations (Mann-Whitney U test, $p = 0.03$).

Programmed cell death was the second largest program dysregulated in both FL and tFL and, thus, presumably in the common ancestor clone. In addition to *BCL2* translocations, detected in 27 of 33 tFL (81.8%) and invariably shared between the two phases ($n = 18$ informative pairs), all t(14;18)-positive cases harbored multiple somatic point mutations within the ~2 kb region downstream of the *BCL2* transcription initiation sites (Figure 3 and S4 and Table S7), reflecting the activity of the AID-dependent SHM process driven by the juxtaposed Ig enhancer (Lohr et al., 2012; Saito et al., 2009).

The *FAS* gene was disrupted in 13 of 39 tFL cases (33.3%) because of inactivating mutations ($n = 4$ of 39 [10.2%]) and genomic deletions ($n = 9$ of 39 [23.1%]), including two focal homozygous events; Figures 3 and S4). In the three affected patients with available pre- and posttransformation biopsy, these lesions were always detectable at FL diagnosis, suggesting their presence in the putative CPC (Table S7). Interestingly, none of the 23 unselected FL exomes harbored *FAS* mutations, giving rise to the possibility that these lesions represent an early marker for transformation. *FAS* encodes for a receptor protein that acts as a major mediator of apoptosis in GCB cells carrying low-affinity and self-reactive antigen receptors (van Eijk et al., 2001). With the exception of one amino acid substitution removing the initiating methionine (M1T), all *FAS* mutations (Y232*, P217_splice, and D317V) cluster in exons 7 to 9, which encode for the protein intracytoplasmic tail (Table S7). This domain is required for the assembly of the death-inducing signaling complex, and its truncation will result in the functional loss of normal *FAS* signaling by a dominant-negative effect (Siegel et al., 2000), as documented in patients with autoimmune lymphoproliferative syndrome (Holzelova et al., 2004). A deleterious effect was also predicted for the D317V amino acid change

on the basis of the PolyPhen 2 algorithm (Table S7). *FAS* was identified as a relevant target of genomic deletions by two independent algorithms, including GISTIC (Figure S5 and Table S6) and MutComFocal.

Altogether, these findings identify the disruption of pathways affecting chromatin modifier functions and resistance to apoptosis as recurrent lesions common to FL and tFL, and, thus, presumably occurring early during the initial clonal expansion of the putative precursor clone.

Genetic Lesions Specifically Associated with tFL

The most common genomic aberration specifically acquired during progression to tFL was the loss of *CDKN2A/B*, two tumor suppressor genes whose protein products (p14-ARF, p16-INK4A, and p15-INK4B) play major roles as negative regulators of cell-cycle G1 progression and as stabilizers of the tumor suppressor p53 (Sherr, 2004). Overall, 46.1% of tFL cases ($n = 18$ of 39) carry genomic aberrations affecting these loci, including 17 CN losses ($n = 6$ heterozygous and 11 homozygous; Figure 4A) and a nonsense C72* mutation combined with cnLOH (Figure S4 and Table S7). In most deleted cases ($n = 10$ of 17), the loss of genetic material encompassed ≤ 3 genes, identifying a minimal common region smaller than 10 kb and exquisitely restricted to the *CDKN2A/B* locus. Biallelic *CDKN2A/B* alterations were never present at FL diagnosis, indicating a specific role during transformation (Figure 4B).

The loss of *CDKN2A/B* may impinge on different biological programs, including DNA damage responses (via the p14-ARF/p53 pathway) and cell-cycle regulation (via the RB/p16 tumor-suppressive pathway). As expected, immunohistochemical staining for p16 expression confirmed its complete loss in the neoplastic lymphocytes of all biallelically deleted tFL cases

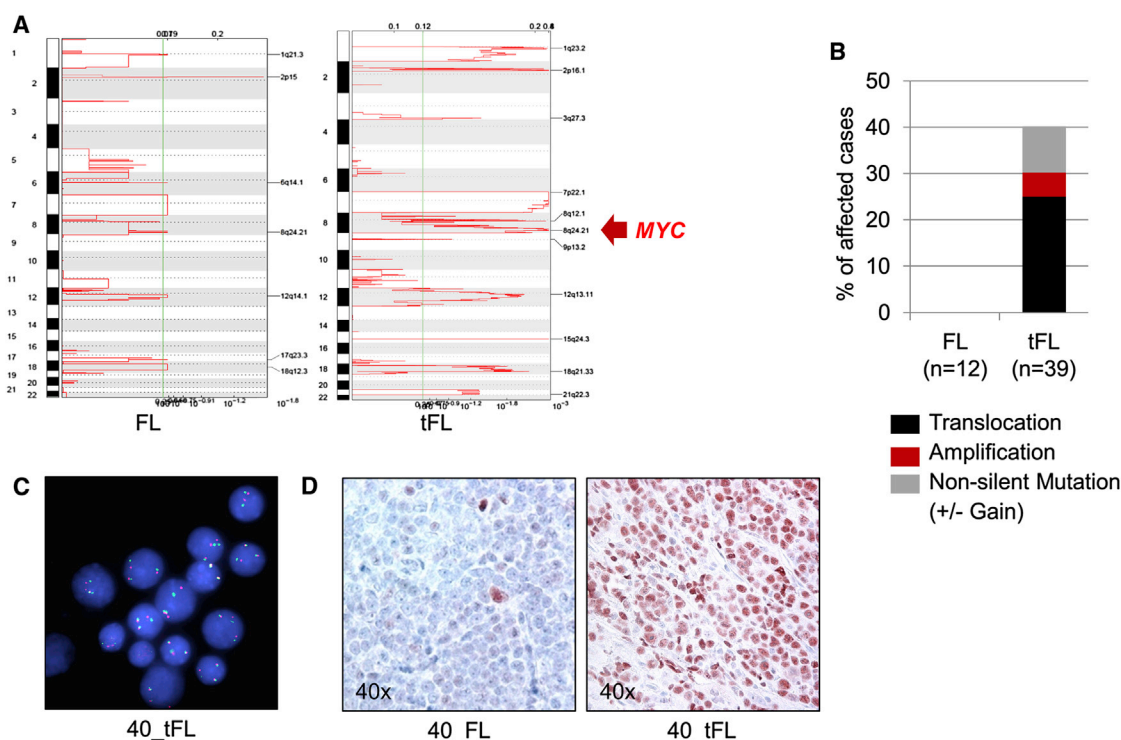


Figure 5. Recurrent Genetic Lesions of MYC in tFL

(A) GISTIC analysis of CN gains in FL and tFL cases (see also Figure S5).

(B) Percentage of cases carrying genetic lesions of MYC in FL and tFL.

(C) FISH analysis with a MYC break-apart probe in tFL #40.

(D) IHC analysis of MYC expression in the pre- and posttransformation biopsy of the same patient.

(n = 8 of 8); however, two of three monoallelically deleted cases and a significant proportion of wild-type (WT) tFL cases (n = 5 of 15 [33.3%]) were also p16-negative (data not shown), suggesting the involvement of epigenetic mechanisms of inactivation. Although the limited number of cases prevents statistical analysis, *CDKN2A/B* biallelic lesions tend to be mutually exclusive with biallelic deletions and/or mutations of *TP53*, observed in 7 of 39 tFL cases (17.9%) but absent at FL diagnosis (n = 3 informative cases; Figure 4C). These observations suggest that *CDKN2A/B* loss may contribute to FL transformation by affecting both cell-cycle regulation and p53-dependent DNA damage responses, thus promoting genomic instability. Consistent with this hypothesis, patients exhibiting dysregulation of the ARF/p53 axis via biallelic alterations of *CDKN2A/B* and/or *TP53* were characterized by a significantly higher number of CNAs in comparison to patients that harbor WT alleles (average n = 45.0 versus 20.5, Mann Whitney U test, p = 0.03; Figure 4D).

Genetic lesions deregulating *MYC*, namely chromosomal translocations (n = 6 of 24 tFL cases with available FISH data [25.0%]), copy-number gains and/or amplifications (n = 13 of 39 [33.3%]), and point mutations reflecting the activity of ASHM (Figures 5 and S4 and Table S7) were the second most common tFL-specific lesions. Although low copy-number gains could be occasionally observed in the original FL biopsy, high CN amplifications, translocations, and point mutations were

either completely absent (n = 13 cases) or only detected in a minor subclone within the dominant FL population (patient #8). Deregulated *MYC* oncogenic activity may provide multiple advantages to the cancer cell through its pleiotropic function in cell growth, metabolism, and genetic instability.

Also enriched at transformation were biallelic mutations and/or deletions encompassing *B2M* (n = 5 of 39) and *CD58* (n = 2 of 39), two genes involved in the control of immune recognition by cytotoxic T lymphocytes and natural killer cells, respectively, and previously shown to be recurrently inactivated in de novo DLBCL (Challa-Malladi et al., 2011) (Figures 3 and S4 and Tables S6 and S7). *B2M* genomic aberrations were specifically acquired or selected at transformation (n = 3 informative cases) and, accordingly, *B2M* as well as *CD58* were not mutated in 23 unselected FL exomes analyzed, implicating escape from immune surveillance as a contributor to the transformation process.

Finally, multiple point mutations, small deletions, and duplications were identified in the 5' sequences of several recognized ASHM target genes, including *PIM1*, *PAX5*, *RhoH/TTF*, *MYC*, *BCL7A*, *CIITA*, and *SOCS1* (overall, 34 of 39 cases [87.1%]; Figure 6). These lesions display typical features of AID-mediated activity (Figure S7) and, depending on the genomic configuration of the involved locus, were variably distributed in coding and/or noncoding sequences. These changes were not detected in the pretransformation biopsy, indicating that they had been

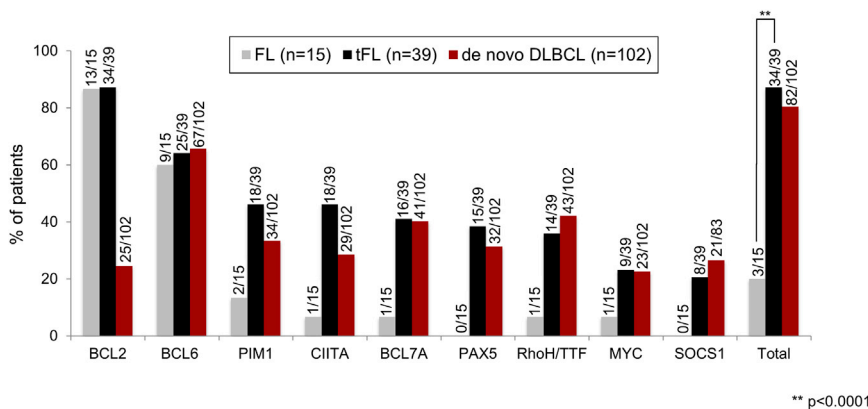


Figure 6. ASHM Is Associated with Transformation

Proportion of cases harboring mutations in known ASHM targets; the BCL6 intron 1 region, a physiologic target of SHM in GC B cells, and the BCL2 sequences, which accumulate mutations in translocated alleles under the influence of the juxtaposed *IGH* promoter, are shown as controls. The three mutated FL cases harbored one single event each.

specifically acquired or selected at transformation (Figure 6). These data point to a malfunction of SHM occurring late in the disease, although the prolonged exposure of the precursor cell to the potentially deleterious environment of the GC reaction may also favor the accumulation of lesions.

Non-Phase-Specific Lesions

A number of genetic alterations were variably observed as shared or phase-specific events, suggesting heterogeneous contributions to disease pathogenesis. Consistent with previous reports (Lossos and Gascoyne, 2011), *TNFRSF14*, encoding for a member of the TNF receptor superfamily that signals to T cells with stimulatory or inhibitory effects depending on the ligand, was disrupted in 22 of 39 tFL (56.4%) and 9 of 17 FL cases (52.9%) due to a combination of truncating mutations ($n = 14$, distributed in 11 tFL and 3 FL cases), genomic deletions ($n = 15$ tFL and 6 FL cases), and cnLOH ($n = 4$ tFL cases and 1 FL case; Figure S4 and Table S7). Importantly, all mutated tFL cases had lost the residual WT allele because of deletion or cnLOH. Although multiple other candidate genes are encompassed by the large heterozygous deletions affecting chromosomal region 1p36.31, this typical pattern of biallelic inactivation documented the specific involvement of *TNFRSF14* in these lesions (Figure S4).

STAT6, a DNA binding transcription factor implicated in IL4 and IL13-mediated responses, was the target of heterozygous somatic point mutations in 9 of 39 (23.1%) tFL cases. Analysis of the diagnostic FL biopsy revealed the presence of the same mutation in three of five available cases, whereas the remaining two had acquired this lesion at transformation; *STAT6* mutations were observed in unselected FL cases (1 of 23 [4.3%]), consistent with a role in both disease phases. Notably, all *STAT6* mutations cluster within the protein DNA binding domain, identifying a mutational hotspot of possible functional relevance at residue D419, which was substituted to G (four cases), H (one case), and Y (one case; Figure S4 and Table S7), as previously reported in primary mediastinal B cell lymphoma (Ritz et al., 2009).

Several genes encoding for core histones were also frequently mutated in both tFL ($n = 11$ of 39 [28.2%]) and FL ($n = 6$ of 12 [50%]) cases, *HIST1H1E* being the most common target ($n = 7$ of 39 tFL [17.9%], 2 of 12 FL, [16.6%], and 2 of 23 unselected FL cases [8.7%]; Tables S5 and S7). In addition, > 10%

of tFL cases harbored loss-of-function (nonsense and frameshift) mutations disrupting the *ARID1A* ($n = 7$ of 39 [17.9%]) and *ARID1B* ($n = 1$ of 39 [2.6%]) genes, two components of the SWI-SNF chromatin remodeling complex, which was recently shown to take part in maintaining the pluripotency of stem cells as well as in reprogramming somatic cells (Ho et al., 2009).

Finally, a number of recurrent lesions are expected to interfere with signaling pathways that are triggered in response to engagement of the BCR or CXCR4 receptors; these include oncogenic mutations of *CARD11* and *CD79B* (4 of 39 tFL [10.3%] and 3 of 39 tFL cases [7.7%], respectively), alterations in negative and positive regulators of NF- κ B (*TNFAIP3*, biallelically lost in 6 of 39 tFL cases, and TRAF2- and Nck-interacting kinase [*TNFK*], mutated in 4 of 39 cases), truncating mutations of *GNA13* ($n = 8$ of 39 tFL [20.5%]) versus 1 of 23 unselected FL cases [4.3%]), and point mutations of *FOXO1* ($n = 6$ of 39 tFL [15.4%]) versus 0 of 23 unselected FL cases; Tables S5 and S7); deregulation of the latter two genes may impinge on the PI3K pathway as well as RhoGTPase responses, and has been observed in de novo DLBCL (Morin et al., 2013; Trinh et al., 2013). Altogether, these results point to several signaling pathways that are recurrently dysregulated in FL and/or tFL, although their individual contribution is not specifically restricted to a discrete phase of disease pathogenesis (see the Discussion).

The Genomic Landscape of tFL Is Unique But More Related to GCB-DLBCL

In order to determine whether tFL and de novo DLBCL represent pathogenically different diseases, we compared the genomic profile of the 39 tFL cases to that of 102 de novo DLBCL cases representative of the two major molecular subtypes; i.e., GCB and activated B cell (ABC) DLBCL (Alizadeh et al., 2000). When analyzed by unsupervised hierarchical clustering with the frequency of aberrations at 34 informative targets, tFL appears to be closer to GCB- than ABC-DLBCL (see the Experimental Procedures). Common features of the two diseases include the presence of *BCL2* rearrangements, *REL* amplifications, *EZH2*, *GNA13*, and *TNFRSF14* mutations (Morin et al., 2010, 2011; Pasqualucci et al., 2011b) along with the absence of typical ABC-DLBCL-specific aberrations (*MYD88* mutations and *PRDM1* inactivation) (Mandelbaum et al., 2010; Morin et al., 2013; Ngo et al., 2011; Pasqualucci et al., 2006) (Figure 7). However, unique combinations of genetic lesions were found in tFL, which are otherwise never observed in GCB-DLBCL (e.g., biallelic

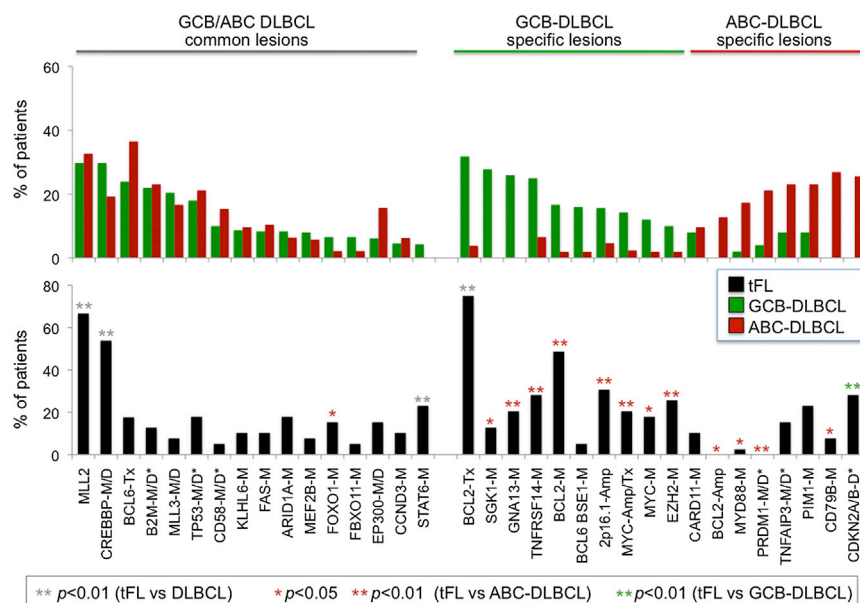


Figure 7. tFL Displays a Unique Genomic Profile that Partially Overlaps with GCB-DLBCL

Percentage of cases carrying the indicated genetic lesions in de novo DLBCL (top: green bars, GCB-DLBCL; n = 50; red bars, ABC-DLBCL; n = 52) and tFL (bottom: black bars; n = 39). Asterisks indicate statistically significant differences (one-tailed Fisher's exact test; *p < 0.05, **p < 0.01). Mutation frequencies for *SGK1* and *GNA13* are derived from Morin et al., 2011. M, mutation (missense, nonsense, frameshift, and splice-site); D, deletion; G, gain; AMP, amplification; M/D*, biallelic inactivation; Tx, translocation; BSE1, binding site in exon 1.

deletions of *CDKN2A/B*; 28.2% of tFL cases versus 0% of GCB-DLBCL, p < 0.01). Moreover, tFL tends to be enriched in alterations that are generally less frequent in de novo DLBCL, such as *STAT6* mutations (23.1% versus 2.4%), *ARID1A* mutations (17.9% versus 7.2%), and *FAS* mutations and/or deletions (33.3% versus 16.7%). Particularly, although observed in both tFL and de novo DLBCL, aberrations of *MLL2*, *CREBBP*, and *BCL2* were significantly enriched in cases derived from FL transformation in comparison to GCB-DLBCL (p = 0.02, 0.0006, and 0.0001), suggesting that at least a subset of GCB-DLBCL arises from the precursor cell postulated by the results described above. In conclusion, the genome of tFL appears more similar to GCB-DLBCL but, overall, is unique in comparison to both subtypes of de novo DLBCL.

DISCUSSION

This study reports a comprehensive characterization of the coding genome of tFL, an aggressive disease with dismal prognosis, the pathogenesis of which has been incompletely understood so far. Our goal was to take advantage of systematic, genome-wide approaches in order to address key questions related to this aggressive condition that have remained unanswered or had been previously only studied by examining individual genes. These include: (1) tracing the evolutionary history of the dominant clone during transformation from indolent FL, (2) providing an assessment of the range and frequency of genetic aberrations that are associated with this event, (3) identifying genomic changes as potential genetic drivers of transformation, and (4) elucidating the relationship between DLBCL deriving from FL transformation and DLBCL arising de novo.

The first finding of our study is that, although all FL-tFL sample pairs have a clear clonal relationship, the dominant tFL clone arises in most patients from a mutated CPC through the acquisition of independent genetic events, consistent with divergent

evolution. The existence of this CPC cannot be physically demonstrated, but it can be postulated on the basis of the presence of a set of lesions that are shared between FL and tFL, and is consistent with previous studies based on the analysis of the rearranged Ig genes (Carloti et al., 2009) as well as with recent work on FL progression (Green et al., 2013). Our study does not address the topography of the tFL clone and/or the intraclonal architecture of these tumors, which require the backtracking of tFL-specific lesions in the diagnostic FL sample by ultrahigh deep sequencing analysis (Ding et al., 2012; Walter et al., 2012; Welch et al., 2012). Thus, additional studies will be needed in order to clarify whether the tFL clone temporally developed after FL diagnosis or whether it can be already detected as a minor subclone within the FL diagnostic sample. Notably, the two most prominent programs deregulated in the precursor clone—epigenetic modifications and resistance to apoptosis—represent actionable targets, and a number of drugs are already being tested in the clinic (e.g., BCL2 inhibitors, histone deacetylase inhibitors, and EZH2 inhibitors) (Sawas et al., 2011). If these lesions are essential to the survival of the fully transformed tumor cells and if the tFL clone is not already present at FL diagnosis, then the development of combination regimens incorporating drugs that specifically target this group of alterations in early stages of FL may lead to the elimination of the precursor clone, possibly preventing transformation.

Our study reveals that, although the genome of tFL is significantly more complex in comparison to FL, no unifying genetic lesion is selected during transformation to DLBCL. Nonetheless, the recurrent alteration of genes involved in the control of cell-cycle progression (*CDKN2A/B* and *MYC*) and DNA damage responses (alternative biallelic loss of *TP53* and *CDKN2A*) suggest that a loss of genetic stability and deregulated proliferation are critical steps in tFL development. Lesions affecting *CDKN2A/B* have been observed in several terminally aggressive lymphoid malignancies, including relapsed acute lymphoblastic leukemia (Mullighan et al., 2008), Richter syndrome (Fabbri et al., 2013), and ABC-DLBCL (Lenz et al., 2008), consistent with a central role in the acquisition of a more aggressive phenotype. The existence of tFL cases that lack p16 expression despite being devoid of *CDKN2A* genetic lesions suggests an even broader

involvement of this pathway through alternative epigenetic mechanisms (e.g., promoter hypermethylation or posttranscriptional modifications). Furthermore, the inactivation of epigenetic modifiers such as CREBBP/EP300 and MLL2 may also contribute to the deregulation of cell growth and DNA damage responses by interfering with p53 acetylation and activation (Pasqualucci et al., 2011a).

The genomic complexity of tFL appears to be remarkably high in respect to other hematologic malignancies, as exemplified by the presence of numerous CNAs and the evidence of ASHM. The latter may represent a major mechanism for transformation, as previously suggested for de novo DLBCL (Pasqualucci et al., 2001). Interestingly, although the physiologic SHM process is known to operate in FL (Bahler and Levy, 1992), ASHM was only observed in the dominant clone of tFL, pointing to a disruption occurring late during evolution to or selection of a more aggressive disease. Although the analysis presented here did not uncover any apparent lesions in genes that are directly involved in SHM, the alteration of histone marks owing to genetic lesions in histone modification and chromatin remodeling enzymes may induce chromatin conformation changes that favor the accessibility of nonphysiologic genomic target regions to the AID mutator. Thus, additional studies interrogating the entire genome as well as the epigenome of this cancer will be necessary in order to conclusively address this question.

Finally, and consistent with observations obtained by gene expression profiling, our data highlight tFL as a distinct disease that, although more similar to GCB-DLBCL, harbors unique combinations of oncogenic and tumor suppressor lesions in comparison to de novo DLBCL. Previous studies have suggested that transformation proceeds through two distinct pathways: one characterized by a high-proliferation signature, and a second where T cell and follicular dendritic-associated genes predominate (Davies et al., 2007; Lossos et al., 2002). We did not observe statistically significant mutual exclusion between genetic lesions affecting these two classes of mutated genes (e.g., *MYC* and *CDKN2A/B* versus *TNFRSF14* and *STAT6*). However, integrated genomic and transcriptional profiling of larger cohorts of patients will be needed in order to address this question. The unique tFL genomic landscape, combining alterations that are specific to GCB- and ABC-DLBCL as well as lesions that are uncommon in de novo DLBCL, could be responsible, at least in part, for its poor response to standard anti-DLBCL regimens. The results herein suggest the potential usefulness of combining current immunochemotherapeutic regimens that harness the proliferative and genetically unstable phenotype of tFL with more specific approaches targeting some of the pathways recurrently altered in tFL, as it could be the case for CDK4/6 activation deriving from the loss of *CDKN2A* or NF- κ B activation deriving from *TNFAIP3* deletions and/or mutations.

EXPERIMENTAL PROCEDURES

Study Panel

A discovery panel of 24 sequential fresh-frozen biopsies (12 pairs) obtained at FL diagnosis and at transformation to DLBCL (Table S1) and a screening panel of 27 tFL samples were selected on the basis of the criteria described in the

Supplemental Information, and were used for WES and SNP6.0 array analysis. Matched normal DNA was available for 4 of the 12 discovery pairs and was documented to lack contaminating tumor cells by PCR amplification of the clonally rearranged tumor-associated Ig genes as well as by SNP array analysis of the corresponding loci. Using these 12 pairs, we estimate a 99% probability of detecting mutations that affect genes at 30% prevalence and 93% probability for genes at 20% prevalence.

Whole-Exome Capture and Next-Generation Sequencing

Purified high-molecular-weight genomic DNA (~3 μ g) from the 12 FL and 39 tFL samples (n = 12 from the discovery panel and 27 from the screening panel) was enriched in protein-coding sequences with the Agilent SureSelect Human 51Mb All Exon v4 Kit (Agilent Technologies) according to the manufacturer's protocol. The resulting target-enriched pool was normalized and combined (four-plex) before high-throughput paired-end (2 $\times$ 100 bp) sequencing was performed on the Illumina HiSeq2000 System at Centillion Biosciences. The analysis produced on average 67.5 million passed-filter paired-end reads per sample (range, 51.7 to 111.6; Table S2). After filtering for duplicate reads (defined as reads with identical start and orientation), sequences were aligned to the reference human genome hg19 assembly (GRCh37) with the Burrows-Wheeler Aligner tool (version 0.5.9). The mean coverage depth (i.e., the mean number of reads covering the target exome of a haploid reference) was 80.8 $\times$ (range, 45.8 to 119.4) with an average of 89.3% of the captured region covered at >10 $\times$ (range, 83.8 to 93.4) and 73.9% covered at >30 $\times$ (range, 53.9 to 84.5; Table S2). Sequence variants, including nucleotide substitutions and small insertions and/or deletions, were obtained independently for each tumor and normal sample with the Statistical Algorithm for Variant Identification (Trifonov et al., 2013b) and were independently validated by conventional Sanger sequencing, as described in the Supplemental Experimental Procedures.

Dominant Clone Analysis

For the purpose of reconstructing the history of clonal evolution during FL transformation to DLBCL, we first estimated the percentage of tumor cells in the biopsy on the basis of the inferred CN value at the clonally rearranged Ig loci (i.e., the region of intrachromosomal deletional recombination at 14q32, 2p11, and 11q22). Then, the allelic frequency of each SNV was corrected for the fraction of tumor cells in the biopsy by calculating its expected frequency, the 95% CIs from its observed frequency, the total depth at the variant position, and the tumor content of the sample, assuming a binomial distribution.

Mutations were classified as clonal if the fraction of variant reads (upon correction for the percentage of tumor cells in the specimen) was > 20 and were classified as subclonal otherwise. Because this analysis focuses on the history of the dominant tumor clone, only SNVs that were clonally represented in at least one disease phase were considered. Then, SNVs were assigned to one of the following three categories: (1) shared mutations (i.e., mutations that are detected in the major clone of both diseases phases), (2) FL-specific mutations (i.e., mutations present in the major clone of the FL phase and either completely absent or present at subclonal levels in the paired tFL specimen, provided the difference in the corrected frequencies between the two phases was statistically significant at $p < 0.05$ [as explained below] and after excluding that its absence in the paired tFL specimen was not due to CN loss or cnLOH involving the same genomic region), and (3) tFL-specific mutations (i.e., mutations present in the major clone of the tFL phase and absent or subclonal in the FL biopsy provided the difference in the corrected frequencies between the two phases was statistically significant at $p < 0.05$ and could not be explained by CN loss or cnLOH in the latter). To assess whether the difference in frequencies between the two phases was significant, we considered a binomial distribution and calculated the probability of observing the variant in the pre-transformation (or posttransformation) phase, given the coverage depth at that position, the observed frequency of variant reads in the paired sample, and the estimated tumor content of the samples. WES and SNP6.0 analysis were used to exclude that the absence of FL-specific SNVs in the longitudinal tFL biopsy was due to recombination events, such as CN losses or cnLOH affecting the same region; the few mutations belonging to this category were conservatively classified as "shared" (in cases with paired normal DNA) or were excluded (in cases without paired normal DNA). Mutations that were

clonally represented in the tFL phase but could be detected in a small fraction of reads in the FL phase (with $p < 0.05$) were defined as DLBCL-specific, given that they most likely reflect a pre-existing smaller clone within the FL major clone, whereas mutations that were clonally represented in the FL phase but were also detected in a small fraction of the tFL population (with $p < 0.05$) were considered as FL-specific, given that they may reflect residual FL cells amidst the DLBCL clone, which were not selected during progression and expansion.

High-Density SNP Array Analysis, Sequencing Analysis of ASHM Target Genes, FISH, and Immunohistochemistry Analysis

High-density SNP array analysis, sequencing analysis of ASHM target genes, FISH, and immunohistochemistry analysis were all performed as previously described, and their detailed protocols can be found in the [Supplemental Information](#).

ACCESSION NUMBERS

The Affymetrix SNP Array 6.0 data and the whole-exome sequencing data from the 12 FL and 39 tFL cases have been deposited in NCBI database of Genotypes and Phenotypes under accession number phs000328.v2.p1.

SUPPLEMENTAL INFORMATION

Supplemental Information contains Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2013.12.027>.

ACKNOWLEDGMENTS

We thank G. Fabbri for discussions, V. Miljkovic for help with the SNP array hybridization, and the Molecular Pathology Shared Resource and the Molecular Cytogenetics and Epigenetics Shared Resource of the Herbert Irving Comprehensive Cancer Center at Columbia University for histology service and cytogenetics service, respectively. We also thank R. Feldman and R. Mei for expert assistance with the whole-exome capture and sequencing, which were completed at Centrillion Biosciences. Automated DNA sequencing was performed at GENEWIZ. This work was supported by the NIH (RO1-CA37295 to R.D.-F., RO1-CA172492-01 to L.P., and RO1-CA136537 to S.N.M.), a Specialized Center of Research grant from the Leukemia & Lymphoma Society (to R.D.-F.), the Northeast Biodefense Center (U54-AI057158), and the National Library of Medicine (1R01LM010140-01 to R.R.), the AIRC Special Program Molecular Clinical Oncology – 5 per mille (contract 10007 to G.G. and G.I.), and the Cariplo Foundation (to G.G.). L.P. is on leave from the Institute of Hematology, University of Perugia Medical School. M.F. was enrolled in the PhD program in Clinical and Experimental Medicine at Amedeo Avogadro University of Eastern Piedmont and was supported in part by the Novara AIL.

Received: August 11, 2013

Revised: November 27, 2013

Accepted: December 16, 2013

Published: January 2, 2014

REFERENCES

Akasaka, T., Lossos, I.S., and Levy, R. (2003). BCL6 gene translocation in follicular lymphoma: a harbinger of eventual transformation to diffuse aggressive lymphoma. *Blood* 102, 1443–1448.

Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A., Boldrick, J.C., Sabet, H., Tran, T., Yu, X., et al. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403, 503–511.

Bahler, D.W., and Levy, R. (1992). Clonal evolution of a follicular lymphoma: evidence for antigen selection. *Proc. Natl. Acad. Sci. USA* 89, 6770–6774.

Bergsagel, P.L., and Kuehl, W.M. (2013). Degree of focal immunoglobulin heavy chain locus deletion as a measure of B-cell tumor purity. *Leukemia* 27, 2067–2068.

Carlotti, E., Wrench, D., Matthews, J., Iqbal, S., Davies, A., Norton, A., Hart, J., Lai, R., Montoto, S., Gribben, J.G., et al. (2009). Transformation of follicular lymphoma to diffuse large B-cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone. *Blood* 113, 3553–3557.

Challa-Malladi, M., Lieu, Y.K., Califano, O., Holmes, A.B., Bhagat, G., Murty, V.V., Dominguez-Sola, D., Pasqualucci, L., and Dalla-Favera, R. (2011). Combined genetic inactivation of $\beta 2$ -Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell* 20, 728–740.

Davies, A.J., Rosenwald, A., Wright, G., Lee, A., Last, K.W., Weisenburger, D.D., Chan, W.C., Delabie, J., Braziel, R.M., Campo, E., et al. (2007). Transformation of follicular lymphoma to diffuse large B-cell lymphoma proceeds by distinct oncogenic mechanisms. *Br. J. Haematol.* 136, 286–293.

Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., et al. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481, 506–510.

Fabbri, G., Rasi, S., Rossi, D., Trifonov, V., Khiabani, H., Ma, J., Grunn, A., Fangazio, M., Capello, D., Monti, S., et al. (2011). Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J. Exp. Med.* 208, 1389–1401.

Fabbri, G., Khiabani, H., Holmes, A.B., Wang, J., Messina, M., Mullighan, C.G., Pasqualucci, L., Rabadan, R., and Dalla-Favera, R. (2013). Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J. Exp. Med.* 210, 2273–2288.

Green, M.R., Gentles, A.J., Nair, R.V., Irish, J.M., Kihira, S., Liu, C.L., Kela, I., Hopmans, E.S., Myklebust, J.H., Ji, H., et al. (2013). Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood* 121, 1604–1611.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl. Acad. Sci. USA* 106, 5181–5186.

Holzova, E., Vonaarbourg, C., Stolzenberg, M.C., Arkwright, P.D., Selz, F., Priour, A.M., Blanche, S., Bartunkova, J., Vilmer, E., Fischer, A., et al. (2004). Autoimmune lymphoproliferative syndrome with somatic Fas mutations. *N. Engl. J. Med.* 351, 1409–1418.

Korbel, J.O., and Campbell, P.J. (2013). Criteria for inference of chromothripsis in cancer genomes. *Cell* 152, 1226–1236.

Kridel, R., Sehn, L.H., and Gascoyne, R.D. (2012). Pathogenesis of follicular lymphoma. *J. Clin. Invest.* 122, 3424–3431.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218.

Lenz, G., Wright, G.W., Emre, N.C., Kohlhammer, H., Dave, S.S., Davis, R.E., Carty, S., Lam, L.T., Shaffer, A.L., Xiao, W., et al. (2008). Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl. Acad. Sci. USA* 105, 13520–13525.

Lo Coco, F., Gaidano, G., Louie, D.C., Offit, K., Chaganti, R.S., and Dalla-Favera, R. (1993). p53 mutations are associated with histologic transformation of follicular lymphoma. *Blood* 82, 2289–2295.

Lohr, J.G., Stojanov, P., Lawrence, M.S., Auclair, D., Chapuy, B., Sougnez, C., Cruz-Gordillo, P., Knoechel, B., Asmann, Y.W., Slager, S.L., et al. (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. USA* 109, 3879–3884.

Lossos, I.S., and Gascoyne, R.D. (2011). Transformation of follicular lymphoma. *Best Pract. Res. Clin. Haematol.* 24, 147–163.

- Lossos, I.S., Alizadeh, A.A., Diehn, M., Warnke, R., Thorstenson, Y., Oefner, P.J., Brown, P.O., Botstein, D., and Levy, R. (2002). Transformation of follicular lymphoma to diffuse large-cell lymphoma: alternative patterns with increased or decreased expression of c-myc and its regulated genes. *Proc. Natl. Acad. Sci. USA* 99, 8886–8891.
- Mandelbaum, J., Bhagat, G., Tang, H., Mo, T., Brahmachary, M., Shen, Q., Chadburn, A., Rajewsky, K., Tarakhovsky, A., Pasqualucci, L., and Dalla-Favera, R. (2010). BLIMP1 is a tumor suppressor gene frequently disrupted in activated B cell-like diffuse large B cell lymphoma. *Cancer Cell* 18, 568–579.
- Martinez-Climent, J.A., Alizadeh, A.A., Segraves, R., Blesa, D., Rubio-Moscardo, F., Albertson, D.G., Garcia-Conde, J., Dyer, M.J., Levy, R., Pinkel, D., and Lossos, I.S. (2003). Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations. *Blood* 101, 3109–3117.
- Montoto, S., and Fitzgibbon, J. (2011). Transformation of indolent B-cell lymphomas. *J. Clin. Oncol.* 29, 1827–1834.
- Morin, R.D., Johnson, N.A., Severson, T.M., Mungall, A.J., An, J., Goya, R., Paul, J.E., Boyle, M., Woolcock, B.W., Kuchenbauer, F., et al. (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* 42, 181–185.
- Morin, R.D., Mendez-Lago, M., Mungall, A.J., Goya, R., Mungall, K.L., Corbett, R.D., Johnson, N.A., Severson, T.M., Chiu, R., Field, M., et al. (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* 476, 298–303.
- Morin, R.D., Mungall, K., Pleasance, E., Mungall, A.J., Goya, R., Huff, R.D., Scott, D.W., Ding, J., Roth, A., Chiu, R., et al. (2013). Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* 122, 1256–1265.
- Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., et al. (2007). Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446, 758–764.
- Mullighan, C.G., Phillips, L.A., Su, X., Ma, J., Miller, C.B., Shurtleff, S.A., and Downing, J.R. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* 322, 1377–1380.
- Ngo, V.N., Young, R.M., Schmitz, R., Jhavar, S., Xiao, W., Lim, K.H., Kohlhammer, H., Xu, W., Yang, Y., Zhao, H., et al. (2011). Oncogenically active MYD88 mutations in human lymphoma. *Nature* 470, 115–119.
- Pasqualucci, L., Neumeister, P., Goossens, T., Nanjangud, G., Chaganti, R.S., Küppers, R., and Dalla-Favera, R. (2001). Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412, 341–346.
- Pasqualucci, L., Compagno, M., Houldsworth, J., Monti, S., Grunn, A., Nandula, S.V., Aster, J.C., Murty, V.V., Shipp, M.A., and Dalla-Favera, R. (2006). Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J. Exp. Med.* 203, 311–317.
- Pasqualucci, L., Dominguez-Sola, D., Chiarenza, A., Fabbri, G., Grunn, A., Trifonov, V., Kasper, L.H., Lerach, S., Tang, H., Ma, J., et al. (2011a). Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* 471, 189–195.
- Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O., et al. (2011b). Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* 43, 830–837.
- Pinyol, M., Cobo, F., Bea, S., Jares, P., Nayach, I., Fernandez, P.L., Montserrat, E., Cardesa, A., and Campo, E. (1998). p16(INK4a) gene inactivation by deletions, mutations, and hypermethylation is associated with transformed and aggressive variants of non-Hodgkin's lymphomas. *Blood* 91, 2977–2984.
- Ritz, O., Guiter, C., Castellano, F., Dorsch, K., Melzner, J., Jais, J.P., Dubois, G., Gaulard, P., Möller, P., and Leroy, K. (2009). Recurrent mutations of the STAT6 DNA binding domain in primary mediastinal B-cell lymphoma. *Blood* 114, 1236–1242.
- Rossi, D., Berra, E., Cerri, M., Deambrogi, C., Barbieri, C., Franceschetti, S., Lunghi, M., Conconi, A., Paulli, M., Matolcsy, A., et al. (2006). Aberrant somatic hypermutation in transformation of follicular lymphoma and chronic lymphocytic leukemia to diffuse large B-cell lymphoma. *Haematologica* 91, 1405–1409.
- Rossi, D., Trifonov, V., Fangazio, M., Brusca, A., Rasi, S., Spina, V., Monti, S., Vaisitti, T., Arruga, F., Famà, R., et al. (2012). The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J. Exp. Med.* 209, 1537–1551.
- Saito, M., Novak, U., Piovani, E., Basso, K., Sumazin, P., Schneider, C., Crespo, M., Shen, Q., Bhagat, G., Califano, A., et al. (2009). BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc. Natl. Acad. Sci. USA* 106, 11294–11299.
- Sander, C.A., Yano, T., Clark, H.M., Harris, C., Longo, D.L., Jaffe, E.S., and Raffeld, M. (1993). p53 mutation is associated with progression in follicular lymphomas. *Blood* 82, 1994–2004.
- Sawas, A., Diefenbach, C., and O'Connor, O.A. (2011). New therapeutic targets and drugs in non-Hodgkin's lymphoma. *Curr. Opin. Hematol.* 18, 280–287.
- Sherr, C.J. (2004). Principles of tumor suppression. *Cell* 116, 235–246.
- Siegel, R.M., Frederiksen, J.K., Zacharias, D.A., Chan, F.K., Johnson, M., Lynch, D., Tsien, R.Y., and Lenardo, M.J. (2000). Fas preassociation required for apoptosis signaling and dominant inhibition by pathogenic mutations. *Science* 288, 2354–2357.
- Swerdlow, S.H., Campo, E., Harris, N.L., Jaffe, E.S., Pileri, S.A., Stein, H., Thiele, J., and Vardiman, J.W. (2008). WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. International Agency for Research on Cancer (Geneva: World Health Organization).
- Trifonov, V., Pasqualucci, L., Dalla Favera, R., and Rabadan, R. (2013a). MutComFocal: an integrative approach to identifying recurrent and focal genomic alterations in tumor samples. *BMC Syst. Biol.* 7, 25.
- Trifonov, V., Pasqualucci, L., Tiacci, E., Falini, B., and Rabadan, R. (2013b). SAVI: a statistical algorithm for variant frequency identification. *BMC Syst. Biol.* 7 (Suppl 2), S2.
- Trinh, D.L., Scott, D.W., Morin, R.D., Mendez-Lago, M., An, J., Jones, S.J., Mungall, A.J., Zhao, Y., Schein, J., Steidl, C., et al. (2013). Analysis of FOXO1 mutations in diffuse large B-cell lymphoma. *Blood* 121, 3666–3674.
- van Eijk, M., Defrance, T., Hennino, A., and de Groot, C. (2001). Death-receptor contribution to the germinal-center reaction. *Trends Immunol.* 22, 677–682.
- Walter, M.J., Shen, D., Ding, L., Shao, J., Koboldt, D.C., Chen, K., Larson, D.E., McLellan, M.D., Dooling, D., Abbott, R., et al. (2012). Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* 366, 1090–1098.
- Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The origin and evolution of mutations in acute myeloid leukemia. *Cell* 150, 264–278.