UNIVERSITÀ DEL PIEMONTE ORIENTALE

School of Medicine

Department of Translational Medicine

Ph.D. Program in Medical Sciences and Biotechnology

XXXVI Cycle

MOLECULAR CLUSTERING AND FRAGMENTOMICS ON CIRCULATING TUMOR DNA IMPROVE OUTCOME PREDICTION IN DIFFUSE LARGE B CELL LYMPHOMA

Tutor:

Chiar.mo Prof. Gianluca GAIDANO

Coordinator:

Prof.ssa Marisa GARIGLIO

Candidate: **Donatella TALOTTA**

Matricula: **20028771**

Academic Year 2022/2023

MED/15 MALATTIE DEL SANGUE

Table of Contents

ABSTRACT **3**

1. INTRODUCTION **4**

1.1 Diffuse large B cell lymphoma (DLBCL) *4*

1.2 Molecular pathogenesis of DLBCL *5*

1.3 Genetic lesions associated with GCB-DLBCL *8*

1.4 Genetic lesions associated with ABC-DLBCL *9*

1.5 Novel DLBCL molecular subtypes *10*

1.6 Liquid biopsy and ctDNA analysis in DLBCL *16*

1.7 Cell-free DNA fragmentation in DLBCL *17*

1.8 DLBCL treatment and prognostic factors *18*

2. AIMS OF THE STUDY **20**

3. MATERIALS AND METHODS **21**

3.1 Patients and biological samples *21*

3.2 Separation of granulocytes from peripheral blood (PB) *21*

3.3 Lymph node tissue processing for gDNA extraction *21*

3.4 Extraction of tumor and normal gDNA *22*

3.5 Extraction of tumor gDNA from formalin-fixed paraffin-embedded (FFPE) *22*

3.6 Plasma cell-free DNA (cfDNA) extraction *22*

3.7 DNA quantification and fragmentation *23*

3.8 Library design for hybrid selection *23*

3.9 CAPP-seq library preparation and sequencing *23*

3.10 Bionformatic Pipeline and Analysis *24*

3.11 Copy-number analysis on ctDNA *25*

3.12 ctDNA quantitation and molecular groups classification *25*

3.13 Analysis of cfDNA fragments *25*

3.14 Statistical analysis *26*

4. RESULTS **27**

4.1 Patient characteristics *27*

4.2 Molecular profile of the DLBCL cohort *28*

4.3 Comparison of molecular clustering between ctDNA and tissue biopsy *30*

4.4 Integration of ctDNA levels and molecular clusters identified on ctDNA *32*

4.5 Fragmentome analysis *34*

5. DISCUSSION **36**

REFERENCES **38**

ABSTRACT

Diffuse large B-cell lymphoma (DLBCL) can be classified into molecular clusters that reflect the genetic and clinical heterogeneity of the disease. These subtypes have been identified on tissue biopsy that might not accurately capture the anatomical variability of DLBCL. The analysis of ctDNA in DLBCL is an accurate and non-invasive approach to retrieve molecular information not otherwise detectable in tissue biopsies. Moreover, the quantification of baseline ctDNA stratifies DLBCL outcomes, but its integration with molecular clustering has not been evaluated yet. The aims of the study were to evaluate the role of ctDNA in detecting DLBCL molecular clusters and to assess whether molecular clustering combined with ctDNA levels may improve outcome prediction. A cohort of 166 newly diagnosed DLBCL patients treated with R-CHOP and provided with ctDNA from plasma and germline gDNA from granulocytes (N=77/166 also with gDNA from LNF biopsies), were enrolled in the study. The LyV4.0 CAPP-seq assay consisting of a panel of 59 genes relevant to DLBCL and the LymphGen clustering tool were used. Clustering analysis allowed to classify 42.8% of patients based on ctDNA mutational profile. More precisely, 13.3% patients were classified as EZB, 11.4% as BN2, 8.4% as MCD, 6.0% as ST2 and 3.6% as A53. MCD patients exhibited a shorter PFS compared to other patients (4-years PFS: 40% *vs* 66.6%, p=0.032). The LymphGen tool assigned 46.5% of patients on the tissue biopsy. The concordance rate between ctDNA and LNF molecular clustering was 95.8%. Clusters ST2 and BN2 were associated with an excellent outcome after R-CHOP. Patients with ctDNA load < $2.5\log_{10}$hGE and identified as ST2/BN2 showed superior outcome compared to patients with ctDNA load >$2.5\log_{10}$hGE and not assigned to ST2/BN2 clusters (4-years PFS and OS: 77.4% and 94.4% *vs* 39.9% and 46.5%, p=0.017 and p=0.002, respectively). Compared to ctDNA levels only, the addition of BN2/ST2 cluster improved the C statistics of the model (0.62 vs 0.59 for PFS and 0.68 *vs* 0.63 for OS). The assignment to cluster BN2/ST2 maintained an independent association with an improved PFS (HR 0.37, 95% CI 0.14-0.95, p=0.04) and OS (HR 0.07, 95% CI 0.01-0.55, p=0.01) when adjusted for ctDNA levels, IPI and COO. Finally, patients with no detectable mutations on ctDNA or not assigned to a specific cluster, were stratified based on their fragmentome. Subnucleosomal group had less favourable outcome compared to nucleosomal group (4-years PFS and OS: 25.9% and 44.4% *vs* 76.9% and 87.3%; both p<0.001). The prognostic impact of subnucleosomal cfDNA was maintained in multivariate analysis both for PFS (HR 3.12, 95% CI 1.36-7.15, p=0.007) and OS (HR 3.06, 95% CI 1.17-8.02, p=0.023) when adjusted for ctDNA levels.

1. INTRODUCTION

1.1 Diffuse large B cell lymphoma (DLBCL)

Diffuse large B cell lymphoma (DLBCL) is the most prevalent lymphoid malignancy among adults, accounting for up to 35% of non-Hodgkin lymphomas (Swerdlow *et al.*, 2016). Approximately 60% of patients achieve enduring remission with the standard front-line rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) immunochemotherapy. Conversely, the relapsed/refractory subset of DLBCL patients is not easily identified *a priori* and remains an unmet clinical need (Sehn *et al.*, 2021). Progress in dissecting DLBCL biology has enabled the classification of patients into different molecular subtypes, each distinguished by unique pathogenetic mechanisms and clinical outcomes. DLBCL arises due to the malignant transformation of mature B cells that have undergone the germinal center (GC) reaction (Pasqualucci *et al.*, 2018). GCs are dynamic compartments that form when B cells are challenged by a foreign antigen, and represent the primary site for clonal expansion and antibody affinity maturation (De Silva *et al.*, 2015; Mesin *et al.*,2016). Gene expression profiling (GEP) identified two major subtypes of DLBCL: germinal center B cell-like (GCB) DLBCL, which derives from GC light zone B cells, and activated B cell-like (ABC) DLBCL, which originates from a later stage of GC differentiation when B cells are committed to plasmablastic differentiation (Alizadeh *et al*., 2000) (Figure 1).

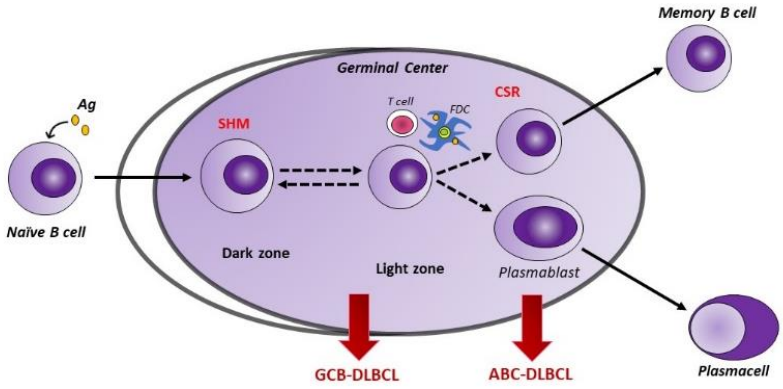

Figure 1. Cell-of-origin of DLBCL subtypes

In addition, 15-30% of cases remain unclassified and rapresent a unique entity with peculiar molecular and clinical features (Alizadeh *et al*., 2000).

Consistent with their putative cell-of-origin (COO), GCB-DLBCLs display high-level expression of the master regulator B cell lymphoma/leukemia-6 (BCL-6) and harbour hypermutated immunoglobulin genes with ongoing somatic hypermutation, whereas ABC-DLBCLs show activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) and B cell receptor (BCR) signalling pathways and upregulation of genes required for plasmatic differentiation (Shaffer *et al.*,2012). Stratifying DLBCL risk based on the cell-of-origin holds prognostic significance in the context of R-CHOP treatment. GCB-DLBCL patients have a more favourable clinical outcome, whereas ABC-DLBCL patients fail more frequently R-CHOP therapy (Figure 2) (Rosenwald *et al.*, 2002; Thieblemont *et al.*, 2011; Scott *et al.*, 2015; Dunleavy *et al.*, 2014). Therefore, the 2016 revision of the World Health Organization (WHO) classification of lymphoid neoplasms requires the identification of molecular subgroups of DLBCL (Swerdlow *et al.*, 2016).

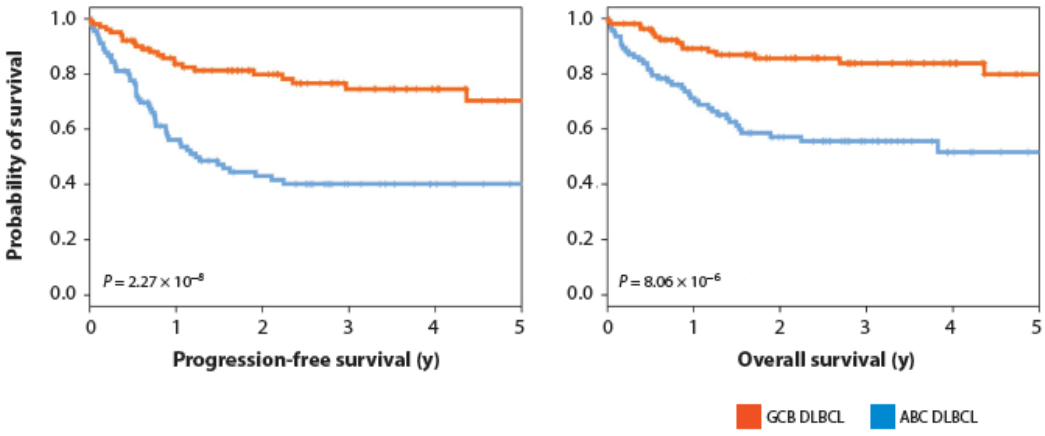

Figure 2. Clinical outcome of GCB-DLBCL and ABC-DLBCL patients treated with R-CHOP

1.2 Molecular pathogenesis of DLBCL

DLBCL pathogenesis is associated with a consistent set of genetic abnormalities, with some being common to both GCB- and ABC-DLBCL, while others are specific to each molecular subtype. Shared genetic lesions in both GCB and ABC-DLBCL encompass: *i)* modifications of histone/chromatin modifiers; *ii)* dysregulation of BCL-6 activity; *iii)* evasion from immune surveillance, and *iv)* additional mutations, including somatic copy number alterations (SCNAs) and structural variants (SVs).

i) Alterations of histone/chromatin modifiers

Monoallelic and biallelic somatic mutations of lysine-specific methyltransferase 2D (*KMT2D*) gene are found in 30% of DLBCL cases, representing the single most frequent genetic aberration linked to this disease (Morin *et al.,* 2011; Pasqualucci *et al.,* 2011). The *KMT2D* gene encodes for a member of the SET1 family of histone methyltransferases, primarily responsible for trimethylating the lysine at position 4 of histone H3 (H3K4) thus inducing an active chromatin conformation (Shilatifard *et al.,* 2012). *KMT2D* mutations comprise mostly truncating events and missense mutations that impair the protein enzymatic function by removing the C-terminal cluster of conserved domains, including the SET domain (Morin *et al.,* 2011; Pasqualucci *et al.,* 2011). Deletion of *KMT2D* in mouse pre-GC B cells results in a significant expansion of GC B cells, supporting the notion that *KMTD2* inactivation is an early event (Zhang *et al.,* 2015). In one-third of DLBCL samples, somatic mutations and/or deletions affect the acetyltransferase genes coding for CREB binding protein (*CREBBP*) and, less frequently, the E1A binding protein P300 (*EP300*) (Pasqualucci *et al.,* 2011). These enzymes function as pleiotropic regulators of gene expression, catalyzing the addition of acetyl groups to specific lysine residues in histone and nonhistone proteins, including tumor protein p53 (TP53) (Brooks *et al.,* 2011; Goodman *et al.,* 2000). DLBCL-associated mutations and small insertion or deletion (indels) disrupt the function of the CREBBP/EP300 proteins by either removing the histone acetyltransferase domain or introducing amino acid changes within this domain, resulting in reduced affinity for Acetyl-CoA (Pasqualucci *et al.,* 2011).

ii) Deregulation of BCL-6 activity

Approximately one-third of DLBCL of cases carries chromosomal translocations that prevent BCL-6 downregulation by positioning the intact coding domain of the gene downstream to heterologous promoter sequences derived from chromosomal partners, leading to deregulation of *BCL-6* expression by a mechanism known as “promoter substitution” (Iqbal *et al.,* 2007; Ye *et al.,* 1995). As a consequence, these chromosomal translocations induce constitutive *BCL-6* expression leading to the pathological maintenance of the GC phenotype, including tolerance to DNA damage and block of terminal differentiation (Pasqualucci *et al.,* 2018). The *BCL-6* gene can also undergo alterations due to somatically acquired point mutations, some of which disrupt a negative autoregulatory loop that the BCL-6 protein employs to control its own transcription (Pasqualucci *et al.,*1998; Shen *et al.,*1998; Pasqualucci *et al.,* 2003; Wang *et al.,* 2002).

In addition to direct alterations involving the *BCL-6* locus, various other genetic abnormalities disrupt the expression of BCL-6 through alternative or indirect mechanisms. Deleterious mutations of *CREBBP/EP300* impair acetylation-mediated inactivation of the BCL-6 transrepressive function (Pasqualucci *et al.,* 2011; Bereshchenko *et al.,* 2002). Approximately 15% of cases exhibit somatic mutations in the myocyte enhancer binding factor 2B (*MEF2B*) (Morin *et al.,* 2011; Ying *et al.,* 2013). *MEF2B* mutations enhance its transactivator function and consequently *BCL-6* expression (Ying *et al.,* 2013). Finally, 5% of DLBCL cases display loss-of-function mutations and/or deletions of F-box only protein 11 (*FBXO11*), a ubiquitin adaptor protein that normally targets BCL-6 for proteosomal degradation (Duan *et al.,* 2012; Schneider *et al.,* 2016).

iii) Evasion from immune surveillance

Approximately 60% of DLBCL samples exhibit a deficiency in expressing the major histocompatibility complex (MHC) class I, attributable to various genetic and epigenetic mechanisms. These mechanisms encompass: *i)* deletions affecting the *B2M* locus, encoding for the beta-2-microglobulin invariant subunit crucial for the formation of the human leukocyte antigen class I (HLA-I) complex on the cell surface; *ii)* point mutations; and *iii)* genomic loss of the *HLA* loci. Addionally, the lack of expression or aberrant cytoplasmic localization of the B2M/HLA-I protein represent key genetic and epigenetic mechanisms contributing to immune evasion (Challa-Malladi *et al.,* 2011).

iv) Other lesions

In approximately 20% of DLBCL cases, *TP53* mutations have been observed and proven to independently indicate a poorer prognosis (Karube *et al.,* 2018; Young *et al.,* 2007; Ichikawa *et al.,* 1997; Kerbauy *et al.,* 2004; Xu-Monette *et al.,* 2012). ypically resulting in loss-of-function, *TP53* mutations can occur in the DNA binding domain or other locations; notably, mutations in the DNA binding domain have a more pronounced adverse prognostic impact (Karube *et al.,* 2018; Young *et al.,* 2007; Xu-Monette *et al.,* 2012). Forkhead box protein O1 (*FOXO1*) mutations have been detected in 8% to 10% of all DLBCL cases. These mutations involve amino acid alterations clustered around a phosphorylation site crucial for serine/threonine kinase (AKT)-mediated nuclear-cytoplasmic translocation. This prevents the inactivation of FOXO1 in response to phosphoinositide 3-kinase (PI3K) signalling (Trinh *et al.,* 2013). FOXO1, a transcription factor pivotal in B cell differentiation, has its activity negatively regulated by PI3K-AKT, as highlighted by Dominguez-Sola *et al.,* 2015.

1.3 Genetic lesions associated with GCB-DLBCL

Genetic abnormalities that have been predominantly associated with GCB-DLBCL comprise: *i)* chromosomal translocations of B cell lymphoma/leukemia-2 (*BCL*-2) and myelocytomatosis viral oncogene homolog (*MYC*) genes; *ii)* mutations of the enhancer of zeste homolog 2 (*EZH2*) methyltransferase and *iii)* mutations of the tumor necrosis factor-receptor superfamily member 14 (*TNFRSF14*).

i) Chromosomal translocations of BCL-2 and MYC

BCL-2 is a key antiapoptotic molecule expressed in most tissues but notably absent in the GC, reflecting the need of GC B cells to maintain a inherent proapoptotic program (Saito *et al.,* 2009; Iqbal *et al.,* 2004). Approximately 30% of the GCB-DLBCL cases exhibit the t(14;18) translocation, causing *BCL-2* coding exons to be juxtaposed under the control of the immunoglobulin (IG) locus, resulting in its constitutive expression. Dysregulation of BCL-2 has been associated with an unfavorable outcome, particularly in conjunction with *MYC* deregulation (Barrans *et al.,* 2003). The *MYC* gene encodes a transcription factor that controls numerous biological functions, including proliferation, cell growth, telomerase activity, energy metabolism, differentiation, and apoptosis (Conacci-Sorrell *et al.,* 2014). In 10% to 14% of GCB-DLBCLs, *MYC* is ectopically and constitutively expressed, often as the result of chromosomal translocations that fuse its intact coding domain with the IG heavy or light chains loci (Karube *et al.,* 2015; Dalla-Favera *et al.,* 2010). The presence of *MYC* translocations has been associated with a poorer prognosis in DLBCL (Barrans *et al.,* 2010). In 5% to 10% of DLBCL cases, chromosomal translocations involving both *MYC* and *BCL-2* coexist (Burotto *et al.,* 2016).

ii) Mutations of EZH2 methyltransferase

EZH2 encodes a SET-domain histone methyltransferase that is responsible for trimethylating the lysine 27 residue of histone H3 (H3K27me3) (Czermin *et al.,* 2002). Around 22% of GCB-DLBCL cases display heterozygous mutations in the *EZH2* gene. In most instances, these mutations involve the substitution of a single evolutionarily conserved residue (Y641) within the protein SET domain, enhancing its ability to catalyse the addition of H3K27me3 mark (Morin *et al.*, 2010; Sneeringer *et al.*, 2010).

iii) Mutations of TNFRSF14

TNFRSF14 encodes for a member of the tumor necrosis factor-receptor superfamily that is expressed in both T and B cells and can deliver opposing signals based on its specificity for diverse ligands (Steinberg *et al.,* 2011). Deletions and mutations of *TNFRSF14*, including missense (~50%), nonsense (~40%), and frameshift (2.5%) events confined to the exons encoding for its ectodomain, are recurrently found in DLBCL and segregate with the GCB subtype (30% of cases) (Boice *et al.,* 2016). One mechanism underlying the tumorigenic effect of *TNFRSF14* loss is the inhibition of cell-cell interactions between this receptor and its ligand B- and T-lymphocyte attenuator (BTLA), which induces a tumor-supportive microenvironment (Boice *et al.,* 2016).

1.4 Genetic lesions associated with ABC-DLBCL

The core biology of ABC-DLBCL is defined by alterations leading to constitutive activation of NF-κB pathway, the expression of which is required for ABC-DLBCL survival (Pasqualucci *et al.,* 2018). These lesions include: *i)* mutations in the B cell receptor (BCR) signalling pathway; *ii)* mutations of myeloid differentiation primary response 88 (*MYD88*) and, *iii)* mutations of tumor necrosis factor alpha induced protein 3 (*TNFAIP3*).

i) Mutations in the BCR signalling pathway

ABC-DLBCL exhibits a persistent "chronic active" state of BCR signalling, sustained by genetic alterations affecting proximal members of the pathway (Davis *et al.,* 2010). In 21% of cases, these alterations are gain-of-function mutations in the immunoreceptor tyrosine-based activation motifs of cluster of differentiation 79B *(CD79B*), which uphold BCR signalling by mitigating negative feedback (Davis *et al.,* 2010). In 9% of cases, mutations involve the gene enconding the caspase recruitment domain family member 11 *(CARD11*) (Lenz *et al.,* 2008), a component of the "signalosome" complex crucial for the proper transduction of BCR signalling (Thome *et al.,* 2004). These events cluster in the exons that encode for the protein coiled-coil domain, thereby enhancing the ability of *CARD11* to transactivate NF-κB target genes (Davis *et al.,* 2010; Knies *et al.,* 2015).

ii) Mutations of MYD88

Approximately 30% of ABC-DLBCLs harbour mutations that result in the hotspot L265P substitution within the hydrophobic core of the MYD88 Toll/interleukin-1 receptor (TIR)-domain. This adaptor molecule plays a pivotal role in transmitting signals from the Toll-like receptor (TLR) to NF-κB transcription complex (Ngo *et al.*, 2011; Lam *et al.*, 2008). Specifically, the L265P substitution triggers the activity and phosphorylation of interleukin-1 receptor-associated kinase 4 (IRAK4) through the spontaneous formation of a protein complex involving interleukin 1 receptor-associated kinase 1 (IRAK1) and IRAK4. Consequently, this complex can activate NF-κB and the Janus kinase/signal transducers and activators of transcription (JAK/STAT3) responses. (Rhyasen *et al.*, 2015).

iii) Mutations of TNFAIP3

Nearly 30% of ABC-DLBCL cases display biallelic truncating mutations and/or focal deletions inactivating the *TNFAIP3* gene, which encodes a dual function ubiquitin-modification enzyme involved in the negative regulation of NF-κB responses triggered by TLR and BCR signalling (Compagno *et al.*, 2009; Kato *et al.*, 2009; Boone *et al.*, 2004). The inactivation of TNFAIP3 may play a role in lymphomagenesis by diminishing excessively prolonged NF-κB responses, thereby contributing to the development of lymphoma (Compagno *et al.*, 2009; Kato *et al.*, 2009).

1.5 Novel DLBCL molecular subtypes

The advent of gene expression profiling (GEP), employed to characterize the two major COO subtypes - GCB-DLBCL and ABC-DLBCL - along with a third subtype comprising "unclassified" cases, marked the initial advancements in establishing a molecular diagnosis for DLBCL subtypes (Alizadeh *et al.*, 2000; Rosenwald *et al.*, 2002). The application of this classification has proven valuable in gaining insights into the different responses of DLBCL patients to target therapies (Wilson *et al.*, 2015). Nevertheless, the COO differentiation, which offers a phenotypic description rather than a genetic one, does not entirely account for the heterogeneous responses to chemotherapy and targeted therapy (Wright *et al.*, 2020). Multiplatform genomic analyses that integrate SCNAs, chromosomal translocations, recurrent mutations, and GEP resulted in a novel genetic subclassification for this disease.

Remarkably, two independent studies revealed a very similar genetic taxonomy when subdividing DLBCL tumors into distinct subtypes, characterized by distinct outcomes and pathway dependencies (Chapuy *et al.,* 2018; Schmitz *et al.,* 2018). To characterize molecular subtypes, Chapuy and colleagues analyzed 304 DLBCL biopsies to identify potential cancer driver genes. The integration of recurrent mutations, low-frequency alterations, SCNAs and SVs identified, resulted in five DLBCL clusters termed C1-C5, each exhibiting a distinct genetic signature and clinical outcomes, along with a small cluster lacking any detectable alterations (termed C0) (Weber and Schmitz, 2022).

Conversely, the parallel genomic analysis of DLBCL utilized the GenClass algorithm to group 574 tumors into genetic subtypes. The algorithm begins with an initial set of genetic aberrations and iteratively explores all possible rearrangements of cases into subgroups, optimizing for genetic distinctiveness (Schmitz *et al.,* 2018). The genetic alterations defining the foundational classes of this taxonomy were notably more prevalent in either ABC, GCB, or unclassified DLBCL. This methodology identified four genetic subtypes, labeled MCD, BN2, N1, and EZB. These designations correspond to the seed alterations that characterize the genetic subgroups: *MYD88*^{L265P} and *CD79B* mutations in MCD, *BCL6* translocations and *NOTCH2* mutations in BN2, *NOTCH1* mutations in N1, and *EZH2* mutations and *BCL2* translocations in EZB (Weber and Schmitz, 2022).

In a subsequent study that extended the GenClass algorithm, genomic data from over 1200 DLBCL cases introduced two novel genetic subgroups: A53 defined by cases exhibiting *TP53* inactivation coupled with aneuploidy, and ST2 for tumors characterized by mutations in serum/glucocorticoid regulated kinase 1 (*SGK1*) and tet methylcytosine dioxygenase 2 (*TET2*) genes (Wright *et al.*, 2020).

MCD (Cluster C5)

The MCD genetic subtype (cluster C5) is primarily composed of ABC-DLBCLs and is characterized by significantly inferior survival compared to other genetic subgroups. Genetic hallmarks of the MCD subtype are *MYD88*^{L265P} and *CD79B* mutations, which cooperatively activate NF-κB signaling (Phelan *et al.,* 2018). MCD tumors frequently harbour deletions in the *9p21.3*/cyclin dependent kinase inhibitor 2A (*CDKN2A*), which encodes the cell cycle inhibitor p16, thus accelerating cell proliferation (Monti *et al.,* 2012). Additionally, frequent chromosomal amplification of the *BCL2* locus leads to the upregulation of the anti-apoptotic protein BCL2, contributing to sustained viability.

Approximately 70% of MCD DLBCL cases have genetic aberrations affecting immune regulator genes. These include inactivation of MHC class I genes, hindering antigen presentation; gene fusions increasing expression of *CD274* and *PDCD1LG2* (encoding PD-L1 and PD-L2), reducing T cell activation; and mutations/deletions of *CD58*, causing diminished NK cell activation (Challa-Mallad *et al.,* 2011).

A53 (Cluster C2)

The A53 subtype (cluster C2) is characterized by biallelic inactivation of *TP53* and *17p*, *9p21.3*/*CDKN2A* and *13q14.2*/retinoblastoma-associated protein (*RB1*) copy loss, which perturb chromosomal stability and cell cycle. Additionally, A53 tumors frequently deactivate tumor protein p53 binding protein 1 (*TP53BP1*), a gene responsible for sensing DNA damage and working in conjunction with p53 to prevent genomic instability (Morales *et al.,* 2006). Notably, lymphomas of the A53 genetic subtype exhibit alterations in multiple chromosomal regions, including copy number changes across entire arms of various chromosomes. Specific focal deletions target tumor protein 73 (*TP73)* and inhibitor of growth protein 1 (*ING1*) loci (*1p36.32* and *13q34* respectively), both crucial components of the p53 signaling pathway (Tallen *et al.,* 2014). The A53 subtype shows enrichment for homozygous deletions and truncating mutations affecting the MHC class I subunit, *B2M*, offering a mechanism for evading immune surveillance (Challa-Mallad *et al.,* 2011). ABC-DLBCL is the predominant type within the A53 genetic subtype. In terms of survival, A53 shows inferior outcomes compared to the BN2 group but does not significantly differ from MCD or N1.

N1 subtype

The N1 subtype, which does not have an equivalent in the molecular classification using consensus clustering, is defined by gain-of-function mutations in *NOTCH1* that eliminate all or part of the C-terminal PEST domain. These *NOTCH1* variants share functional similarities with those found in chronic lymphocytic leukemia (CLL) and mantle cell lymphoma (MCL) but differ from the membrane-proximal *NOTCH1* mutations observed in T-cell acute lymphoblastic leukemia (Rosati *et al.,* 2018). Almost exclusively composed of ABC-DLBCLs, the N1 subtype shows significantly poorer survival outcomes compared to patients with BN2 or ABC tumors not genetically classified.

BN2 (Cluster C1)

The BN2 subtype (cluster C1) is identified by specific genetic features, including B cell lymphoma 6 (*BCL6*) translocations, activating mutations of *NOTCH2*, and the inactivation of *SPEN*, a nuclear repressor that counteracts NOTCH-dependent gene activation. This subgroup also exhibits mutations in genes regulating B cell receptor (BCR) signaling, such as protein kinase C beta (*PRKCB*) and *BCL10*, along with inactivating lesions in TNF alpha induced protein 3 (*TNFAIP3*) and TNFAIP3-interacting protein 1 (*TNIP1*), encoding inhibitors of the NF-κB pathway (Young *et al.*, 2015). BN2 comprises a combination of unclassified and ABC-DLBCL cases, with a smaller proportion of GCB-DLBCLs. Interestingly, BN2 predominantly represents DLBCL cases that were unclassified by gene expression profiling. In terms of survival, BN2 demonstrates favorable outcomes compared to MCD and N1 when considering all cases.

ST2 (Cluster C4)

The ST2 molecular subtype (cluster C4) is characterized by mutations in *SGK1* and *TET2*. TET2, an epigenetic regulatory enzyme, plays a role in catalyzing the hydroxylation of methylated cytosine residues in DNA. ST2 tumors exhibit inactivating mutations in *TET2*, aligning with its role in promoting germinal centers when disabled in mice (Dominguez *et al.*, 2018). Mutations in SGK1, a kinase in the AKT family activated by various cellular responses, likely modulate PI3K signaling in these tumors. While *SGK1* is a target for somatic hypermutation (SHM), the preferential acquisition of loss-of-function mutations suggests its potential function as a tumor suppressor in ST2. Lesions defining ST2 subtype, including inactivation of suppressor of cytokine signaling-1 (*SOCS1*), a negative regulator of JAK signaling, activating mutations in *STAT3*, and inactivation of dual specificity phosphatase 2 (*DUSP2*), a phosphatase controlling STAT3 activity, indicate JAK/STAT signaling as a likely target (Lu *et al.*, 2015). The ST2 genetic subtype is predominantly composed of GCB-DLBCLs and characterized by a favorable survival outcome.

EZB (Cluster C3)

The EZB genetic subgroup (cluster C3) is predominantly comprised of GCB-DLBCL cases. Consequently, the distinctive features characterizing the EZB subtype include *BCL2* translocations, *2p16.1*/*REL* locus amplification, phosphatase and tensin homolog (*PTEN*) inactivation, and activating mutations of the histone methyltransferase *EZH2*.

Notably, the genetic signature of this subtype also encompasses mutations in various epigenetic regulators such as *KMT2D*, *CREBBP*, *EP300*, and AT-rich interactive domain-containing protein 1A (*ARID1A*) (Chapuy *et al.,* 2018; Schmitz *et al.,* 2018, Wright *et al.,* 2020). The disruption of epigenetic homeostasis in EZB tumors combined to mutations in genes controlling germinal center B cell development, such as *IRF8, MEF2B*, *S1PR2*, and *GNA13,* facilitate the adoption of a GC B cell transcriptional profile (Brescia *et al.,* 2018; Muppidi *et al.,* 2014; Wang *et al.,* 2019). Among GCB cases, EZB demonstrates inferior survival compared to all non-EZB patients.

The biological and molecular features of DLBCL genetic subtypes are summarized in Table 1.

Genetic subtype (GenClass/LymphGen)	**Consesuns clustering (Chapuy *et al.*, 2018)**	**Prevalence (Schmitz *et al.*, 2018)**	**Cell-of-origin**	**Distinctive genetic lesions**	**Affected onco-genic pathways**
MCD	C5	14%	ABC 99% UC 1% GCB 0%	*MYD88*^{L265P}, *CD79B*, *CD79A*, *CDKN2A*, *CD274*, *PDCD1LG2*, *CD58*, *HLA-B*, *-A*, *-C* *BCL2*, *PRDM1*, *SPIB*, *IRF4*	BCR-dependent NF-κB signaling Cell cycle control Immune evasion Apoptosis B cell differentiation
N1	-	3%	ABC 94% UC 6% GCB 0%	*NOTCH1* *ID3*, *BCOR*	NOTCH1 signaling B cell differentiation
A53	C2	7%	ABC 76% UC 5% GCB 19%	*TP53*, *TP53BP*, *TP73*, *ING1* *B2M*	Cell cycle regulation DNA damage response Immune evasion
BN2	C1	16%	ABC 40% UC 42% GCB 18%	*NOTCH2*, *SPEN* *PRKCB*, *BCL10*, *TNFAIP3*, *TNIP1* *CD70* *CCND3*	NOTCH2 signaling BCR-dependent NF-κB pathway Immune evasion Cell cycle control
ST2	C4	5%	ABC 22% UC 22% GCB 56%	*SOCS1*, *DUSP2*, *STAT3* *SGK1*, *P2RY8*	JAK/STAT signaling PI3K/AKT signaling
EZB	C3	13%	ABC 5% UC 9% GCB 86%	*EZH2*, *KMT2D*, *CREBBP*, *EP300*, *ARID1A* *PTEN*, *S1PR2*, *GNA13* *BCL2*, *FAS* *IRF8*, *MEF2B*, *REL*	Epigenetic deregulating PI3K signaling Apoptosis GC B-cell differentiation

Table 1. Biological features of DLBCL molecular subtypes

1.6 Liquid biopsy and ctDNA analysis in DLBCL

Limitations in accessing fresh tumor material from DLBCL tissue biopsies has prevented the rapid translation of DLBCL gene mutations into prognostic or predictive tools for clinical practice (Rossi *et al.,* 2017). Additionally, serial sampling of tumors to monitor the development of drug-resistant mutations necessitates a rebiopsy of the DLBCL, a procedure that may not be routinely feasible in clinical practice (Rossi *et al.,* 2017). Therefore, exploring alternative accessible sources of tumor DNA may help to complement the molecular diagnostic analyses that are routinely conducted on formalin-fixed paraffine-embedded (FFPE) tissue biopsies (Rossi *et al.,* 2017). The availability of tumor-specific DNA simply in plasma, serum or other body fluids, without the necessity of an invasive tumor biopsy, led to the concept of "liquid biopsy" in lymphomas lacking leukaemic involvement (Spina *et al.,* 2019).

Accessing tumor-derived cell-free DNA (cfDNA), known as circulating tumor DNA (ctDNA), through the bloodstream facilitates real-time monitoring of disease genetics, presenting opportunities for diverse applications in hematological malignancies (Talotta *et al.,* 2023). ctDNA serves as a comprehensive representation of the entire tumor heterogeneity, allowing the circumvention of anatomical biases associated with tissue biopsies. This enables the reconstruction of the complete cancer clonal architecture and the identification of dormant resistant clones present in inaccessible tumor sites (Jahr *et al.*, 2001).

The levels of cfDNA vary greatly across different lymphoma subtypes and reflect tumor type, clinical stage, and tumor burden. Overall, the mean concentration of cfDNA in patients with untreated DLBCL is significantly higher than in healthy individuals (26.9 ng/ml and 12.1 ng/ml, respectively) (Hohaus *et al.,* 2009; Rossi *et al.,* 2019). The use of real-time PCR for quantifying ctDNA at the time of diagnosis can identity patients with elevated levels, that are linked to disease characteristics indicative of aggressive disease and poor prognosis (Hohaus *et al.,* 2009). Quantification of cfDNA has been extensively used to track tumor clonotypic immunoglobulin heavy chain (IGHV) rearrangement for minimal residual disease (MRD) monitoring in DLBCL patients (Hohaus *et al.,* 2009; Kurtz *et al.,* 2015; Roschewski *et al.,* 2016).

During the surveillance periods preceding relapse, high-throughput sequencing of immunoglobulin genes (Ig-HTS) from plasma ctDNA showed enhanced specificity and comparable sensitivity when compared to the combination of positron emission tomography and computed tomography (PET/CT) (Kurtz *et al.,* 2015).

Given its high specificity, Ig-HTS from plasma has potential clinical utility for surveillance after complete remission (Kurtz *et al.*, 2015). Genotyping on plasma cfDNA is shown to be as accurate as genotyping on tissue biopsy for detecting somatic mutations, with the added advantages of identifying mutations not otherwise detectable in the biopsy tissue due to anatomical constraints and offering real-time, non-invasive tracking of clonal evolution and treatment-resistant clones in DLBCL (Rossi *et al.*, 2017). While plasma cfDNA provides a complementary source of tumor DNA compared to tissue biopsy, it captures genetic information from all disease sites. However, it may miss a proportion of small subclonal mutations. Overall, liquid biopsy is not a substitute for tumor biopsy but rather offers complementary information in DLBCL (Rossi *et al.*, 2017).

1.7 Cell-free DNA fragmentation in DLBCL

Quantifying ctDNA levels, kinetics, and MRD depends on the detection of somatic mutations. Cutting-edges sequencing approaches can be employed to capture ctDNA fragmentation profiles, allowing for the investigation of mutation-independent translational applications of ctDNA. Tumor cells release cfDNA fragments into the bloodstream through various mechanisms, such as apoptosis, necrosis, and secretion, in the form of short double-stranded DNA molecules with sizes ranging between 40-200 base pairs (bp) (Stroun *et al.*, 2001; Fleischhacker *et al.*, 2007). The patterns of fragmentation closely resemble the occupancy maps found in the chromatin proteome and epigenetic signatures (Snyder *et al.*, 2016). Despite ongoing investigations, the precise physiological mechanisms controlling the fragmentation and molecular characteristics of cfDNA are not well-understood.

Recent research suggests that cfDNA fragmentation is a nonrandom process linked to nucleosome positioning. The generation of cfDNA fragments occurs in two primary processes: *(i)* intracellular DNA fragmentation, mediated by DNA fragmentation factor subunit beta (DFFB), a nuclease that induces blunt double-stranded breaks in open internucleosomal DNA regions, as well as intracellular deoxyribonuclease 1 like 3 (DNASE1L3) and other apoptosis-related nucleases, and *(ii)* extracellular DNA fragmentation facilitated by serum DNASE1L3 and deoxyribonuclease 1 (DNASE1). The latter contributes to the production of short T-end fragments measuring less than 150 bp (Serpas *et al.*, 2019; Han *et al.*, 2020).

The study of cfDNA fragmentation patterns in liquid biopsy, also known as "fragmentomics," and its correlation with clinical outcomes has emerged as an active research area in recent years (Ding *et al.,* 2022). The concept of cfDNA fragmentomics was initially introduced in 2015 (Ivanov *et al.,* 2015), followed by the development of several computational and experimental methods to assess cfDNA fragmentation patterns in plasma (Liu *et al.,* 2022). Typically, the fragment length of cfDNA exhibits peaks around 166 bp or multiples, supporting the indication that apoptosis may be the primary mechanism for cfDNA release. Cancer patients present ctDNA molecules with a shorter size distribution compared to the background cfDNA, predominantly originating from hematopoietic cells (Jiang *et al.,* 2015; Underhill *et al.,* 2016; Mouliere *et al.,* 2018; Cristiano *et al.,* 2019; Underhill *et al.,* 2021; Esfahani *et al.,* 2022).

Recent findings suggest that ctDNA fragment lengths in lymphoma patients can vary individually and correlate with the disease stage. Additionally, these fragmentation patterns have been shown to predict outcomes in DLBCL, featuring fragmentome disparities as a promising novel prognostic biomarker that may complement mutation-based MRD detection (Meriranta *et al.,* 2022).

1.8 DLBCL treatment and prognostic factors

The standard front-line therapy for patients with DLBCL is R-CHOP (Tilly *et al.,* 2015). Rituximab, an antibody targeting the CD20 protein predominantly found on B cells and numerous lymphoma cells, is combined with chemotherapy agents such as cyclophosphamide, doxorubicin, vincristine, and prednisone (Kwak *et al.,* 2012; Chiappella *et al.,* 2013). This therapeutic approach results in the cure of approximately 60-70% of DLBCL patients. Nevertheless, a subset of around 30-40% may experience relapse or, in some cases, show resistance to R-CHOP therapy (Li *et al.,* 2018). R-CHOP failures are primarily attributed to either primary refractoriness or relapse after reaching a complete response (CR) (Coiffier *et al.,* 2016). Relapsed DLBCL is distinguished by the emergence of new lesions following CR, whereas refractory DLBCL is defined as the failure of <50% of lesions to decrease in size after the initial treatment, according to the criteria established by Cheson *et al.,* in 2007.

MYD88^{L265P} mutation constitutively activates MYD88 signalling, which in turn triggers the alternative NF-κB pathway, bypassing the BCR-dependent NF-κB activation, and imparting resistance to ibrutinib (Wilson *et al.,* 2015). However, when *MYD88*^{L265P} mutation occurs concurrently with mutations in *CD79A* or *CD79B*, it renders tumors sensitive to ibrutinib (Wilson *et al.,* 2015). Conversely, mutations in *CARD11* or components of the NF-κB signalling lead to the constitutive activation of the NF-κB signalling pathway independently of the upstream BCR signalling pathway, potentially conferring resistance to BCR inhibitors (Wilson *et al.,* 2012).

MYC is rearranged in 5-15% of DLBCL patients and is frequently associated with *BCL-2* translocation, or less commonly, *BCL-6* translocation, leading to the formation of "double hit" (*i.e.* rearrangements of *MYC* with either *BCL-2* or *BCL-6*) or "triple hit" (*i.e.* rearrangements of *MYC* with *BCL-2* and *BCL-6)* lymphomas, respectively (Green *et al.*, 2012). Patients with "double-hit" or "triple-hit" B cell lymphomas exhibit an aggressive clinical course, limited response to R-CHOP standard chemotherapy, and elevated rates of relapse (Rosenthal *et al.,* 2017).

In 2019, Rushton and colleagues gathered samples from 134 patients experiencing relapse or refractory (R/R) conditions, participants in three clinical trials. They conducted a combination of exome sequencing and target panel sequencing on ctDNA extracted from both plasma samples and tissue biopsies, focusing on genes associated with lymphoma. The analysis revealed an enrichment of mutations in six genes - *TP53*, *KMT2D*, interleukin 4 receptor (*IL4R*), *HVCN1*, *RB1*, and membrane spanning 4-domains A1 (*MS4A1*) - among R/R patients. *TP53* and *KMT2D* mutations were initially present at the time of diagnosis, but their frequency increased during relapse.

Notably, the presence of *TP53* and *KMT2D* at the time of diagnosis correlated with shorter survival (Rushton *et al.,* 2019). *MS4A1* encodes CD20, the target of rituximab, and mutations in this gene destabilize a common transmembrane helix, leading to reduced membrane expression. Strikingly, *MS4A1* mutations were absent at the time of diagnosis but increased during relapse, predisposing individuals to a diminished efficacy of anti-CD20 salvage chemoimmunotherapy regimens (Rushton *et al.,* 2019).

2. AIMS OF THE STUDY

DLBCL displays significant heterogeneity in its genetic profile, and recent genomic studies have unveiled distinct molecular subtypes associated with different outcomes. Liquid biopsy, here intended as the analysis of circulating tumor DNA (ctDNA), is claimed as a minimally invasive approach to genotype DLBCL in a biopsy-free manner. The evaluation of ctDNA levels proves to be a powerful tool for predicting outcomes at both the initial diagnosis and throughout the R-CHOP therapy course. Additionally, the emerging field of fragmentomics, focusing on ctDNA fragment size analysis, has been proposed as a novel prognostic biomarker for DLBCL. Patients exhibiting shorter cfDNA fragment sizes, likely reflecting the elevated apoptotic rates of DLBCL cells, are associated with poorer survival compared to those with longer fragment sizes. Building upon these findings, the objectives of this study include:

i) Assessing the role of plasma ctDNA as a non-invasive tool to molecularly classify DLBCL patients;

ii) Investigating whether the integration of ctDNA levels with molecular clustering enhances the accuracy of predicting outcomes for DLBCL patients;

iii) Examining whether the cfDNA fragmentation profile can provide additional insights to enhance outcome prediction for patients not categorized within a specific molecular cluster.

3. MATERIALS AND METHODS

3.1 Patients and biological samples

A real-life multicenter cohort of 166 DLBCL patients was included in the study. Patients had a diagnosis of DLBCL not otherwise specified (NOS) and were treated with R-CHOP based therapy. All cases were provided with: *i)* ctDNA from plasma and *ii)* germline gDNA extracted from granulocytes for comparative purposes. In 77 DLBCL cases, tumor gDNA extracted from fresh frozen lymph node cells or formalin-fixed paraffin-embedded (FFPE) lymph node biopsies were also analyzed. The extraction of cfDNA from plasma aimed at obtaining at least 30 nanograms (ng), corresponding to ~5000 genome equivalents (GE), to allow a stoichiometric sensitivity for the detection of 0.1% mutated alleles. Patients provided informed consent in accordance with 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 Università del Piemonte Orientale (study number CE 120/19).

3.2 Separation of granulocytes from peripheral blood (PB)

Granulocytes were isolated from peripheral blood (PB) through Ficoll gradient density centrifugation to obtain normal germline genomic DNA (gDNA). Peripheral blood (PB) was mixed in a 1:2 ratio with a physiological solution (NaCl 0.9%) and then subjected to centrifugation at 1800 revolutions per minute (rpm) for 25 minutes using Sigma Diagnostic™ Histopaque®-1077 Cell Separation Medium (Sigma-Aldrich, St. Louis, MO, USA). This process resulted in the separation of granulocytes and mononuclear cells (including monocytes and lymphocytes).

3.3 Lymph node tissue processing for gDNA extraction

The separation of tumor cells from fresh lymph nodes was performed by processing the material into a Petri dish with Roswell Park Memorial Institute (RPMI) medium (Sigma-Aldrich, St. Louis, MO, USA) and then gently scraping the tumor tissue surface to detach the cells. Subsequently, the sample underwent centrifugation at 1500 rpm for 10 minutes to facilitate the separation of tumor cells for subsequent DNA extraction.

3.4 Extraction of tumor and normal gDNA

Tumor from fresh LNF cells and normal gDNA from PB granulocytes were both extracted by using the "salting out" protocol (Miller *et al.*, 1988). Cells were lysed using a lysis buffer (Tris-HCl 1M, pH 8.2, NaCl 5M, EDTA 0.5M), sodium dodecyl sulphate (SDS) 20% and digested with 20 mg/ml of proteinase enzyme (pronase E). Samples were incubated at 37°C overnight in a shaking incubator. Proteins were precipitated with 6M NaCl, and subsequently discarded after centrifugation at 3200 rpm for 20 minutes. DNA was isolated by precipitation with pure ethanol and the lactescent "jellyfish" of DNA, formed as a result of the addition of ethanol, was recovered with glass loops and washed three times in 75% ethanol. The excess of ethanol was evaporated, and the DNA was dissolved with TE Buffer (Tris-HCl 1M, pH 8.2, EDTA 0.5M).

3.5 Extraction of tumor gDNA from formalin-fixed paraffin-embedded (FFPE)

Tumor gDNA was isolated from formalin-fixed paraffin-embedded (FFPE) diagnostic tissue biopsies containing >70% of tumor cells, as estimated by morphology and immunohistochemistry, by using the truXTRAC FFPE DNA – column purification kit (Covaris®,Woburn, MA, USA). FFPE tissues biopsies were rehydrated with tissue SDS Buffer and transferred into microTUBE Screw-Cap. The samples were processed by using M220 focused-ultrasonicator (Covaris®,Woburn, MA, USA) to dissociate the paraffin and rehydrate the tissues for 5 minutes at 20°C (20% duty factor, 75% peak incident power, 200 cycles per burst). Proteins were digested by using proteinase K for 10 seconds (20% duty factor, 75% peak incident power, 200 cycles per burst) and with the incubation at 56°C overnight. Afterwards, samples were incubated for 1 hour at 80°C for crosslink reversing and then DNA was extracted and purified with columns.

3.6 Plasma cell-free DNA (cfDNA) extraction

PB samples were collected in Streck Cell-Free DNA BCT® tubes (Streck, Nebraska, USA) and centrifuged at 800 rcf for 10 minutes at 4°C to separate plasma from cells. Plasma was then further centrifuged at 13.000 rpm for 10 minutes at 4°C to pellet and remove any remaining cells and stored at -80°C until DNA extraction. cfDNA was extracted from 2-3 ml aliquots of plasma immediately after thawing by using Maxwell® RSC LV ccfDNA Kit Custom (Promega Corporation, Madison, WI, USA) and quantified by Quantus Fluorometer using

QuantiFluor dsDNA System (Promega Corporation, Madison, WI, USA). The quality of the extracted cfDNA was assessed by 2100 Bioanalyzer Instrument (Agilent Technologies, St. Clara, CA, USA).

3.7 DNA quantification and fragmentation

Tumor and germline gDNA were quantified using the Quant-iT^{TM} PicoGreen dsDNA Assay kit (ThermoFisher Scientific, Eugene, OR, USA). PicoGreen selectively binds the double helix of DNA thus allowing a precise estimation of double-stranded DNA (dsDNA) amount. The fluorimetric reading was performed using the Infinite F200 fluorometer (TECAN, Männedorf, Switzerland) using the Magellan software. The fluorimetric readings were obtained at a wavelength of 485 nm in absorption and 530 nm in emission. For quantification a standard curve was prepared using a DNA of known concentration and performing serial 1:2 scalar dilutions. Quant-iT^{TM} PicoGreen dsDNA Assay kit was used at the 1:200 dilutions. Tumor and germline gDNA from tissues were sheared through sonication with M220 focused-ultrasonicator (Covaris® Woburn, MA, USA) before library construction to obtain 250/300 base pairs (bp) fragments. The size of the fragments was checked by using the 2100 Bioanalyzer Instrument. The library preparation of physiologically fragmented cfDNA samples did not include any additional shearing of DNA.

3.8 Library design for hybrid selection

A targeted resequencing gene panel, including coding exons and splice sites of 59 genes that are recurrently mutated in DLBCL and in other B cell malignancies, has been specifically designed for this project. In addition, the gene panel included genes mapping in genomic regions affected by copy-number aberrations previously reported in DLBCL (Chapuy *et al.*,2018; Schmitz *et al.,* 2018; Lacy *et al.,* 2020), for a total of 109 genes (target region size: 199,430 bp).

3.9 CAPP-seq library preparation and sequencing

The LyV4.0 CAncer Personalized Profiling by deep Sequencing (CAPP-Seq) assay was used for the study (Newman *et al.,* 2014; Newman *et al.,* 2016). The next generation sequencing (NGS) libraries were constructed using the KAPA HyperPrep Library Preparation Kit (KAPA Biosystems, Wilmington, MA, USA), and subjected to hybrid capture using the custom KAPA HyperChoice probes (Roche, Basilea, Switzerland). Multiplexed libraries were

sequenced using 300-bp paired end runs on MiSeq and 150-bp paired end runs on NextSeq550 platforms (Illumina, San Diego, CA, USA). The number of the libraries loaded in the MiSeq/NextSeq550 instruments was tailored to obtain at least a coverage > 2000× in > 80% of the region of interest. A background error-suppressed approach was used for variant calling. The limit of quantification of the LyV4.0 CAPP-seq assay was 0.09%, which represented the analytical background noise threshold over which the assay produced a signal distinguishable from "blank". The analytical sensitivity of the LyV4.0 CAPP-seq was 0.1%, representing the smallest detectable allele frequency (Rossi *et al.*, 2017; Diop *et al.*, 2020; Spina *et al.*, 2018).

3.10 Bionformatic Pipeline and Analysis

FASTQ sequencing reads were deduped by using the FastUniq v1.1. Then, the deduped FASTQ sequencing reads were locally aligned to the hg19 version of the human genome assembly using the BWA v.0.6.1 software with the default setting, and sorted, indexed, and assembled into a mpileup file, using SAMtools v.1. Somatic SNVs and small insertions and deletions (indels) were identified in tumor/normal paired samples using the somatic function of VarScan2. Multiple comparisons test was performed to compare the variant allele frequency (VAF) *versus* the mean allele frequency in unpaired normal gDNA samples to filter out variants below the base-pair resolution background frequencies across the target region. Only variants that had a significant call in Z-test were retained (Bonferroni adjusted $p \leq 6.11 \times 10^{-8}$). Mutations were annotated by using wANNOVAR (https://wannovar.wglab.org/). *TP53* variants that were scored as SNPs according to the International Agency for Research on Cancer TP53 database; https://tp53.isb-cgc.org/), intronic variants, mapping >2 bp before the start or after the end of coding exons, and synonymous variants were then filtered out. Splice-acceptor and splice-donor variants were annotated by using the SeattleSeq Annotation 138 tool (http://snp.gs.washington.edu/SeattleSeqAnnotation138). Among the remaining variants, only protein truncating variants (i.e., indels, stop codons and splice site mutations), as well as nonsynonymous variants, not included in the dbSNP 138 and reported as somatic in the COSMIC v96 database (https://cancer.sanger.ac.uk/cosmic), were retained. All the variants were visualized using IGV (Integrative Genomics Viewer) software (v2.8.0).

3.11 Copy-number analysis on ctDNA

Copy-number analysis was performed using CNVkit software toolkit (version 0.9.10) in Python 3.11.4 (Talevich *et al.*, 2016). CNVkit infers CNAs from targeted capture sequencing data. Sequencing data (cfDNA BAM-files) from CAPP-seq pipeline, including both on-target and off-target reads, were used as input. The target bin size was set to default. A reference file was built from 31 MRD-negative samples; read depths were median-centered and bias-corrected (GC-content, sequence repeats, target density) to produce normalized $\log_2$ read-depth values for each bin. To infer copy-number changes, "observed" normalized $\log_2$ read-depth values from patient samples were subtracted from the "expected" values in the reference file and bins segmented using circular binary segmentation with the "drop-low-coverage" option (Venkatraman *et al.,* 2007; Olshen *et al.,* 2011). GISTIC algorithm was applied on the segmented data using the GISTIC2.0 GenePattern module (version 6.15.30) to identify statistically significant CNAs (Beroukhim *et al.,* 2007). The thresholds used in our analysis were 0.1 for amplifications and 0.05 for deletions. Regions with q-values below 0.25 are considered significant and are reported. GISTIC output files were processed and summarized using maftools R package (version 2.10.05) (Mayakonda *et al.*, 2018)

3.12 ctDNA quantitation and molecular groups classification

ctDNA levels were reported as haploid genome equivalents per ml of plasma (hGE/ml), determined as the product of total cfDNA concentration and the mean allele fraction of somatic mutations. This value was expressed as a base-10 logarithm ($\log_{10}$ hGE/ml). The validated 2.5 hGE/ml threshold was used to classify patients into low or high ctDNA levels (Kurtz *et al.*, 2018). The LymphGen probabilistic classification tool was used to classify our DLBCL cases into the recently described genetics subtypes. We consider a patient assigned to a specific cluster when the tool considered the case as ‘‘core’’ or ‘‘extended’’ subtype (Wright *et al.*, 2020).

3.13 Analysis of cfDNA fragments

Fragment length data were extracted from targeted cfDNA BAM-files using the *GenomicAlignments* (version 1.30.0) and ‘GenomicRanges’ (version 1.46.1) R packages. As the percentage of fragments with a length exceeding 400 is nearly negligible, fragments between 1 and 400 in length were considered as the first pre-processing step. To characterize the fragment-size distribution, initially, a matrix representing the fragments length across all

cfDNA samples was generated. The resulting matrix was used to calculate the proportion of each fragment length to the total number of fragments for each sample. Short fragments were designated as those with a length of ≤ 150 bp, while long fragments were characterized by lengths exceeding 150 bp. The K-means algorithm was applied to cluster patients while considering just short fragments as features. Elbow method was used to determine the optimal number of clusters (in this case two). A Wilcoxon test was conducted to visualize the mean distribution of fragment sizes within the two identified clusters.

3.14 Statistical analysis

Medical statistical analysis was performed using R packages *survival* (version 3.5-5) and *survminer* (version 0.4.9). Primary endpoints of survival analysis were progression free survival (PFS) and overall survival (OS). PFS was evaluated from the time of treatment start to the date of progression or death, and OS was evaluated from the time of treatment start to the date of death. Time to event outcomes (PFS and OS) were estimated using the Kaplan-Meier method and compared between groups using the Log-Rank test. To evaluate the associations between molecular features, we used Chi-square test and Fisher exact test. Statistical significance was defined as p value < 0.05.

4. RESULTS

4.1 Patient characteristics

A total of 166 newly diagnosed DLBCL patients treated with curative intent R-CHOP-based therapy were included in the study. The median age of the study cohort was 67 years, 118 (71.1%) presented a stage of III/IV, 89 (53.6%) presented a IPI ≥3. The clinical characteristics of the studied cohort are reported in Table 2.

Characteristics	Values
Median Age	66.9 (IQR 52.6-75.9)
IPI	
0-1	44 (26.5%)
2	33 (19.9%)
3	44 (26.5%)
4-5	45 (27.1%)
COO	
GC	56 (33.7%)
Non-GC	80 (48.2%)
Not available	30 (18.1%)
LDH above ULN	
Yes	61 (42.1%)
No	84 (57.9%)
Extranodal sites	
Yes	88 (65.7%)
No	46 (34.3%)

Table 2. Patients baseline characteristics. IPI: International Prognostic Index; COO: cell-of-origin; GCB: Germinal Center B- cell like; non-GCB: non-Germinal Center B-cell like; LDH: lactate dehydrogenase; ULN: Upper Limit Normal

After a median follow-up of 49.1 months, the median PFS and OS were not reached. PFS and OS at 4-years were 64.3% and 76.1%, respectively (Figure 3).

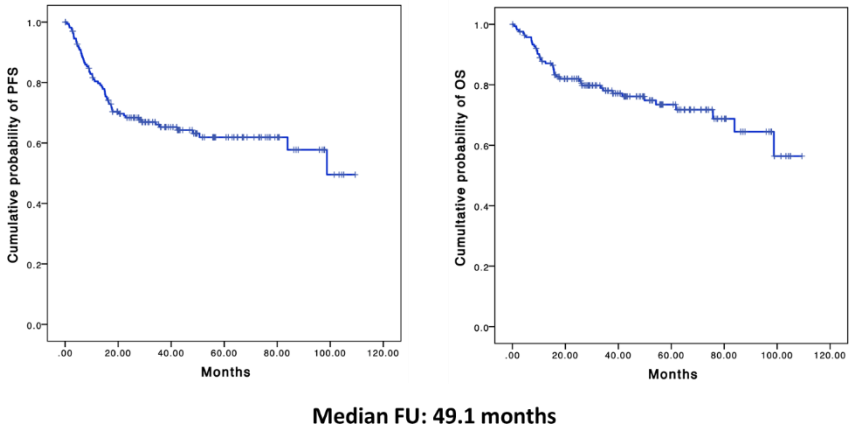

Figure 3. Kaplan-Meier estimates of PFS and OS of the studied DLBCL cohort (N=166)

4.2 Molecular profile of the DLBCL cohort

Overall, on ctDNA the most frequently mutated genes were *KMT2D* in 32 (19.3%) patients, *PIM1* in 29 (17.5%), *MEF2B* in 24 (14.5%) followed by *TP53* in 23 (13.9%) and *CREBBP* in 22 (13.3%) patients. Using the LymphGen tool on ctDNA, 71 (42.8%) patients have been assigned to a specific molecular cluster. More precisely, 22 (13.3%) patients were classified as EZB, 19 (11.4%) as BN2, 14 (8.4%) as MCD, 10 (6.0%) as ST2 , and 6 (3.6%) as A53 subtype, mainly defined by *TP53* mutations (4/6 | 66.7%) and SCNAs.

The 22 DLBCL cases assigned to cluster EZB harbored frequent mutations in *BCL2* (9/22 | 40.9%) gene, chromatin modifers, *KMT2D* (6/22 | 27.3%), *CREBBP* (10/22 | 45.5%), and *EZH2* (4/18 | 18.2%). These tumors also had alterations of the B cell transcription factors *MEF2B* (11/22 | 50%) and *IRF8* (5/22 | 22.7%), and indirect modifiers of BCR and PI3K signaling, *TNFSF14* (4/18 | 18.2%) and *GNA13* (4/18 | 18.2%). In addition, these tumors exhibit a mechanism of inactivating *PTEN* through focal *10q23.31/PTEN* loss (11/22 | 50%). Consistent with this finding, EZB cluster was enriched of GCB type.

DLBCL cases classified as BN2 harbored mutations of the NF-kB pathway members *BCL10* (8/19 | 42.1%) and *TNFAIP3* (3/19 | 15.8%). Notably, BN2 cases exhibited genetic features of immune evasion, such as inactivating mutations in *B2M* (6/19 | 31.6%) and *CD70* (8/19 | 42.1%).

As expected, the majority of MCD cases with available COO designations (12 of 14) were ABC-DLBCL type (11/12 | 91.6%). These tumors exhibited frequent mutations in *CD79B* (8/14 | 57.1%) and *MYD88* (4/14 | 28.6%). In the majority of cases, *MYD88* mutations selectively involved L265P. Additional alterations linked to ABC-DLBCLs, including gain of *19q13.41* (8/14 | 57.1%) and inactivation of *PRMD1* (11/14 | 78.6%), were observed in this cluster. Additional mutations characterizing MCD cluster were identified on *BTG1* (6/14 | 42.9%), *PIM1* (5/14 | 35.7%) and *ETV6* (3/14 | 21.4%) genes.

The 10 DLBCL cases assigned to ST2 subtype were characterized by mutations in H1 linker and core histone genes (*HIST1H1E*, *HIST1H1C*, *HIST1H1D* and *HIST1H2AM/AC/2BK/2BC*), immune evasion molecules (*CD58* and *CD83*), BCR/PI3K signaling intermediates (*SGK1* and *RHOA*), NF-kB modifiers (*NFKBIA* and *CARD11*), and RAS/JAK/STAT pathway members (*STAT3* and *BRAF*).

A total of 3 focal regions of copy gain and 17 focal regions of copy loss were detected. The number of genes in focal peaks varied from 7 (*2q22.3 loss*) to 595 (*1p36.32 loss*) genes, with a median of 35 and 157 genes for amplifications and deletions, respectively.

Significant genes perturbated by mutations and SCNAs included *TNFRSF14* (*1p36.32*), *TNFAIP3* and SGK1(*6q23.3*), *PTEN* and *FAS* (*10q23.31*), *CREBBP*(*16p13.3*), *EP300*(*22q13.31*) and *TP63*(*3q27.3*). The complete molecular characterization of the analyzed cohort is reported in Figure 4.

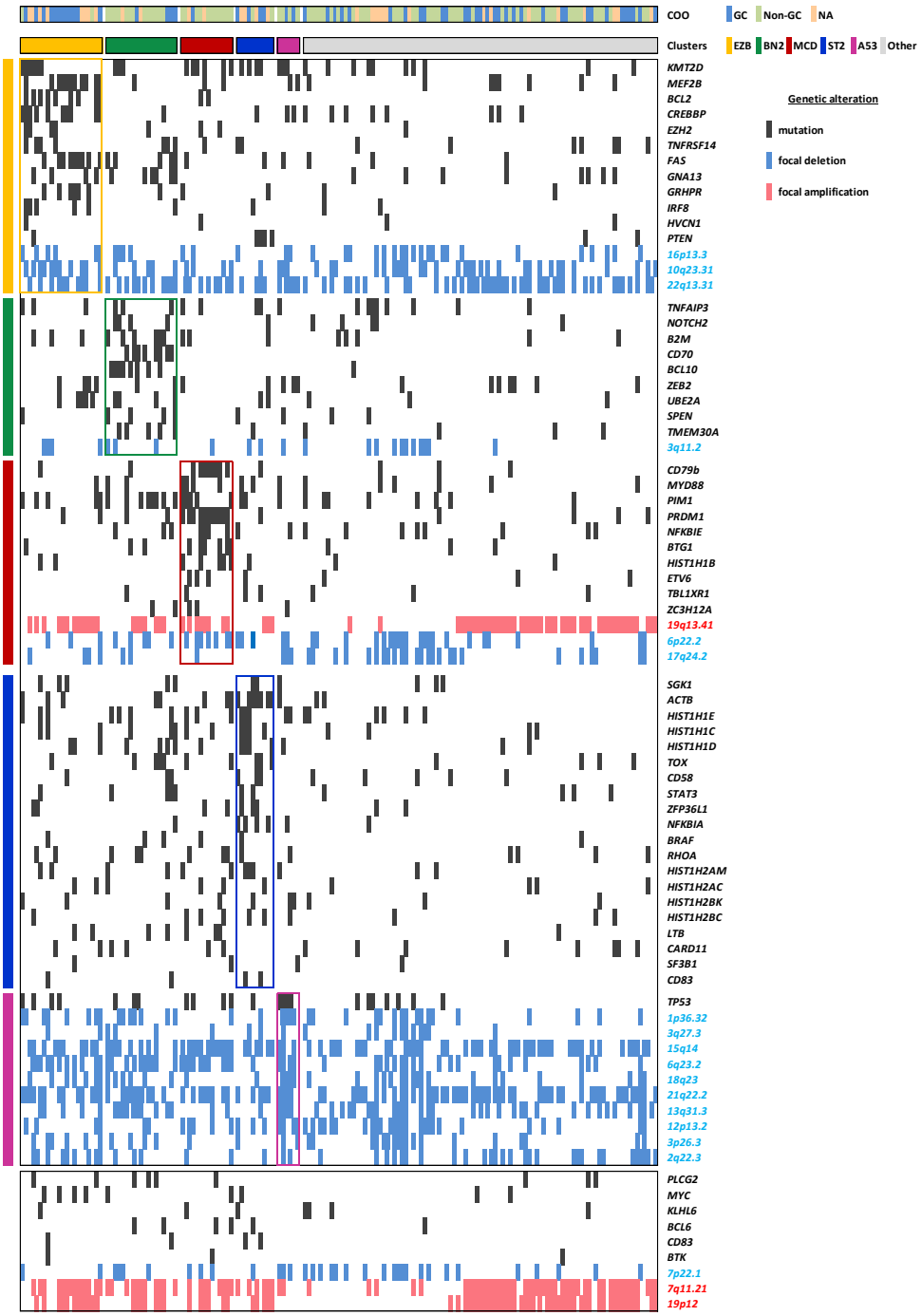

Figure 4. Molecular characterization of DLBCL patients. Overall landscape of genetic subtypes based on the LymphGen tool derived from 166 DLBCL analyzed on ctDNA. Each column represents an individual patient, each row represents one genomic abnormality. Cell-of-origin and molecular clusters are plotted above the heatmap.

From the clinical perspective, MCD patients showed a shorter PFS and OS compared to other patients. The 4-years PFS of MCD patients was 40% compared to 66.6% for other patients (p=0.032) (Figure 5). No other clusters predicted PFS and OS in the overall population.

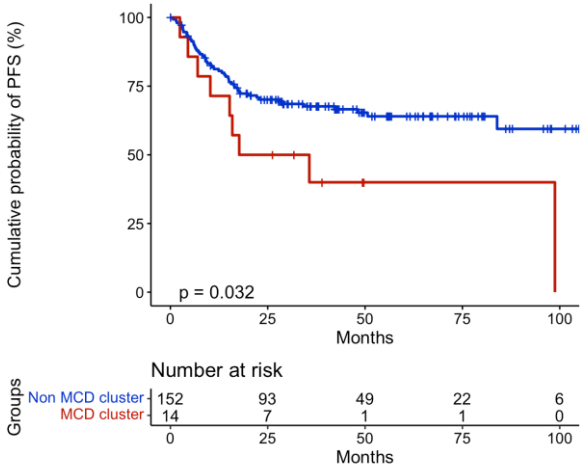

Figure 5. Kaplan-Meier estimates of PFS for MCD cluster

4.3 Comparison of molecular clustering between ctDNA and tissue biopsy

Analyzing 77 DLBCL cases provideds with lymph node biopsies, we aimed to compare the molecular cluster assignments obtained from both ctDNA and tissue biopsy. Due to the unavailability of SCNAs data on FFPE, clusters were defined by using gene mutations only. Consequently, the A53 cluster could not be evaluated in this comparative analysis. The mutational analysis of ctDNA and LNF compartments, resulted in 40.3% (27/67) of cases assigned to a specific molecular cluster on liquid biopsy. Among the classified patients, 25.9% (7/27) were assigned to the EZB subtype, 33.3% (9/27) to MCD, 18.5% (5/27) were identified as BN2, and 18.5% (5/27) as ST2. One patient (3.7%) was defined as genetic composite BN2/ST2 because harboured features of both subtypes. Overall, 12.9% patients (10/77) have not be assigned to a specific molecular cluster because they did not harbor on ctDNA mutations affecting genes included in our target panel.

The LymphGen algorithm allowed to assign 46.5% (33/71) of patients on the tissue biopsy. Precisely, 46.5% (33/71) cases were assigned to a specific molecular cluster on the LNF, among which 24.2% (8/33) as EZB subtype, 27.3% (9/33) as MCD subtype, 18.2% (6/33) of patients were classified as BN2 subtype, 8/33 (24.2%) as ST2 subtype, 1 patient (3.03%) was classified as BN2/ST2 subtype, and another one (3.03%) as MCD/ST2 subtype.

In the group of 24 patients assigned to a specific molecular cluster, there was a 95.8% concordance rate between ctDNA and lymph node biopsy results (Figure 6).

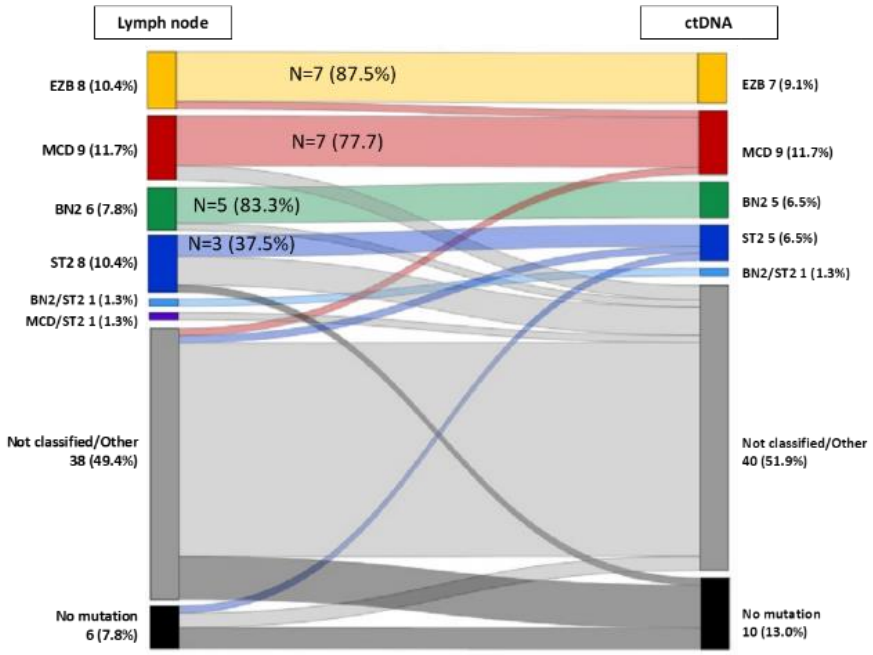

Figure 6. Sankey plot showing the comparison of the LymphGen algorithm applied to 77 DLBCL cases by using mutational data obtained from LNF and ctDNA analyses.

From a clinical perspective, patients assigned to ST2 and BN2 clusters on both ctDNA and LNF compartment were associated with an excellent outcome after R-CHOP compared to other (Figure 7).

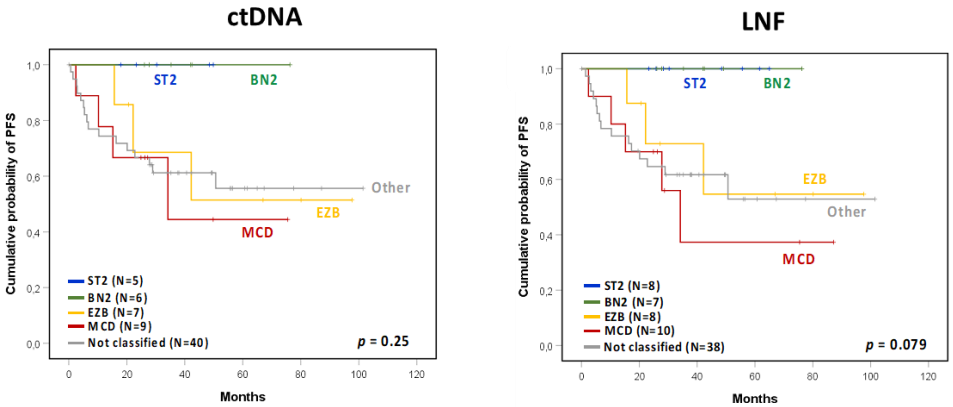

Figure 7. Kaplan-Meier estimates of PFS for the different molecular cluster assigned on ctDNA and LNF

4.4 Integration of ctDNA levels and molecular clusters identified on ctDNA

Subsequently, our aim was to integrate ctDNA levels with the molecular clusters identified on ctDNA to leverage two distinct biomarkers from liquid biopsy. Using the well-validated cutoff of 2.5log$_{10}$hGE/ml, patients with ctDNA levels < 2.5log$_{10}$hGE/ml demonstrated a 4-years PFS and OS of 71.7% and 85.7%, respectively, compared to 50.3% and 61.0% for patients with ctDNA levels > 2.5log$_{10}$hGE/ml (p=0.0018 and p=0.0017, respectively) (Figure 8).

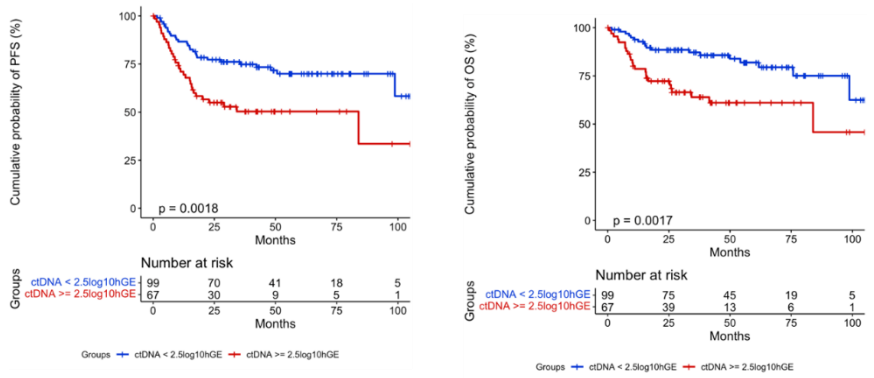

Figure 8. Kaplan-Meier estimates of PFS and OS based on ctDNA load cutoff

Through recursive partitioning, patients with ctDNA levels lower than 2.5log$_{10}$hGE/ml were not further stratified by a specific molecular cluster. In contrast, patients with ctDNA load higher than 2.5log$_{10}$hGE/ml were significantly further stratified by assignment to cluster ST2/BN2 (p=0.038). Among patients with ctDNA levels > 2.5log$_{10}$hGE/ml, those belonging to the ST2/BN2 cluster exhibited a 4-years PFS and OS of 77.4% and 94.4%, respectively, compared to 39.9% and 46.5% for other molecular subgroups (p=0.017 and p=0.002) (Figure 9).

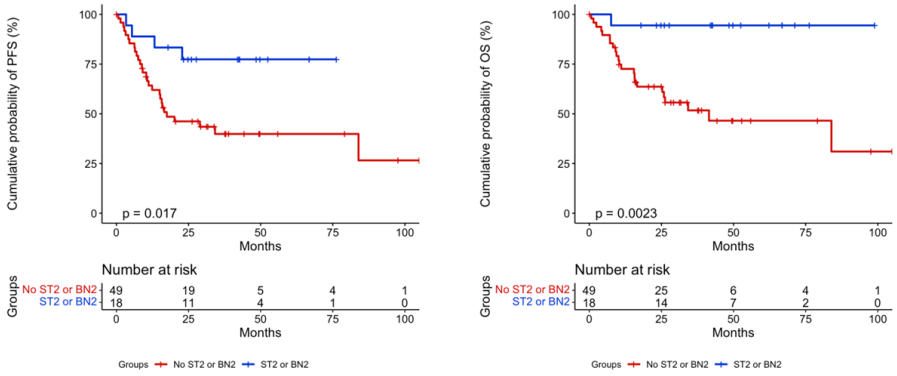

Figure 9. Kaplan-Meier estimates of PFS and OS in patients with ctDNA levels ≥ 2.5 log hGE/mL according to ST2/BN2 molecular cluster. Patients assigned to ST2 or BN2 in blue. Patients not assigned to ST2 or BN2 in red.

Consequently, ctDNA levels and ST2/BN2 clusters were combined to enhance the outcome prediction for DLBCL patients. Low-risk patients (N=49), characterized by ctDNA levels < 2.5log$_{10}$hGE/ml and/or BN2/ST2 cluster, demonstrated a 4-years PFS and OS of 73.9% and 87.1%, respectively. High-risk patients (N=117), characterized by ctDNA levels > 2.5log$_{10}$hGE/ml and no BN2/ST2 cluster, exhibited a 4-years PFS and OS of 39.9% and 46.5%, respectively (Figure 10).

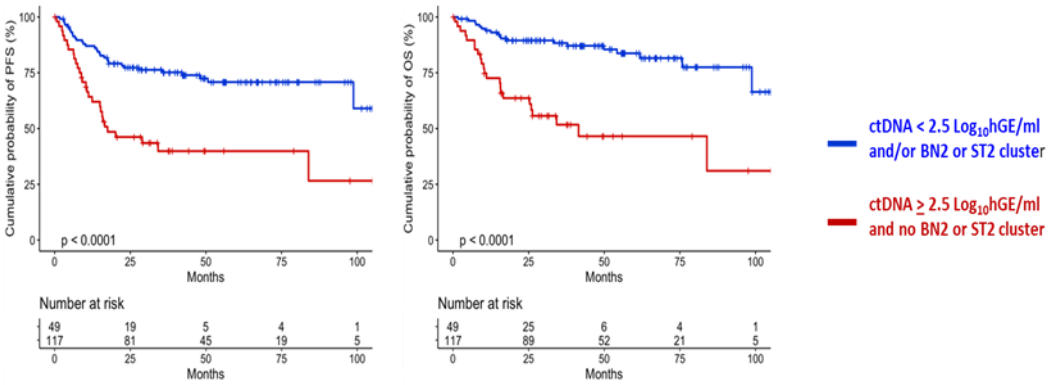

Figure 10. Kaplan-Meier estimates of PFS and OS according to ctDNA levels and ST2/BN2 cluster. Low-risk patients are represented by the blue line. High-risk patients are represented by the red line.

In comparison to ctDNA levels only, the addition of BN2/ST2 cluster improved the C statistics of the model (0.62 *vs* 0.59 for PFS and 0.68 *vs* 0.63 for OS). In multivariate analysis, the assignment to cluster BN2 or ST2 maintained an independent association with improved PFS (HR 0.37, 95% CI 0.14-0.95, p=0.04) and OS (HR 0.07, 95% CI 0.01-0.55, p=0.011) when adjusted for ctDNA levels, IPI, and COO (Figure 11).

Multivariate analysis for PFS

Variable	Hazard ratio	95% CI	p
Cluster BN2 and ST2		0.37 (0.14, 0.95)	0.04
ctDNA>=Log10hGE		3.52 (1.80, 6.88)	<0.001
NonGC		1.64 (0.86, 3.12)	0.14
IPI>2		1.65 (0.82, 3.32)	0.16

Multivariate analysis for OS

Variable	Hazard ratio	95% CI	p
Cluster BN2 and ST2		0.07 (0.01, 0.55)	0.01
ctDNA>=Log10hGE		4.47 (2.03, 9.84)	<0.001
NonGC		1.65 (0.76, 3.59)	0.21
IPI>2		2.15 (0.90, 5.10)	0.08

Figure 11. Multivariate analysis for PFS and OS adjusted for ctDNA levels, IPI and COO. IPI: International Prognostic Index; COO: cell-of-origin

4.5 Fragmentome analysis

Unsupervised machine learning was employed to analyze a total of 166 cases of DLBCL, resulting in the identification of two distinct patient subgroups characterized by unique fragmentation patterns. Among DLBCL patients, 73.5% (N=122) exhibited a median fragment length of 197bp, comparable to that of healthy donors, and were classified as nucleosomal cfDNA. The remaining 26.5% of DLBCL cases (N=44) exhibited shorter fragment lengths with a median of 175bp, referred to as sub-nucleosomal cfDNA (Figure 12).

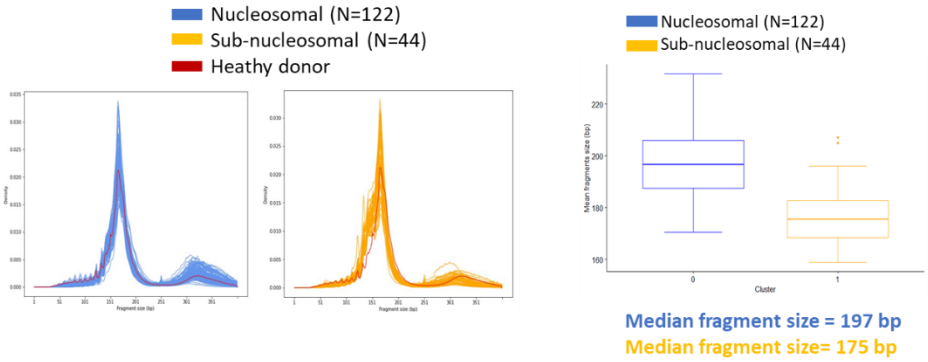

Figure 12. Density plot (left) of the mean cfDNA fragment-size distribution of nucleosomal group (N=122, in blue), of sub-nucleosomal group (N=44, in orange) and of healthy donor (in red) overlaid. Box plot (right) showing the different fragment size in nucleosomal (in blue) and sub-nucleosomal (in orange) cfDNA patients.

Sub-nucleosomal cfDNA was found to be significantly associated with an IPI ≥ 2 (p=0.013), elevated LDH levels (p=0.001), and ctDNA levels exceeding 2.5 Log_{10}hGE/ml (p=0.001), while no association was observed with the COO (p=0.698) (Figure 13).

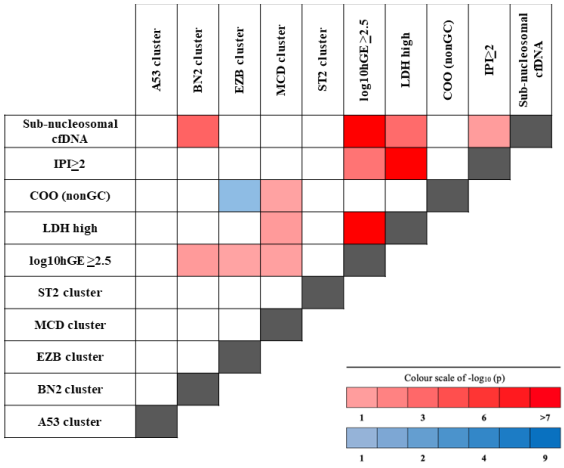

Figure 13. Correlation map comparing the cfDNA fragmentation profile with DLBCL clinical and molecular features. The red color scale of -log_{10} (p) points to a co-occurrence; the blue color scale of -log_{10} (p) points to exclusivity between two variables. The intensity of the color corresponds to the strength of the correlation.

Focusing on 95 DLBCL patients not assigned to a specific cluster or lacking detectable mutations in ctDNA, we sought to assess whether the fragmentomic profile could refine prognosis. Notably, patients with sub-nucleosomal cfDNA exhibited a 4-years PFS and OS of 25.9% and 44.4%, respectively, in contrast to 76.9% and 87.3% for nucleosomal cfDNA patients (both p<0.001) (Figure 14).

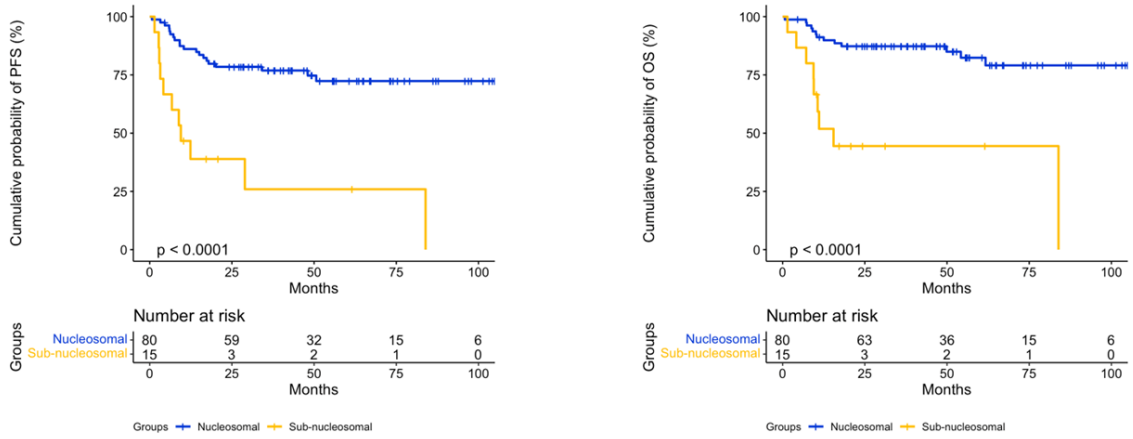

Figure 14. Kaplan-Meier estimates of PFS and OS according cfDNA fragmentation patterns in patients not classified in any molecular cluster (N=95). Nucleosomal cfDNA patients (N=80) are represented by the blue line and sub-nucleosomal patients (N=15) are represented by the orange line.

Notably, the prognostic impact of sub-nucleosomal cfDNA persisted in multivariate analysis for both PFS (HR 3.12, 95% CI 1.36-7.15, p=0.007) and OS (HR 3.06, 95% CI 1.17-8.02, p=0.023), even after adjusting for ctDNA levels (Figure 15).

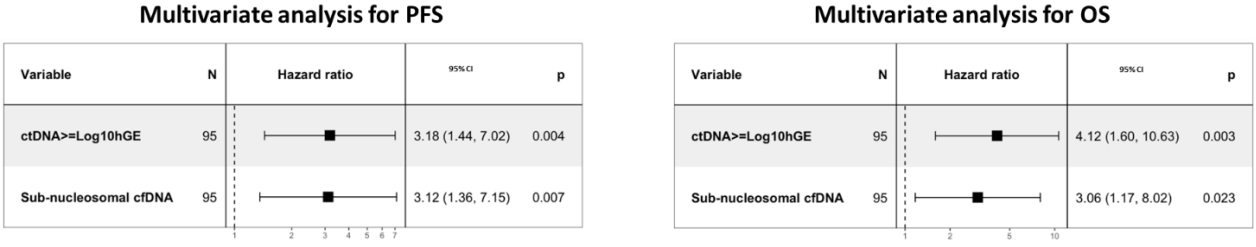

Figure 15. Multivariate analysis of PFS and OS including ctDNA levels and the fragmentomic profile in patients not assigned to any molecular cluster.

5. DISCUSSION

The findings of the present study indicate that: *i)* the levels of ctDNA significantly predict outcomes in a real-world cohort of DLBCL patients uniformly treated with R-CHOP; *ii)* ctDNA can reliably categorize DLBCL patients into different molecular clusters based on gene mutations and SCNAs; *iii)* combining ctDNA levels with molecular clusters improves the prediction of outcomes for DLBCL patients; and *iv)* in cases not assigned to any molecular cluster, the fragmentomic profile of cfDNA identifies a subgroup of patients with shorter cfDNA fragments, characterized by a dismal outcome. Overall, liquid biopsy emerges as a tool that can be employed to identify various disease biomarkers with potential clinical relevance. The analysis of ctDNA successfully assigned a consistent proportion of DLBCL patients to specific molecular clusters in a minimally invasive manner. The percentage of patients classified into different clusters using the LymphGen tool aligns with previous reports, approximately 50% of cases (Rivas-Delgado A *et al.,* 2021; Wilson *et al.,* 2021; Wright *et al.,* 2020). In the subset of 77 cases with available tissue biopsy, the assignment of patients to specific molecular clusters showed similar rates between ctDNA and tissue biopsy. Among the 24 patients assigned to a specific cluster, the concordance rate between ctDNA and lymph node biopsy was 95.8%. Furthermore, the outcomes identified through molecular cluster analysis on tissue biopsy and ctDNA were comparable, underscoring ctDNA as a reliable tool for molecular classification of DLBCL. Current treatment options for DLBCL generally result in durable remission, but few baseline biomarkers accurately predict long-lasting remission or early relapse. While ctDNA levels predict outcomes at baseline, its integration with other molecular features has not been extensively evaluated (Kurtz *et al.,* 2018). Through recursive partitioning, ST2 and BN2 patients exhibited a very good clinical outcome despite elevated ctDNA levels, suggesting that the integration of these biomarkers (*i.e.* ST2 and BN2 cluster and ctDNA levels) improves outcome prediction for DLBCL patients. This evidence supports the idea that DLBCL's heterogeneity can be better understood through an integrated analysis of different disease markers. BN2 and ST2 clusters are characterized by alterations in NOTCH and NF-kB pathway components, as observed in low-grade marginal zone lymphomas (MZLs) in previous studies (Rossi *et al.,* 2012; Spina *et al.,* 2016). These tumors might have undergone occult transformation before diagnosis, or could have potentially originated *de novo* from a common extrafollicular B cell precursor with shared genetic characteristics.

Major constituents of MCD cluster, including *MYD88*^{L265P} and *CD79B* mutations, along with additional mutations involving *ETV6*, *PIM1* and *BTG1*, closely resemble those recently identified in primary CNS and testicular lymphomas (Chapuy *et al.,* 2016). This suggests a potential association of MCD with extranodal tropism. In comparison to EZB tumors, ST2 DLBCLs generally lacked changes in chromatin-modifying enzymes. However, they frequently displayed mutations in H1 linker histones and other core histones, previously observed in follicular lymphoma (FL) (Krysiak et al., 2017; Li et al., 2014). Alterations in H1 linker and core histones may enhance mutation rates by opening chromatin, exposing DNA to the ongoing activation-induced cytidine deaminase (AID) activity in the GC environment, thereby facilitating the malignant transformation of B cells.

Since the LymphGen tool cannot classify a significant fraction of patients, the fragmentation profile was employed to further classify those not assigned to any molecular cluster. This profile, proposed as a novel DLBCL prognostic biomarker (Meriranta *et al.,* 2022), revealed that patients with shorter fragment sizes, indicative of high apoptotic rates in DLBCL cells, had inferior survival. In patients not assigned to any molecular cluster, those with shorter fragment sizes, specifically in the sub-nucleosomal group, showed a worse outcome, independent in multivariate analysis when adjusted for ctDNA levels. The real-life nature of the present study introduces potential selection bias, excluding cases with very aggressive presentations requiring therapy before sampling. However, the PFS and OS of the cohort align with previous studies, suggesting minimal confounding effects. Another limitation of this study rely on the absence of the *NOTCH1* gene in the NGS target panel, precluding the identification of the N1 cluster. Additionally, the targeted sequencing approach did not evaluate *BCL2* and *BCL6* translocations, which could have improved cluster assignment. Nevertheless, the study reliably classified DLBCL patients into specific molecular clusters, reflecting previous reports on tissue biopsy. Overall, liquid biopsy in DLBCL could be exploited to better understand the disease by looking at three different layers of the DLBCL biology, namely ctDNA levels, molecular clusters and fragmentomic profile. The integration of these layers can potentially help identify patients who would benefit the most from the initial chemo-immunotherapy and those who might need a second-line therapy sooner. Future studies on DLBCL focusing more on combining mutational and fragmentation features of ctDNA could influence the design of clinical trials and personalized treatment strategies using liquid biopsy.

REFERENCES

Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000 Feb 3;403(6769):503-11.

Barrans S, Crouch S, Smith A, Crouch S, Smith A, Turner K, Owen R, Patmore R, Roman E, Jack A. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B cell lymphoma treated in the era of rituximab. *J Clin Oncol*. 2010;28(20):3360-3365.

Barrans SL, Evans PA, O'Connor SJ, Kendall SJ, Owen RG, Haynes AP, Morgan GJ, Jack AS. The t(14;18) is associated with germinal center-derived diffuse large B cell lymphoma and is a strong predictor of outcome. *Clin Cancer Res.* 2003;9(6):2133-2139.

Bereshchenko OR, Gu W, Dalla-Favera R. Acetylation inactivates the transcriptional repressor BCL6. *Nat Genet.* 2002;32(4):606-613.

Beroukhim R, Getz G, Nghiemphu L, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc Natl Acad Sci U S A.* 2007;104(50):20007-20012.

Boice M, Salloum D, Mourcin F, Sanghvi V, Amin R, Oricchio E, Jiang M, Mottok A, Denis-Lagache N, Ciriello G, Tam W, Teruya-Feldstein J, de Stanchina E, Chan WC, Malek SN, Ennishi D, Brentjens RJ, Gascoyne RD, Cogné M, Tarte K, Wendel HG. Loss of the HVEM Tumor Suppressor in Lymphoma and Restoration by Modified CAR-T Cells. *Cell.* 2016 Oct 6;167(2):405-418.e13.

Boone DL, Turer EE, Lee EG, Ahmad RC, Wheeler MT, Tsui C, Hurley P, Chien M, Chai S, Hitotsumatsu O, McNally E, Pickart C, Ma A. The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses [published correction appears in Nat Immunol. 2005;6(1):114]. *Nat Immunol.* 2004;5(10):1052-1060.

Brescia P, Schneider C, Holmes AB, Shen Q, et al. MEF2B Instructs Germinal Center Development and Acts as an Oncogene in B Cell Lymphomagenesis. *Cancer Cell*. 2018;34(3):453-465 e9.

Brooks CL, Gu W. The impact of acetylation and deacetylation on the p53 pathway. *Protein Cell.* 2011;2(6):456-462.

Burotto M, Berkovits A, Dunleavy K. Double hit lymphoma: from biology to therapeutic implications. *Expert Rev Hematol.* 2016;9(7):669-678.

Challa-Malladi M, Lieu YK, Califano O, Holmes AB, Bhagat G, Murty VV, Dominguez-Sola D, Pasqualucci L, Dalla-Favera R. Combined genetic inactivation of β2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell*. 2011 Dec 13;20(6):728-40.

Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, Lawrence MS, Roemer MGM, Li AJ, Ziepert M, Staiger AM, Wala JA, Ducar MD, Leshchiner I, Rheinbay E, Taylor-Weiner A, Coughlin CA, Hess JM, Pedamallu CS, Livitz D, Rosebrock D, Rosenberg M, Tracy AA, Horn H, van Hummelen P, Feldman AL, Link BK, Novak AJ, Cerhan JR, Habermann TM, Siebert R, Rosenwald A, Thorner AR, Meyerson ML, Golub TR, Beroukhim R, Wulf GG, Ott G, Rodig SJ, Monti S, Neuberg DS, Loeffler M, Pfreundschuh M, Trümper L, Getz G, Shipp MA. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med.* 2018 May;24(5):679-690.

Chapuy B, Roemer MG, Stewart C, Tan Y, Abo RP, Zhang L, Dunford AJ, Meredith DM, Thorner AR, Jordanova ES, Liu G, Feuerhake F, Ducar MD, Illerhaus G, Gusenleitner D, Linden EA, Sun HH, Homer H, Aono M, Pinkus GS, Ligon AH, Ligon KL, Ferry JA, Freeman GJ, van Hummelen P, Golub TR, Getz G, Rodig SJ, de Jong D, Monti S, Shipp MA. Targetable genetic features of primary testicular and primary central nervous system lymphomas. *Blood*. 2016 Feb 18;127(7):869-81.

Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, Coiffier B, Fisher RI, Hagenbeek A, Zucca E, Rosen ST, Stroobants S, Lister TA, Hoppe RT, Dreyling M, Tobinai K, Vose JM, Connors JM, Federico M, Diehl V; International Harmonization Project on Lymphoma. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007 Feb 10;25(5):579-86.

Chiappella A, Tucci A, Castellino A, Pavone V, Baldi I, Carella AM, Orsucci L, Zanni M, Salvi F, Liberati AM, Gaidano G, Bottelli C, Rossini B, Perticone S, De Masi P, Ladetto M, Ciccone G, Palumbo A, Rossi G, Vitolo U; Fondazione Italiana Linfomi. Lenalidomide plus cyclophosphamide, doxorubicin, vincristine, prednisone and rituximab is safe and effective in untreated, elderly patients with diffuse large B cell lymphoma: a phase I study by the Fondazione Italiana Linfomi. *Haematologica.* 2013 Nov;98(11):1732-8.

Coiffier B, Sarkozy C. Diffuse large B cell lymphoma: R-CHOP failure-what to do?. *Hematology Am Soc Hematol Educ Program.* 2016;2016(1):366–378.

Compagno M, Lim WK, Grunn A, Nandula SV, Brahmachary M, Shen Q, Bertoni F, Ponzoni M, Scandurra M, Califano A, Bhagat G, Chadburn A, Dalla-Favera R, Pasqualucci L. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B cell lymphoma. *Nature.* 2009;459(7247):717-21.

Conacci-Sorrell M, McFerrin L, Eisenman RN. An overview of MYC and its interactome. *Cold Spring Harb Perspect Med*. 2014;4(1):a014357.

Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, Jensen SØ, Medina JE, Hruban C, White JR, Palsgrove DN, Niknafs N, Anagnostou V, Forde P, Naidoo J, Marrone K, Brahmer J, Woodward BD, Husain H, van Rooijen KL, Ørntoft MW, Madsen AH, van de Velde CJH, Verheij M, Cats A, Punt CJA, Vink GR, van Grieken NCT, Koopman M, Fijneman RJA, Johansen JS, Nielsen HJ, Meijer GA, Andersen CL, Scharpf RB, Velculescu VE. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature*. 2019 Jun;570(7761):385-389.

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell.* 2002;111(2):185-196.

Dalla-Favera R, Pasqualucci L. Molecular Genetics of Lymphomas. In: Mauch PM, Armitage JO, Coiffier B, Dalla-Favera R, Harris NL, eds. Non-Hodgkin's Lymphoma. Philadelphia: *Lippincott Williams & Wilkins*; 2010:825-843.

Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, Kohlhammer H, Lamy L, Zhao H, Yang Y, Xu W, Shaffer AL, Wright G, Xiao W, Powell J, Jiang JK, Thomas CJ, Rosenwald A, Ott G, Muller-Hermelink HK, Gascoyne RD, Connors JM, Johnson NA, Rimsza LM, Campo E, Jaffe ES, Wilson WH, Delabie J, Smeland EB, Fisher RI, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Pierce SK, Staudt LM. Chronic active B cell-receptor signalling in diffuse large B cell lymphoma. *Nature.* 2010 Jan 7;463(7277):88-92.

De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol*. 2015 Mar;15(3):137-48.

Ding SC, Lo YMD. Cell-Free DNA Fragmentomics in Liquid Biopsy. *Diagnostics* (Basel). 2022 Apr 13;12(4):978.

Diop F, Moia R, Favini C, Spaccarotella E, De Paoli L, Bruscaggin A, Spina V, Terzi-di-Bergamo L, Arruga F, Tarantelli C, Deambrogi C, Rasi S, Adhinaveni R, Patriarca A, Favini S, Sagiraju S, Jabangwe C, Kodipad AA, Peroni D, Mauro FR, Giudice ID, Forconi F, Cortelezzi A, Zaja F, Bomben R, Rossi FM, Visco C, Chiarenza A, Rigolin GM, Marasca R, Coscia M, Perbellini O, Tedeschi A, Laurenti L, Motta M, Donaldson D, Weir P, Mills K, Thornton P, Lawless S, Bertoni F, Poeta GD, Cuneo A, Follenzi A, Gattei V, Boldorini RL, Catherwood M, Deaglio S, Foà R, Gaidano G, Rossi D. Biological and clinical implications of *BIRC3* mutations in chronic lymphocytic leukemia. *Haematologica*. 2020 Jan 31;105(2):448-456.

Dominguez PM, Ghamlouch H, Rosikiewicz W, Kumar P, Béguelin W, Fontán L, Rivas MA, Pawlikowska P, Armand M, Mouly E, Torres-Martin M, Doane AS, Calvo Fernandez MT, Durant M, Della-Valle V, Teater M, Cimmino L, Droin N, Tadros S, Motanagh S, Shih AH, Rubin MA, Tam W, Aifantis I, Levine RL, Elemento O, Inghirami G, Green MR, Figueroa ME, Bernard OA, Aoufouchi S, Li S, Shaknovich R, Melnick AM. TET2 Deficiency Causes

Germinal Center Hyperplasia, Impairs Plasma Cell Differentiation, and Promotes B-cell Lymphomagenesis. *Cancer Discov*. 2018 Dec;8(12):1632-1653.

Dominguez-Sola D, Kung J, Holmes AB, Wells VA, Mo T, Basso K, Dalla-Favera R. The FOXO1 Transcription Factor Instructs the Germinal Center Dark Zone Program. *Immunity*. 2015 Dec 15;43(6):1064-74.

Duan S, Cermak L, Pagan JK, Rossi M, Martinengo C, di Celle PF, Chapuy B, Shipp M, Chiarle R, Pagano M. FBXO11 targets BCL6 for degradation and is inactivated in diffuse large B cell lymphomas. *Nature.* 2012 Jan 5;481(7379):90-3.

Dunleavy K, Wilson WH. Appropriate Management of Molecular Subtypes of Diffuse Large B cell Lymphoma. *Oncology* (Williston Park). 2014 Apr;28(4):326-34.

Esfahani MS, Hamilton EG, Mehrmohamadi M, et al. Inferring gene expression from cell-free DNA fragmentation profiles. *Nat Biotechnol.* 2022;40(4):585-597.

Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer--a survey. *Biochim Biophys Acta*. 2007;1775(1):181–232.

Goodman RH, Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* 2000;14(13):1553-1577.

Green TM, Young KH, Visco C, Xu-Monette ZY, Orazi A, Go RS, Nielsen O, Gadeberg OV, Mourits-Andersen T, Frederiksen M, Pedersen LM, Møller MB. 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.* 2012; 30(28): 3460–346–-7.

Han DSC, Ni M, Chan RWY, Chan VWH, Lui KO, Chiu RWK, Lo YMD. The Biology of Cell-free DNA Fragmentation and the Roles of DNASE1, DNASE1L3, and DFFB. *Am J Hum Genet.* 2020 Feb 6;106(2):202-214.

Hohaus S, Giachelia M, Massini G, Mansueto G, Vannata B, Bozzoli V, Criscuolo M, D'Alò F, Martini M, Larocca LM, Voso MT, Leone G. Cell-free circulating DNA in Hodgkin's and non-Hodgkin's lymphomas. *Ann Oncol.* 2009;20(8):1408–13.

Ichikawa A, Kinoshita T, Watanabe T, Kato H, Nagai H, Tsushita K, Saito H, Hotta T. Mutations of the p53 gene as a prognostic factor in aggressive B cell lymphoma. *N Engl J Med*. 1997 Aug 21;337(8):529-34.

Iqbal J, Greiner TC, Patel K, Dave BJ, Smith L, Ji J, Wright G, Sanger WG, Pickering DL, Jain S, Horsman DE, Shen Y, Fu K, Weisenburger DD, Hans CP, Campo E, Gascoyne RD, Rosenwald A, Jaffe ES, Delabie J, Rimsza L, Ott G, Müller-Hermelink HK, Connors JM, Vose JM, McKeithan T, Staudt LM, Chan WC; Leukemia/Lymphoma Molecular Profiling Project.

Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B cell lymphoma. *Leukemia.* 2007 Nov;21(11):2332-43.

Iqbal J, Sanger WG, Horsman DE, Rosenwald A, Pickering DL, Dave B, Dave S, Xiao L, Cao K, Zhu Q, Sherman S, Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Ott G, Müller-Hermelink HK, Delabie J, Braziel RM, Jaffe ES, Campo E, Lynch JC, Connors JM, Vose JM, Armitage JO, Grogan TM, Staudt LM, Chan WC. BCL2 translocation defines a unique tumor subset within the germinal center B cell-like diffuse large B cell lymphoma. *Am J Pathol*. 2004;165(1):159-66.

Ivanov M, Baranova A, Butler T, Spellman P, Mileyko V. Non-random fragmentation patterns in circulating cell-free DNA reflect epigenetic regulation. *BMC Genomics*. 2015;16 Suppl 13(Suppl 13):S1.

Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, Knippers R. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* 2001;61(4):1659–65.

Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, Wong GL, Chan SL, Mok TS, Chan HL, Lai PB, Chiu RW, Lo YM. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc Natl Acad Sci U S A*. 2015 Mar 17;112(11):E1317-25.

Karube K, Campo E. MYC alterations in diffuse large B cell lymphomas. *Semin Hematol.* 2015;52(2):97-106.

Karube K, Enjuanes A, Dlouhy I, Jares P, Martin-Garcia D, Nadeu F, Ordóñez GR, Rovira J, Clot G, Royo C, Navarro A, Gonzalez-Farre B, Vaghefi A, Castellano G, Rubio-Perez C, Tamborero D, Briones J, Salar A, Sancho JM, Mercadal S, Gonzalez-Barca E, Escoda L, Miyoshi H, Ohshima K, Miyawaki K, Kato K, Akashi K, Mozos A, Colomo L, Alcoceba M, Valera A, Carrió A, Costa D, Lopez-Bigas N, Schmitz R, Staudt LM, Salaverria I, López-Guillermo A, Campo E. Integrating genomic alterations in diffuse large B cell lymphoma identifies new relevant pathways and potential therapeutic targets. *Leukemia.* 2018;32(3):675-684.

Kato M, Sanada M, Kato I, Sato Y, Takita J, Takeuchi K, Niwa A, Chen Y, Nakazaki K, Nomoto J, Asakura Y, Muto S, Tamura A, Iio M, Akatsuka Y, Hayashi Y, Mori H, Igarashi T, Kurokawa M, Chiba S, Mori S, Ishikawa Y, Okamoto K, Tobinai K, Nakagama H, Nakahata T, Yoshino T, Kobayashi Y, Ogawa S. Frequent inactivation of A20 in B cell lymphomas. *Nature.* 2009;459(7247):712-716.

Kerbauy FR, Colleoni GW, Saad ST, Regis Silva MR, Correa Alves A, Aguiar KC, Albuquerque DM, Kobarg J, Seixas MT, Kerbauy J. Detection and possible prognostic relevance of p53 gene mutations in diffuse large B cell lymphoma. An analysis of 51 cases and review of the literature. *Leuk Lymphoma.* 2004 Oct;45(10):2071-8.

Knies N, Alankus B, Weilemann A, Tzankov A, Brunner K, Ruff T, Kremer M, Keller UB, Lenz G, Ruland J. Lymphomagenic CARD11/BCL10/MALT1 signaling drives malignant B cell proliferation via cooperative NF-κB and JNK activation. *Proc Natl Acad Sci U S A.* 2015 Dec 29;112(52):E7230-8.

Krysiak K, Gomez F, White BS, Matlock M, Miller CA, Trani L, Fronick CC, Fulton RS, Kreisel F, Cashen AF, Carson KR, Berrien-Elliott MM, Bartlett NL, Griffith M, Griffith OL, Fehniger TA. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. *Blood*. 2017 Jan 26;129(4):473-483.

Kurtz DM, Green MR, Bratman SV, Scherer F, Liu CL, Kunder CA, Takahashi K, Glover C, Keane C, Kihira S, Visser B, Callahan J, Kong KA, Faham M, Corbelli KS, Miklos D, Advani RH, Levy R, Hicks RJ, Hertzberg M, Ohgami RS, Gandhi MK, Diehn M, Alizadeh AA. Noninvasive monitoring of diffuse large B cell lymphoma by immunoglobulin high-throughput sequencing. *Blood.* 2015;125(24):3679-87.

Kurtz DM, Scherer F, Jin MC, et al. Circulating Tumor DNA Measurements As Early Outcome Predictors in Diffuse Large B-Cell Lymphoma. *J Clin Oncol*. 2018;36(28):2845-2853.

Kwak JY. Treatment of diffuse large B cell lymphoma. *Korean J Intern Med.* 2012;27(4):369–377.

Lacy SE, Barrans SL, Beer PA, Painter D, Smith AG, Roman E, Cooke SL, Ruiz C, Glover P, Van Hoppe SJL, Webster N, Campbell PJ, Tooze RM, Patmore R, Burton C, Crouch S, Hodson DJ. Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. *Blood*. 2020 May 14;135(20):1759-1771.

Lam LT, Wright G, Davis RE, Lenz G, Farinha P, Dang L, Chan JW, Rosenwald A, Gascoyne RD, Staudt LM. Cooperative signaling through the signal transducer and activator of transcription 3 and nuclear factor-{kappa}B pathways in subtypes of diffuse large B cell lymphoma. *Blood*. 2008 Apr 1;111(7):3701-13.

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, Gascoyne RD, Connors JM, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Fisher RI, Chan WC, Staudt LM. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science.* 2008;319(5870):1676-1679.

Li H, Kaminski MS, Li Y, Yildiz M, Ouillette P, Jones S, Fox H, Jacobi K, Saiya-Cork K, Bixby D, Lebovic D, Roulston D, Shedden K, Sabel M, Marentette L, Cimmino V, Chang AE, Malek SN. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood*. 2014 Mar 6;123(10):1487-98.

Li S, Young KH, Medeiros LJ. Diffuse large B-cell lymphoma. Pathology. 2018 Jan;50(1):74-87

Lu D, Liu L, Ji X, Gao Y, Chen X, Liu Y, Liu Y, Zhao X, Li Y, Li Y, Jin Y, Zhang Y, McNutt MA, Yin Y. The phosphatase DUSP2 controls the activity of the transcription activator STAT3 and regulates TH17 differentiation. *Nat Immunol.* 2015 Dec;16(12):1263-73.

Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747-1756.

Meriranta L, Alkodsi A, Pasanen A, Lepistö M, Mapar P, Blaker YN, Jørgensen J, Karjalainen-Lindsberg ML, Fiskvik I, Mikalsen LTG, Autio M, Björkholm M, Jerkeman M, Fluge Ø, Brown P, Jyrkkiö S, Holte H, Pitkänen E, Ellonen P, Leppä S. Molecular features encoded in the ctDNA reveal heterogeneity and predict outcome in high-risk aggressive B-cell lymphoma. *Blood*. 2022 Mar 24;139(12):1863-1877.

Mesin L, Ersching J, Victora GD. Germinal Center B Cell Dynamics. *Immunity*. 2016;45(3):471–482.

Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988 Feb 11;16(3):1215.

Monti S, Chapuy B, Takeyama K, Rodig SJ, Hao Y, Yeda KT, Inguilizian H, Mermel C, Currie T, Dogan A, Kutok JL, Beroukhim R, Neuberg D, Habermann TM, Getz G, Kung AL, Golub TR, Shipp MA. Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. *Cancer Cell.* 2012 Sep 11;22(3):359-72.

Morales JC, Franco S, Murphy MM, Bassing CH, Mills KD, Adams MM, Walsh NC, Manis JP, Rassidakis GZ, Alt FW, Carpenter PB. 53BP1 and p53 synergize to suppress genomic instability and lymphomagenesis. *Proc Natl Acad Sci U S A.* 2006 Feb 28;103(9):3310-5.

Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, Paul JE, Boyle M, Woolcock BW, Kuchenbauer F, Yap D, Humphries RK, Griffith OL, Shah S, Zhu H, Kimbara M, Shashkin P, Charlot JF, Tcherpakov M, Corbett R, Tam A, Varhol R, Smailus D, Moksa M, Zhao Y, Delaney A, Qian H, Birol I, Schein J, Moore R, Holt R, Horsman DE, Connors JM, Jones S, Aparicio S, Hirst M, Gascoyne RD, Marra MA. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B cell lymphomas of germinal-center origin. *Nat Genet*. 2010 Feb;42(2):181-5.

Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, Johnson NA, Severson TM, Chiu R, Field M, Jackman S, Krzywinski M, Scott DW, Trinh DL, Tamura-Wells J, Li S, Firme MR, Rogic S, Griffith M, Chan S, Yakovenko O, Meyer IM, Zhao EY, Smailus D, Moksa M, Chittaranjan S, Rimsza L, Brooks-Wilson A, Spinelli JJ, Ben-Neriah S,

Meissner B, Woolcock B, Boyle M, McDonald H, Tam A, Zhao Y, Delaney A, Zeng T, Tse K, Butterfield Y, Birol I, Holt R, Schein J, Horsman DE, Moore R, Jones SJ, Connors JM, Hirst M, Gascoyne RD, Marra MA. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature.* 2011 Jul 27;476(7360):298-303.

Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, Mair R, Goranova T, Marass F, Heider K, Wan JCM, Supernat A, Hudecova I, Gounaris I, Ros S, Jimenez-Linan M, Garcia-Corbacho J, Patel K, Østrup O, Murphy S, Eldridge MD, Gale D, Stewart GD, Burge J, Cooper WN, van der Heijden MS, Massie CE, Watts C, Corrie P, Pacey S, Brindle KM, Baird RD, Mau-Sørensen M, Parkinson CA, Smith CG, Brenton JD, Rosenfeld N. Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci Transl Med.* 2018 Nov 7;10(466):eaat4921.

Muppidi JR, Schmitz R, Green JA, Xiao W, Larsen AB, Braun SE, An J, Xu Y, Rosenwald A, Ott G, Gascoyne RD, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Vaidehi N, Staudt LM, Cyster JG. Loss of signalling via Gα13 in germinal centre B-cell-derived lymphoma. *Nature*. 2014 Dec 11;516(7530):254-8.

Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, Shrager JB, Loo BW Jr, Alizadeh AA, Diehn M. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med.* 2014 May;20(5):548-54.

Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, Stehr H, Liu CL, Bratman SV, Say C, Zhou L, Carter JN, West RB, Sledge GW, Shrager JB, Loo BW Jr, Neal JW, Wakelee HA, Diehn M, Alizadeh AA. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotechnol*. 2016 May;34(5):547-555.

Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, Kohlhammer H, Xu W, Yang Y, Zhao H, Shaffer AL, Romesser P, Wright G, Powell J, Rosenwald A, Muller-Hermelink HK, Ott G, Gascoyne RD, Connors JM, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Fisher RI, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Staudt LM. Oncogenically active MYD88 mutations in human lymphoma. *Nature.* 2011;470(7332):115-9.

Olshen AB, Bengtsson H, Neuvial P, Spellman PT, Olshen RA, Seshan VE. Parent-specific copy number in paired tumor-normal studies using circular binary segmentation. *Bioinformatics*. 2011;27(15):2038-2046.

Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B cell lymphoma. *Blood.* 2018 May 24;131(21):2307-2319.

Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, Kasper LH, Lerach S, Tang H, Ma J, Rossi D, Chadburn A, Murty VV, Mullighan CG, Gaidano G, Rabadan R, Brindle PK, Dalla-Favera R. Inactivating mutations of acetyltransferase genes in B cell lymphoma. *Nature*. 2011 Mar 10;471(7337):189-95.

Pasqualucci L, Migliazza A, Basso K, Houldsworth J, Chaganti RS, Dalla-Favera R. Mutations of the BCL6 proto-oncogene disrupt its negative autoregulation in diffuse large B cell lymphoma. *Blood*. 2003;101(8):2914-2923.

Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, Chaganti RS, Klein U, Küppers R, Rajewsky K, Dalla-Favera R. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci* U S A. 1998; Sep 29;95(20):11816-21.

Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, Wells VA, Grunn A, Messina M, Elliot O, Chan J, Bhagat G, Chadburn A, Gaidano G, Mullighan CG, Rabadan R, Dalla-Favera R. Analysis of the coding genome of diffuse large B cell lymphoma. *Nat Genet*. 2011 Jul 31;43(9):830-7.

Phelan JD, Young RM, Webster DE, Roulland S, Wright GW, Kasbekar M, Shaffer AL 3rd, Ceribelli M, Wang JQ, Schmitz R, Nakagawa M, Bachy E, Huang DW, Ji Y, Chen L, Yang Y, Zhao H, Yu X, Xu W, Palisoc MM, Valadez RR, Davies-Hill T, Wilson WH, Chan WC, Jaffe ES, Gascoyne RD, Campo E, Rosenwald A, Ott G, Delabie J, Rimsza LM, Rodriguez FJ, Estephan F, Holdhoff M, Kruhlak MJ, Hewitt SM, Thomas CJ, Pittaluga S, Oellerich T, Staudt LM. A multiprotein supercomplex controlling oncogenic signalling in lymphoma. *Nature*. 2018 Aug;560(7718):387-391.

Rhyasen GW, Starczynowski DT. IRAK signalling in cancer. *Br J Cancer.* 2015;112(2):232–237.

Rivas-Delgado A, Nadeu F, Enjuanes A, Casanueva-Eliceiry S, Mozas P, Magnano L, Castrejón de Anta N, Rovira J, Dlouhy I, Martín S, Osuna M, Rodríguez S, Simó M, Pinyol M, Baumann T, Beà S, Balagué O, Delgado J, Villamor N, Setoain X, Campo E, Giné E, López-Guillermo A. Mutational Landscape and Tumor Burden Assessed by Cell-free DNA in Diffuse Large B-Cell Lymphoma in a Population-Based Study. *Clin Cancer Res.* 2021 Jan 15;27(2):513-521.

Rosati E, Baldoni S, De Falco F, Del Papa B, Dorillo E, Rompietti C, Albi E, Falzetti F, Di Ianni M, Sportoletti P. NOTCH1 Aberrations in Chronic Lymphocytic Leukemia. *Front Oncol*. 2018 Jun 27;8:229.

Roschewski M, Dunleavy K, Pittaluga S, Moorhead M, Pepin F, Kong K, Shovlin M, Jaffe ES, Staudt LM, Lai C, Steinberg SM, Chen CC, Zheng J, Willis TD, Faham M, Wilson WH. Circulating tumor DNA and CT monitoring in patients with untreated diffuse large B cell lymphoma: a correlative biomarker study. *Lancet Oncol.* 2015 May;16(5):541-9.

Roschewski M, Staudt LM, Wilson WH. Dynamic monitoring of circulating tumor DNA in non-Hodgkin lymphoma. *Blood.* 2016;127(25):3127-3132.

Rosenthal, A, Younes, A. High grade B cell lymphoma with rearrangements of MYC and BCL2 and/or BCL6: double hit and triple hit lymphomas and double expressing lymphoma. *Blood Rev.* 2017; 31: 37–42.

Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, López-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T, Staudt LM for the Lymphoma/Leukemia Molecular Profiling Project. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B cell lymphoma. *N Engl J Med.* 2002;346(25):1937-1947.

Rossi D, Diop F, Spaccarotella E, Monti S, Zanni M, Rasi S, Deambrogi C, Spina V, Bruscaggin A, Favini C, Serra R, Ramponi A, Boldorini R, Foà R, Gaidano G. Diffuse large B-cell lymphoma genotyping on the liquid biopsy. *Blood*. 2017 Apr 6;129(14):1947-1957.

Rossi D, Spina V, Bruscaggin A, Gaidano G. Liquid biopsy in lymphoma. *Haematologica* (2019) 104:648–52.

Rossi D, Trifonov V, Fangazio M, Bruscaggin A, Rasi S, Spina V, Monti S, Vaisitti T, Arruga F, Famà R, Ciardullo C, Greco M, Cresta S, Piranda D, Holmes A, Fabbri G, Messina M, Rinaldi A, Wang J, Agostinelli C, Piccaluga PP, Lucioni M, Tabbò F, Serra R, Franceschetti S, Deambrogi C, Daniele G, Gattei V, Marasca R, Facchetti F, Arcaini L, Inghirami G, Bertoni F, Pileri SA, Deaglio S, Foà R, Dalla-Favera R, Pasqualucci L, Rabadan R, Gaidano G. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med.* 2012 Aug 27;209(9):1537-51.

Rushton C, Alcaide M, Cheung M, Thomas N, Arthur S, Michaud N, Daigle S, Davidson J, Bushell K, Yu S, Jain M, Shepherd L, Crump M, Mann K, Kuruvilla J, Assouline S, Johnson N, Scott DW, Morin RD. Identifying Mutations Enriched in Relapsed-Refractory Dlbcl to Derive Genetic Factors Underlying Treatment Resistance. *Hematol. Oncol*. 2019;37, :35–36.

Rushton C, Arthur SE, Alcaide M, Cheung M, Thomas N, Hilton LK, Michaud NR, Daigle S, Davidson J, Bushell KR, Yu S, Jain MD, Shepherd LE, Crump M, Mann KK, Kuruvilla J, Assouline S, Cleary K, Cragg MS, Scott DW, Johnson N, Morin RD. Recurrent Patterns of Clonal Evolution in Relapsed-Refractory DLBCL Following Treatment with R-CHOP. *Blood.* 2019:134 (Supplement_1): 921.

Saito M, Novak U, Piovan E, Basso K, Sumazin P, Schneider C, Crespo M, Shen Q, Bhagat G, Califano A, Chadburn A, Pasqualucci L, Dalla-Favera R. BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc Natl Acad Sci* U S A. 2009 Jul 7;106(27):11294-9.

Schneider C, Kon N, Amadori L, Shen Q, Schwartz FH, Tischler B, Bossennec M, Dominguez-Sola D, Bhagat G, Gu W, Basso K, Dalla-Favera R. FBXO11 inactivation leads to abnormal germinal-center formation and lymphoproliferative disease. *Blood.* 2016;128(5):660-666.

Scott DW, Mottok A, Ennishi D, Wright GW, Farinha P, Ben-Neriah S, Kridel R, Barry GS, Hother C, Abrisqueta P, Boyle M, Meissner B, Telenius A, Savage KJ, Sehn LH, Slack GW, Steidl C, Staudt LM, Connors JM, Rimsza LM, Gascoyne RD. Prognostic Significance of Diffuse Large B cell Lymphoma Cell-of-origin Determined by Digital Gene Expression in Formalin-Fixed Paraffin-Embedded Tissue Biopsies. *J Clin Oncol*. 2015 Sep 10;33(26):2848-56.

Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. *N Engl J Med*. 2021 Mar 4;384(9):842-858.

Serpas L, Chan RWY, Jiang P, Ni M, Sun K, Rashidfarrokhi A, Soni C, Sisirak V, Lee WS, Cheng SH, Peng W, Chan KCA, Chiu RWK, Reizis B, Lo YMD. *Dnase1l3* deletion causes aberrations in length and end-motif frequencies in plasma DNA. *Proc Natl Acad Sci U S A.* 2019 Jan 8;116(2):641-649.

Shaffer AL 3rd, Young RM, Staudt LM. Pathogenesis of Human B Cell Lymphomas. *Annual Review of Immunology* 2012; 30:1, 565-610.

Shen HM, Peters A, Baron B, Zhu X, Storb U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science.* 1998;280(5370):1750-1752.

Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem.* 2012; 81:65-95.
Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, Richon VM, Copeland RA. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B cell lymphomas. *Proc Natl Acad Sci* U S A. 2010 Dec 7;107(49):20980-5.

Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell*. 2016 Jan 14;164(1-2):57-68.

Spina V, Bruscaggin A, Cuccaro A, Martini M, Di Trani M, Forestieri G, Manzoni M, Condoluci A, Arribas A, Terzi-Di-Bergamo L, Locatelli SL, Cupelli E, Ceriani L, Moccia AA, Stathis A, Nassi L, Deambrogi C, Diop F, Guidetti F, Cocomazzi A, Annunziata S, Rufini V, Giordano A, Neri A, Boldorini R, Gerber B, Bertoni F, Ghielmini M, Stüssi G, Santoro A, Cavalli F, Zucca E, Larocca LM, Gaidano G, Hohaus S, Carlo-Stella C, Rossi D. Circulating tumor DNA reveals genetics, clonal evolution, and residual disease in classical Hodgkin lymphoma. *Blood*. 2018 May 31;131(22):2413-2425.

Spina V, Rossi D. Liquid biopsy in tissue-born lymphomas. *Swiss Med Wkly*. 2019 Jan 23;149:w14709.

Spina V, Khiabanian H, Messina M, Monti S, Cascione L, Bruscaggin A, Spaccarotella E, Holmes AB, Arcaini L, Lucioni M, Tabbò F, Zairis S, Diop F, Cerri M, Chiaretti S, Marasca R, Ponzoni M, Deaglio S, Ramponi A, Tiacci E, Pasqualucci L, Paulli M, Falini B, Inghirami G, Bertoni F, Foà R, Rabadan R, Gaidano G, Rossi D. The genetics of nodal marginal zone lymphoma. *Blood*. 2016 Sep 8;128(10):1362-73.

Steinberg MW, Cheung TC, Ware CF. The signaling networks of the herpesvirus entry mediator (TNFRSF14) in immune regulation. *Immunol Rev*. 2011;244(1):169-187.

Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clin Chim Acta.* 2001 Nov;313(1-2):139-42.

Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, Advani R, Ghielmini M, Salles GA, Zelenetz AD, Jaffe ES. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016 May 19;127(20):2375-90.

Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol*. 2016;12(4):e1004873.

Tallen G, Riabowol K. Keep-ING balance: tumor suppression by epigenetic regulation. *FEBS Lett.* 2014;588(16):2728–42.

Talotta D, Almasri M, Cosentino C, Gaidano G, Moia R. Liquid biopsy in hematological malignancies: current and future applications. *Front Oncol.* 2023 Apr 20;13:1164517.

Thieblemont C, Briere J, Mounier N, Voelker HU, Cuccuini W, Hirchaud E, Rosenwald A, Jack A, Sundstrom C, Cogliatti S, Trougouboff P, Boudova L, Ysebaert L, Soulier J, Chevalier C, Bron D, Schmitz N, Gaulard P, Houlgatte R, Gisselbrecht C. The Germinal Center/Activated B cell Subclassification Has a Prognostic Impact for Response to Salvage Therapy in Relapsed/Refractory Diffuse Large B cell Lymphoma: A Bio-CORAL Study *J Clin Oncol*. 2011; 29(:31), :4079-4087.

Thome M. CARMA1, BCL-10 and MALT1 in lymphocyte development and activation. *Nat Rev Immunol*. 2004;4(5):348-359.

Tilly H, Gomes da Silva M, Vitolo U, Jack A, Meignan M, Lopez-Guillermo A, Walewski J, André M, Johnson PW, Pfreundschuh M, Ladetto M; ESMO Guidelines Committee. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2015 Sep;26 Suppl 5:v116-25.

Trinh DL, Scott DW, Morin RD, Mendez-Lago M, An J, Jones SJ, Mungall AJ, Zhao Y, Schein J, Steidl C, Connors JM, Gascoyne RD, Marra MA. Analysis of FOXO1 mutations in diffuse large B cell lymphoma. *Blood.* 2013 May 2;121(18):3666-74.

Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, Gligorich KM, Rostomily RC, Bronner MP, Shendure J. Fragment Length of Circulating Tumor DNA. *PLoS Genet.* 2016 Jul 18;12(7):e1006162.

Underhill HR. Leveraging the fragment length of circulating tumour dna to improve molecular profiling of solid tumour malignancies with next-generation sequencing: a pathway to advanced non-invasive diagnostics in precision oncology? *Mol Diagn Ther* (2021) 25:389–408.

Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics*. 2007;23(6):657-663.

Wang H, Jain S, Li P, Lin JX, et al. Transcription factors IRF8 and PU.1 are required for follicular B cell development and BCL6-driven germinal center responses. *Proc Natl Acad Sci U S A.* 2019;116(19):9511–20.

Wang X, Li Z, Naganuma A, Ye BH. Negative autoregulation of BCL-6 is bypassed by genetic alterations in diffuse large B cell lymphomas. *Proc Natl Acad Sci U S A.* 2002 Nov 12;99(23):15018-23.

Weber T, Schmitz R. Molecular Subgroups of Diffuse Large B Cell Lymphoma: Biology and Implications for Clinical Practice. *Curr Oncol Rep.* 2022 Jan;24(1):13-21.

Wilson WH, Gerecitano JF, Goy A, Vos SD, Kenkre VP, Barr PM, Blum KA, Shustov AR, Advani RH, Lih J, Williams MP, Schmitz R, Yang Y, Pittaluga S, Wright GE, Kunkel LA, McGreivy J, Balasubramanian S, Cheng M, Moussa D, Buggy JJ, Staudt LM. The bruton's tyrosine kinase (BTK) inhibitor, ibrutinib (PCI-32765), has preferential activity in the ABC subtype of relapsed/refractory de novo diffuse large B cell lymphoma (DLBCL): Interim results of a multicenter, open-label, phase 2 study. *Blood.* 2012; 120(21): 686.

Wilson WH, Wright GW, Huang DW, Hodkinson B, Balasubramanian S, Fan Y, Vermeulen J, Shreeve M, Staudt LM. Effect of ibrutinib with R-CHOP chemotherapy in genetic subtypes of DLBCL. *Cancer Cell.* 2021 Dec 13;39(12):1643-1653.e3.

Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G, Lih CJ, Williams PM, Shaffer AL, Gerecitano J, de Vos S, Goy A, Kenkre VP, Barr PM, Blum KA, Shustov A, Advani R, Fowler NH, Vose JM, Elstrom RL, Habermann TM, Barrientos JC, McGreivy J, Fardis M, Chang BY, Clow F, Munneke B, Moussa D, Beaupre DM, Staudt LM. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med* 21, 922–926 (2015).

Wright GW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, Wang JQ, Schmitz R, Morin RD, Tang J, Jiang A, Bagaev A. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell*. 2020; 37(4):551-68.

Xu-Monette ZY, Wu L, Visco C, Tai YC, Tzankov A, Liu WM, Montes-Moreno S, Dybkaer K, Chiu A, Orazi A, Zu Y, Bhagat G, Richards KL, Hsi ED, Zhao XF, Choi WW, Zhao X, van Krieken JH, Huang Q, Huh J, Ai W, Ponzoni M, Ferreri AJ, Zhou F, Kahl BS, Winter JN, Xu W, Li J, Go RS, Li Y, Piris MA, Møller MB, Miranda RN, Abruzzo LV, Medeiros LJ, Young KH. Mutational profile and prognostic significance of TP53 in diffuse large B cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. *Blood.* 2012 Nov 8;120(19):3986-96.

Ye BH, Chaganti S, Chang CC, Niu H, Corradini P, Chaganti RS, Dalla-Favera R. Chromosomal translocations cause deregulated BCL6 expression by promoter substitution in B cell lymphoma. *EMBO J.* 1995;14(24):6209-6217.

Ying CY, Dominguez-Sola D, Fabi M, Lorenz IC, Hussein S, Bansal M, Califano A, Pasqualucci L, Basso K, Dalla-Favera R. MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. *Nat Immunol*. 2013 Oct;14(10):1084-92.

Young KH, Weisenburger DD, Dave BJ, Smith L, Sanger W, Iqbal J, Campo E, Delabie J, Gascoyne RD, Ott G, Rimsza L, Müller-Hermelink HK, Jaffe ES, Rosenwald A, Staudt LM, Chan WC, Greiner TC. Mutations in the DNA-binding codons of TP53, which are associated with decreased expression of TRAILreceptor-2, predict for poor survival in diffuse large B cell lymphoma. *Blood*. 2007 Dec 15;110(13):4396-405.

Young RM, Shaffer AL 3rd, Phelan JD, Staudt LM. B cell receptor signaling in diffuse large B cell lymphoma. *Semin Hematol.* 2015 Apr;52(2):77-85.

Zhang J, Dominguez-Sola D, Hussein S, Lee JE, Holmes AB, Bansal M, Vlasevska S, Mo T, Tang H, Basso K, Ge K, Dalla-Favera R, Pasqualucci L. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med.* 2015 Oct;21(10):1190-8.