UNIVERSITY OF PIEMONTE ORIENTALE "AMEDEO AVOGADRO"

PhD in Medical Sciences and Biotechnologies

XXX cycle

Department of Health Sciences (DISS)
SSD: MED/08

Targeted therapies in lung adenocarcinomas: new methodologies and new markers

PhD thesis by:
EPISTOLIO SAMANTHA

Director of PhD programme: Prof. Marisa Gariglio
Supervisor: Prof. Renzo Boldorini
Advisor: Dr. Milo Frattini

Academic Year 2016/2017

Table of contents

List of abbreviations

AC: adenocarcinoma

AL: lysis buffer

ALK: anaplastic lymphoma receptor tyrosine kinase

APC: adenomatous polyposis coli gene

ATP: adenosine triphosphate

ATS: american thoracic society

B-ALL: acute lymphocytic leukemia

B-CLL: B-cell chronic lymphocytic leukemia

BCR: breakpoint cluster region

BCT: blood collection tubes

BEAMing: Beads, emulsion, amplification and magnetics

cDNA: complementary DNA

cfDNA: cell-free DNA

CR: cysteine-rich domain

CRC: colorectal cancer

CRD: cysteine-rich frizzled domain

Ct: threshold cycle

CTCs: circulating tumor cells

ctDNA: circulating tumor DNA

Da: dalton

DAPI: 4',6-diamidino-2-phenylindole

DCR: disease control rate

ddPCR: droplet digital PCR

DS: direct sequencing

ECM: extracellular matrix

e.g.: example given

EGF: epidermal growth factor

EGFR: epidermal growth factor receptor

EMA: european medicines agency

EML4: echinoderm microtubule-associated protein-like 4

ERS: european respiratory society

ex: exon

FDA: food and drug administration

FISH: fluorescent in situ hybridization

Fw: forward

GDP: guanosine diphosphate

GTP: guanosine triphosphate

HE: hematoxylin and eosin

HER2: human epidermal growth factor receptor 2

IASLC: international association for the study of lung cancer

i.e.: id est

Ig-like: immunoglobulin-like

IHC: immunohistochemistry

KRAS: Kirsten Rat sarcoma viral oncogene homolog

L: large EGF binding domain

LCC: large cell carcinoma

LOD: limit of detection

M: molarity

MCL: mantle cell leukemia

MET: tyrosine-protein kinase Met

Mg: magnesium

min: minute

miRNA: microRNA

mplx: multiplex assay

n: number

NCCN: national comprehensive cancer network

NE: not evaluable

ng: nanogram

NGS: next-generation sequencing

NRAS: neuroblastoma RAS viral oncogene homolog

NSCLC: non-small-cell lung cancer

nt: nucleotide

ORR: objective response rate

OS: overall survival

p: p value

PCR: polymerase chain reaction

PD-1: programmed cell death protein-1

PD-L1: programmed cell death protein-1 ligand

PFS: progression free survival

PK: proteinase K

PRD: proline-rich frizzled domain

Pt: platinum

Q: quencher

qPCR: quantitative PCR

R: reporter

rcf: relative centrifugal force

real-time qPCR: real-time quantitative PCR

RFU: relative fluorescence units

ROR: receptor tyrosine kinase-like orphan receptors

RS: relative survival

RT: room temperature

Rv: primer reverse

s: second

SCC: squamous cell carcinoma

SCLC: small-cell lung cancer

simpl: simplex assay

SRD: serine rich domain

T-DM1: trastuzumab emtansine

TK: tyrosine kinase

TKI: tyrosine kinase inhibitor

TNM: tissue-node-metastasis classification

TRD: threonine rich domain

TTF-1: thyroid transcription factor-1

UV: ultraviolet

V: variant

VEGF: vascular endothelial growth factor

vs: versus

WHO: world health organization

wt: wild-type

WTB: wt BaseBlocker^{TM}

µL: microliter

µm: micrometer

1stG: first-generation

2ndG: second-generation

3rdG: third-generation

ABSTRACT

Lung cancer is the leading cause of tumor related death worldwide. Lung AdenoCarcinoma (AC) histotype is the most diffused, corresponding to 60% of Non-Small-Cell Lung Cancers (NSCLCs), and its treatment is, generally, based on resection combined with platinum-doublet based chemotherapy. However chemotherapeutic treatment harbours some limitations such as the lack of specificity for tumor cells and the frequent and severe dose-limiting toxicities. In order to solve these limitations, last years have seen the development of targeted therapies, treatments capable to specifically target the main molecular alterations driving tumorigenesis. The most diffused targeted therapies in lung ACs are small-molecule tyrosine kinase inhibitors (TKIs) acting against Epidermal Growth Factor Receptor (EGFR) and Anaplastic Lymphoma receptor tyrosine Kinase (ALK) alterations. More recently, targeted drugs against *ROS1*, Human Epidermal growth factor Receptor 2 gene (*HER2*), and *BRAF* have been introduced in the clinical setting, whereas the therapies against Kirsten Rat sarcoma viral oncogene homolog (*KRAS*) mutations, the most frequent in lung AC, are still under development. These data highlight that the molecular characterization of histological or cytological samples is strictly necessary to define a possible response to targeted drugs. In last two years a new important advancement was accepted by FDA: the characterization of molecular markers, especially *EGFR*, in liquid biopsies. These analyses have many advantages compared to tissue biopsies testing: they are less invasive, they better reflect tumor heterogeneity and they can be repeated serially overtime. Liquid biopsies are fundamental for patients experiencing secondary resistance to EGFR inhibitors but can be used in upfront patients as well, especially in the cases without tissue material available.

The aim of my project was to improve the clinical care of patients affected by lung AC, by enhancing the characterization of *EGFR* mutations in one side and by characterizing the expression of a new, putative marker, the tyrosine kinase-like orphan receptor-1 (ROR1) in the other side. In order to improve the determination of *EGFR* in lung AC specimens, we decided to test and validate a new methodology characterized by high sensitivity, simplicity of execution and uniformity in the interpretation of results. The need of a more sensitive kit is due to the fact that lung AC samples are generally represented by small biopsies characterized by poor quantity and quality of cell tumor content. In addition, cancer cells are very often dispersed in a high quantity of normal cells, leading the biopsy difficult to be considered representative of the whole tumor. In particular, in collaboration with a Danish company (PentaBase ApS), I developed new real-time based assays (SensiScreen®) containing reagents capable to increase specificity and sensitivity of *EGFR* mutation detection. Simplex assays (for G719A, G719C, G719S, S768I, T790M, L858R, L861Q mutations) and multiplex assays

(for G719 mutations, exon 19 deletions or exon 20 insertions), specific for tissue and liquid biopsies, were validated on plasmids and on different dilution points of DNA extracted from mutated cell lines. The validation revealed a limit of detection (LOD) that is between 0.1% and 1%, corresponding to the detection of until 1 copy of mutant DNA in a wild-type (wt) background. Afterwards, the new methodology was tested on two retrospective lung AC patient cohorts: the first one characterized by 471 tissue samples and the second one by 61 plasma, 5 serum and 39 paired tissue specimens. Before SensiScreen® application the first cohort was characterized for *EGFR, KRAS, BRAF, HER2* mutations by Direct Sequencing (DS), for *ALK* and *ROS1* gene rearrangements by Fluorescent In Situ Hibridization (FISH) and for thyroid transcriptional factor-1 (TTF-1) expression by immunohistochemistry (IHC); whereas the second cohort was evaluated for *EGFR* by different methodologies (DS; cobas®, Roche; Ion Torrent®, Thermo Fisher Scientific; TheraScreen®, QIAGEN). In the first and second cohort the percentage of alterations of the various markers are in line with the data reported in the literature for Caucasian patients. In cohort one, by applying the two-tailed Fisher's exact test for the comparison among the molecular alterations, we found a nearly mutual exclusivity between *EGFR* and *KRAS* mutations and a tendency of *BRAF* or *KRAS* wt sequences in TTF-1 positive cases. Using the new developed assays, we confirmed the *EGFR* mutations found by DS and by the other aforementioned methodologies. Most importantly, in the first cohort, we identified 14 additional mutations compared to DS and in the second cohort two new mutated cases compared to TheraScreen® (QIAGEN) and cobas® (Roche). These two patients, classified as L858R negative by TheraScreen® (QIAGEN) and cobas® (Roche) techniques on DNA from plasma, were subjected to a subsequent tissue biopsy for the molecular characterization. In tissue cobas® (Roche), TheraScreen® (QIAGEN) and SensiScreen® assays demonstrated the presence of the L858R change. So we can conclude that the application of these highly sensitive assays may lead to address a higher number of patients to EGFR-inhibitors (that are more efficient in *EGFR*-mutant cases with respect to standard chemotherapies) and may also prevent the use of tissue biopsies, that can be the source of side effects for the patients. These conclusions suggest the adoption of the new methodology in the laboratories of molecular pathology because SensiScreen® is a faster, easy-to-use and highly sensitive method for *EGFR* characterization in tissue and plasma samples.

The new assay enlarges the number of lung AC patients characterized by *EGFR* alterations, however nearly 70% of patients do not harbour mutations in *EGFR* or in the other molecular

markers for which a targeted therapy has been already developed. As a consequence, as second aim of my PhD, we evaluated a promising marker for lung AC, ROR1. This protein may represent a good candidate for new drugs, currently studied in tumour cell lines, because it is not expressed in adult normal tissues but only in cancer and because a study in literature reports its association with TTF-1, the typical AC marker. We investigated ROR1 and miR-382 (a microRNA, miRNA, involved in ROR1 inhibition) by TaqMan real-time PCR assays in a third retrospective cohort of 102 lung AC patients, for which we have got pathological and clinical data. In this cohort, we characterized *EGFR*, *KRAS*, *BRAF*, *HER2* by DS, *ALK*, *ROS1* by FISH, TTF-1 by IHC and we found alterations in percentages similar to those reported in the literature. Our analyses revealed ROR1 overexpression in 28.6% and miR-382 expression in 48.1% out of the evaluable cases. In addition, by applying the two-tailed Fisher's exact test, we found no correlations between ROR1 or miR-382 and the molecular-clinical-pathological data or with the overall survival and the progression free survival, calculated on the basis of the chemotherapeutic treatments. Moreover, we found no significant statistical correlation between miR-382 and ROR1.

To conclude, in our project we tried to improve the care of patients affected by lung cancer, firstly by the validation of a new more sensitive real-time kit for plasma and tissue, able to enlarge the number of *EGFR* mutant patients, and secondly by the evaluation of ROR1 expression, a new marker for targeted therapies.

1. INTRODUCTION

1.1) Lung cancer: epidemiology and etiology

Despite the improvements in therapies and in clinical analyses seen in the last decades, lung cancer remains the leading cause of tumor related death worldwide. Indeed it is the neoplastic disease with the highest mortality. Death rates are nearly the same in males and females, accounting respectively for 27% and 26% out of all deceases associated to cancer (Cancer.org. Cancer facts and figures 2016. American Cancer Society; 2016). The mortality from this neoplastic disease exceeds that of breast, prostate and colorectal cancers combined (Miller KD et al, 2016). This is mainly due to the relevant metastatic potential that makes lung cancer difficult to be diagnosed before the advanced stage.

In addition to the high mortality, it is characterized by an incidence that, in both men and women, is the second one after sex related tumors (i.e. prostate and breast cancers) (Travis WD et al, 2004; Siegel R et al, 2015; Travis WD et al, 2016). Lung cancer incidence is higher in men than in women but in the last years male incidence has decreased, mainly thanks to the diffusion of anti-smoking campaigns. Indeed both smoking prevention and smoking cessation can lead to a significant reduction in lung cancer development (as well as in the one of other cancers). In countries characterized by anti-smoking campaigns, lung cancers incidence is declining in men and is reaching a plateau in women (IARC website: https://www.iarc.fr/). In 2013, men highest incidence rates were observed in North America, East Asia, Central-Eastern and Southern Europe (ranging from 48.5 to 56.5 new cases per 100,000 inhabitants). While in less developed countries, the highest rates were seen in West Asia, South Africa and the Caribbean (ranging from 25.7 to 32.2 new cases per 100,000 inhabitants). In women the worldwide incidence rates of lung cancer are lower than those for men and in 2013 the highest rates were seen in North America and in Northern Europe (ranging from 35.8 to 37 new cases per 100,000 inhabitants) (Ridge CA et al, 2013).

The overall survival (OS) for lung cancer is poor: the EUROCARE 5 study reports that, in European countries, lung cancer survival is low with a relative survival (RS) of 39% and 13% at 1 and 5 years after diagnosis, respectively (Francisci S et al, 2015). Recent data from 1999 to 2006 show that women have better survival compared with men across all ages, irrespective of the histologic subtype. Indeed, in 2006, the 5-year survival rate for women with lung cancer was 19% compared to 14% for men (Ridge CA et al, 2013). Moreover, OS is higher for squamous cell carcinoma and adenocarcinomas than for small and large cell carcinoma (Francisci S et al, 2015). Lung cancer carcinogenesis is related to interactions of

both genetic and epigenetic factors leading to an uncontrolled growth of abnormal cells in lung tissue.

In most of the patients the main driver is tobacco smoking, indeed the geographic and temporal patterns reflect tobacco consumption during last decades. In addition to tobacco smoking, there are also other factors (e.g. hormonal imbalance; viruses; genetic factors; genetic individual susceptibility; exposure to asbestos, arsenic, radon, non-tobacco-related polycyclic aromatic hydrocarbons and other environmental aspects) which have been proposed to predispose to lung cancer development. All these factors can cause genetic and global transcriptome changes resulting in cells characterized by aberrant pathways activation that can persist long term leading to dysplasia and clonal patches. Afterward, other additional changes bring to early stage cancer, angiogenesis, metastasis and, at the end, the increasing invasion brings to advanced lung cancer (Herbst RS et al, 2008).

1.2) Lung cancer: classification

Lung cancers classification should be made following 2015 World Health Organization (WHO) guidelines and International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society **(**IASLC/ATS/ERS) recommendations.

Lung cancers, on the basis of morphology, can be subdivided into Non-Small-Cell Lung Cancers (NSCLCs) and Small-Cell Lung Cancers (SCLCs). The percentage of lung cancers classified as NSCLCs is equal to 80-85% (Miller KD et al, 2016). Tissue morphology, analyzed by a pathologist, and immunohistochemistry (IHC) staining allow the classification of NSCLCs as AdenoCarcinoma (AC), Squamous Cell Carcinoma (SCC) or Large Cell Carcinoma (LCC). In particular Thyroid Transcription Factor-1 (TTF-1) positivity by IHC is typical of ACs, p40 or p63 IHC positivity is related to SCCs and no staining pattern is linked to LCCs. Among NSCLCs, AC represents the histotype with the highest rate (60%). SCC and LCC constitute 25% and 15% of all NSCLCs, respectively (Herbst RS et al, 2008). During the last 25 years, the distribution of histological types of NSCLC has changed: in the USA, SCC, which was predominant, decreased and AC has increased in both genders becoming, nowadays, the most diffuse histotype. In Europe, similar trends have occurred in men, while, in women, both SCC and AC are still increasing (Forman D et al, 2013). The new 2015 WHO classification recommends to test by IHC not only tumor resection but also biopsies, with a limited panel of IHC assays (e.g. only TTF-1 and p40 test).

Concerning classification, on the basis of morphological features, pathologists are used to define lung cancer with a grade that describes the differentiation of the tumor: G1 indicates a well differentiated cancer, G2 a moderately differentiated cancer, G3 a poorly differentiated cancer, G4 an undifferentiated cancer and GX a cancer in which grade of differentiation cannot be assessed. The criteria revised in 2015 advise to exclude grade classification in biopsies. Another way to characterize lung cancer is the TNM classification, where T indicates the extent of the primary tumor, N the absence or presence and the extent of regional lymph nodes metastases and M the absence or presence of distant metastases (Travis WD et al, 2004; Goldstraw P et al, 2015; TNM classification for lung cancer proposed by the IASLC: 8th edition). The addition of numbers to these letters indicates the extent of the malignant disease (Table 1.1, 1.2 and 1.3).

T		
T0	No primary tumor	
Tis	Carcinoma in situ (SCC or AC)	
TX	T status cannot be assessed	
T1	Tumor ≤ 3cm	
	$T1a_{(mi)}$	Minimally invasive Adenocarcinoma
	$T1a_{SS}$	Superficially spreading tumor in central airways
	$T1a_{\leq 1}$	Tumor ≤ 1cm
	$T1b_{>1-2}$	Tumor > 1cm but ≤ 2cm
	$T1c_{>2-3}$	Tumor > 2cm but ≤ 3cm
T2	Tumor > 3cm but ≤ 5cm	
	$T2_{Visc\ Pl}$	Tumor involving visceral pleura
	$T2_{Centr}$	Tumor involving main bronchus (not carina), atelectasis to hilum
	$T2a_{>3-4}$	Tumor > 3cm but ≤ 4cm
	$T2b_{>4-5}$	Tumor > 4cm but ≤ 5cm
T3	$T3_{>5-7}$	Tumor extension > 5cm but ≤ 7cm
	$T3_{Inv}$	tumor invading chest wall, pericardium, phrenic nerve
	$T3_{Satell}$	tumor invading separate tumor nodules in the same lobe
	$T3_{>7}$	Tumor > 7cm
T4	$T4_{Inv}$	Tumor invading mediastinum, diaphragm, heart, great vessels, recurrent laryngeal nerve, carina, trachea, esophagus, spine
	$T4_{Ipsi\ Nod}$	Tumor nodules in a different ipsilateral lobe

Table 1.1: Definitions for T descriptor (Detterbeck FC et al, 2017).

N	
N0	No regional node metastasis
NX	N status cannot be assessed
N1	Metastasis in ipsilateral pulmonary or hilar nodes
N2	Metastasis in ipsilateral mediastinal/subcranial nodes
N3	Metastasis in contralateral mediastinal/hilar or supraclavicular nodes

Table 1.2: Definitions for N descriptor (Detterbeck FC et al, 2017).

M		
M0	No distant metastasis	
M1	Tumor with distant metastasis(es)	
	$M1a_{Pl\ Dissem}$	Malignant pleural/pericardial effusion
	$M1a_{Contr\ Nod}$	Pleural/pericardial nodules or separate tumor nodules in a contralateral lobe
	$M1b_{Single}$	Single extrathoracic metastasis
	$M1c_{Multi}$	Multiple extrathoracic metastases (one or more organs)

Table 1.3: Definitions for M descriptor (Detterbeck FC et al, 2017).

Each patient can be described by two classifications, the first one is a pretreatment clinical classification (cTNM) and the second one is a postsurgical histopathological classification (pTNM) (Goldstraw P et al, 2015; TNM classification for lung cancer proposed by the IASLC: 8th edition).

Oncologists can define lung cancer stage applying the TNM classification (Table 1.1, 1.2 and 1.3), indeed the combination between T, N and M descriptors permits the classification of lung tumors in Stage 0, Stage I (i.e. IA; IB), Stage II (i.e. IIA; IIB), Stage III (i.e. IIIA; IIIB; IIIC) and Stage IV (i.e. IV; IVA and IVB) (Detterbeck FC et al, 2017).

1.3) AC: general characteristics and molecular data

As aforementioned, AC is the most common histotype among lung cancers. Patients affected by this histotype are frequently women and Asian. Indeed, AC represents 42% of lung cancers in women and 28% in men. ACs are anatomically characterized by neoplastic cells producing

mucus and organized in different structures, such as beads, ducts or solid clumps (Figure 1.1). Usually, this histotype is identified as a grey nodule with irregular borders that blend in the lung parenchyma located near visceral pleura (Corrin B, 2000). ACs patterns could be clearly present or not clearly present but they can be supported by special stains analyses (e.g. TTF-1 IHC positivity) (Figure 1.1B).

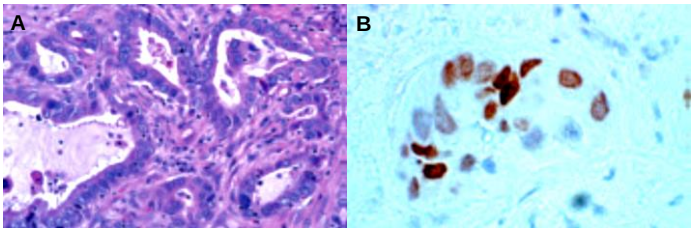

Figure 1.1: A) Hematoxylin and eosin (HE) staining of lung AC cells organized in an acinar structure. The cells exhibit hyperchromatic nuclei in a fibroblastic stroma. **B) Positive immunohistochemical staining for TTF-1 in AC tissue** (IARC/WHO classification; https://ww.iarc.fr/).

AC frequently leads to the development of distant metastases in liver, bone, central nervous system, adrenal glands through blood, and in loco regional lymph nodes through the lymphatic system (Corrin B, 2000). As aforementioned, AC histotype is typically identified by the histochemical marker TTF-1.

In lung cancers, the most important molecular markers are the following genes: *EGFR* (Epidermal Growth Factor Receptor), *KRAS* (Kirsten Rat sarcoma viral oncogene homolog)**,** *BRAF*, *ALK* (Anaplastic Lymphoma receptor tyrosine Kinase), *ROS1* and *HER2* o *ErbB2* (Human Epidermal growth factor Receptor 2) (Herbst RS and Lippman AM, 2007; Bos M et al, 2013; De la Bellacasa PR et al, 2013) (Figure 1.2).

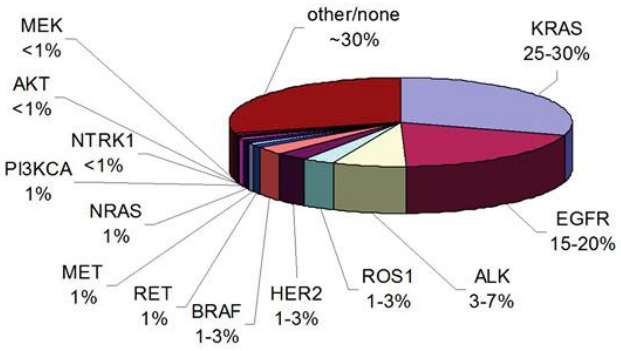

Figure 1.2: Pie chart representing the frequency of oncogenic alterations in lung AC (Gerber DE et al, 2014).

1.3.1 EGFR

The *EGFR* gene is located on the short arm of chromosome 7 (region 7p11) and encodes for a tyrosine kinase (TK) receptor of 170kDa. This transmembrane glycoprotein binds to the epidermal growth factor (EGF) and the binding induces receptor dimerization and tyrosine autophosphorylation leading to the creation of specific binding sites for the protein members of RAS-RAF-MEK or PI3K-Akt-mTOR pathways (Mitsudomi T and Yatabe Y, 2010) (Figure 1.3 and 1.4).

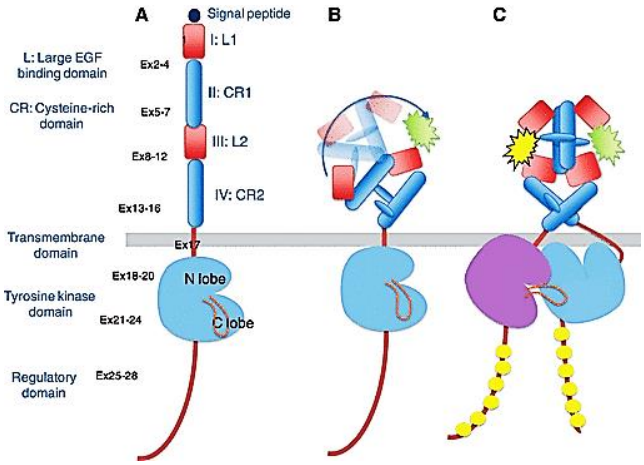

Figure 1.3: EGFR protein: A) Structure; B) Activation; C) Dimerization by ligand binding (Mitsudomi T and Yatabe Y, 2010). Abbreviations: CR, cysteine-rich domain; EGF, epidermal growth factor; ex, exon; L, large EGF binding domain.

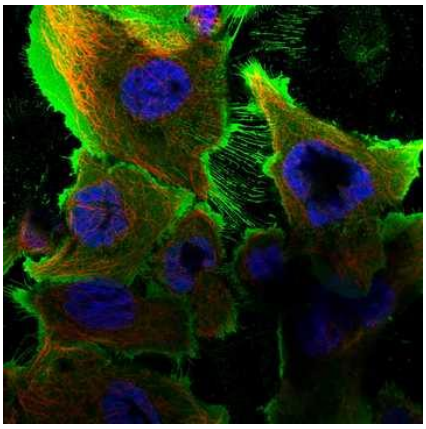

Figure 1.4:Picture of a direct immunofluorescence experiment conducted on a lung cell line and showing the transmembrane localization of EGFR receptor. In this figure EGFR receptor specific antibodies are targeted with green signal, cells nuclei are colored with DAPI (4',6-diamidino-2-phenylindole) staining (blue) and microtubules are colored with the binding of a specific antibody against tubulin (red) (The Human Protein Atlas).

Therefore, ligand binding to the receptor activates one of the two aforementioned pathways related to transcriptional activation of genes involved in cellular proliferation, differentiation, metastasis, angiogenesis and programmed cell-death (Koudelakova V et al, 2013).

EGFR activating alterations are observed in 15-20% of NSCLC among Caucasian patients and in 40% of NSCLC patients among Asians (Banno E et al, 2016; Ko B. et al, 2017). Mutations in *EGFR* are present in 20% of Caucasian patients with AC histotype and are represented by alterations occurring in exons 18, 19, 20 and 21 (Gerber DE et al, 2014). The most widespread alterations of *EGFR* in lung cancers are point mutations (e.g. G719C, G719S, G719A in exon 18; T790M in exon 20; L858R, L861Q, L861R in exon 21), deletions (e.g. E746_A750del, L747_T751del and L747_P753del in exon 19) and insertions (e.g. D770_N771insNPG, D770_N771insSVQ and D770_N771insG in exon 20). Exons 19 and 21 mutations account for 85-90% of all the *EGFR* mutated cases and the incidence of de novo exon 20 insertions is 3.4% (Rosell R et al, 2012; Noronha V et al, 2017). *EGFR* mutations mainly occur in women, non-smokers of Asian origin.

1.3.2 KRAS

The *KRAS* gene is located on the short arm of chromosome 12 (region 12p12) and encodes for a GTPase protein of 21kDa, involved in the intracellular transductional pathways regulated by TK receptors which are related to apoptosis, cell growth and proliferation. KRAS protein activity is influenced by the exchange between GDP (guanosine triphosphate) and GTP (guanosine triphosphate). The binding of GTP to KRAS leads to the activation of this protein

and the recruitment of RAF proteins. KRAS binding to RAF leads to the phosphorylation of MAP2K and MAP2K-2 and, consequently, to the activation of MAPK (Mitogen-Activated Protein Kinase).

In lung cancer, *KRAS* alterations are represented by point mutations, mainly occurring in exon 2 (codons 12-13) and rarely in exon 3 (codons 59-61) or 4 (codon 146) (Koudelakova V et al, 2013). They are identified in 10-40% of AC patients and are generally mutually exclusive with *EGFR* genetic alterations (Gerber DE et al, 2014). *KRAS* mutation occurs predominantly in smoker patients.

1.3.3 BRAF

BRAF gene is situated on the long arm of chromosome 7 (region 7q34) and encodes for a protein of 94 kDa belonging to the RAF family of serine/threonine protein kinases. This protein plays an important role in regulating the MAPK signaling pathway, which affects cell division, differentiation, and secretion.

In lung cancer, *BRAF* alterations are point mutations located in exons 11-15 and they have been found in 2-4% of ACs (Herbst RS et al, 2008; Michaloglou C et al, 2008; Gerber DE et al, 2014). Half of them are represented by the V600E change (exon 15). The remaining 50% of mutations are mainly G469A and G466V (40%) in exon 11 and D594G (10%) in exon 15 (Michaloglou C et al, 2008). *BRAF* mutations mainly occur in heavy smokers.

1.3.4 ALK

ALK gene is located on the short arm of chromosome 2 (region 2p23) and encodes for a TK receptor of 140kDa which belongs to the insulin receptor superfamily. It plays an important role in the development of the brain and exerts its effects on specific neurons in the nervous system.

This gene has been found to be rearranged in NSCLC and in a series of cancers, including anaplastic large cell lymphomas and neuroblastoma (Rikova K et al, 2007). Chromosomal rearrangements are the most common genetic alterations in *ALK* gene, resulting in creation of multiple fusion genes essential for cancer development. The most common rearrangement is the *EML4/ALK* (Echinoderm microtubule-associated protein-like 4/*ALK*) represented by a paracentric inversion between these two genes, both situated on chromosome 2. This inversion characterizes 3-7% of AC cases and it leads to the chimeric *EML4-ALK* gene. The *EML4-ALK* gene is the result of a chromosome rearrangement between the DNA sequence coding for N-terminal portion of EML4 and the sequence coding for ALK TK domain.

Rearrangements involving *EML4* and *ALK* genes are characterized by a huge number of variants; literature reports at least 15 variants of *ALK* inversion, with variant one (exons 1-13), variant two (exons 1-20) and variant three (exons 1-6) being the most common (Figure 1.5). *ALK* gene can be characterized also by rare rearrangements (e.g. *ALK*/*KIF5B*, *ALK*/*TGF*, *ALK*/*KLC1* and *ALK*/*HIP1*), representing only 1% of the rearranged cases (Takeuchi T et al, 2006; Rikova K et al, 2007) (Figure 1.5). *ALK* rearrangements are usually detected in never or light smokers, non-Asians and young women.

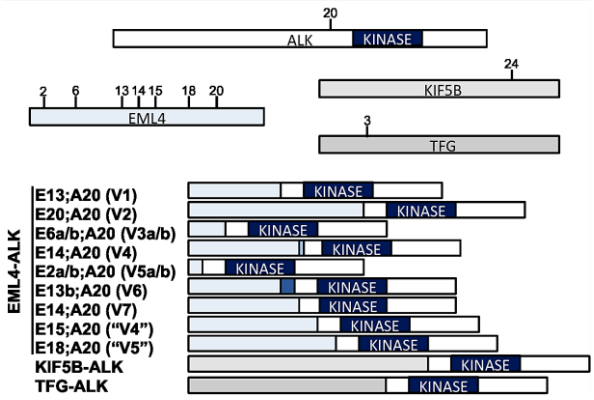

Figure 1.5: Chart representing some ALK fusion variants. This figure reports different variants of ALK/EML4 rearrangements and the rare fusions of ALK with other genes (e.g. KIF5B, TFG) (My Cancer Genome website). Abbreviations: V, variant.

1.3.5 ROS1

ROS1 gene is situated on the long arm of chromosome 6 (region 6q21-22) and encodes for a TK receptor belonging to the insulin receptor superfamily whose molecular weight is 82kDa. ROS1 plays an important role in the regulation of PI3K-Akt-mTOR and RAS-RAF-MAPK pathways, which are related to cell proliferation and differentiation. *ROS1* alterations in cancers are generally characterized by rearrangements and in lung cancer the most common ones are *ROS1*/*FIG*, *ROS1*/*SLC34A2*, *ROS1*/*CD74*, *ROS1*/*EZR*, *ROS1*/*LRIG3*, *ROS1*/*SDC4* and *ROS1*/*TPM3*, cumulatively identified in about 1% of all NSCLC cases and in 1-3% of AC subtype (Gerber DE et al, 2014). These alterations are mutually exclusive with *ALK* rearrangements and are associated with young, female patients (Yoshida A et al, 2013).

1.3.6 HER2

The *HER2* gene is situated on the long arm of chromosome 7 (region 17q21-17q22) and encodes for a protein belonging to the EGFR family. This protein seems not to have a ligand

binding domain and therefore should not bind any growth factor, indeed an HER2 growth factor has not been found yet. However, it tightly binds other ligand-bound EGF receptor family members to form a heterodimer, stabilizing ligand binding and enhancing kinase-mediated activation of downstream signaling pathways, such as those involving RAS-RAF-MEK, PI3K-Akt-mTOR and JAK-STAT axis. The activation of these pathways leads to the promotion of cellular proliferation and angiogenesis (Cappuzzo F et al, 2005; Garrido Castro AC and Felip E, 2013; Martin V et al, 2013).

In primary lung cancer, *HER2* alterations are mainly represented by mutations (usually characterized by small nucleotide insertions) occurring in exon 20 and identified in 1-3% of all the ACs (Gerber DE et al, 2014). *HER2* mutations, like *EGFR*, mainly occur in non-smokers women.

1.4) Therapies

Surgical resection is the first approach in patients affected by early staged AC; in case of advanced stage, tumor resection must be combined with chemotherapy.

Currently, the most diffused treatment for lung cancer is platinum (cis-Pt or carboPt)-based doublet chemotherapy from four to six cycles. The doublet is characterized by the combination of Pt and a couple of cytotoxic molecules (Gemcitabine, Paclitaxel, Docetaxel, Vinorelbine or Pemetrexed) (Cufer T et al, 2013; Shepherd FA et al, 2013). Several studies have compared Pt-based doublets containing these cytotoxic molecules and survival rates have been similar in all the trials. In 2008, the administration of Pt doublet chemotherapy brought to the reduction of the mortality rate of nearly 23% (Cufer T et al, 2013). A clinical trial of 2009 showed that, in patients treated with four cycles of chemotherapy, without progression, a maintenance cure with a single cytotoxic molecule leads to the increase of OS and Progression Free Survival (PFS) (Ciuleanu T et al 2009). Recent data demonstrate that the administration of bevacizumab (Avastin®; F. Hoffmann-La Roche Ltd, Basel, CH), a monoclonal antibody against the VEGF (Vascular Endothelial Growth Factor), in second and third line and in combination with Pt doublet chemotherapy leads to an advantage in lung cancer treatment. Indeed, the ECOG 4599 trial showed that bevacizumab administered in combination with carboPt and Paclitaxel may increase the PFS and OS of nearly two months (Sandler A et al, 2006; Zarogoulidis K et al, 2013).

Radiotherapy can also be recommended before surgery to reduce tumor dimensions and after surgery to reduce the risk of local relapse (Shepherd FA et al, 2013).

1.5) Targeted therapies

Chemotherapies are limited by their lack of specificity and by frequent and potentially severe dose-limiting toxicities. Therefore, there is an urgent need for more effective, better-tolerated treatments that specifically target the process pivotal to tumorigenesis and metastasis. Consequently, in recent years there was a rapid progress in the development of new treatment strategies for advanced NSCLC, in particular with the introduction of molecular targeted therapies (i.e. chemotherapies able to specifically target molecular alterations occurring in a given tumour) (Kaneda H et al, 2013). The most diffused targeted drugs in lung cancers are the small-molecule tyrosine kinase inhibitors (TKIs) acting against *EGFR* and *ALK* alterations. However last years have seen also the birth of phase I, II and III studies based on targeted therapies against rare AC mutations (e.g. alterations in *ROS1*, *HER2*, *BRAF* and *KRAS*) (Tsao AS et al, 2016) (Figure 1.6).

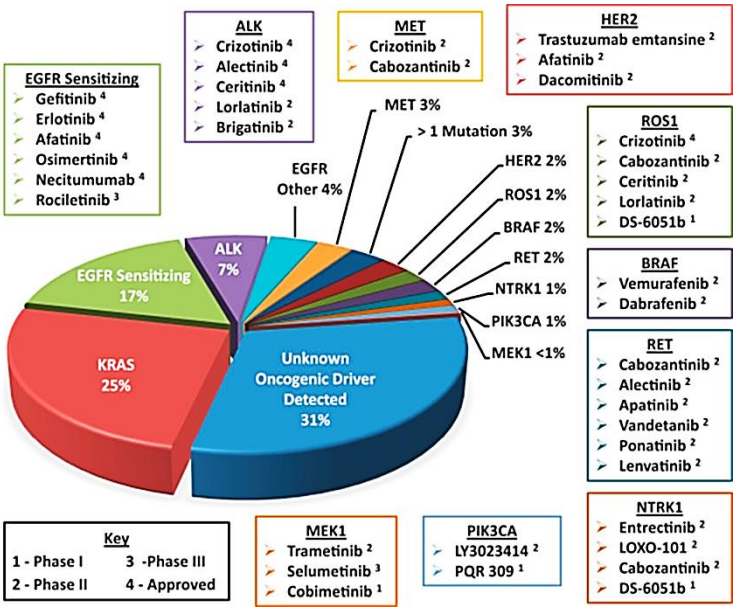

Figure 1.6: Pie chart representing the frequency of molecular alterations in various driver oncogenes in lung AC and the current available drugs against these oncogenic proteins (Tsao AS et al, 2016).

1.5.1 EGFR-targeted therapies

Last years have seen the introduction of targeted therapies against EGFR, the most effective are the TKIs. These drugs specifically target EGFR in lung ACs characterized by activating alterations in this gene.

Other targeted therapies available are, beside EGFR TKIs, cetuximab (Erbitux®; Bristol-Myers Squibb, New York, USA) and panitumumab (Vectibix®; AMGEN, CA, USA), two anti-EGFR antibodies directed against EGFR receptor binding domain. However trials demonstrated that patients affected by lung cancer do not show a clear response to the treatment with anti-EGFR antibodies (Mukohara T et al, 2005; Rossi A et al, 2008).

Consequently, on these bases, molecular therapies against EGFR in lung ACs are mainly characterized by TKIs. In 2009, the IPASS study demonstrated that in patients with *EGFR* mutations the treatment with TKIs gives a better response compared to standard chemotherapy regimens (Mok TS et al, 2009). Current guidelines, such as those of the National Comprehensive Cancer Network (NCCN) and the European Society for Medical Oncology, recommend the testing of *EGFR* mutations in all patients with advanced non-squamous NSCLC before initiation of first-line therapy. This recommendation is based on phase III randomized controlled trials demonstrating improvements in Response rates (RR) and PFS when *EGFR* mutated patients are treated with TKIs, compared to chemotherapy administration (Batson S et al, 2017).

To date, three generations of TKIs against EGFR have been developed and approved for the treatment of mutated patients. The first generation (1^{st}G) TKIs, gefitinib (Iressa®, ZD1839; AstraZeneca, Cambridge, UK) and erlotinib (Tarceva®; Roche), are small molecules acting as reversible inhibitors of EGFR through their competitive binding to the TK intracellular domain in proximity of the ATP/Mg^{2+}(adenosine triphosphate/magnesium^{2+}) binding site. This binding inhibits the TK activity enhanced by activating mutations of *EGFR*. In 2015, in the USA, gefitinib was finally introduced as first-line therapy for patients harbouring mutations in *EGFR*, after a phase IV study demonstrating an objective response rate (ORR) of 70%, a medium PFS of 9.7 months and a medium OS of 19 months (Douillard J et al, 2014). Concerning erlotinib, in 2012 the EURTAC randomised phase III trial demonstrated that this TKI gives a better OS versus standard chemotherapy as first-line treatment for European patients with *EGFR* mutations (Rosell R et al, 2012).

On these bases, in 2013, in Europe and USA erlotinib was approved as first-line therapy in the metastatic setting after the results of EURTAC trial (Rosell R et al, 2012).

The most diffused genetic alterations in *EGFR* correlated with sensitivity to 1stG TKIs are all the deletions occurring in exon 19 (e.g. E746_A750del), point mutations in exon 21 (e.g. L858R, L861R and L861Q), exon 18 (e.g. G719A, G719C and G719S) and exon 20 (e.g. V765A, T783A) (Banno E et al, 2016; Saad N et al, 2017). There is little evidence whether different deletions in exon 19 are associated with different therapeutic response and clinical outcomes under TKI therapies. In preliminary studies the TKIs versus EGFR seem to be more effective against NSCLCs with E746del than those with E747del (Lee VH et al, 2013; Kaneda T et al, 2014). The efficacy of 1stG TKI in patients harbouring the S768I mutation (in exon 20) is till debated (Yang JC et al, 2015; Banno E et al, 2016; Chen K et al, 2017). On the contrary, all the rearrangements (e.g. D770_N771insNPG, D770_N771insSVQ, D770_N771insG) and some point mutations (T790M, V796L, N771T) in exon 20 are associated with resistance to 1stG TKIs (Figure 1.7). Last alterations are classified as primary resistance, happening before the therapy. Furthermore there are also mechanisms of secondary resistance happening after the administration of the treatment. Indeed, despite the fact that 60-80% of patients show an initial response to EGFR TKIs, in the majority of cases there is the development of disease progression related to the acquired resistance happening within one to two years after treatment initiation (Jänne PA et al, 2015).

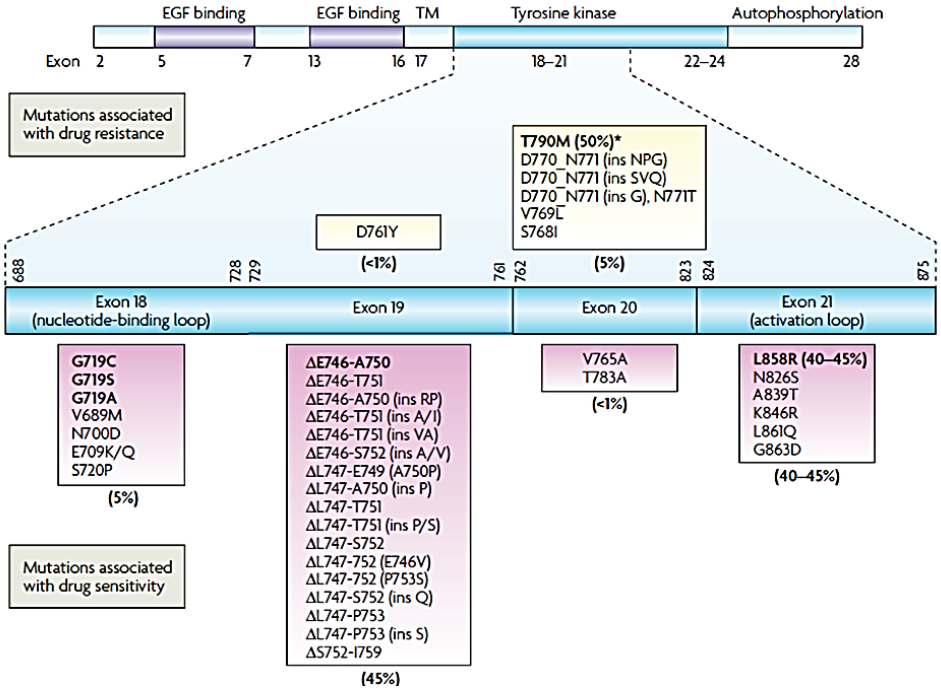

Figure 1.7: *EGFR* mutations associated with response or resistance to gefitinib and erlotinib in lung AC (Sharma VS et al, 2007).

The most common molecular mechanism inducing secondary resistance is the mutation T790M in exon 20. Indeed, T790M is found in 50-70% of ACs resistant to TKIs (Murray S et al, 2012; Santarpia M et al, 2017). Other mechanisms of resistance are: *HER2* and *MET* gene amplification, transformation to SCC, *PIK3CA* gene mutation and activating mutations in *RAS* or *BRAF* genes; representing respectively 12-15%, 5-11%, 5%, 1% and 0.5 % of resistant cases (Yun C-H et al, 2008; Takezawa K et al, 2012) (Figure 1.8).

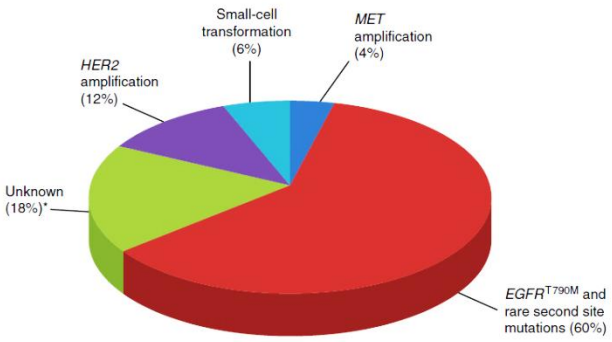

Figure 1.8: Mechanisms of resistance to EGFR TKIs. *EGFR* T790M mutation is the major cause of TKIs resistance but many other mechanisms could be present even if in minor rates (Takezawa K et al, 2012).

To overcome TKIs resistance, some trials have studied the efficacy of the treatment with second generation (2^{nd}G) inhibitors; such as afatinib (Gilotrif® or Giotrif®; Boehringer Ingelheim, Germany), dacomitinib (PF-00299804; Pfizer, New York, USA) and neratinib (HKI-272; Puma Biotechnology, Los Angeles, CA, USA) (Kaneda H et al, 2013; Saad N et al, 2017). Second generation TKIs can overcome resistance to 1^{st}G TKIs by an irreversible binding to EGFR TK domain; in addition, uncommon EGFR mutations (i.e. alterations representing a little part of EGFR mutated patients) like L861Q, S768I and mutations in *EGFR* exon 18 are more sensitive to 2^{nd}G TKIs than 1^{st}G TKIs (Yang JC et al, 2015). Afatinib has been studied and compared to the other treatments (1^{st}G TKIs or different chemotherapy regiments) in the LUX-Lung trials and, after the demonstration of its efficacy, in 2013 it was approved by the U.S. Food and Drug Administration (FDA) as first-line treatment for metastatic *EGFR*-mutated NSCLC (Sequist LV et al, 2013; Wu Y-L et al, 2014; Ko B et al, 2017). Recent studies report that 2^{nd}G TKIs efficacy can be compromised by acquired mutation in the cysteine 797 residue (C797) of EGFR receptor, the position of the irreversible covalent bond between these molecular drugs and EGFR (Chong CR et al, 2013;

Gainor JF et al, 2013). In addition, 2ndG TKIs have a limited effect against T790M because their high rate of toxicities limit the availability to administer doses sufficient to effectively inhibit this mutation (Santarpia M et al, 2017).

To solve this problem, there was the introduction of third generation (3rdG) TKIs, mutant selective irreversible inhibitors. Third generation TKIs, such as osimertinib (Tagrisso®, AZD9291; Astra Zeneca) and rociletinib (CO-1686; Clovis Oncology, CO, USA) are effective in patients with NSCLC harbouring specific mutations. Osimertinib is an irreversible TKI targeting the T790M mutation with a covalent binding to C797 in the EGFR ATP binding site. Phase I AURA trial tested the safety and efficacy of osimertinib as second line in NSCLC patients showing progression after previous treatment with EGFR TKI administration (Janne PA et al, 2015). This study demonstrated that, in the subgroup of T790M-positive patients, osimertinib showed high activity with an ORR of 61% and a median PFS of 9.6 months. In comparison, in patients with no detectable *EGFR* T790M, the ORR was 21% and PFS was 2.8 months. Subsequently, the Phase II AURA trial, testing advanced or metastatic NSCLC patients with T790M, confirmed the high activity of osimertinib at the dose of 80 mg daily, single administration (Santarpia et al, 2017). On these bases, in 2015, the FDA approved the treatment with osimertinib in patients with a T790M-positive NSCLC whose disease had progressed after the administration of other TKIs. The approval by European Commission was received in 2016 (Mitsudomi T et al, 2015).

Rociletinib is another covalent inhibitor of *EGFR* mutations that does not affect exon 20 insertions but inhibits exon 19 deletions, L858R, and T790M mutations. Patients enrollment for ongoing rociletinib studies, including the Phase III TIGER-3 trial, is finished but the manufacturer has withdrawn its application for regulatory approval by FDA and EMA (European Medicines Agency). Rociletinib and others 3rdG TKIs are still object of severe clinical trials. In patients showing disease progression and treated with 1stG or 2ndG TKIs, the NCCN panel recommends to continue the same TKI with local treatment if there is local progression and to add chemotherapy to TKI or switch to 3rdG TKIs in case of T790M mutation (NCCN website: http://www.nccn.org/).

The efficacy of 3rdG TKIs on T790M-mutant patients can be compromised by new acquired mutations, like alterations occurring in the codon 797. Indeed, the most important mutation that is correlated with 3rdG TKIs resistance is C797S (Figure 1.9). This missense mutation brings to the loss of the potential for covalent bond formation in 797 position in the kinase binding site (Niederst MJ et al, 2015; Song H-N et al, 2016). A study based on cell lines showed that the position of C797S and T790M mutations on *EGFR* alleles influences the

response to molecular targeted therapies. Indeed, when C797S and T790M are in trans (on different alleles) there could be a positive response to a combined 1stG/3rdG TKIs therapy; in contrast, when these mutations are in cis (on the same allele) there is no response to any EGFR TKI treatment alone and in combination (Niederst MJ et al, 2015). The importance of the configuration of C797S and T790M in the response to therapy has been reported also in patients (Sacher AG et al, 2016).

Furthermore, a not negligible fraction of T790M mutated patients may acquire resistance to 3rdG TKIs without harbouring mutations in the codon 797. Ongoing studies havedemonstrated that, in these cases, the resistance to 3rdG TKIs could be due to other *EGFR*alterations that interfere with the drug binding (e.g. L798I,L718Q, L692V,L844V, E709K;*EGFR*gene amplification), to alterations in other genes that activate by pass pathways (e.g. *HER2*, *MET*, *FGFR1* gene amplification; *PIK3CA* and *RAS/MAPK* pathway activating mutations; *PTEN* deletion; *IGFBP3* loss) or to phenotypic alterations (e.g. SCLC transformation; epithelial mesenchymal transition) (Walter AO et al, 2013; Kim TM et al, 2015; Piotrowska Z et al, 2015; Planchard D et al, 2015; Chabon JJ et al, 2016; Ortiz-Cuaran S et al, 2016; Park JH et al, 2016) (Figure 1.9). In addition, 3rdG TKIs resistance could be due to the reduction or disappearance of T790M clones determined by the selective pressure caused by the treatment with 3rdG TKIs (Minari R et al, 2016).

Osimertinib resistance due to T790M reduction or disappearance and the resistance due to alterations in genes of the bypass pathways can be overcome using existing methods like exchange to/addition of a 1stG TKI or concurrent combination therapy of an inhibiting alternative pathway (Uchibori K et al, 2017).

To date, we have no clinically available strategy to care patients characterized by C797S/T790M/activating mutation (triple-mutation). However a study reports that brigatinib (AP26113; Ariad Pharmaceuticals, MA, USA), an ALK-TKI in clinical development, could have efficacy in the care of triple mutated cases (Uchibori K et al, 2017). Furthermore, the efficacy of brigatinib seems to be enhanced markedly by its combination with anti-EGFR antibodies (cetuximab and panitumumab) (Uchibori K et al, 2017). In the future, fourth generation (4thG) TKIs will be developed. Indeed, in 2016 the combination of EAI045 (Chemscene, New York, USA), an experimental 4thG TKI, with cetuximab was effective in mouse models of lung cancer driven by *EGFR* L858R/T790M and by *EGFR* L858R/T790M/C797S (Jia Y et al, 2016).

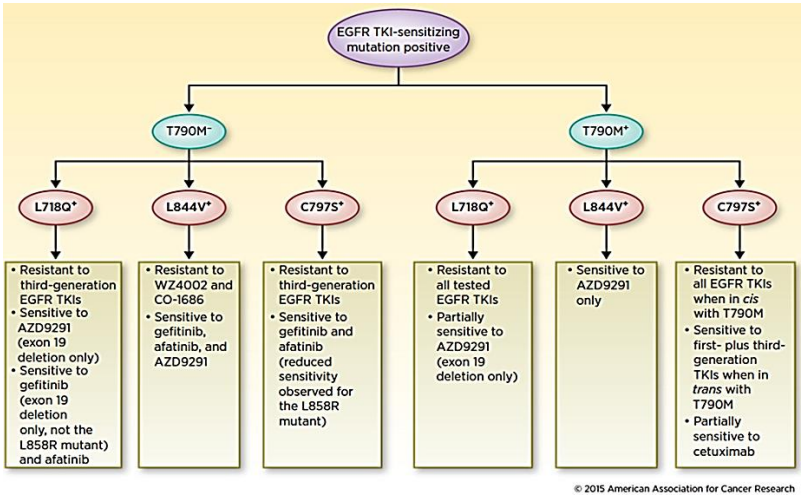

Figure 1.9: Schematic diagram showing sensitivity of cells with primary EGFR TKI-sensitizing mutations (in purple) in the presence or absence of the secondary *EGFR* T790M mutation (in turquoise) and with different tertiary mutations (in red) (Ayeni D et al, 2015). Abbreviations: TKIs, tyrosine kinase inhibitor.

Beside monotherapies based on TKIs administration, recent clinical trials are investigating the effect of combination therapies, like the aforementioned brigatinib plus anti-EGFR antibodies treatment, for highly resistant patients. Recently, BELIEF Phase II trial investigated the combination of erlotinib plus bevacizumab and demonstrated a prolonged medium PFS of 16 months compared with the 9.7 months for erlotinib monotherapy in T790M mutated patients (Seto T et al, 2014). In support of these data, preclinical evidence suggests that erlotinib resistance may be associated with a rise of VEGF levels. In addition, a study assessed gefitinib plus chemotherapy and demonstrated an advantage for combination therapy over gefitinib monotherapy in *EGFR* mutant cases (Batson S et al, 2017). To conclude, in another Phase II trial the combination of afatinib with cetuximab resulted in a RR of 30% and a medium PFS of 4.7 months in heavily pretreated patients (Janjigian YY et al, 2014).

1.5.2 ALK and ROS1-targeted therapies

Other molecular therapies for NSCLC treatment are the TKIs acting in *ALK*-rearranged patients. The first targeted therapy developed against *ALK* alterations was crizotinib (Xalkori®; Pfizer), a dual ALK and MET TKI able to inhibit ALK through the interaction with its ATP binding site, leading to the block of downstream molecular pathways (Bergethon K et al, 2012). PROFILE 1001, a phase I trial testing crizotinib in advanced-stage *ALK* positive NSCLC, demonstrated an ORR equal to 57% and a PFS equal to 72%. In addition, PROFILE 1005, a phase II study that enrolled patients with advanced, previously treated,

ALK-positive NSCLC, showed an ORR of 60% and a PFS of 8.1 months after crizotinib administration (Kim DW et al, 2012). Later, two phase III trials, PROFILE 1007 and PROFILE 1014 demonstrated the better efficacy of crizotinib compared to chemotherapy in patients with *ALK* rearrangements. More in detail, crizotinib administration showed a PFS of 7.7 months compared to 3 months in patients treated with standard chemotherapy in PROFILE 1007, and a PFS of 10.9 months compared to the 7 months of the chemotherapy arm in PROFILE 1014 (Leprieur EG et al, 2016; Novello S et al, 2016).

Crizotinib was approved in 2011 by FDA and in 2012 by EMA for the clinical care of previously treated *ALK*-positive NSCLC. In 2013, crizotinib received approval by FDA also as first-line treatment for patients with *ALK*-positive NSCLC and in 2015 the EMA approved the same indication.

Unluckily patients treated with crizotinib can experience disease progression after 8-12 months through primary or acquired resistance. Various mechanism of resistance have been identified, including secondary *ALK* mutations, *ALK* fusion gene amplifications and alterations in alternative signaling pathways (e.g. *KIT* amplification, *EGFR* or *KRAS* mutations) (Doebele RC et al, 2012; Katayama R et al, 2012). In 30% of cases, the resistance to crizotinib is associated with a secondary mutation in the kinase domain of ALK that interferes with the drug binding or ATP affinity (Passaro A et al, 2016). Several resistance *ALK* mutations have been identified but the most common are L1196M and G1269. Other, rarer *ALK* mutations leading to crizotinib resistance are: C1156Y, F1174L, L1152R, S1206Y, I1171T, V1180L, D1203N (Toyokawa G and Seto T, 2015).

Recently, new 2ndG ALK TKIs have been developed to overcome resistance to crizotinib, such as ceritinib (Zykadia®; Novartis, CH and USA) and alectinib (Alecensa®; Roche).

In preclinical models, ceritinib is nearly twenty times more efficient than crizotinib in both crizotinib-sensitive and crizotinib-resistant tumors. Ceritinib was evaluated in ASCEND-1 trial, a phase I trial including patients with *ALK*-positive NSCLC, in both ALK inhibitor naïve and ALK inhibitor pretreated cases. The ORR was 72% in ALK inhibitor naïve patients compared with 56% in crizotinib pretreated patients, in addition the mean PFS was 18.4 months and 6.9 months respectively (Kim DW et al, 2016). In 2014 on these results FDA approved ceritinib for patients with advanced or metastatic *ALK*-positive NSCLC progressing to crizotinib.

Alectinib is another ALK TKI developed to overcome crizotinib resistance. This drug is efficient in crizotinib naïve and in crizotinib-resistant *ALK*-positive NSCLC, as recently demonstrated. Two phase I/II trials were conducted in order to evaluate the efficacy and

safety of this drug. The first one, named AF-001JP, was conducted in Japan on ALK-positive, crizotinib naïve NSCLC. AF-001JP in phase II showed an ORR of 93.5% and a PFS at twelve months of 83%. The second trial, AF-002JG, was conducted in USA on patients with *ALK*-positive NSCLC who progressed on or were intolerant to crizotinib. This study in phase II demonstrated an ORR to alectinib equal to 55%. Furthermore, ALEX and J-ALEX trials demonstrated the superiority of alectinib over crizotinib as first-line treatment in patients with *ALK*-positive advanced NSCLC and not treated with ALK inhibitor (Ou SH et al, 2016; Shaw AT et al, 2016; Nokihara et al, 2016). On these bases, alectinib was approved by FDA in 2015 for the treatment of patients with advanced *ALK*-positive NSCLC whose disease worsened after, or who could not tolerate, the administration of crizotinib. In addition, it got a conditional approval by the EMA in February 2017 for the same indication. This means that additional studies are awaited to confirm a positive benefit-risk ratio.

Some patients can develop secondary *ALK* mutations bringing to resistance to ceritinib (e.g. G1202R, F1174C and F1174V) or to alectinib (e.g. G1202R, V1180L, I1171T and I1171N) (Passaro A et al, 2016). In particular, the G1202R mutation increases significantly after treatment with 2ndG ALK TKIs and seems to induce resistance to both crizotinib and 2ndG ALK TKIs. Moreover, mutations affecting the codon 1171 bring to resistance to alectinib but are sensitive to ceritinib.

In order to overcome 2ndG ALK TKI resistance, last year has seen the development of 3rdG ALK TKIs: brigatinib (Alunbrig^{TM}, AP26113; ARIAD) and lorlatinib (PF-6463922; Pfizer).

Brigatinib overcomes crizotinib and 2ndG ALK TKIs resistance showing activity especially against L1196M and G1202R mutations. In May 2017, the FDA approved brigatinib for patients with metastatic NSCLC characterized by alterations in *ALK* gene and whose cancer has progressed during their initial therapy (Passaro A et al, 2016).

Lorlatinib is a 3rdG ALK TKI that is nearly 10-fold more effective against wild-type (wt) *EML4-ALK* and nearly 40-fold more active against the L1196M point mutation compared to crizotinib. Preliminary data show an ORR of 50% in patients with *ALK* rearrangement treated with this drug (Passaro A et al, 2016).

Some ALK TKIs are useful against *ROS1* alteration as well: indeed, patients harbouring *ROS1* rearrangement and treated with crizotinib demonstrated an ORR of 57% and a disease control rate (DCR) of 80% after two months (Shaw AT et al, 2016). Crizotinib is a more potent inhibitor of ROS1 than of ALK, leading to more effective target inhibition and more durable response (Gainor JF and Shaw AT, 2013; Shaw AT et al, 2016). In 2016, the FDA

approved crizotinib for the treatment of patients with *ROS1*-positive, advanced NSCLC (Zhao Z et al, 2017).

Furthermore, ceritinib, brigatinib and lorlatinib can also be used against *ROS1* alterations and could be useful when a secondary resistance to 1stG or 2ndG ROS1 TKIs appears (Tsao AS et al, 2016). In addition, MET inhibitors could be effective against *ROS1* rearrangements as well: this seems to be the case of foretinib (GSK1363089, Exelixis) and cabozantinib (Cabometyx®, Exelixis), that are currently studied in undergoing clinical trials for the treatment of prostate, bladder, ovarian, brain, melanoma, breast, pancreatic, hepatocellular and lung cancers (Davare MA et al, 2013; Gainor JF and Shaw AT, 2013).

1.5.3 HER2, BRAF and KRAS-targeted therapies

Besides EGFR and ALK TKIs, researchers are going to investigate if there could be other molecular markers associated with specific targeted therapies that may help the care of NSCLC patients. For example, HER2 is an important molecular marker in gastric and breast carcinomas. In these cancers, *HER2* molecular alterations are associated with sensibility to trastuzumab (Herceptin®; Genentech, CA, USA), a monoclonal antibody that inhibits HER2 through the binding of the extracellular receptor domain. After the discovery of trastuzumab, a plethora of clinical trials demonstrated that a novel antibody–drug conjugate named trastuzumab emtansine (T-DM1, Kadcyla®; Roche) is more efficient in breast cancer than trastuzumab single agent (Hurvitz SA et al, 2013). T-DM1, made of trastuzumab and a potent cytotoxic drug connected via a stable linker to the anti-HER2 antibody, was approved by FDA on 2013 for breast cancers resistant to trastuzumab and characterized by *HER2* gene amplification.

The first clinical trials concerning HER2 targeted therapy in lung cancers were conducted in 2005 but they did not show a better OS in patients mutated in *HER2* and treated with trastuzumab (Cappuzzo F et al, 2005).

However patients with lung cancer and *HER2* mutations seem to give response also to the TKIs against EGFR (e.g. afatinib, dacomitinib, neratinib). Indeed, in 2015, a group of *HER2* mutated patients extracted from the European EUHER2 cohort and treated with chemotherapy and molecular drugs (afatinib, dacomitinib, neratinib or trastuzumab) showed good response to targeted therapies. In particular in these *HER2* mutant patients, the ORR and PFS after trastuzumab administration were 50.9% and 4.8 months respectively (Mazières J et al, 2016).

To date, in NSCLC, the FDA has not approved yet the targeted therapies against HER2 because there is the need of more data on this issue.

Besides HER2, also BRAF could become another molecular marker in NSCLC. Indeed, dabrafenib (Tafinlar®; Novartis Pharma) and vemurafenib (Zelboraf®; Genentech), are BRAF inhibitors that are currently administered to *BRAF*-mutant melanomas with great success (Spagnolo F et al, 2015). In 2011, the FDA approved vemurafenib for the treatment of patients with unresectable or metastatic melanoma characterized by *BRAF* V600E mutation. In addition dabrafenib was approved by FDA in 2013 as a single agent for treatment of *BRAF* V600E mutation-positive unresectable or metastatic melanoma. Furthermore, more recently, some studies demonstrated that, in melanoma, the combination of dabrafenib and trametinib (Mekinist®; GlaxoSmithKline, UK and USA), a MEK inhibitor, is more effective in *BRAF* mutated patients compared to the administration of dabrafenib as a single agent (Paik PK et al, 2011). In 2014, on these bases, the FDA granted accelerated approval for the combined use of trametinib and dabrafenib in order to treat patients with unresectable or metastatic melanoma with *BRAF* V600E or V600K mutation.

These drugs seem to be associated with an increment of OS in *BRAF*-mutant NSCLCs as well, especially in patients harbouring the V600E mutation. Following these promising preliminary data, the efficacy of targeted therapies against *BRAF* mutation in NSCLC has been evaluated in some studies (Paik P K et al, 2011; Gautschi O et al, 2013). Experiments on cell lines described that vemurafenib is effective in *BRAF* V600E mutated cells (HCC364) and not effective in non-*V600E* BRAF mutated cells (H1755), conversely trametinib is effective in both conditions. In addition the combination of vemurafenib and trametinib better promotes apoptosis than trametinib alone in HCC364 and in H1755 cell lines (Joshi M et al, 2015). In 2015, a retrospective study on a cohort of patients affected by lung AC, harbouring *BRAF* mutations and extracted from the EURAF cohort, demonstrated an ORR equal to 53% and a DCR equal to 85% (Gautschi O et al, 2015). In 2017, the FDA approved the combination of dabrafenib and trametinib for the treatment of *BRAF* V600E mutation-positive metastatic NSCLC (National Cancer Institute, National Institutes of Health website: https://www.cancer.gov/about-cancer).

Finally also KRAS could be a potential, new molecular marker in lung AC. The prognostic and predictive role of KRAS is controversial, even though the majority of USA studies indicate KRAS as a negative predictive marker of response to EGFR TKIs (Eberhard DA et al, 2005; Pao W et al, 2005; Massarelli E et al, 2007; Mao C et al, 2010; Garrido-Castro AC and Felip E, 2013). In addition, *KRAS* mutations are a negative predictive marker of response to systemic chemotherapy; indeed AC patients characterized by *KRAS* alterations have a worse PFS and a worse RR after chemotherapy administration compared to the cases with a

KRAS wt sequence (Wei P et al, 2016). However another work proposed *KRAS* mutations as a factor able to sensitize tumors to pemetrexed (Moran DM et al, 2014).

Nowadays, no targeted therapies against KRAS have been approved, but in a phase II clinical trial of 2013, selumetinib (AZD6244, AstraZeneca), a MEK inhibitor, in combination with docetaxel showed preliminary promising results in previously treated advanced NSCLC (Jänne PA et al, 2013). On this basis, new ongoing studies are testing other MEK inhibitors efficacy in *KRAS* mutated patients.

Moreover, KRAS is associated to a worst prognosis. Indeed the majority of studies demonstrate that *KRAS* mutations are an independent predictor of poor prognosis in patients with advanced NSCLC (Sun JM et al, 2013).

1.6) Liquid biopsies testing

Last years have seen the introduction of liquid biopsies testing, a methodology that is based on the analyses of non-solid biological sample (e.g. blood, urine, saliva and cerebral spinal fluid). One of the most important applications of liquid biopsies testing is the characterization of molecular markers in fluids from patients affected by cancer (Crowley E et al, 2013; Diaz LA and Bardelli A, 2014).

1.6.1 Circulating tumor DNA (ctDNA)

The majority of liquid biospies analyses are conducted on plasma and they are based on the characterization of the small portion of cell-free DNA (cfDNA) represented by circulating tumor DNA (ctDNA). ctDNA represents less than 0.5% of the cfDNA found in the blood and it is released in the blood stream from apoptotic, necrotic or living tumor cells (Malapelle U et al, 2016). In addition, it can be produced by circulating tumor cells (CTCs), cells derived from the primary tumor or metastasis that can be found in the blood (Cheng F et al, 2016). The fragment size of ctDNA is still undetermined because it depends on the cellular process causing its release in the circulation. However, it is esteemed that ctDNA size is minor than cfDNA dimensions (i.e. around 166bp) (Sholl LM et al, 2016; Sorber L et al, 2016). ctDNA half-life is less than two hours and its quantity correlates with tumor volume (Bidard FC et al, 2014). Many studies highlight an increase in ctDNA levels when the tumor disease progresses and a decline in case of resective surgery and/or successful medical therapy (Bettegowda C et al, 2014; Diaz LA and Bardelli A, 2014; Pereira E et al, 2015; Sorber L et al, 2016).

The quantity of detectable ctDNA is variable on the basis of the tumor type (Bettegowda C et al, 2014). The tumors or metastases producing the highest quantity of ctDNA originate in lung, bladder, colorectum, stomach, esophagus and ovaries. On the contrary, tumor and metastasis involving brain and bones generally do not produce detectable ctDNA (Bettegowda C et al, 2014); because of the presence of the ematoencefalic barrier and the poor vascularization of the bones.

1.6.2 Liquid biopsies applications and ctDNA testing

ctDNA shows the characteristic mutations of the corresponding primary tumor or metastasis (Sholl LM et al, 2016). Recently, on this basis,ctDNA analyses brought to the development of many diagnostic and treatment applications. Firstly, ctDNA analyses can be used, after and during the treatment, to detect the presence of resistance mechanisms, to predict response to the therapy or to monitor cancer progression (Figure 1.10). Secondly, ctDNA analyses can also be used for early cancer detection, subtyping, prognosis or to detect mutations in patients in which tumor biopsy is not feasible (Diaz LA and Bardelli A, 2014; Cheng F et al, 2016).

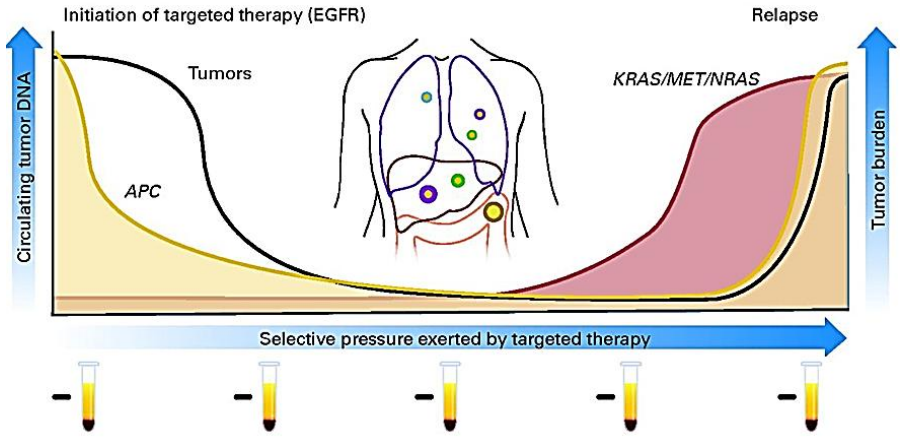

Figure 1.10: Detection of tumor-specific DNA mutations in the blood of patients to monitor response and relapse to targeted therapies. This figure represents a patient with metastatic colorectal cancer. After treatment with anti EGFR monoclonal antibody, the patient experiences a clinical response and has a corresponding decrease in *APC* mutation level, further indicating a decrease in tumor burden. Continuous monitoring of plasma DNA shows the emergence of *KRAS* and *NRAS* mutations and/or *MET* amplification, indicative of the emergence of multiple different resistance clones. Interestingly cfDNA quantity decreases when there is response to therapy and increases when clinical resistance happens (Diaz LA and Bardelli A, 2014). Abbreviations: *APC*, adenomatous polyposis coli gene; MET, tyrosine-protein kinase Met; NRAS, neuroblastoma RAS viral oncogene homolog.

The methodologies for ctDNA testing must be optimized for high sensitivity because the genomic material from liquid biopsies is highly fragmented (Sholl LM et al, 2016; Sorber L et al, 2016). Several assays with high sensitivity have been developed in order to analyze ctDNA for the presence of particular biomarkers. The main techniques are: real-time quantitative polymerase chain reaction (real-time qPCR), droplet digital polymerase chain reaction (ddPCR), beads emulsion amplification and magnetics quantitative polymerase chain reaction (BEAMing qPCR) and next-generation sequencing (NGS) (Sorber L et al, 2016). To date, EMA and FDA do not specify which is the better methodology to analyze ctDNA. However, the EMA recommends to apply a reliable test with high sensitivity (EMA website: http://www.ema.europa.eu).

Liquid biopsies are less invasive and better reflect tumor heterogeneity than tissue biopsies and several studies on mutations detection in coupled peripheral blood and tissues samples demonstrated a good correlation of the results obtained with this two kind of materials (Bai H et al, 2009; Rosell R et al, 2009; Goto K et al, 2012; Yam I et al, 2012).

1.6.3 ctDNA in lung cancer

The first studies concerning cfDNA testing in plasma obtained from patients affected by tumor were based on the characterization of mutations in lung primary tumor and metastasis (Rosell R and Karachaliou N, 2016; Sholl LM et al, 2016).

Initially, in NSCLC, analyses of plasma were focused on the detection of EGFR mutations associated to the resistance to EGFR-targeted TKIs (Rosell R and Karachaliou N, 2016; Zheng D et al, 2016; Zhu Y et al, 2017). In particular, the majority of analyses for the detection of EGFR TKIs resistance are based on the T790M characterization because 50-60% of secondary resistance to TKIs is due to this mutation (Rosell R and Karachaliou N, 2016; Zheng D et al, 2016).

Recent guidelines consider equal the analyses on plasma and on tissue for the detection of the T790M and therefore recommend to perform plasma analyses for the screening of this mutation before tissue re-biopsy in patients showing relapse after treatment with EGFR TKIs (Oxnard GR et al, 2016). In case of detection of the T790M mutation in plasma, the patient can be directly adressed to the administration of osimertinib whereas in case of not detection of T790M in plasma, clinicians must perform at least a tissue biopsy, on which the molecular characterization must be carried on. In case of the presence of T790M mutation in the tissue biopsy the patient can be treated with osimertinib, in contrast absence of the T790M mutation

also in the tissue specimen addresses the patient to other chemotherapies not involving osimertinib. This scheme is reported in Figure 1.11.

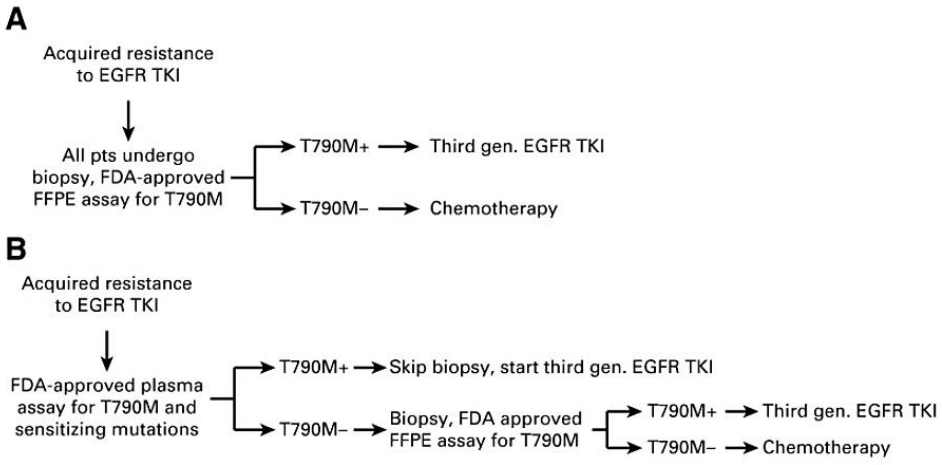

Figure 1.11: Proposed paradigm of plasma genotyping for *EGFR* T790M mutation in patients with acquired resistance to EGFR TKIs. A) In the conventional paradigm, all patients undergo a re-biopsy for T790M genotyping. B) Recommended new paradigm in which plasma genotyping for T790M is used as a screening test before tissue biopsy. Tissue biopsy will be done only in patients with no T790M detected in plasma (Oxnard GR et al, 2016). Abbreviations: FDA, food and drug administration; TKI, tyrosine kinase inhibitor.

As regards the techniques to be used to perform the molecular characterization, in 2016 the FDA approved cobas®, a real-time polymerase chain reaction (real-time PCR) assay developed by Roche, for the analysis of T790M mutation in plasma. However, other technologies can be used as well.

Liquid biopsies are less invasive than tissue biopsies so in the future this feature will permit continuative molecular monitoring of lung cancer in order to detect, beside the T790M mutation, the appearance of other mutations related to resistance to 1stG, 2ndG and 3rdG EGFR TKIs (e.g. C797S) (Rosell R and Karachaliou N, 2016) (Figure 1.12).

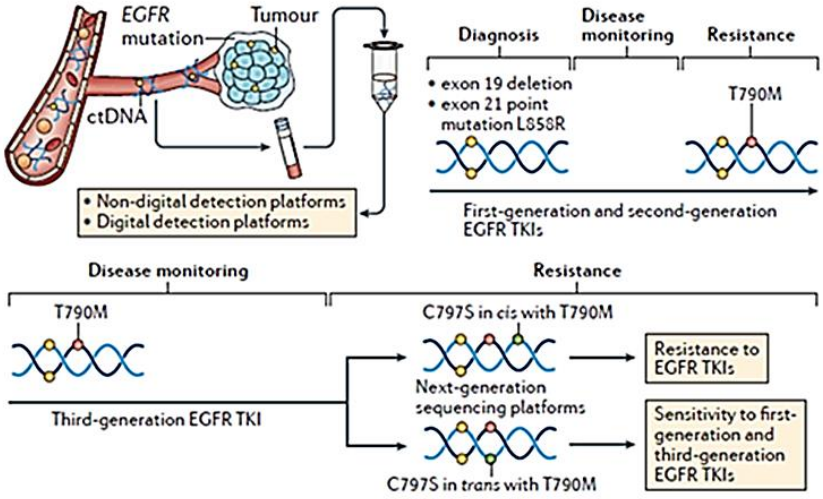

Figure 1.12: Analyses of ctDNA in liquid biopsies for disease monitoring and for detection of mutations involved in TKIs resistance (e.g. T790M and C797S) (Rosell R and Karachaliou N, 2016). Abbreviations: ctDNA, circulating tumor DNA; EGFR, epidermal growth factor receptor; TKI, tyrosine kinase inhibitor.

1.7) ROR1 and miR-382 in cancer

The receptor tyrosine kinase-like orphan receptors (ROR1 and ROR2) are transmembrane proteins of the receptor TK family. ROR proteins display a common structure: an extracellular domain consisting of an immunoglobulin-like (Ig-like) motif, a Cysteine-Rich frizzled Domain (CRD), a Kringle domain and an intracellular domain characterized by a TK domain, a Proline-Rich frizzled Domain (PRD) and a Serine/Threonine-Rich Domain (S/TRD1 o 2) (Figure 1.13).

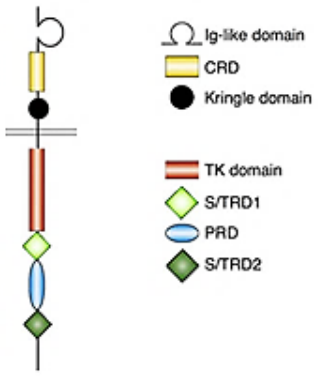

Figure 1.13: Structure of ROR proteins. The extracellular portion contains the Ig-like domain, the CRD (Cysteine-Rich frizzled Domain) and the Kringle domain. The intracellular portion contains the TK domain (Tyrosine Kinase domain), PRD (Proline-Rich frizzled Domain) and S/TRD (Serine/Threonine-Rich Domain) (Rebagay G et al, 2012).

ROR proteins are named “orphan receptors” because their endogenous ligand has not been discovered yet. However, recent data demonstrate that a putative ligand of ROR1 could be WNT, a signaling protein. Indeed, real-time qPCR experiments demonstrate that ROR receptors are highly expressed in breast cancer cell lines when also WNT proteins can be detected in these cells (Klemm F et al, 2011; Anastas JN et al, 2013).

ROR receptors are expressed at high levels during embryo development playing an important role in skeletal and neural organogenesis but they are not expressed in normal adult tissues (Rebagay G et al, 2012).

ROR1 is upregulated in cancer, in particular in B-cell Chronic Lymphocytic Leukemia (B-CLL), B-cell Acute Lymphocytic Leukemia (B-ALL) and Mantle Cell Leukemia (MCL) (Baskar S et al, 2008; DaneshManesh AH et al, 2008; Shabani M et al, 2008). In 2012, ROR1 overexpression was observed also in breast cancers by IHC (Zhang S et al, 2012). In addition, it was reported that ovarian cancer patients with high expression levels of ROR1 had a higher rate of relapse and a shorter mean survival than ovarian cancers patients expressing low or negligible levels of ROR1. On these bases, ROR1 may act as a novel prognostic marker in ovarian cancers (Dave H et al, 2012; Zhang H et al, 2014; Zhang S et al, 2014; Tan H et al, 2015).

ROR1 oncogenic activity is related to its expression in cancer, bringing to the activation of PI3K-Akt-mTOR pathway and consequently to the enhancement of cellular migration, proliferation and survival. Indeed, literature reports a co-expression between ROR1 expression and proteins localized in invadopodia and in extracellular matrix (ECM) (Rebagay G et al, 2012).

Recently, one study reported a correlation between ROR1 expression and miR-382, a microRNA (miRNA) that is down-regulated in ovarian cancer (Tan H et al, 2016). It was found that ROR1 is upregulated in human ovarian cancer tissues whereas miR-382 is downregulated. The overexpression of ROR1 in human ovarian cancer cell lines promoted cell invasion, while this effect was reversed by overexpression of miR-382. Consequently, these results demonstrate that, in ovarian cancer cell lines, miR-382 directly binds ROR1 thus inhibiting cell migration and invasion (Tan H et al, 2016).

1.7.1 ROR1 in lung AC

In lung AC, ROR1 expression was proposed to be associated with TTF-1, an important histochemical marker of this histotype. Indeed, two studies have demonstrated that TTF-1

may induce ROR1 expression and consequently may activate the pro-survival PI3K-Akt-mTOR pathway (Yamaguchi T et al, 2012; Zhang S et al, 2012).

Moreover, it has been demonstrated that the inhibition of ROR1 by silencing RNA (siRNA) promotes apoptosis over survival in lung AC cell lines. This response due to ROR1 inhibition highlights that its activation in lung AC is related to an increase in tumor cell survival (Figure 1.14). Interestingly, ROR1 repression inhibits lung AC irrespective of the *EGFR* status (Yamaguchi T et al, 2012).

In lung tissues, ROR1 seems to mediate survival signals, at least in part, by two mechanisms: ROR1 kinase-dependent c-Src mediated signaling and ROR1 kinase-indipendent sustainment of EGFR-ERBB3-PI3K signaling (Yamaguchi T et al, 2012).

In 2012, a study demonstrated that ROR1 is expressed, at protein level by IHC, in about 77% of NSCLC (Zhang S et al, 2012).

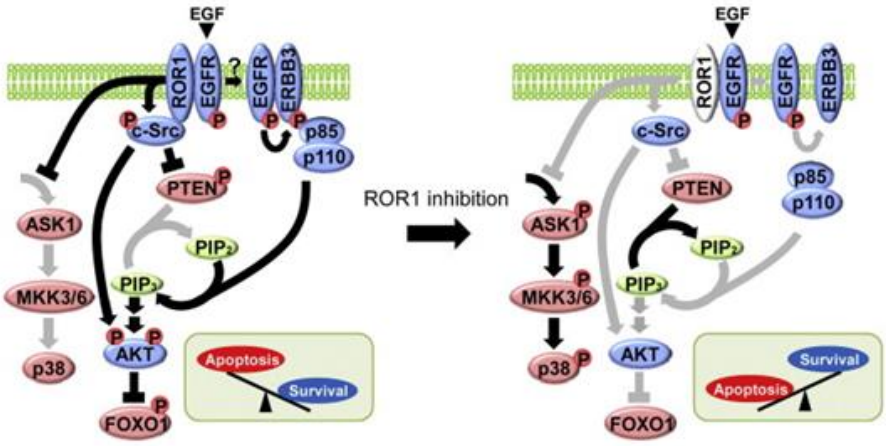

Figure 1.14: Proposed model of ROR1 playing a key role in sustaining a favourable balance between pro-survival and pro-apoptotic signaling. ROR1 inhibition promotes apoptosis over cell survival (Yamaguchi T et al, 2012).

1.7.2 ROR1 and targeted therapies

Recent studies indicate ROR1 as a new potential molecular marker for targeted therapies. ROR1 is a good candidate for the formulation of new targeted drugs because it is not expressed in adult normal tissues but only in cancer cells. In order to block ROR1 TK activity, there are two possible approaches. The first one is represented by the administration of monoclonal antibodies able to block the ROR1 binding domain; the second one is the treatment with TKIs recognizing ROR1 TK domain (Gentile A et al, 2011; Rebagay G et al, 2012). The development of TKIs against ROR1 could be promising because literature reports

aminoacid substitutions in TK domain that modify ROR1 activity; as a consequence these aminoacids could be a good target for the developmentof ROR1 TKIs. Concerning antibodies, in vitro experiments have reported that they do not induce the expected apoptosis in lymphoma cell lines (Baskar S et al, 2008). Nevertheless, 5 different antibodies tested on leukemia cell lines gave better results because they caused cytotoxicity in cancer cell but not in normal tissues. Furthermore, cells treated with these monoclonal antibodies showed a better response to rituximab (Mabthera®; Roche), a targeted therapy against the protein CD20. Both treatments are being studied and tests on B-ALL cell lines showed that ROR1 inhibition makes cells more sensitive to the treatment with dasatinib (Sprycel®; Bristol-Myers Squibb), a Src inhibitor (Bicocca VT et al, 2011).

ROR1 expression can also be predictive of response to chemotherapy or to EGFR TKIs. In last years ROR1 expression has been discovered to have a differential effect on the outcome to erlotinib and to chemotherapy in *EGFR*-mutant NSCLC patients. High ROR1 expression limits PFS in erlotinib-treated patients with T790M mutations so ROR1 targeted therapies could enhance the efficacy of the treatment. In contrast, high ROR1 expression seems to confer longer PFS to chemotherapy (Karachaliou N et al, 2014).

2. AIM

The aim of my PhD project is to improve the clinical care of patients affected by lung AC enhancing the characterization of molecular markers for targeted therapies and characterizing the expression of a new putative marker for targeted therapies, ROR1. Firstly we proposed to find new methodologies for the analysis of AC molecular markers that are essential for the administration of molecular targeted therapies. The need of new assays is essentially due to the particular features of lung AC samples available for molecular diagnosis. Indeed, lung AC specimens are generally constituted by poor quantity and poor quality of tumor material. Very often only small biopsies are available and, with the current methodology (i.e. direct sequencing, DS), it is really difficult to estimate the molecular profile accurately, because the sensitivity of these assays is too low. Furthermore, cancer cells are often dispersed in a high quantity of normal cells, leading these biopsies cases difficult to be considered representative of the tumor. In order to solve all the problems caused by these limits, we decided to test and optimize new methodologies with higher sensitivity that will permit to identify a larger number of mutations or genetic alterations in patients affected by lung AC. Consequently these new assays will permit to enlarge the number of cases that could benefit from targeted therapy. The new methodologies that we used are real-time PCR assays named SensiScreen® (developed by the Danish company PentaBase ApS located in Odense C, Denmark) that are based on oligonucleotides with higher specificity, higher sensibility and higher replicability than DS for mutations detection in *EGFR*, *KRAS* and *BRAF* genes. Furthermore we proposed to test another real-time-based PentaBase kit, focused on the analysis of *EGFR* T790M mutation in blood samples. This test is extremely important because T790M mutation is the main reason of secondary resistance to EGFR TKIs and because this assay is specific for plasma analyses permitting a less invasive approach, than tissue biopsies, for the characterization of this alteration.

Finally, since about 50% of AC cases display a normal gene status sequences for the aforementioned markers, we decided to study ROR1 expression in patients affected by lung AC in order to define if it could represent a new target for the formulation of new molecular targeted therapies. We proposed to evaluate its expression and to see a potential correlation between this marker, the most relevant lung cancers molecular alterations (i.e. *EGFR, KRAS*, *BRAF*, *HER2, ALK* and *ROS1*) and the clinical-pathological features. In addition, we evaluated also the association between ROR1 and miR382 expression, a miRNA involved in its regulation. This possible new marker could be important for all the patients that cannot be treated with the current therapies against alterations in the standard known molecular markers.

3. PATIENTS, MATERIALS AND METHODS

3.1) Patients

The study population consists of three different cohorts characterized by samples from lung AC patients.

The first cohort includes 471 cases that have been characterized for *EGFR*, *KRAS*, *BRAF* and *HER2* by DS, for *ALK* and *ROS1* by FISH and for *EGFR* by SensiScreen®, a real-time PCR kit developed by our institute and PentaBase ApS. In these samples, the analyses were done on formalin-fixed, paraffin-embedded (FFPE) blocks containing small biopsies or resections of primary lung AC or lung AC metastases. Pathological data are available for some patients of this cohort.

The second cohort consists of 61 plasma, 5 serum and 39 tissue samples from patients affected by lung AC. In some cases plasma and tissue are associated to the primary tumor of the same patient. All these cases have been characterized by SensiScreen® liquid biopsy kit for T790M mutation in *EGFR* exon 20 (PentaBase ApS) and by SensiScreen® tissue kit (PentaBase ApS) for the other *EGFR* main mutations. In addition, Ion Torrent® (IOT) Oncomine cell-free nucleic acids assay (Thermo Fisher Scientific; Waltham, Massachusetts, USA), TheraScreen® (a real-time PCR assay developed by QIAGEN; Chatsworth, CA, USA) or DS results were available for some of these patients. To conclude our comparison analyses, we tested 42 samples of the second cohort also by cobas® (a real-time PCR assay developed by Roche).

The third cohort includes 102 samples that have been analyzed for *EGFR*, *KRAS*, *BRAF* and *HER2* by DS; for *ALK*, *ROS1* by FISH; for ROR1 and miR-382 expression by TaqMan real-time PCR (Applied Biosystems; Foster City, CA, USA). In addition to pathological data in this group of patients we have clinical data concerning the type of treatment, the PFS and the OS. In this cohort we analyzed the material from FFPE blocks of the primary tumor.

3.2) Mutational status of *EGFR*, *KRAS*, *BRAF* and *HER2* by DS

Genomic DNA was extracted from six 7-µm FFPE tumor tissue sections using the QIAamp Mini kit (QIAGEN), according to the manufacturer's instructions, and the DNA was amplified by Polymerase Chain Reaction (PCR) experiments. The tumor area was selected by a pathologist and when the tumor cells were less than 70%, the FFPE section was macro-

dissected. We searched for point mutations, deletions or insertions in *EGFR* exon 18 (including codons 709 and 719), exon 19 (including codons from 746 to 753), exon 20 (including codons 768, 770, 771, 776 and 790) and exon 21 (including codons 858 and 861) (Table 3.1).

In addition we investigated *KRAS* point mutations in exon 2 (including codons 12 and 13), *BRAF* point mutations in exon 11 (including codons 466 and 469) and exon 15 (including codon 600) (Table 3.1). Furthermore, we analyzed *HER2* genetic alterations paying attention to insertions in exon 20 (Table 3.1). The DS of PCR products was based on the Sanger method and was done using a 3130 Genetic Analyzer (Applied Biosystems). Finally, the electropherograms were analyzed with the appropriate software (SeqScape Software Version 2.5^{TM}, Applied Biosystems). Each sequence reaction was performed at least twice, starting from independent PCR reactions in order to confirm the DNA sequence.

Gene	Annealing temperature	Primer Sequence	PCR cycles
EGFR 18 Fw	58°C	5′-TCCAGCATGGTGAGGGCTGAG-3′	40
EGFR 18 Rv	58°C	5′-GGCTCCCCACCAGACCATG-3′	40
EGFR 19 Fw	58°C	5′-TGGGCAGCATGTGGCACCATC-3′	40
EGFR 19 Rv	58°C	5′-AGGTGGGCCTGAGGTTCAG-3′	40
EGFR 20 Fw	58°C	5′-CCTCCTTCTGGCCACCATGCG-3′	40
EGFR 20 Rv	58°C	5′-CATGTGAGGATCCTGGCTCC-3′	40
EGFR 21 Fw	58°C	5′-CCTCACAGCAGGGTCTTCTC-3′	40
EGFR 21 Rv	58°C	5′-CCTGGTGTCAGGAAAATGCT-3′	40
KRAS 2 Fw	55◦C	5′-TGGTGGAGTATTTGATAGTGTA-3′	45
KRAS 2 Rv	55◦C	5′-CATGAAAATGGTCAGAGAA-3′	45
BRAF 11 Fw	52°C	5′-TCCCTCTCAGGCATAAGGTAA-3′	45
BRAF 11 Rv	52°C	5′-CGAACAGTGAATATTTCCTTTGAT-3′	45
BRAF 15 Fw	52°C	5′-TCATAATGCTTGCTCTGATAGGA-3′	45
BRAF 15 Rv	52°C	5′-GGCCAAAAATTTAATCAGTGGA-3′	45
HER2 20 Fw	57◦C	5′-CCATACCCTCTCAGCGTA-3′	40
HER2 20 Rv	57◦C	5′-GCTCCGGAGAGACCTGCAA-3′	40

Table 3.1: Primers used for PCR reactions. This table illustrates the annealing temperature, the primer sequence and the number of PCR cycles that were applied for the amplification of the different genes. Abbreviations: Fw, forward; Rv, reverse; PCR, polymerase chain reaction.

3.3) *ALK* and *ROS1* gene status by FISH

FISH was performed on 4-µm formalin-fixed, paraffin-embedded tissue sections treated using the Paraffin Pretreatment kit II (Pretreatment Reagent VP 2000, Abbott Molecular AG; Baar, Switzerland) according to the manufacturer's instructions. The *ALK* FISH assay was done using the LSI *ALK* Dual Colour Break Apart Rearrangement Probe (Abbott Vysis; Illinois, North Chicago, USA) and the *ROS1* FISH assay was performed with the SPEC *ROS1* Dual Colour Break Apart Probe (Zytovision; Bremerhaven, Germania). The signals were evaluated with a fluorescent automated microscope (Zeiss© Axioplan 2 Imaging, Oberkochen, Germany) equipped with a 100W UV lamp; an AxioCam camera (Zeiss© AxioCam MRm) and single, double, triple band pass filters. In addition, the positive cases were analyzed also by the Bioview© duet3 technology: an automatic platform for FISH fluorescent signal detection characterized by a fluorescent microscope, a software controlling filter, a software controlling focus and a digital camera. .

ALK probe hybridizes in the 2p23 region and it is characterized by a dual colour (Spectrum Green and Spectrum Orange) break-apart methodology that permits to recognize *ALK* downstream and upstream sequences of the *ALK* usual breakpoints (Figure 3.1). The analyses are done counting cell by cell and observing the number of coloured fluorescent signals in the tissue.

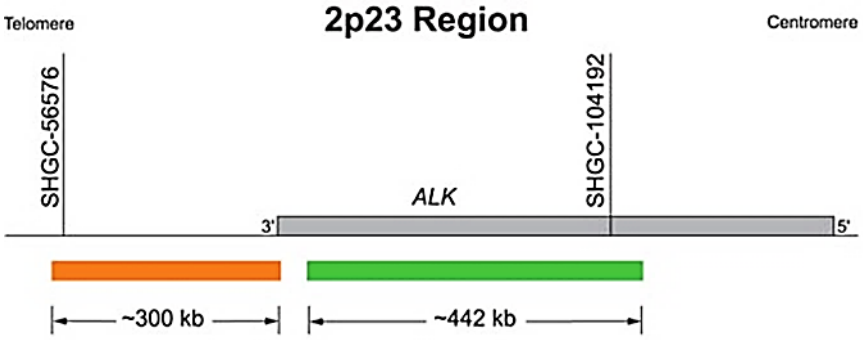

Figure 3.1: Schematic representation of the LSI *ALK* Dual Colour Break Apart Rearrangement Probe (Abbott Vysis). Abbreviations: ALK, Anaplastic Lymphoma receptor tyrosine Kinase.

When the tumor cell is characterized by two fusion signals (i.e. two yellow signals) there is no *ALK* rearrangement. On the contrary, if the red signal is separated by the green one, the tumor cell carries the *ALK* rearrangement. In this case, if there is an intrachromosomal translocation

(e.g. *EML4*/*ALK* inversion) the signals are less than 2/3 diameters distant. On the contrary, if the distance is more than 3 diameters, there is an interchromosomal translocation.

ROS1 probe hybridizes in the 6q22 region and it is a dual color break-apart probe like the aforementioned *ALK* probe. The green signal hybridizes next to the breakpoint cluster region (BCR) and the red one is distal (Figure 3.2). The interpretation of the signals is the same as that of *ALK*.

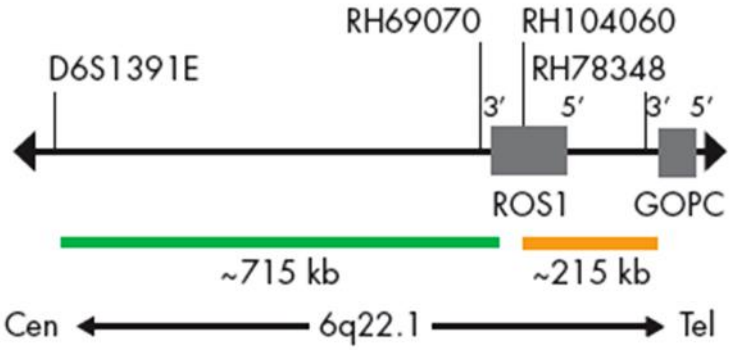

Figure 3.2: Schematic representation of SPEC *ROS1* Dual Colour Break Apart Probe (Zytovision). Abbreviations: Cen, Centromere; Tel, Telomere.

For both ALK and ROS1 analyses, a minimum of 100 morphologically clear, non-overlapping nuclei from at least 8-10 areas were scored for each tumor. In the case of small biopsies, the minimum considered for interpretation was 50 cells. Only experiments with at least 90% hybridization efficiency were considered. A tissue is considered positive for *ALK* or *ROS1* rearrangements if this alteration is observed in at least 15% of cells (Yamaguchi T et al, 2012).

3.4) TTF-1 immunohistochemistry

TTF-1 expression, the principal lung AC marker, was tested by immunohistochemistry (IHC). We used a mouse monoclonal antibody against TTF-1 (M3575©, Dako; Glostrup, Denmark). The presence of TTF-1 is demonstrated when some cells of the samples assume nuclear coloration.

TTF-1 expression is reported on medical reports with values indicating the intensity, the distribution and the percentage of positive cells.

3.5) SensiScreen®

EGFR genetic alterations were detected also by SensiScreen®, real-time PCR assays developed in a collaboration between our institute (Istituto Cantonale di Patologia, Locarno) and a Danish company (PentaBase ApS).

3.5.1 SensiScreen® validation on mutated cell lines

SensiScreen® assays development started with sensitivity studies on plasmids comprising sequences with different *EGFR* mutations (i.e. G719A, G719C, G719S, 746_750del, 746_752del, 747_750del, 747_751del, 747_753del, 767_768ins, 769_770ins, 770_771ins, 772_773ins, 773_774ins, 774_775ins, S768I, T790M, L858R, L861Q). Afterwards, sensitivity assays were conducted on DNA extracted from lung cell lines harbouring some *EGFR* mutations: NCI-H1650 (ATCC; Manassas, VA, USA; catalogue number: CRL-5883™) carrying E746_A750del in exon 19, HCC4006 (ATCC; catalogue number: CRL-2871™) carrying L747_A750del and A750P in exon 19, NCI-H1975 (ATCC; catalogue number: CRL-5908™) carrying T790M in exon 20 and L858R in exon 21 (Table 3.2). The lung cell line HSAEC1-KT (ATCC; catalogue number: CRL-4050™) was purchased in order to obtain wt DNA. The cell lines were subcultured in appropriate media according to the manufacturer's instructions (ATCC) and genomic DNA was isolated using the QIAmp Mini kit (QIAGEN). Seven percentages of mutated cell line DNA in fixed amounts of wt DNA (10%, 1%, 0.5%, 0.1%, 0.05%, 0.01% and 0%) were tested by SensiScreen® assay and DS. The analyses on cell lines and on plasmids permitted to define the sensitivity of these assays by the description of the limit of detection (LOD).

After validation on DNA extracted from cell lines, the testing proceeded with the validation on DNA extracted from FFPE tissues of patients that has been previously characterized by DS.

Cell line	*EGFR* mutation	Origin
NCI-H1650 (ATCC; CRL-5883™)	p.E746_A750delELREA (c.2235_2249del15)	Lung AC
HCC4006 (ATCC; CRL-2871™)	p.L747_A750>P (c.2239_2248TTAAGAGAAG>C) + p.A750P (c. 2248G>C)	Lung AC
NCI-H1975 (ATCC; CRL-5908™)	p.T790M (c.2369C>T) + p.L858R (c.2573T>G)	Lung AC
HSAEC1-KT (ATCC; CRL-4050™)	none (wt)	Lung AC

Table 3.2: Cell lines tested for SensiScreen® validation. Abbreviations: AC, adenocarcinoma.

3.5.2 Mutational status of *EGFR* by SensiScreen®

SensiScreen® contains Suprimer™ (i.e. DNA primers modified with pentabases that will improve affinity, sensitivity, specificity and that will reduce primer-dimer formation); HydrolEasy™ probes (fluorescent-labeled probes able to recognize a specific mutated sequence of the EGFR gene) and wt BaseBlockers™ (WTB) (probes with a sequence that binds specifically to the wild-type DNA region, blocking the amplification of the wild-type DNA during PCR process) (Christensen UB and Pedersen EB, 2002; Christensen UB et al, 2004). WTBs block the amplification of the wt sequence and allow the amplification of the mutant variant.

In every SensiScreen® assay, the reaction mixture is made of 300-900nM of each Suprimer™, 200nM of the probe, specific for each mutation, and 1000-5000nM of WTB.

qPCR was performed using 50ng of genomic DNA. The thermocycling conditions were: 3´(minutes, min) of initial activation of the hotstart taq-polymerase at 95°C, followed by 40 cycles of a 2-step PCR with a 10´´ (seconds, s) denaturation step at 95°C and 30´´ extension step at 55°C (Riva A et al, 2017) (Figure 3.3).

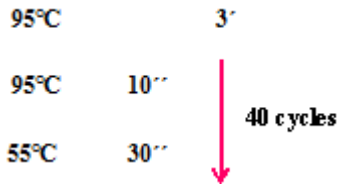

Figure 3.3: SensiScreen® thermocycling conditions for *EGFR* characterization (Riva A et al, 2017).

Fluorescence was measured at the end of each extension step. The qPCR threshold cycle (Ct) of normalized fluorescence was used for the evaluation of the data. Ct is defined as the number of cycles where a fluorescence signal crosses the threshold. In order to make data analyses independent from the type of instrument used, the threshold was defined as 10% of the signal strength of the reference assay at cycle 45. For all valid samples ($23< Ct_{reference}<36$), a ΔCt value was calculated by taking the Ct value of the mutation-specific assay and subtracting the Ct value of the reference assay: $\Delta Ct = Ct_{mutation} - Ct_{reference}$.

Patient samples analysed by SensiScreen® assays were classified as positive for a given mutation if the $Ct_{mutation}$ was ≤ 38 and the ΔCt was ≤9.

SensiScreen® assays have been developed for the analyses of three mutations in *EGFR* exon 18 (i.e. G719A, G719C and G719S), for the most common deletions in *EGFR* exon 19 (e.g. 746_750del, 746_752del, 747_751del and 747_753del), for the most diffuse insertions in *EGFR* exon 20 (e.g. 767_768ins, 769_770ins, 770_771ins, 772_773ins, 773_774ins, 774_775ins), for two mutations in *EGFR* exon 20 (i.e. S768I and T790M) and for two mutations in *EGFR* exon 21 (i.e. L858R and L861Q). SensiScreen® assays have been created in simplex version for G719A, G719C, G719S, S768I, T790M, L858R, L861Q mutations and in multiplex version for G719 mutations, exon 19 deletions or exon 20 insertions. Exon 20 insertions multiplex is made of two reactions and G719 or exon 19 deletions multiplex by only one reaction.

3.6) Plasma analyses

Whole blood of patients affected by lung AC has been collected in Cell-Free DNA BCT® tubes (Streck; Omaha, Nebraska, USA), containing preservative reagents. These tubes are designed to minimize cfDNA degradation and to prevent the release of DNA from blood cells. After blood collection, the samples are reversed ten times and blood is stored at room temperature (RT) until plasma separation.

3.6.1 Plasma separation

Plasma separation is done within 48h after blood collection in Streck tubes. The mechanism is based on two centrifugation steps at 3000xg for 15´ at RT. After centrifugation, the plasma component is transferred into a new, sterile tube and stored at -80°C until the DNA extraction and molecular characterization.

3.6.2 ctDNA analyses by cobas®

cfDNA extraction was done in 2mL of plasma obtained from patients affected by lung AC using cobas® cfDNA sample preparation kit (Roche) according to the manufacturer's instructions. After extraction, cfDNA was analyzed by cobas® EGFR Mutation Test v2 (CE-IVD) (Roche) following the standard protocol (available from: http://egfrmutationtestv2.roche.com/). The cobas® EGFR Mutation Test v2 is a real-time PCR test that identifies 42 mutations in *EGFR* exons 18, 19,20 and 21 (e.g. L858R, L861Q, exon 19 deletions and T790M). Amplification results are detected by cobas® 4800 system, software v2.1 or higher. The LODs for the *EGFR* mutations range from 2 to 13.4% (for details, see http://egfrmutationtestv2.roche.com/).

3.6.3 ctDNA analyses by SensiScreen® assays

cfDNA was extracted from 500µL of plasma using QIAamp Mini kit (QIAGEN) according to the manufacturer's instructions after an initial treatment at 57°C for 10´ with 500µL of lysis buffer (AL) and 50µL of proteinase K (PK).

The extracted cfDNA was tested by SensiScreen® T790M liquid biopsy assay (PentaBase ApS) in order to identify the presence of the T790M mutation in the blood obtained from patients affected by lung AC. SensiScreen® T790M liquid biopsy assay was developed by PentaBase ApS and our institute. This kit follows the same methodology aforementioned in paragraph 3.5 but it is characterized by a higher sensitivity compared to the SensiScreen® assay for T790M analyses on tissues.

In addition, the analyses on cfDNA from plasma were done also for the other *EGFR* mutations for which we developed the SensiScreen® assays in tissue (paragraph 3.5). Concerning these mutations, *EGFR* characterization of cfDNA was done using the same assays applied in tissues because they have enough sensitivity for the test on liquid biopsies.

3.7) ROR1 expression analyses

RNA extraction was done using the RNeasy FFPE kit (QIAGEN) according to the manufacturer's instructions. In particular it was applied on two 10-µm FFPE tissue sections for each paired normal and tumor sample from the same patient.The tumor and normal area were selected by a pathologist. When the tumor cells in the cancerous tissue were less than

70%, the FFPE section was macro-dissected. 500 ng of RNA from each paired tumor and normal tissue was retro-transcribed in complementary DNA (cDNA) using the Superscript Vilo Master mix III (Invitrogen; Carlsbad, CA, USA). After RNA extraction and retro-transcription in cDNA, ROR1 expression was evaluated by real-time PCR assays. To quantify the amplified cDNA, we used a TaqMan fluorescent probe (Applied Biosystems) recognizing the target gene (i.e. ROR1) and a TaqMan probe marked with a different fluorochrome recognizing a reference gene (i.e. the RN18S1 housekeeping gene, which encodes for the 18S rRNA). The reference gene is an internal control that must always be expressed at the same level in both normal and cancer tissues. TaqMan probes are characterized by a quencher (Q), a fluorochrome located on 3′-end, and a reporter (R), a fluorochrome located on 5′-end. During the real-time PCR amplification process, the probe recognizes the denaturated cDNA strand in ROR1 and DNA polymerase cleaves the probe, bringing to the split of Q and R. This division results in a fluorescent signal that is proportional to the number of DNA molecules obtained because the Q cannot mask R fluorescence when they are far away.

For the amplification, we used 100ng and the test was repeated three times for each sample.

The data were analyzed considering the threshold cycle in both cancer and normal tissues for each sample. We used the Livak method (Livak KJ and Schmittgen TD, 2001), that consists in the calculation of the $2^{-\Delta\Delta Ct}$ value. ΔΔCt is the difference between the sample ΔCt and the control ΔCt. Sample ΔCt is the difference between the Ct of the target gene (ROR1) and the Ct of the reference gene in tumor tissue. Control ΔCt is the difference between the Ct of the target gene (ROR1) and the Ct of the reference gene in normal tissue.

ΔCt sample: Ct tumoral with ROR1 probe - Ct tumoral with RN18S1 probe

ΔCt control: Ct normal with ROR1 probe - Ct normal with RN18S1 probe

ΔΔCt: ΔCt sample - ΔCt control

We fixed the value of 1 as cut-off, so in all the the samples with $2^{-\Delta\Delta Ct}$ value > 1 ROR1 has been considered overexpressed.

3.8) miR-382 expression analyses

miRNAs extraction, retro-transcription and real-time were done by applying RecoverAll™ Total Nucleic Acid Isolation Kit and TaqMan Advanced miRNA Assays (Applied Biosystems) according to the manufacturer's instructions. For retro-transcription, we used 10ng of RNA and a 1:10 dilution of the cDNA was used after retro-transcription, in order to use 5µL of the dilution in the real-time PCR experiment. The probes for real-time testing were two: hsa-miR-382-5p (Applied Biosystems) for miR-382 expression testing and hsa-miR-451a (Applied Biosystems) for the evaluation of the housekeeping gene. For miR-382 analyses, we applied again Livak method, as previously described (see 3.7), using the housekeeping gene miR-451a and comparing the threshold cycle for each sample in both cancer and normal tissues. A sample was considered as overexpressing miR-382 when $2^{-\Delta Ct}$ value was > 1.

3.9) Statistical analyses

The comparisons among genetic alterations in *EGFR*, *KRAS*, *BRAF*, *HER2*, *ALK*, *ROS1* genes and the association of them with clinical-pathological characteristics were evaluated through the two-tailed Fisher's exact test (available from: http://in-silico.net/statistics/Fisher_exact_test). We set a value of statistical significance equal to p=0.05.

This test has been applied also for the comparison between ROR1, miR-382 expression and for the association of them with clinical-pathological characteristics and the other molecular markers data.

In addition, ROR1 and miR-382 were compared also by the determination of the Spearman correlation coefficient and of the Pearson correlation coefficient. These coefficients need to be near to 0 to demonstrate absence of correlation; in contrast, a value near to -1 or to 1 demonstrate an opposite or concordant correlation, respectively.

4. RESULTS

4.1) Patients

In our project we enrolled three cohorts of patients affected by lung AC.
The first one was characterized for *EGFR*, *KRAS*, *BRAF* and *HER2* by DS; for *ALK* and *ROS1* by FISH and for *EGFR* by SensiScreen® assay (PentaBase ApS). The second cohort was analyzed by SensiScreen® liquid biopsy kit for the analysis of the T790M mutation (PentaBase ApS) and by SensiScreen® tissue kit for the other *EGFR* mutations (PentaBase ApS). In some samples we already had the molecular data obtained by IOT®, (Thermo Fisher Scientific) or TheraScreen® (QIAGEN). In addition, we characterized 42 samples of this cohort also by cobas® (Roche). The third cohort was characterized for *EGFR*, *KRAS*, *BRAF* and *HER2* by DS, for *ALK* and *ROS1* by FISH, and for ROR1 and miR-382 expression by real-time PCR.

The first cohort is represented by 471 cases, 262 men (55.6%) and 209 women (44.4%), with a mean age at diagnosis of 68 years (range: 38-88 years). The primary tumor was analyzed in 381 samples whereas in 90 cases the DNA was extracted from lung AC metastases characterized, in particular, by breast metastases (one patient), intestinal metastases (four patients), adrenal glands metastases (five patients), skin metastases (seven patients), liver metastases (ten patients), lymph nodes metastases (eighteen patients), brain metastases (twenty-four patients) and bone metastases (twenty-one patients).
Histopathological T staging is available in 205 cases: 87 are classified as pT1, 81 as pT2, 26 as pT3, 10 as pT4 and 1 as pTX (respectively 42.4%, 39.5%, 12.7%, 4.9% and 0.5% out of the available data). In addition, the N descriptor is available in 132 samples: 73 are classified as pN0, 23 as pN1, 30 as pN2 and 6 as pNX (respectively 55.3%, 17.4%, 22.7% and 4.6% out of the available data). M descriptor is equal to M1 in three patients (Table 4.1).
Concerning tumor grading, the datum is available for 304 patients: 41 are classified as G1, 91 as G2, 157 as G3, 5 as G1/G2, 8 as G2/G3 and 2 as G4 (respectively 13.6%, 29.9%, 51.6%, 1.6%, 2.6% and 0.7% out of the available data) (Table 4.1).

The second cohort is characterized by 61 plasma, 5 serum and 39 tissue samples. In 9 cases, blood was collected before the treatment with EGFR TKIs (up-front patients) and in 52 cases after their administration. All the serum samples were obtained from patients after treatment with TKIs. Concerning tissues, 23 patients are up-front and 16 cases were previously treated with TKIs.

In 22 patients we succeeded to collect both plasma and tissue associated samples:
- 16 associated tissue pre-treatment and plasma post-treatment samples
- 1 associated tissue post-treatment and plasma post-treatment samples
- 5 associated tissue post- plus pre-treatment and plasma post-treatment samples

In 5 patients we analyzed paired plasma and serum samples.

The 39 tissue samples are from patients with a mean age at diagnosis of 63 years (range: 35-84 years). In this group, there are 18 men and 21 women (46.2% and 53.8% out of the tissue samples) (Table 4.2).

The 66 liquid biopsies samples are from patients with a mean age at diagnosis of 66 years (range: 39-87 years). In particular, the patients from which we obtained liquid biopsies are subdivided in 32 men and 34 women (48.5% and 51.5% out of this group) (Table 4.3).

In this cohort we do not have information concerning the histopathological staging and the grade of differentiation.

The third cohort is represented by 102 cases, 48 men (47.1%) and 54 women (52.9%), with a mean age at diagnosis of 64 years (range: 40-86 years). For all the patients of this group we analyzed the primary tumor and, if possible, we selected also the normal tissue in order to do the comparisons necessary for the the analysis of ROR1 and miR-382 by real-time experiments.

The histopathological T descriptor is available in 71 patients: 31 are defined as T1, 22 as T2, 18 as T3 (respectively 43.6%, 31%, and 25.4% out of the available data). N staging is described in 58 cases: 29 are N0, 8 are N1, 20 are N2 and 1 is N3 (respectively 50%, 13.8%, 34.5% and 1.7% out of the available data). Only one patient is M1 (Table 4.4).

The differentiation grade of the tumor is available in 89 samples: 15 are classified as G1, 3 as G1/G2, 32 as G2, 5 as G2/G3, 33 as G3 and 1 as G4 (respectively 16.8%, 3.4%; 36%, 5.6%; 37.1% and 1.1% out of the available data) (Table 4.4).

This cohort has been selected for ROR1 study because we have clinical data concerning the treatment, the PFS and the OS. In particular 32 patients (31.4%) underwent surgery, 17 (16.7%) surgery followed by chemotherapy, 16 (15.7%) surgery followed by chemotherapy and radiotherapy, 22 (21.6%) chemotherapy with palliative purpose, 2 (1.9%) radiotherapy with palliative purpose, 11 (10.8%) chemotherapy followed by radiotherapy (specifically six with palliative purpose; five with curative intent) and 2 (1.9%) received TKIs vs EGFR (specifically afatinib).

In patients treated with chemotherapy the PFS is equal to 9 months whereas OS is equal to 11 months in case of palliatiave intent and 40 months in case of curative intent.
These data are not available for targeted therapies because only two patients received TKIs.

Cohort 1			
Patients characteristics (n=471)		**Number of cases**	**Pecentage (%)**
Age	≥ mean (68)	229	48.6
	< mean (68)	242	51.4
Gender	male	262	55.6
	female	209	44.4
Staging: T descriptor (n=205)	T0	0	-
	TX	1	0.5
	T1	87	42.4
	T2	81	39.5
	T3	26	12.7
	T4	10	4.9
Staging: N descriptor (n=132)	N0	73	55.3
	NX	6	4.6
	N1	23	17.4
	N2	30	22.7
	N3	0	-
	N4	0	-
Differentiation (n=304)	GX	0	-
	G1	41	13.6
	G1/G2	5	1.6
	G2	91	29.9
	G2/G3	8	2.6
	G3	157	51.6
	G4	2	0.7

Table 4.1: Cohort 1. Clinical-pathological characteristics. Abbreviations: n, number.

Cohort 2			
Tissue samples			
Patients characteristics (n=39)		**Number of cases**	**Pecentage (%)**
Age	≥ mean (63)	15	38.5
	< mean (63)	24	61.5
Gender	male	18	46.2
	female	21	53.8

Table 4.2: Cohort 2, tissue samples. Clinical-pathological characteristics. Abbreviations: n, number.

Cohort 2			
Liquid biopsies (plasma and serum) samples			
Patients characteristics (n=66)		**Number of cases**	**Pecentage (%)**
Age	≥ mean (66)	31	47.0
	< mean (66)	35	53.0
Gender	male	32	48.5
	female	34	51.5

Table 4.3: Cohort 2, liquid biopsies. Clinical-pathological characteristics. Abbreviations: n, number.

Cohort 3			
Patients characteristics (n=102)		**Number of cases**	**Pecentage (%)**
Age	≥ mean (64)	46	45.1
	< mean (64)	56	54.9
Gender	male	48	47.1
	female	54	52.9
Staging T descriptor (n=71)	T0	0	-
	TX	0	-
	T1	31	43.6
	T2	22	31
	T3	18	25.4
	T4	0	-
Staging N descriptor (n=58)	N0	29	50
	NX	0	-
	N1	8	13.8
	N2	20	34.5
	N3	1	1.7
	N4	0	-
Differentiation (n=89)	GX	0	-
	G1	15	16.8
	G1/G2	3	3.4
	G2	32	36
	G2/G3	5	5.6
	G3	33	37.1
	G4	1	1.1
Treatment (n=102)	surgery	32	31.4
	surgery+chemotherapy	17	16.7
	surgery+chemotherapy+radiotherapy	16	15.7
	palliative chemotherapy	22	21.6
	palliative radiotherapy	2	1.9
	chemotherapy+radiotherapy	11	10.8
	afatinib	2	1.9

Table 4.4: Cohort 3. Clinical-pathological characteristics. Abbreviations: n, number.

4.2) Cohort one

4.2.1 Molecular markers characterization by DS

In the first cohort (n= 471) we observed mutations in 185 patients (Figure 4.1), corresponding to 39.2% of the cases. In particular: 67/468 *EGFR* mutated cases, 110/471 *KRAS* exon 2 mutated cases, 8/457 *BRAF* mutated cases and 2/462 *HER2* exon 20 mutated cases, representing 14.3%, 23.3%, 1.7% and 0.4% out of the evaluable samples respectively. In *EGFR*, we identified four mutations in exon 18, twenty-eight in exon 19, fifteen in exon 20 and twenty-four in exon 21. In the totality of the *EGFR* mutations, equal to 71, the different exons (18, 19, 20 and 21) were mutated in 5.6%, 39.4%, 21.2% and 33.8% of the cases, respectively. In *BRAF*, we detected four mutations in both exon 11 and exon 15 respectively. The specific mutations are shown in Table 4.5. In this cohort we found that twelve patients harboured two mutations: one sample is mutated in *EGFR* and *KRAS* (i.e. patient 112 in Table 4.5 is E746_A750 del in *EGFR* exon 19 and G12V in *KRAS* exon 2); one sample is mutated in *EGFR* and *BRAF* (i.e. patient 173 in Table 4.5 is T847I in *EGFR* exon 21 and G469A in *BRAF* exon 11); eight samples are mutated in two different *EGFR* exons (e.g. patient 189 in Table 4.5 is E746_A750del in exon 19 and T790M in exon 20) and two samples have a double mutation in the same *EGFR* exon (i.e. patients 155 and 321 in Table 4.5 have double mutations in *EGFR* exon 19 and exon 21 respectively) (Table 4.5).

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
1	wt	**G12V**	wt	wt
2	wt	**G12C**	wt	wt
3	wt	**G12C**	wt	wt
5	wt	**G13C**	wt	wt
6	wt	**G12C**	wt	wt
7	**L858R (21)**	wt	wt	wt
10	wt	**G12A**	wt	wt
11	wt	**G12C**	wt	wt
13	wt	**G12C**	wt	wt
14	wt	**G13D**	wt	wt
16	wt	**G12C**	wt	wt

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
182	**L861Q (21)**	wt	wt	wt
189	**E746_A750 del (19) + T790M (20)**	wt	wt	wt
190	**E746_A750 del (19) + T790M (20)**	wt	wt	wt
191	wt	wt	**V600E (15)**	wt
193	**L747_A750 del (19)**	wt	wt	wt
194	**E746_S752 del (19)**	wt	wt	wt
195	wt	wt	wt	**E770_A771 ins AYVM**
196	**S752_I759 del (19)**	wt	wt	wt
197	**G719S (18) + S768I (20)**	wt	wt	wt
198	**G719S (18) + S768I (20)**	wt	wt	wt
199	**G719S (18) + S768I (20)**	wt	wt	wt

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
17	wt	wt	wt	**A775_G776 ins YVMA**
18	wt	**G12C**	wt	wt
20	wt	**G12D**	wt	wt
22	wt	**G13R**	wt	wt
28	**L747_S753 del (19)**	wt	wt	wt
29	**L858R (21)**	wt	wt	wt
30	wt	**G12C**	wt	wt
32	wt	**G12D**	wt	wt
36	**G719A (18) + S768I (20)**	wt	wt	wt
37	wt	wt	**V600E (15)**	wt
40	wt	**G12D**	wt	wt
46	wt	wt	**V600E (15)**	wt
47	wt	**G12R**	wt	NE
48	wt	**G12V**	wt	wt
49	wt	**G12C**	wt	wt
50	wt	**G12C**	wt	wt
54	**L858R (21)**	wt	wt	wt
56	**L747P (19)**	wt	wt	wt
59	**E746_A750 del (19)**	wt	wt	wt
63	wt	**G12C**	wt	wt
66	NE	**G12C**	NE	NE
70	**T790M (20)**	wt	wt	wt
80	**E746_A750 del (19)**	wt	NE	NE
81	**P772_H773 ins PR (20)**	wt	NE	NE
82	wt	**G12V**	NE	NE
83	wt	**G12C**	wt	wt
84	wt	wt	**G469V (11)**	wt
88	**E746_A750 del (19)**	wt	wt	wt
89	**T790M (20)**	wt	wt	wt
96	wt	**G12C**	wt	wt
99	wt	**G12A**	wt	wt
101	**T790M (20)**	wt	wt	wt
104	wt	**G12V**	wt	wt
108	**H773_V774 ins NPH (20)**	wt	wt	wt
109	**D770_N771 ins G (20)**	wt	wt	wt
112	**E746_A750 del (19)**	**G12V**	wt	wt
118	**E746_A750 del (19)**	wt	wt	wt
119	wt	wt	**V600E (15)**	wt
120	wt	**G12C**	wt	wt
121	wt	**G12C**	wt	wt

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
202	**H773_V774 ins NPH (20)**	wt	wt	wt
203	**L747_T751 del (19)**	wt	wt	wt
204	**L747_T751 del (19)**	wt	wt	wt
205	**T790M (20) + L858R (21)**	wt	wt	wt
206	**E746_A750 del (19)**	wt	wt	wt
208	wt	**G12A**	wt	wt
209	**E746_A750 del (19)**	wt	wt	wt
215	**D770_N771 ins SVE (20)**	wt	wt	wt
222	wt	wt	**T440A (11)**	wt
223	wt	**G12C**	wt	wt
226	wt	**G12C**	wt	wt
228	wt	**G12V**	wt	wt
230	wt	**G12C**	wt	wt
231	wt	**G12D**	wt	wt
232	wt	**G12D**	wt	wt
233	**E746_A750 del (19)**	wt	wt	wt
236	wt	**G12C**	wt	wt
240	wt	**G13C**	wt	wt
241	**E746_A750 del (19)**	wt	wt	wt
246	**E746_A750 del (19)**	wt	wt	wt
247	**L858R (21)**		wt	wt
253	wt	**G12D**	wt	wt
255	wt	**G12V**	wt	wt
256	**L858R (21)**		wt	wt
258	wt	**G12C**	wt	wt
259	wt	**G12C**	wt	wt
268	wt	**G12C**	wt	wt
273	**E709G (18) + L858R (21)**	wt	wt	wt
275	wt	**G12D**	wt	wt
279	wt	**G12V**	wt	wt
280	wt	**G12C**	wt	wt
283	wt	**G12C**	wt	wt
284	wt	**G12C**	wt	wt
285	wt	**G12A**	wt	wt
294	wt	**G12V**	wt	wt
295	wt	**G12C**	wt	wt
296	wt	**G12V**	wt	wt
301	wt	**G12C**	wt	wt
308	wt	**G12D**	wt	wt
309	wt	**G12C**	wt	wt

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
122	wt	wt	**G469V (11)**	wt
126	wt	**G12C**	wt	wt
127	wt	**G12A**	wt	wt
128	wt	**G12C**	wt	wt
129	wt	**G12C**	wt	wt
130	wt	**G12C**	wt	wt
132	wt	**G12C**	wt	wt
133	wt	**G12C**	wt	wt
134	wt	**G12C**	wt	wt
136	wt	**G12V**	wt	wt
138	wt	**G12C**	wt	wt
139	wt	**G12V**	NE	NE
140	wt	**G12A**	wt	wt
141	wt	**G12C**	wt	wt
142	wt	**G12A**	wt	wt
144	wt	**G12D**	wt	wt
145	**E746_A750 del (19)**	wt	wt	wt
147	wt	**G12C**	NE	wt
148	wt	**G12V**	wt	wt
149	wt	**G12V**	wt	wt
150	**E746_S752 del (19)**	wt	wt	wt
153	wt	**G12V**	wt	wt
154	**T790M (20)**	wt	wt	wt
155	**E746G + L747S (19)**	wt	wt	wt
156	**E746_A750 del (19)**	wt	wt	wt
159	wt	**G12C**	wt	wt
163	wt	**G12C**	wt	wt
164	**S752F (19)**	wt	wt	wt
165	**E746_A750 del (19)**	wt	NE	NE
166	wt	**G12A**	wt	wt
168	wt	**G12C**	wt	wt
169	wt	**G12V**	wt	wt
170	wt	**G12C**	wt	wt
171	wt	**G12C**	wt	wt
172	wt	**G12V**	wt	wt
173	**T847I (21)**	wt	**G469A (11)**	wt
174	**E746_S752 del (19)**	wt	wt	wt
175	wt	**G12R**	wt	wt
176	**L858R (21)**	wt	wt	wt
177	**K745_E746 ins IPVAIK (20)**	wt	wt	wt

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
313	wt	**G12C**	wt	wt
316	wt	**G12D**	wt	wt
317	wt	**G12V**	wt	wt
318	wt	**G12V**	wt	wt
319	wt	**G12C**	wt	wt
320	wt	**G12V**	wt	wt
321	**L858M + L861Q (21)**	wt	wt	wt
324	**L861Q (21)**	wt	wt	wt
325	wt	**G12C**	wt	wt
329	wt	**G12C**	wt	wt
332	wt	**G12C**	wt	wt
333	wt	**G12C**	wt	wt
336	wt	**G12C**	wt	wt
338	wt	**G12A**	wt	wt
340	wt	**G12C**	wt	wt
345	**E746_A750 del (19)**	wt	wt	wt
348	**L858R (21)**	wt	wt	wt
352	**L858R (21)**	wt	wt	wt
353	wt	**G12V**	wt	wt
356	**L858R (21)**	wt	wt	wt
359	wt	**G12D**	wt	wt
361	**L858R (21)**	wt	wt	wt
366	wt	**G12V**	wt	wt
368	wt	**G12V**	wt	wt
371	wt	**G12A**	wt	wt
375	wt	**G12D**	wt	wt
376	**L858R (21)**	wt	wt	wt
377	wt	**G12C**	wt	wt
378	wt	**G12C**	wt	wt
379	wt	**G12D**	wt	wt
381	wt	**G12C**	wt	wt
383	wt	**G12D**	wt	wt
384	wt	**G12D**	wt	wt
385	wt	**G12D**	wt	wt
406	**L858R (21)**	wt	wt	wt
429	**L858R (21)**	wt	wt	wt
440	**L858R (21)**	wt	wt	wt
445	**E746_A750 del (19)**	wt	wt	wt
450	**L861Q (21)**	wt	wt	wt
451	**L858R (21)**	wt	wt	wt

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
178	**E746_S752 del (19)**	wt	wt	NE
181	**L861Q (21)**	wt	wt	wt

n	Cohort 1: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
455	**L858R (21)**	wt	wt	wt

Table 4.5: Cohort 1-samples mutated in *EGFR* (exons 18, 19, 20 and 21), *KRAS* (exon 2), *BRAF* (exons 11 and 15) and *HER2* (exon 20) by Sanger sequencing. This table shows only the patients with mutations. Abbreviations: ex, exon; n, sample number; NE, not evaluable; wt, wild-type. In brackets it is reported the exon that is mutated.

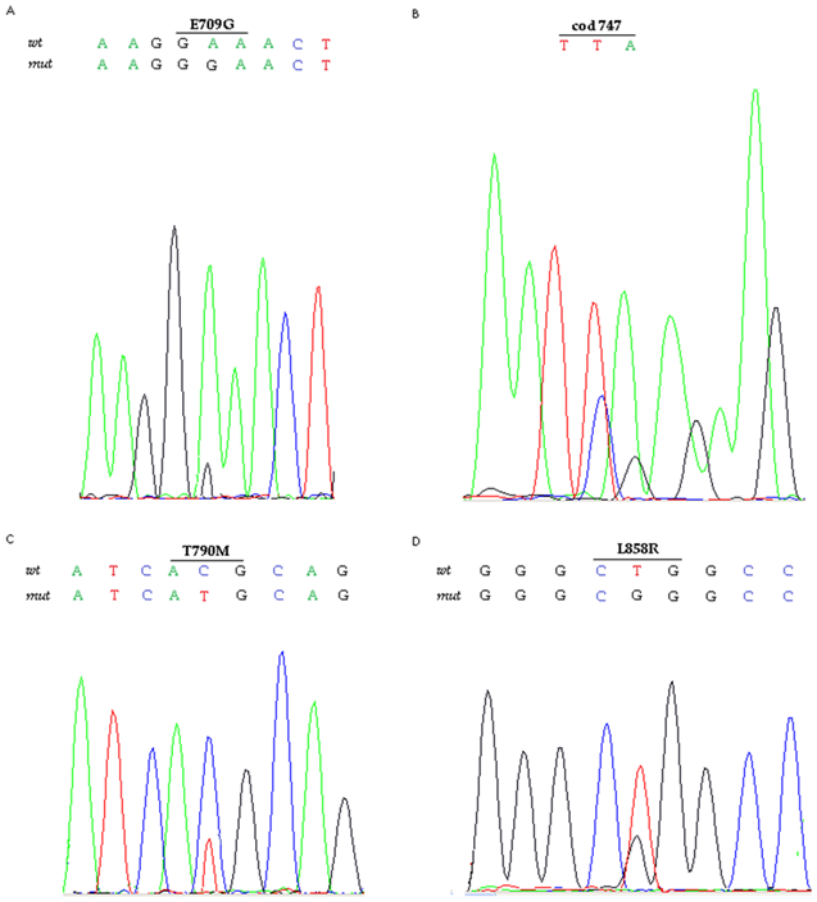

Figure 4.1: Examples of electropherograms obtained by DS. These pictures represent four different *EGFR* mutations corresponding to overlapping peaks: A) The change from GAA codon (position 709) to GAG codon brings to E709G mutation in *EGFR* exon18. B) Sequence of *EGFR* exon 19 in codon 747 containing a deletion from 747 codon to 752 codon (L747_S752del). C) The change from ACG codon (position 790) to ATG codon brings to T790M mutation in *EGFR* exon20. D) The change from CTG (position 858) codon to CGG codon brings to L858R mutation in exon 21.

4.2.2 *ALK* and *ROS1* FISH results

FISH was conducted only in samples for which we had enough material available after the other analyses. In cohort one, we analysed *ALK* in 262 patients and *ROS1* in 261 patients. Fourteen cases were not evaluable for *ALK* and sixteen for *ROS1*. Nine cases are rearranged in *ALK* and one in *ROS1*, representing 3.6% (9/248) and 0.4% (1/245) out of the evaluable samples, respectively (Table 4.6, Figure 4.2 and 4.3).

n	Cohort 1: FISH	
	ALK	*ROS1*
1	-	neg
2	-	neg
3	-	neg
4	-	neg
6	-	neg
7	-	neg
8	-	neg
10	neg	NE
11	-	neg
13	-	neg
14	-	neg
15	-	neg
16	-	neg
18	-	neg
20	-	neg
21	-	neg
22	-	neg
23	-	neg
24	-	neg
25	-	neg
26	-	neg
28	-	neg
29	-	neg
30	-	neg
32	-	neg
33	-	neg
34	-	neg
35	-	neg
36	**pos**	neg
37	-	neg
38	-	neg
39	-	neg
40	neg	neg
41	-	neg
42	-	neg
56	neg	neg
88	neg	-
127	neg	-

n	Cohort 1: FISH	
	ALK	*ROS1*
222	neg	neg
223	neg	neg
224	neg	neg
225	neg	neg
226	neg	neg
227	neg	neg
228	neg	neg
229	neg	neg
230	neg	NE
231	NE	neg
232	neg	neg
233	neg	neg
234	neg	neg
235	neg	neg
237	neg	neg
238	neg	neg
239	neg	neg
240	neg	neg
241	neg	neg
242	neg	neg
243	neg	neg
244	neg	neg
245	neg	neg
246	neg	neg
247	neg	neg
248	neg	neg
249	neg	neg
250	neg	neg
251	neg	neg
252	neg	neg
253	neg	neg
254	neg	neg
255	neg	neg
256	neg	neg
257	NE	neg
258	neg	neg
259	neg	neg
260	neg	neg

n	Cohort 1: FISH	
	ALK	*ROS1*
300	neg	neg
301	neg	neg
302	neg	neg
304	neg	NE
305	neg	neg
306	neg	neg
307	neg	NE
308	neg	neg
309	neg	neg
310	neg	neg
311	neg	neg
312	neg	neg
313	neg	neg
314	neg	neg
315	neg	neg
316	neg	neg
317	neg	neg
318	neg	neg
319	neg	neg
320	neg	neg
321	neg	neg
322	neg	neg
323	neg	neg
324	neg	neg
325	neg	neg
326	neg	neg
327	neg	neg
328	neg	neg
329	neg	neg
330	neg	neg
331	neg	neg
332	neg	neg
333	neg	neg
334	**pos**	neg
335	neg	neg
336	neg	neg
337	neg	neg
338	neg	neg

n	Cohort 1: FISH	
	ALK	*ROS1*
377	neg	neg
378	neg	neg
379	neg	neg
380	neg	neg
381	neg	NE
382	neg	NE
383	neg	neg
384	**pos**	neg
385	neg	neg
386	neg	neg
387	neg	neg
388	neg	neg
389	neg	-
390	neg	neg
391	**pos**	neg
392	neg	neg
393	-	neg
394	-	neg
395	-	neg
396	-	neg
398	neg	-
403	neg	neg
405	**pos**	-
406	neg	-
407	neg	neg
408	neg	-
409	neg	-
410	neg	-
411	neg	neg
412	neg	neg
413	**pos**	-
414	neg	neg
416	neg	neg
417	**pos**	neg
418	neg	-
419	neg	-
420	neg	neg
423	neg	neg

n	Cohort 1: FISH	
	ALK	ROS1
154	neg	-
162	neg	**pos**
166	neg	neg
167	-	neg
168	-	neg
169	-	neg
170	-	neg
171	-	neg
172	-	neg
173	neg	neg
175	neg	-
176	neg	-
177	neg	-
179	neg	-
180	neg	-
181	neg	neg
185	neg	neg
186	neg	neg
188	neg	neg
192	neg	neg
193	neg	-
194	neg	-
195	neg	neg
199	neg	neg
207	neg	neg
208	neg	neg
209	neg	neg
210	neg	neg
211	neg	neg
212	neg	neg
213	neg	neg
214	neg	neg
215	neg	neg
217	NE	neg
218	NE	neg
219	neg	neg
220	neg	neg
221	neg	neg

n	Cohort 1: FISH	
	ALK	ROS1
261	NE	neg
262	NE	neg
263	neg	neg
264	neg	neg
265	neg	neg
266	neg	neg
267	neg	neg
268	neg	neg
269	neg	neg
270	neg	neg
271	neg	neg
272	neg	neg
273	neg	neg
274	neg	neg
275	neg	neg
276	neg	neg
277	neg	neg
278	neg	neg
279	NE	neg
280	neg	neg
281	NE	neg
283	neg	neg
284	neg	neg
285	neg	neg
286	neg	neg
287	neg	neg
288	NE	neg
289	neg	neg
290	neg	neg
291	neg	neg
292	neg	neg
293	neg	neg
294	neg	neg
295	neg	neg
296	neg	neg
297	neg	neg
298	neg	neg
299	neg	neg

n	Cohort 1: FISH	
	ALK	ROS1
339	neg	neg
340	neg	neg
341	neg	neg
342	neg	neg
343	neg	neg
344	neg	neg
345	neg	neg
346	neg	neg
347	neg	neg
348	neg	neg
349	neg	neg
350	neg	neg
351	neg	neg
352	neg	neg
353	neg	neg
354	neg	neg
355	neg	neg
356	neg	neg
357	neg	neg
358	neg	neg
359	neg	neg
360	neg	neg
361	neg	neg
362	neg	neg
363	neg	neg
364	neg	neg
365	neg	neg
366	neg	neg
367	neg	neg
368	neg	neg
369	neg	neg
370	neg	neg
371	neg	neg
372	neg	neg
373	neg	neg
374	neg	neg
375	neg	neg
376	neg	neg

n	Cohort 1: FISH	
	ALK	ROS1
424	neg	neg
425	neg	neg
426	neg	neg
427	neg	neg
429	neg	neg
431	**pos**	-
432	neg	-
433	neg	-
434	neg	neg
436	neg	neg
437	neg	neg
439	neg	-
440	neg	-
441	neg	neg
442	neg	-
443	neg	neg
444	neg	neg
445	neg	-
446	neg	-
449	neg	neg
452	neg	-
454	neg	-
455	neg	-
456	neg	
457	neg	-
458	neg	-
459	neg	-
461	neg	-
462	neg	neg
463	neg	neg
464	neg	-
465	neg	-
466	neg	-
467	neg	neg
468	neg	neg
469	**pos**	neg
470	neg	-
471	neg	-

Table 4.6: Cohort 1-*ALK* and *ROS1* FISH results. Abbreviations: n, sample number; NE; not evaluable; neg; sample negative for ALK or ROS1 rearrangement; pos, sample positive for ALK or ROS1 rearrangement.

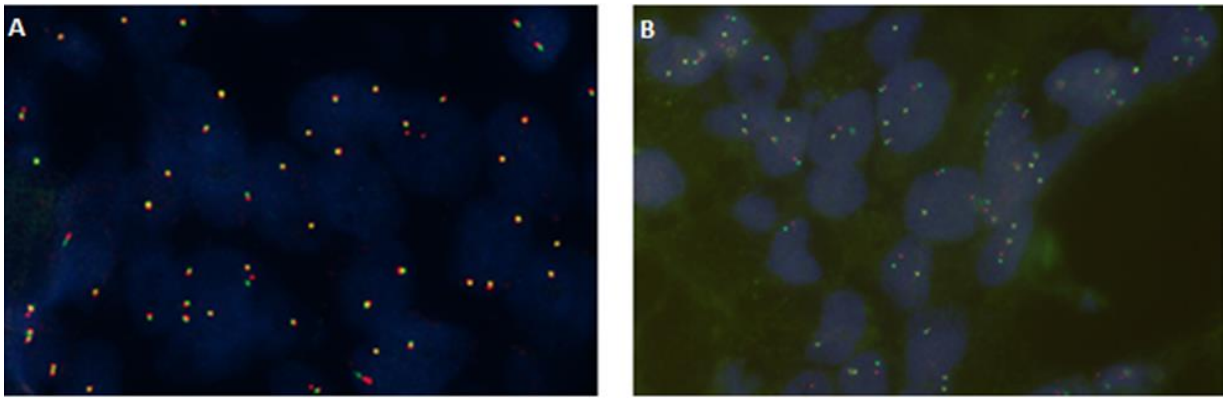

Figure 4.2: FISH analyses for identification of rearrangements in *ALK* gene. A) FISH picture obtained by Bioview© duet3 technology. It represents an *ALK* positive sample: indeed in addition to the normal yellow signals there are also distant red and green fluorescences. B) FISH picture obtained by automated fluorescent microscope (Zeiss© Axioplan 2 Imaging). It represents an *ALK* negative sample: in fact, there are only yellow signals or near red and green fluorescences.

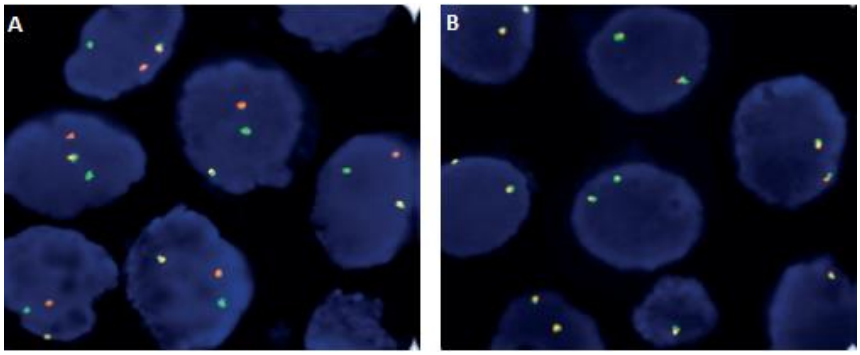

Figure 4.3: FISH analyses for identification of rearrangements in *ROS1* gene. A) FISH picture obtained by automated fluorescent microscope. It represents a *ROS1* positive sample: indeed in addition to the normal yellow signals there are also distant red and green fluorescences. B) FISH picture obtained by automated fluorescent microscope. It represents a *ROS1* negative sample: in fact, there are only yellow signals or near red and green fluorescences. Figures are obtained from ZytoVision brochure.

4.2.3 TTF-1 IHC results

In cohort one the cases with enough residual tissue material were screened for TTF-1 expression by IHC. Consequently we tested 321 samples: 252/321 (78.5%) patients are positive for TTF-1 antibody staining and 69/321 (21.5%) are negative (Table 4.7 and Figure 4.4).

Cohort 1	
n	TTF-1 IHC
1	neg
2	neg
3	**pos**
4	**pos**
5	**pos**
6	neg
7	**pos**
8	neg
9	**pos**
10	**pos**
11	**pos**
12	**pos**
13	**pos**
14	**pos**
15	**pos**
16	**pos**
17	**pos**
18	**pos**
19	**pos**
20	**pos**
21	**pos**
22	neg
23	**pos**
24	neg
25	**pos**
26	**pos**
27	**pos**
28	**pos**
29	**pos**
30	**pos**
31	**pos**
32	**pos**
33	**pos**
34	**pos**
35	**pos**
36	**pos**
37	**pos**
38	**pos**
39	**pos**
40	**pos**
41	**pos**
42	**pos**
43	**pos**
55	**pos**
57	**pos**
63	**pos**
65	**pos**
66	**pos**
67	**pos**
68	**pos**
69	**pos**

Cohort 1	
n	TTF-1 IHC
80	neg
87	**pos**
89	**pos**
90	**pos**
92	**pos**
95	**pos**
97	**pos**
100	**pos**
102	**pos**
105	**pos**
109	**pos**
113	**pos**
114	**pos**
115	**pos**
122	neg
123	**pos**
124	**pos**
126	**pos**
130	neg
135	**pos**
137	**pos**
140	neg
146	**pos**
147	**pos**
148	**pos**
149	**pos**
152	**pos**
154	**pos**
155	**pos**
157	**pos**
159	neg
163	**pos**
165	**pos**
166	**pos**
167	**pos**
168	neg
169	neg
170	neg
171	**pos**
172	**pos**
173	**pos**
174	**pos**
175	neg
177	**pos**
178	**pos**
181	**pos**
183	**pos**
184	**pos**
187	**pos**
192	**pos**
194	**pos**

Cohort 1	
n	TTF-1 IHC
206	**pos**
208	neg
211	**pos**
212	**pos**
213	**pos**
214	**pos**
215	**pos**
216	**pos**
217	**pos**
218	**pos**
219	**pos**
220	neg
222	neg
223	**pos**
224	**pos**
225	**pos**
226	neg
227	neg
228	**pos**
231	**pos**
232	neg
235	**pos**
236	neg
237	neg
238	**pos**
240	**pos**
241	**pos**
243	**pos**
246	**pos**
250	**pos**
251	**pos**
252	**pos**
253	neg
254	**pos**
255	neg
256	**pos**
257	**pos**
258	neg
259	**pos**
260	**pos**
261	neg
262	**pos**
263	neg
264	neg
265	**pos**
266	**pos**
267	**pos**
268	**pos**
269	neg
270	**pos**
271	neg

Cohort 1	
n	TTF-1 IHC
278	**pos**
279	**pos**
281	neg
282	neg
283	neg
284	**pos**
285	**pos**
286	**pos**
289	**pos**
291	**pos**
292	neg
294	**pos**
297	**pos**
298	**pos**
299	**pos**
300	neg
301	neg
302	**pos**
303	**pos**
304	**pos**
306	**pos**
307	**pos**
309	**pos**
310	**pos**
312	**pos**
313	neg
315	neg
316	**pos**
317	**pos**
318	neg
319	**pos**
320	neg
321	**pos**
323	neg
325	**pos**
327	neg
328	neg
329	neg
330	neg
332	**pos**
333	neg
334	**pos**
337	**pos**
338	**pos**
339	**pos**
340	**pos**
341	**pos**
342	**pos**
343	**pos**
344	**pos**
345	**pos**

Cohort 1	
n	TTF-1 IHC
349	**pos**
350	**pos**
351	**pos**
353	**pos**
355	**pos**
356	**pos**
357	**pos**
359	**pos**
360	neg
361	**pos**
363	neg
364	**pos**
365	**pos**
366	**pos**
368	neg
369	neg
370	**pos**
371	**pos**
372	**pos**
373	**pos**
378	**pos**
379	neg
380	neg
382	**pos**
385	**pos**
387	**pos**
388	**pos**
391	**pos**
392	**pos**
393	**pos**
394	**pos**
395	**pos**
396	**pos**
397	**pos**
398	**pos**
399	**pos**
400	**pos**
401	**pos**
402	**pos**
403	**pos**
404	neg
405	**pos**
406	**pos**
407	**pos**
408	**pos**
409	**pos**
410	**pos**
411	**pos**
412	**pos**
413	**pos**
414	**pos**

Cohort 1	
n	TTF-1 IHC
417	**pos**
418	**pos**
419	**pos**
420	**pos**
421	neg
422	neg
423	neg
424	neg
425	neg
426	**pos**
427	**pos**
428	neg
429	**pos**
430	**pos**
431	**pos**
432	**pos**
433	neg
434	**pos**
435	**pos**
436	**pos**
437	**pos**
438	**pos**
439	**pos**
440	**pos**
441	**pos**
442	**pos**
443	**pos**
444	**pos**
445	**pos**
446	**pos**
447	**pos**
448	neg
449	**pos**
450	**pos**
451	**pos**
452	neg
453	neg
454	neg
455	neg
456	neg
457	**pos**
458	**pos**
459	neg
460	**pos**
461	**pos**
462	**pos**
463	**pos**
464	neg
465	**pos**
466	**pos**
467	**pos**

Cohort 1	
n	TTF-1 IHC
72	pos
74	pos
75	pos

Cohort 1	
n	TTF-1 IHC
195	pos
196	pos
199	pos

Cohort 1	
n	TTF-1 IHC
272	pos
275	pos
276	pos

Cohort 1	
n	TTF-1 IHC
346	pos
348	pos

Cohort 1	
n	TTF-1 IHC
415	pos
416	pos

Cohort 1	
n	TTF-1 IHC
468	pos
469	pos

Table 4.7: Cohort 1-TTF-1 IHC results. Abbreviations: n, sample number; neg; sample negative for TTF-1 expression; pos, sample positive for TTF-1 expression.

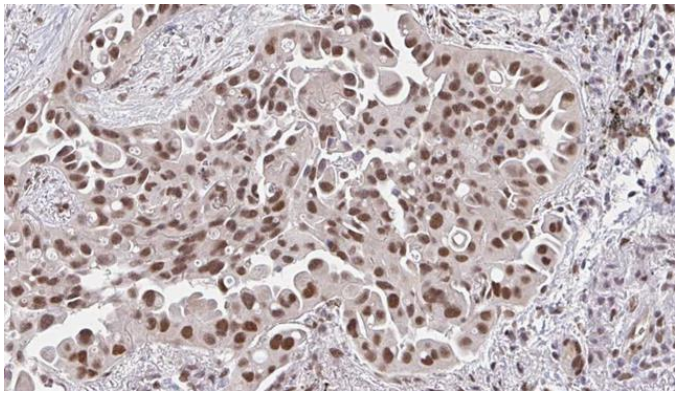

Figure 4.4: Positive TTF-1 IHC staining. Lung AC sample presenting expression of TTF-1 in the cellular nucleus (enlargement: 20X) (Figure obtained from The Human Protein Atlas).

4.2.4 *EGFR* characterization by SensiScreen®

In order to validate SensiScreen® assays, we searched for *EGFR* mutated cell lines in published articles. During the development of SensiScreen®, we repeated several times the sensitivity assays with mutated plasmids and DNA extracted from mutated cell lines, in order to obtain an assay with high sensitivity. After some changes in reagents concentrations and in other main reaction parameters, we succeded in the development of a kit with a LOD between 0.1%, and 1%, corresponding to the detection of until 1 copy of mutant DNA in a wt background. The development of the reagents contained in EGFR SensiScreen® lung was based on the results obtained with KRAS SensiScreen® kit in colorectal cancer. The latter kit is CE IVD and the data obtained for colorectal cancer has already been published by our laboratory, in collaboration with PentaBase ApS (Riva A et al, 2017).

In addition, DS confirmed that DNA extracted from the subcultured lung AC cell lines is really mutated, as found in literature (cell lines mutations are reported in Table 3.2).

After the development of the kits, we proceeded with the validation of *EGFR* SensiScreen® assays on DNA extracted from lung tumoral tissue samples. To do this, we tested

SensiScreen® kit in cohort one and we compared the results with the data obtained by DS. In the first cohort, all the *EGFR* multiplex and simplex assays (i.e. G719A, G719C and G719S in exon 18; exon 19 deletions; exon 20 insertions; S768I and T790M in exon 20; L858R and L861Q in exon 21) confirmed the mutations found by DS (Table 4.8 and Figure 4.5).

More importantly, SensiScreen® assays were able to detect additional mutated cases compared to DS: exon 19 deletions multiplex revealed four new mutated samples (patients number 175, 179, 239 and 287), exon 20 insertions multiplex two new mutated samples (patients number 202 and 315), T790M simplex two new mutated samples (patients number 156 and 256), L858R simplex five new mutated samples (patients number 28, 44, 113, 128 and 310) and L861Q simplex one new mutated sample (patient number 248) (Table 4.8). No additional mutated cases were found by G719 simplex and multiplex and S768I simplex SensiScreen® assays. Samples number 57, 66 and 213, defined as not evaluable by DS, resulted wt with all the developed SensiScreen® kit.

Table 4.8 reports the percentages of the new mutated samples detected by SensiScreen®, compared to DS.

Cohort 1: new mutated samples (DS vs SensiScreen®)				
***EGFR* assay**	**Mutated cases identified by**		**Additional mutated cases identified by SensiScreen®**	
	DS	**SensiScreen®**	**n**	**%**
G719 simpl and mplx (exon 18)	4/468	4/471	0	-
exon 19 deletions mplx	21/468	25/471	4	19%
exon 20 insertions mplx	5/468	7/471	2	40%
T790M simpl (exon 20)	6/468	8/471	2	33%
S768I simpl (exon 20)	4/468	4/471	0	-
L858R simpl (exon 21)	11/468	16/471	5	45%
L861Q simpl (exon 21)	3/468	4/471	1	33%

Table 4.8: Cohort 1-Comparison between DS and SensiScreen® assays. This table reports the number of *EGFR* mutated cases detected by DS or SensiScreen® and the percentages of the new mutated cases with respect to the number of mutated cases by DS. SensiScreen® found 14 additional mutant cases compared to DS. Abbreviations: DS, direct sequencing; mplx, multiplex; n, sample number; simpl, simplex.

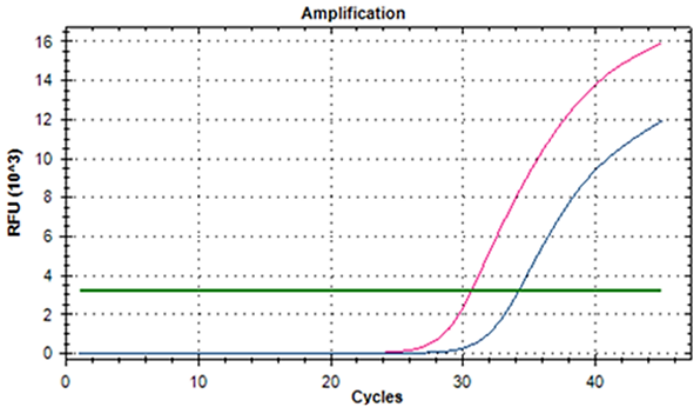

Figure 4.5: Plot resulting from *EGFR* SensiScreen® assays in a mutated patient. The blue curve represents the signal of the mutation (in this case, the T790M mutation) and the pink curve represents the signal of the reference gene. Abbreviations: RFU, relative fluorescence units.

4.3) Cohort two

4.3.1 Plasma analyses

The patients in cohort two (characterized by plasma samples and, in some cases, by the associated tissue or serum sample) were tested by SensiScreen® assays for *EGFR* mutations (i.e. G719A, G719C and G719S in exon 18; exon 19 deletions; exon 20 insertions; S768I in exon 20; L858R and L861Q in exon 21) validated, as aforementioned, in lung tissues and by a SensiScreen® T790M assay especially developed for liquid biopsies. T790M liquid biopsies test has been adapted from tissue assays in order to be more sensible and to be able to detect the highly fragmented and low concentrated DNA contained in plasma samples. This kit has been modified changing the concentration of reagents that are contained in the assays developed for tissue characterization.

In this cohort, the patients were previously characterized by SensiScreen® assays and other methodologies in order to compare the *EGFR* results obtained with different methodologies.

From analyses conducted before our characterization experiments, we have: DS results in samples from 1 to 4 (Table 4.9); IOT® Oncomine cell-free nucleic acids assay data (Thermo Fisher Scientific) in samples 20, 21, 22 and in samples from 86 to 105 (Table 4.9) and TheraScreen® (QIAGEN) characterization in samples from 23 to 85 (Table 4.9).

In addition, in the samples for which we had enough material, or that were not previously characterized by other methodologies different than SensiScreen® (patients from 5 to 22 and patients from 35 to 58), we applied cobas® (Roche) (Table 4.9).

4.3.2 DS, IOT®, TheraScreen® and cobas® results

The results obtained by the other methodologies, necessary for comparisons with SensiScreen® assays, are reported in table 4.9.

Cohort 2: DS, IOT®, TheraScreen® and cobas® *EGFR* (ex) results					
n		**DS**	**IOT®**	**Thera®**	**cobas®**
1	t	**E746_A750 del (19)**	-	-	-
2	p	wt	-	-	-
3	t	**E746_A750 del (19)**	-	-	-
4	p	wt	-	-	-
5	p	-	-	-	**del (19)**
6	p	-	-	-	wt
7	p	-	-	-	wt
8	p	-	-	-	wt
9	p	-	-	-	**del (19)**
10	p	-	-	-	wt
11	p	-	-	-	wt
12	p	-	-	-	wt
13	p	-	-	-	wt
14	p	-	-	-	wt
15	p	-	-	-	wt
16	p	-	-	-	wt
17	p	-	-	-	**del (19) + T790M (20)**
18	p	-	-	-	wt
19	p	-	-	-	wt
20	p	-	wt	-	**wt**
21	p	-	**E746_A750 del (19) + L858R (21)**	-	**del (19) + L858R (21)**
22	p	-	**T790M (20) + L858R (21)**	-	**T790M (20) + L858R (21)**
23	p	-	-	wt	-
24	p	-	-	wt	-
25	p	-	-	wt	-
26	s	-	-	wt	-

Cohort 2: DS, IOT®, TheraScreen® and cobas® *EGFR* (ex) results					
n		**DS**	**IOT®**	**Thera®**	**cobas®**
54	p	-	-	wt	wt
55	p	-	-	wt	wt
56	p	-	-	wt	wt
57	p	-	-	wt	wt
58	p	-	-	wt	wt
59	t	-	-	**del (19) + T790M (20)**	-
60	t	-	-	**del (19) + T790M (20)**	-
61	t	-	-	wt	-
62	t	-	-	wt	-
63	t	-	-	wt	-
64	t	-	-	wt	-
65	t	-	-	**del (19)**	-
66	t	-	-	**del (19)**	-
67	t	-	-	**del (19)**	-
68	t	-	-	**del (19)**	-
69	t	-	-	**del (19)**	-
70	t	-	-	**del (19)**	-
71	t	-	-	**del (19)**	-
72	t	-	-	**del (19)**	-
73	t	-	-	**del (19)**	-
74	t	-	-	**del (19)**	-
75	t	-	-	**del (19)**	-
76	t	-	-	NE	-
77	t	-	-	**L858R (21)**	-
78	t	-	-	**L858R (21)**	-
79	t	-	-	**L858R (21) + S768I (20)**	-

Cohort 2: DS, IOT®, TheraScreen® and cobas® *EGFR* (ex) results					
n		**DS**	**IOT®**	**Thera®**	**cobas®**
27	p	-	-	**L858R (21)**	-
28	s	-	-	**L858R (21)**	-
29	p	-	-	**L858R (21)**	-
30	s	-	-	**L858R (21)**	-
31	p	-	-	**T790M (20) + L858R (21)**	-
32	s	-	-	**T790M (20) + L858R (21)**	-
33	p	-	-	wt	-
34	p	-	-	wt	-
35	p	-	-	**T790M (20)**	**T790M (20)**
36	p	-	-	**T790M (20)**	**T790M (20)**
37	p	-	-	**T790M (20)**	**T790M (20)**
38	p	-	-	**T790M (20)**	**T790M (20)**
39	p	-	-	**T790M (20)**	**T790M (20)**
40	p	-	-	**T790M (20) + L858R (21)**	**T790M (20) + L858R (21)**
41	p	-	-	**T790M (20)**	**T790M (20)**
42	p	-	-	**T790M (20)**	**T790M (20)**
43	p	-	-	wt	wt
44	p	-	-	wt	wt
45	p	-	-	wt	wt
46	p	-	-	NE	NE
47	p	-	-	wt	wt
48	p	-	-	wt	wt
49	p	-	-	wt	wt
50	p	-	-	wt	wt

Cohort 2: DS, IOT®, TheraScreen® and cobas® *EGFR* (ex) results					
n		**DS**	**IOT®**	**Thera®**	**cobas®**
80	t	-	-	**L858R (21) + S768I (20)**	-
81	t	-	-	**L861Q (21)**	-
82	t	-	-	NE	-
83	t	-	-	NE	-
84	t	-	-	NE	-
85	t	-	-	wt	-
86	t	-	wt	-	-
87	t	-	**T790M (20) + L858R (21)**	-	-
88	t	-	**T790M (20) + L858R (21)**	-	-
89	t	-	**E746_A750 del (19)**	-	-
90	t	-	**E746_A750 del (19)**	-	-
91	t	-	**T790M (20) + L858R (21)**	-	-
92	t	-	wt	-	-
93	t	-	**E746_A750 del (19)**	-	-
94	t	-	**E746_A750 del (19) + T790M (20)**	-	-
95	t	-	wt	-	-
96	p	-	**T790M (20) + L858R (21)**	-	-
97	p	-	**T790M (20) + L858R (21)**	-	-
98	p	-	**T790M (20) + L858R (21)**	-	-
99	p	-	wt	-	-
100	p	-	wt	-	-
101	p	-	**T790M (20) + L858R (21)**	-	-
102	p	-	**T790M (20) + L858R (21)**	-	-
103	p	-	**T790M (20) + L858R (21)**	-	-

Cohort 2: DS, IOT®, TheraScreen® and cobas® *EGFR* (ex) results					
n		**DS**	**IOT®**	**Thera®**	**cobas®**
51	p	-	-	wt	wt
52	p	-	-	wt	wt
53	p	-	-	wt	wt

Cohort 2: DS, IOT®, TheraScreen® and cobas® *EGFR* (ex) results					
n		**DS**	**IOT®**	**Thera®**	**cobas®**
104	p	-	**T790M (20) + L858R (21)**	-	-
105	p	-	**T790M (20) + L858R (21)**	-	-

Table 4.9: Cohort 2-*EGFR* results obtained by DS, IOT®, TheraScreen® or cobas® on plasma and some tissue associated samples. By TheraScreen® and by cobas® is not possible to define the specific deletion. Abbreviations: DS, direct sequencing; ex, exon; IOT®, Ion Torrent; n, sample number; NE, not evaluable; p, plasma; s, serum; t, tissue; thera, TheraScreen®; wt, wild-type.

4.3.3 SensiScreen® results

In the second cohort, all the *EGFR* multiplex and simplex tissue kit (i.e. G719A, G719C and G719S in exon 18; exon 19 deletions; exon 20 insertions; S768I in exon 20; L858R and L861Q in exon 21) and the T790M (exon 20) liquid biopsy simplex kit confirmed the mutations found by the other methodologies described in paragraph 4.3.2 (Table 4.9).

In addition, SensiScreen® assays were able to detect additional mutated cases compared to TheraScreen® and cobas®: L858R (exon 21) simplex kit revealed two new mutated sample (patient number 48, 49) (Table 4.10 and 4.11). The five samples (four tissues and one plasma) that were not evaluable by the other methodologies resulted evaluable and wt in SensiScreen® analyses.

Cohort 2 plasma and serum samples: new mutated cases				
***EGFR* assay**	**Mutated cases identified by**		**Additional mutated cases identified by SensiScreen®**	
	Other methodology	**SensiScreen®**	**n**	**%**
G719 simpl and mplx (exon 18)	0/65	0/66	0	-
exon 19 deletions mplx	4/65	4/66	0	-
exon 20 insertions mplx	0/65	0/66	0	-
T790M simpl (exon 20)	17/65	17/66	0	-
S768I simpl (exon 20)	0/65	0/66	0	-
L858R simpl (exon 21)	14/65	16/66	2	14%
L861Q simpl (exon 21)	0/65	0/66	0	-

Table 4.10: Cohort 2-Comparison between the other methodologies and SensiScreen® *EGFR* assays in plasma and serum samples. This table reports the number of mutated cases detected by other methodologies or SensiScreen® and the percentages of the new mutated cases with respect to the total of the mutations found by another technology. SensiScreen® found 2 additional mutant cases compared to TheraScreen®. Abbreviations: mplx, multiplex; n, sample number; simpl, simplex.

Cohort 2 tissue samples: new mutated cases				
***EGFR* assay**	**Mutated cases identified by**		**Additional mutated cases identified by SensiScreen®**	
	Other methodology	**SensiScreen®**	**n**	**%**
G719 simpl and mplx (exon 18)	0/35	0/39	0	-
exon 19 deletions mplx	19/35	19/39	0	-
exon 20 insertions mplx	0/35	0/39	0	-
T790M simpl (exon 20)	6/35	6/39	0	-
S768I simpl (exon 20)	2/35	2/39	0	-
L858R simpl (exon 21)	7/35	7/39	0	-
L861Q simpl (exon 21)	1/35	1/39	0	-

Table 4.11: Cohort 2-Comparison between the other methodologies and SensiScreen® *EGFR* assays in tissue samples. This table reports the number of mutated cases detected by other methodologies or SensiScreen® and the percentages of the new mutated cases. SensiScreen® did not find new additional mutant cases compared to other methodologies but confirms all the results. Abbreviations: mplx, multiplex; n, sample number; simpl, simplex.

4.4) Cohort three

4.4.1 Molecular markers characterization by DS

In the third cohort (n=102) we observed 54 mutated patients (Figure 4.1), corresponding to 52.9% of the cases. *EGFR* alterations were detected in 13/101 patients; *KRAS* mutations in 41/100 patients; *BRAF* mutations in 3/79 patients and *HER2* alterations in 1/81 patients, representing 12.9%, 41%, 3.8% and 1.2% out of the evaluable cases, respectively. One patient is mutated in *EGFR* exon 18, six in exon 19, two in exon 20 and five in exon 21. In the totality of the *EGFR* mutations, equal to fourteen, the different exons are altered in 7.1%, 42.9%, 14.3% and 35.7% of the cases, respectively. Two patients are mutated in *BRAF* exon 11 and one in exon 15. The specific mutations are reported in Table 4.12.

In this cohort we found that five patients harboured two mutations: two sample are mutated in *EGFR* and *KRAS* (i.e. patient 32 in Table 4.12 is D800N in *EGFR* exon 20 and G12C in *KRAS* exon 2; patient 50 in Table 4.12 is Y827F in *EGFR* exon 21 and G12C in *KRAS* exon 2); one sample is mutated in *KRAS* and *BRAF* (i.e. patient 23 in Table 4.12 is G12C in *KRAS* exon 2 and D594G in *BRAF* exon 15); one sample is mutated in *EGFR* and *BRAF* (i.e. patient 94 in Table 4.12 is T847I in *EGFR* exon 21 and G469A in *BRAF* exon 11) and one sample is mutated in two different *EGFR* exons (i.e. patient 59 in Table 4.12 is E746_A750del in exon 19 and T790M in exon 20) (Table 4.12).

n	Cohort 3: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
6	wt	**G12V**	wt	wt
7	wt	**G12C**	wt	wt
8	**E746_A750 del (19)**	wt	NE	NE
10	wt	**G12C**	wt	wt
12	wt	**G12V**	wt	wt
14	wt	**G12V**	wt	wt
15	**E746_A750 del (19)**	wt	NE	NE
17	wt	**G12D**	wt	wt
19	wt	wt	wt	**A775_G 776 ins YVMA**
23	wt	**G12C**	**D594G (15)**	wt
24	wt	**G12C**	wt	wt
28	**E746_A750 del (19)**	wt	wt	wt

n	Cohort 3: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
55	**L747_T751 del (19)**	wt	NE	NE
57	wt	**G12C**	wt	wt
58	wt	**G12D**	wt	wt
59	**E746_A750 del (19) + T790M (20)**	wt	wt	wt
61	wt	**G12C**	wt	wt
63	wt	**G12V**	wt	wt
64	wt	**G12A**	wt	wt
65	wt	**G12D**	wt	wt
71	**L747_S752 del (19)**	wt	wt	wt
73	**L858R (21)**	wt	wt	wt
75	wt	**G12C**	wt	NE
76	wt	**G12C**	wt	wt

n	Cohort 3: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
32	**D800N (20)**	**G12C**	NE	wt
33	wt	**G13R**	wt	wt
35	wt	**G12C**	wt	wt
36	wt	**G12C**	NE	NE
37	wt	**G12D**	wt	wt
39	wt	**G13D**	wt	wt
44	wt	**G12C**	wt	wt
46	wt	**G12A**	wt	wt
48	wt	**G12C**	wt	wt
49	wt	**G12V**	NE	wt
50	**Y827F (21)**	**G12C**	NE	NE
51	**L858R (21)**	wt	wt	wt
52	wt	**G12V**	wt	wt
53	wt	**G13C**	wt	wt
54	wt	**G12C**	NE	wt

n	Cohort 3: samples mutated by Sanger sequencing			
	***EGFR* (ex)**	***KRAS* ex 2**	***BRAF* (ex)**	***HER2* ex 20**
78	wt	**G12D**	wt	wt
79	wt	**G12C**	wt	wt
80	wt	**G13D**	NE	NE
81	wt	**G12C**	wt	wt
85	wt	**G13C**	wt	wt
86	wt	**G12V**	wt	wt
90	**G719A (18)**	wt	wt	wt
91	wt	**G12C**	wt	NE
92	wt	**G12C**	wt	wt
93	wt	wt	**V600E (15)**	wt
94	**T847I (21)**	wt	**G469A (11)**	wt
95	wt	**G12C**	wt	wt
96	**L858R (21)**	NE	NE	NE
97	wt	**G12V**	wt	wt
99	wt	**G12D**	wt	wt

Table 4.12: Cohort 3-samples mutated in *EGFR* (exons 18, 19, 20 and 21), *KRAS* (exon 2), *BRAF* (exons 11 and 15) and *HER2* (exon 20) by Sanger sequencing. This table shows only the patients with mutations. Abbreviations: ex, exon; n, sample number; NE, not evaluable; wt, wild-type. In brackets it is reported the exon that is mutated.

4.4.2 *ALK* and *ROS1* FISH results

In cohort three, *ALK* analyses were done in 31 cases and *ROS1* analyses in 78 cases. Seven cases were not evaluable for *ALK* and twenty for *ROS1*. In this cohort, we found one *ALK* translocation (4.2% out of the evaluable patients) but we did not detect any *ROS1* alterations (Table 4.13 and Figure 4.2).

n	Cohort 3: FISH	
	ALK	***ROS1***
1	neg	-
3	**-**	neg
7	**-**	neg
9	**-**	neg
10	**-**	neg
11	neg	NE
12	**-**	neg
13	**-**	neg
14	neg	-
17	**-**	neg
22	**-**	neg
24	**-**	neg
30	neg	NE

n	Cohort 3: FISH	
	ALK	***ROS1***
39	-	neg
42	neg	-
44	-	neg
46	neg	NE
47	neg	neg
48	**-**	neg
50	**-**	neg
51	**-**	neg
52	**-**	neg
54	**-**	neg
57	**-**	neg
59	**-**	neg
61	**-**	neg

n	Cohort 3: FISH	
	ALK	***ROS1***
66	neg	neg
67	neg	neg
68	**-**	neg
69	**-**	neg
70	**-**	neg
72	neg	neg
73	-	neg
74	neg	neg
75	neg	neg
76	**-**	neg
78	**-**	neg
79	neg	NE
80	neg	neg

n	Cohort 3: FISH	
	ALK	***ROS1***
84	neg	neg
85	neg	neg
86	neg	neg
87	**-**	neg
88	**-**	neg
89	neg	neg
90	**pos**	neg
92	**-**	neg
93	**-**	neg
94	neg	neg
95	**-**	neg
97	**-**	neg
98	**-**	neg

n	Cohort 3: FISH	
	ALK	*ROS1*
31	-	neg
32	-	neg
33	-	neg
35	-	neg

n	Cohort 3: FISH	
	ALK	*ROS1*
62	-	neg
63	-	neg
64	neg	neg

n	Cohort 3: FISH	
	ALK	*ROS1*
81	-	neg
82	neg	neg
83	-	neg

n	Cohort 3: FISH	
	ALK	*ROS1*
99	neg	neg
100	-	neg
102	neg	neg

Table 4.13: Cohort 3-*ALK* and *ROS1* FISH results. Abbreviations: n, sample number; NE; not evaluable; neg; sample negative for *ALK* or *ROS1* rearrangement; pos, sample positive for *ALK* or *ROS1* rearrangement

4.4.3 TTF-1 IHC results

In the cases of cohort three with enough residual tissue material we screened for TTF-1 expression by IHC. Therefore we tested 74 samples. In cohort three, 55/74 (74.3%) are positive for TTF-1 expression, whereas 19/74 cases (25.7%) showed a negative staining to the antibody (Table 4.14 and Figure 4.4).

Cohort 3	
n	TTF-1 IHC
2	**pos**
3	**pos**
7	**pos**
9	**pos**
10	**pos**
11	**pos**
12	neg
13	**pos**
17	**pos**
19	**pos**
22	neg
24	neg
26	neg

Cohort 3	
n	TTF-1 IHC
27	**pos**
29	**pos**
31	**pos**
32	**pos**
33	neg
35	**pos**
39	**pos**
40	**pos**
41	**pos**
43	**pos**
44	**pos**
46	**pos**
47	**pos**

Cohort 3	
n	TTF-1 IHC
48	neg
49	neg
50	neg
51	neg
52	neg
53	**pos**
54	neg
55	neg
56	neg
57	**pos**
59	**pos**
60	**pos**

Cohort 3	
n	TTF-1 IHC
61	neg
62	neg
63	**pos**
64	**pos**
65	**pos**
68	neg
69	**pos**
70	**pos**
71	**pos**
72	**pos**
73	**pos**
74	**pos**

Cohort 3	
n	TTF-1 IHC
75	**pos**
76	**pos**
77	**pos**
78	**pos**
81	neg
82	**pos**
83	**pos**
84	neg
85	neg
86	**pos**
87	**pos**
88	**pos**

Cohort 3	
n	TTF-1 IHC
89	**pos**
90	**pos**
91	**pos**
92	**pos**
93	**pos**
94	**pos**
95	**pos**
97	**pos**
98	**pos**
99	**pos**
100	**pos**
101	**pos**

Table 4.14: Cohort 3-TTF-1 IHC results. Abbreviations: n, sample number; neg; sample negative for TTF-1 expression; pos, sample positive for TTF-1 expression.

4.4.4 ROR1 expression

In the third cohort, we evaluated ROR1 expression, by TaqMan real-time PCR (Applied Biosystems), in 58/102 samples for which we had sufficient tissue material left for the extraction of RNA. Indeed, in some cases, there was not enough tissue for these analyses after

the extraction of the DNA used for DS (paragraph 4.4.1). In addition, in some patients we had not the availability of normal tissue that is necessary for the evaluation of the data.

Two samples (2/58, 1.3% out of the tested samples) resulted not evaluable because the extracted RNA was highly degraded.

The comparison between ROR1 and the housekeeping gene (reference) RN18S in both tumoral and normal tissues was done by applying of the Livak method through the calculation of the $2^{-\Delta\Delta Ct}$ value.

In 16 samples (28.6% out of the evaluable cases) ROR1 resulted overexpressed in tumour compared to normal tissue, presenting a $2^{-\Delta\Delta Ct}$ value > 1. On the contrary, in 40 patients (71.4% out of the evaluable cases) the expression of ROR1 was not significantly different between tumour and normal tissue, indeed $2^{-\Delta\Delta Ct}$ values resulted ≤ 1 (Table 4.15).

Cohort 3: ROR1 results		
n	**$2^{-\Delta\Delta Ct}$**	**ROR1 expression**
2	**1.10**	+
3	0.69	-
7	0.25	-
10	**1.51**	+
12	0.20	-
13	0.72	-
17	0.16	-
19	0.69	-
22	**1.64**	+
24	**3.66**	+
27	**2.50**	+
29	0.57	-
31	0.16	-
35	0.10	-
39	0.15	-
40	0.44	-
43	**1.67**	+

Cohort 3: ROR1 results		
n	**$2^{-\Delta\Delta Ct}$**	**ROR1 expression**
48	0.72	-
51	**1.09**	+
52	0.67	-
53	0.16	-
54	0.10	-
57	**1.03**	+
59	0.50	-
61	0.23	-
62	0.46	-
63	0.99	-
64	0.09	-
68	0.45	-
69	0.11	-
70	0.21	-
71	0.35	-
73	0.71	-
75	0.71	-

Cohort 3: ROR1 results		
n	**$2^{-\Delta\Delta Ct}$**	**ROR1 expression**
78	0.70	-
80	NE	NE
81	0.16	-
82	**1.01**	+
83	0.08	-
85	NE	NE
86	**1.20**	+
87	0.34	-
88	0.35	-
89	0.38	-
90	**1.58**	+
92	**1.21**	+
93	0.29	-
95	0.35	-
97	**1.41**	+
98	**1.97**	+
99	0.21	-

Cohort 3: ROR1 results		
n	**$2^{-\Delta\Delta Ct}$**	**ROR1 expression**
44	0.72	-
46	0.55	-
47	0.75	-

Cohort 3: ROR1 results		
n	**$2^{-\Delta\Delta Ct}$**	**ROR1 expression**
76	0.12	-
77	**1.85**	+

Cohort 3: ROR1 results		
n	**$2^{-\Delta\Delta Ct}$**	**ROR1 expression**
100	**1.54**	+
101	0.51	-

Table 4.15: Cohort 3 - ROR1 expression obtained by TaqMan real-time. The comparison between ROR1 and RN18S (housekeeping gene) in tumour and normal tissue was done applying the Livak method. Abbreviations: Ct, threshold cycle; n, sample number; NE, not evaluable; ROR1, receptor tyrosine kinase-like orphan receptors.

4.4.5 miR-382 expression

In the third cohort, miR-382 expression was evaluated, by TaqMan real-time PCR (Applied Biosystems) in 54/102 samples for which we had sufficient tissue material left for the extraction of RNA. Indeed, in some cases there was not enough tissue for these analyses after the extraction of the DNA used for DS (paragraph 4.4.1) and after the extraction of the RNA used for ROR1 real-time (paragraph 4.4.4). In addition, in some patients we had not the availability of the normal tissue that is necessary for the evaluation of the data. In this cohort two samples (2/54, 3.7% out of the tested samples) resulted not evaluable because the extracted RNA was highly degraded.

Applying the calculation of $2^{-\Delta Ct}$ value we were able to compare miR-382 expression in tumoral and normal tissues. In 25 samples (48.1% out of the evaluable cases) miR-382 resulted overrepresented in tumour compared to normal tissue, presenting a $2^{-\Delta Ct}$ value > 1. On the contrary, in 27 patients (51.9% out of the evaluable cases) the expression of miR-382 was not significantly different between tumour and normal tissue, indeed $2^{-\Delta Ct}$ values resulted ≤ 1 (Table 4.16).

For miR-382 evaluation we considered $2^{-\Delta Ct}$ and not $2^{-\Delta\Delta Ct}$ (obtained comparing the results of the miRNA probe and the housekeeping probe) because the values of the triplicates with the probe of the reference gene were different and could not be considered as reliable or replicable.

Cohort 3: miR-382 results			Cohort 3: miR-382 results			Cohort 3: miR-382 results		
n	**$2^{-\Delta Ct}$**	**miR-382 expression**	**n**	**$2^{-\Delta Ct}$**	**miR-382 expression**	**n**	**$2^{-\Delta Ct}$**	**miR-382 expression**
2	0.10	-	**53**	0.09	-	**80**	1.78	+
3	2.20	+	**57**	1.83	+	**81**	1.63	+
7	7.11	+	**59**	0.20	-	**82**	0.59	-
10	3.07	+	**61**	2.51	+	**83**	1.12	+
12	0.93	-	**62**	0.04	-	**85**	0.21	-
13	1.80	+	**63**	0.11	-	**86**	0.26	-
17	0.07	-	**64**	5.19	+	**87**	0.02	-
22	5.24	+	**68**	0.03	-	**88**	4.28	+
24	0.53	-	**69**	0.07	-	**89**	1.61	+
29	2.41	+	**70**	6.15	+	**90**	0.02	-
31	2.59	+	**71**	NE	NE	**92**	0.06	-
35	1.85	+	**72**	0.72	-	**93**	6.96	+
39	1.37	+	**73**	0.06	-	**95**	5.37	+
44	0.05	-	**74**	1.78	+	**97**	0.84	-
46	0.28	-	**75**	1.53	+	**98**	0.97	-
48	0.10	-	**76**	0.09	-	**99**	1.45	+
51	0.48	-	**77**	1.83	+	**100**	0.14	-
52	1.37	+	**78**	0.20	-	**101**	NE	NE

Table 4.16: Cohort 3 - miR-382 expression obtained by TaqMan real-time. The evaluation of miR-382 expression was made by calculation of the $2^{-\Delta Ct}$ value between tumour tissue and normal tissues. Abbreviations: Ct, threshold cycle; n, sample number; NE, not evaluable.

4.5) Statistical analyses

The data obtained from the analyses of our three cohorts can be compared and can be better understood by applying different statistical analyses.

4.5.1 Correlations between molecular alterations and clinical-pathological data and between the molecular status of the different genes

In the three cohorts, the two-tailed Fisher's exact statistical test between the clinical-pathological characteristics of the patients (sex, age, TNM classification, differentiation

grade) and the molecular data (*EGFR*, *KRAS*, *BRAF*, *HER2* mutations; *ALK* and *ROS1* rearrangements; TTF-1 IHC expression) revealed no correlations.

However the two-tailed Fisher's exact test revealed a statistical, significant correlation between *EGFR* and *KRAS* in cohorts one and three (Table 4.17). In particular the alterations in EGFR and KRAS were nearly mutually exclusive. In addition BRAF and KRAS wt cases tend to be TTF-1 positive. A correlation was considered positive when the p value, calculated with Fisher's test, was minor than 0.05.

Two-tailed Fisher's exact test							
Cohort 1				**Cohort 3**			
p=0.005		***EGFR***		**p=0.02**		***EGFR***	
		mut	**wt**			**mut**	**wt**
KRAS	**mut**	1	109	***KRAS***	**mut**	1	40
	wt	66	292		**wt**	11	48

Two-tailed Fisher's exact test			
Cohort 1			
p=0.0007		***KRAS***	
		mut	**wt**
TTF-1	**pos**	47	183
	neg	26	35

Two-tailed Fisher's exact test			
Cohort 1			
p=0.02		***BRAF***	
		mut	**wt**
TTF-1	**pos**	2	124
	neg	2	57

Table 4.17: Cohort 1 or 3- Correlations demonstrating statistical significance applying the two-tailed Fisher's exact test analyses. A p value minor than 0.05 indicates a correlation between the two compared molecular marker. Abbreviations: EGFR, Epidermal Growth Factor Receptor; KRAS, Kirsten Rat sarcoma viral oncogene homolog; mut, mutated; neg, negative; p, p value; pos, positive; TTF-1, Thyroid Transcription Factor-1 ; wt, wild-type.

4.5.2 Correlation of ROR1 with clinical-pathological data and molecular data of the other molecular markers

In the third cohort, the two-tailed Fisher's exact test was applied to compare ROR1 to clinical-pathological data (sex, age, TNM classification, differentiation grade, treatment, PFS and OS) and no significant correlations were observed.

Furthermore, the comparison between the alterations occurring in the other molecular markers (*EGFR*, *KRAS*, *BRAF*, *HER2* mutations; *ALK* and *ROS1* rearrangements; TTF-1 IHC expression) and ROR1 expression did not show any significant correlation as well, again using the two-tailed Fisher's exact test (Table 4.18).

Two-tailed Fisher's exact test			
p=1		**ROR1**	
		mut	**wt**
KRAS	**mut**	9	16
	wt	11	20

Two-tailed Fisher's exact test			
p=1		**ROR1**	
		mut	**wt**
EGFR	**mut**	2	3
	wt	18	33

Two-tailed Fisher's exact test			
p=1		**ROR1**	
		mut	**wt**
BRAF	**mut**	0	1
	wt	20	35

Two-tailed Fisher's exact test			
p=1		ROR1	
		mut	wt
HER2	mut	0	1
	wt	19	35

Two-tailed Fisher's exact test			
p=1		ROR1	
		mut	wt
ALK	mut	1	0
	wt	2	6

Two-tailed Fisher's exact test			
p=1		ROR1	
		mut	wt
ROS1	mut	0	0
	wt	17	28

Two-tailed Fisher's exact test			
p=1		ROR1	
		mut	wt
TTF-1	pos	16	29
	neg	4	7

Table 4.18: Cohort 3 - Correlation between ROR1 expression and the other molecular markers. In all the cases the p value is major than 0.05 consequently there is no correlation between ROR1 and the alterations or expression of the other genes. Abbreviations: ALK, Anaplastic Lymphoma receptor tyrosine Kinase; EGFR, Epidermal Growth Factor Receptor; KRAS, Kirsten Rat sarcoma viral oncogene homolog; mut, mutated; neg, negative; p, p value; pos, positive; ROR1, receptor tyrosine kinase-like orphan receptors; TTF-1, Thyroid Transcription Factor-1; wt, wild-type.

4.5.3 Association between ROR1 and miR-382

The association between ROR1 and its associated miR-382 was calculated through the Spearman correlation coefficient and the Pearson correlation coefficient. These two values were respectively equal to 0.05 and -0.04, indicating that in our cohort there is neither opposite nor concordant correlation between ROR1 and miR-382 expression because these results are different from 1 and -1 and are close to 0. Also by applying the two-tailed Fisher's exact test there is no correlation between their expression, indeed the p value resulted > 0.05 (i.e. 0.76) (Table 4.19).

Two-tailed Fisher's exact test			
p=0.76		ROR1	
		> 1	≤ 1
miR-382	> 1	7	15
	≤ 1	10	15

Table 4.19: Cohort 3 - Correlation between ROR1 and miR-382 expression. The p value obtained by the two-tailed Fisher's exact test is major than 0.05 consequently there is no correlation between ROR1 and miR-382. We considered values ≤ 1 as negative expression and > 1 as positive expression. Abbreviations: p, p value; ROR1, receptor tyrosine kinase-like orphan receptors.

5. DISCUSSION

Nowadays the care of patients affected by lung cancer represents one of the biggest challenges in medicine because this tumor is characterized by the highest mortality rate (equal to 27% of all cancer deaths) and by an incidence that is second only to sex related tumors (Travis WD et al, 2004; Siegel R et al, 2015; Cancer.org. Cancer facts and figures 2016. American Cancer Society; Travis WD et al, 2016). These statistical data highlight the relevance of the development of new treatments, especially in AC, the most common histotype among lung cancers (Corrin B, 2000; Herbst RS et al, 2008). To date, in patients with advanced lung AC, tumor resection is combined with platinum (cis-Pt or carbo-Pt)-doublet based chemotherapy, but these drugs are limited by their lack of specificity for tumor tissues and by frequent and potentially severe dose-limiting toxicities.

To solve these limitations, in recent times, there has been the development of molecular targeted therapies (Kaneda H et al, 2013). In lung AC, the approved targeted therapies are small-molecule, tyrosine kinase inhibitors (TKIs) acting against *EGFR*, *ALK* and *ROS1* alterations; furthermore, last years have also seen the evaluation, in several clinical trials, of therapies against other mutations (e.g. alterations in *HER2*, *BRAF* and *KRAS* genes) (Tsao AS et al, 2016).

Despite all the successful trials based on targeted therapies, a huge number of cases cannot benefit from the administration of these specific drugs. Indeed, about 50% of lung AC patients display a normal gene sequence for the genes that are currently used or under investigation as predictive markers of response to targeted therapies; in particular only up to 20% of patients carry *EGFR* mutations and, as a consequence, may benefit from 1^{st}-2^{nd} generation TKIs. After treatment with EGFR inhibitors, a specific mechanism of acquired resistance occurs in 50-60% of cases: the creation of the T790M mutation. However, Pharma companies have developed new, specific drugs against this alteration, such as osimertinib (Tagrisso®; AZD9291; Astra Zeneca).

The low rate of lung AC patients that can be treated by the currently approved TKIs is also related to the fact that, sometimes, mutations in the aforementioned genes are missed because of the poor quantity and quality characterizing the majority of lung AC tissue specimens. Indeed, with the current methodologies (e.g. direct sequencing, DS), is really difficult to estimate the molecular profile of the low concentrated material because the sensitivity of this assay is insufficient. Furthermore, the mutations can be skipped due to the fact that lung ACs are an extremely heterogeneous type of tumor, and therefore only a small fraction of cells could show a specific mutation.

The need to find a solution to all these limitations created the basis for my PhD project. In particular, the aim of my research activities was to improve the care of patients affected by lung AC following two strategies: firstly, we tried to enhance the characterization of molecular markers for targeted therapies and, secondly, we characterized the expression of a new putative marker for targeted therapies, ROR1. In this way we hope to enlarge the group of patients that can avoid standard chemotherapies by addressing them to specific targeted therapies.

We improved the characterization of molecular markers for EGFR-targeted therapies through the development and the validation, in collaboration with a Danish company (PentaBase), of a new real-time PCR based assay (EGFR SensiScreen®). This methodology was created in two versions (for DNA obtained from tissue samples and for DNA obtained from liquid biopsies) and the new assays were compared to other methodologies (DS, IOT, TheraScreen® and cobas®) in terms of sensitivity and specificity.

In cohort one, characterized by 471 DNA samples extracted from tissue, we analyzed *EGFR*, *KRAS*, *BRAF* and *HER2* by DS; *EGFR* by SensiScreen®; *ALK*, *ROS1* by FISH and TTF-1 by IHC. The second cohort, characterized by 61 tissue, 39 plasma and 5 serum samples was characterized by cobas® (Roche) for *EGFR* alterations, by SensiScreen® liquid biopsy kit for T790M mutation in *EGFR* exon 20 (PentaBaseApS) and by SensiScreen® tissue kit (PentaBaseApS) for the other *EGFR* mutations. In some cases of cohort two, we also had *EGFR* mutation data from previous TheraScreen® and IOT analyses.

In the first cohort, the characterization of the main lung AC molecular markers and of TTF-1 was necessary to evaluate if our cohort can be considered representative of a standard population. In particular, we detected *EGFR*, *KRAS*, *BRAF* and *HER2* mutations in 14.3%, 23.3%, 1.7% and 0.4% of cases, respectively. In particular, in the group of *EGFR* mutant patients, the different exons (18, 19, 20 and 21) are altered in 5.6%, 39.4%, 21.2% and 33.8% of cases, respectively. By FISH, we found *ALK* and *ROS1* rearrangements in 3.6% and 0.4% of patients, respectively, and TTF-1 nuclear expression was observed in 78.5% of cases. All these rates are superimposable with those reported in the literature for Caucasian population (Rosell R et al, 2012).

Statistical analyses confirmed that *EGFR* and *KRAS* are nearly mutually exclusive (only one case harbours mutations in both genes). In addition, we observed a significant association of *KRAS/BRAF* wild-type sequences and TTF-1 expression: to the best of our knowledge, it is the first time that such correlation has been reported (Gerber DE et al, 2014).

In cohort one is also highlighted the presence of T790M, the most important mechanism of resistance to EGFR TKIs (Murray S et al, 2012; Santarpia M et al, 2017): in particular, 4 cases harbor the T790M mutation alone, demonstrating that T790M can be a mechanism of primary resistance as well. These data confirm the necessity to analyze EGFR exon 20 also in patients that have not already been treated by EGFR TKIs because they could show primary resistance.

Moreover a deeper exam of the results highlights that, in cohort one, the S768I mutation is always associated with other *EGFR* mutations. This association could have a biological significance but, until now, literature reported no explanation to this event and the few reported data concerning this mutation are contradictory (Yang JC et al, 2015; Banno E et al, 2016). Some papers report that patients harbouring the S768I demonstrate a good response to 2ndG TKIs like afatinib (Gilotrif® or Giotrif®; Boehringer Ingelheim, Germany), whereas other articles report that the clinical benefit obtained from EGFR TKIs is higher in NSCLC patients with common *EGFR* mutations than in those with uncommon *EGFR* mutation types (e.g. S768I) (Yang JC et al, 2015; Banno E et al, 2016; Chen K et al, 2017). The difficult interpretation of S768I influence on TKIs benefit could be associated with the fact that this alteration is mainly detected in concomitance with other mutations which define the response to targeted therapies and hide the effect of this alteration. However, we could hypothesize that S768I in trans with the other mutation could give lower response to TKIs. The evidence that, in our cohort, all the patients with the S768I change carry another EGFR mutation, strongly suggests the S768I is insufficient for a full activation of EGFR.

In addition, a small, but not negligible subgroup of our cohort one (0.85% of cases) carries two different oncogene alterations: this finding demonstrates that, in general, oncogene mutations are mutual exclusive in lung cancer, but there are real cases showing a double activation, where the efficacy of targeted therapies may be altered due to this unusual molecular pattern. In fact, EGFR and ALK therapies displays great efficacy when the EGFR or ALK activation, respectively, is detected alone; whereas it is still debated what happens when more mutations are simultaneously present. To date, only few studies have investigated this issue. Literature reports that patients showing concomitant *KRAS* and *EGFR* mutations experience a poorer clinical outcome after treatment with EGFR TKIs (i.e. erlotinib) when compared to cases characterized by *EGFR* mutations only (Eberhard DA, 2005). In addition, another study describes that patients harbouring concomitant *EGFR* mutation and *ALK* rearrangement might still benefit from ALK inhibitors regarding disease stabilization; however, the administration of EGFR TKIs is associated with less efficacy in comparison to

samples mutated only in *EGFR* gene (Sahnane N et al, 2015). The same study shows that the efficacy of TKIs against ALK is unclear in patients carrying a simultaneous *KRAS* mutation and *ALK* rearrangement: the authors show patients who benefit from crizotinib, but also patients who are primarily resistant to such therapy. All these data highlight that double alterations in *EGFR*, *KRAS* and *ALK* genes are not only induced by drug administration, but they may occur in the primary tumor, probably due to tumor heterogeneity, thus reinforcing the importance of the determination of a wide range of genes with NGS and FISH to better define the real efficacy of targeted therapies. Indeed, in the future, other known and unknown genes alterations (e.g. *BRAF* mutations) could determine and influence, beside *KRAS*, the response to these drugs.

Afterwards we used the first cohort for the comparison of DS with the new, more sensitive real-time PCR-based assay, SensiScreen®, in order to improve the molecular diagnosis of *EGFR* mutations. To characterize the new assays, at first we performed sensitivity studies on plasmids comprising sequences with different *EGFR* mutations and then on DNA extracted from *EGFR* mutant cell lines. By applying changes in the concentration and in the chemical structures of SensiScreen® reagents, we succeeded in the development of a new *EGFR* real-time PCR assay characterized by a LOD between 0.1% and 1% (corresponding to the detection of until 1 copy of mutant DNA in a wt background). This value means that we will be able to identify the presence of a specific mutation also in tumor samples where mutant clones are highly dispersed (for tumor heterogeneity or for normal cells infiltration). After validation, we tested the new *EGFR* assay on cohort one. The samples were tested by G719A, G719C, G719S, S768I, T790M, L858R, L861Q simplex kits and by G719 mutations, exon 19 deletions, exon 20 insertions multiplex kits. The new assays confirmed the *EGFR* mutations previously detected by DS and, more importantly, identified 14 additional mutations: in details, 4 new exon 19 deletions, 2 exon 20 insertions, 2 T790M mutations in exon 20, 5 L858R mutations and 1 L861Q mutation in exon 21. The overall percentage of *EGFR* mutations detected by the new assays corresponds to a global increase of 26% of mutated cases. The new methodology did not find additional G719 (exon 18) and S768I (exon 20) mutant cases. Three samples, resulting not evaluable by DS, were defined as wt by SensiScreen® demonstrating that this new real-time PCR assay is able to identify mutations also in samples containing DNA fragments of little dimensions. This feature is related to the minor size of amplicons obtained by the application of this new assay.

The additional mutations discovered by the new real-time PCR assays can be explained by the different LOD of the two methodologies (10-20% for DS vs 0.1-1% for SensiScreen®) and

this evidence suggests the adoption of these new assays in the laboratories of molecular pathology. Indeed, 12 patients of our cohort were addressed to standard chemotherapies instead of EGFR-targeted therapies, which are more efficient in *EGFR*-mutant cases.

We can notice that two new mutant samples (number 156 and 256; Table 4.5) are characterized by another *EGFR* mutation; this indicates that the discrepancy between DS and PentaBase assays sometimes could be due to the tumor heterogeneity since the new methodology, based on a better sensitivity, is able to detect both mutations in the tumor specimen even if one of the alterations is representative of only a little percentage of cancer cells. These two cases resulted, respectively, 746_750del and L858R mutant by DS but the new methodology found also T790M mutation. As a consequence of DS results these patients were treated by 1stG TKIs even if T790M confers resistance. Unfortunately, clinical data for these patients are not available.

In last years, SensiScreen® has been developed and validated also for *BRAF*, *KRAS* and *NRAS* genes in cohorts of colorectal cancer (CRC) and melanoma patients (Riva A et al, 2017). Also with these kits we found, by comparing the results to Sanger sequencing data, new CRC and melanoma mutated samples demonstrating that the higher sensitivity is an advantage of all SensiScreen® assays. Other peculiarities of these new methodologies are the speed and simplicity of execution and the uniformity in the interpretation of results between the different assays of the aforementioned genes.

To sum up, all the advantages of SensiScreen® lung can give benefit to the care of lung AC patients. First of all, the detection of new mutated samples enlarges the cohort of patients that could benefit from the *EGFR* predictive role of response to EGFR TKIs. Indeed the patients wt by DS, but mutated by PentaBase kit, would be not treated with targeted therapies against EGFR or, in contrast, they would be treated with TKIs even if they harbor a mutation of resistance (e.g. T790M). Secondly, patients resulting not evaluable by DS but evaluable by SensiScreen® can avoid a re-biopsy for a new determination of *EGFR* molecular status.

Beside SensiScreen® and DS, a huge number of methodologies are available on the market to characterize *EGFR* mutations in DNA extracted from tissue biopsies. Many studies report a comparison between different approaches and these data could be compared to the new PentaBase kit. An example is the use of cobas® kit for tissue (Roche). A study describes that cobas® (Roche) showed high concordance rates between three laboratory-developed tests (i.e. PCR clamp, PCR invader, Cycleave assays) and permits to find new mutated cases (Nakamura H et al, 2017). However, cobas® (Roche) is less sensitive than PentaBase assays even if they are based on the same methodology (i.e. real-time PCR). This fact is particularly

clear if we observe the different LOD for the detection of T790M mutation in tissue samples (2-3% for cobas® vs 0.1-1% for SensiScreen®). Another assay that could be used for *EGFR* characterization is ddPCR. This methodology has a sensitivity equal to 0.1% comparable to the one of SensiScreen®. Literature reports that ddPCR has a superior analytical performance and equivalent or higher clinical performance than cobas® (Roche) (Kim SS et al, 2018). Consequently, also the new PentaBase assays should be better than cobas® (Roche) for *EGFR* characterization in tumor tissues. Moreover, SensiScreen® can be defined as the better approach because it is easier, faster, and cheaper than ddPCR, a technique that requires highly trained personnel.

All these high-sensitive techniques, including PentaBase, permit to detect very small percentages of *EGFR* mutant cells in the specimens. The presence of a small amount of DNA mutant copies in the sample could be due to either tumor heterogeneity, or to the fact that the majority of lung AC samples are small biopsies or to the possible oversights happening during the tumor area selection. The real clinical value of little clones is still debated. However, some studies report the importance of the determination of allelic alterations, also in a small amount of clones, for an accurate definition of the response to targeted therapies. Indeed, literature reports that the application of highly sensitive methods for *KRAS* mutation detection (e.g. mutant enriched PCR or engineered mutant enriched PCR) may improve the identification of anti-EGFR antibodies resistance, at least in CRC patients (Molinari F et al, 2011).

The previous characterization of *EGFR* by SensiScreen® was done on DNA extracted from FFPE tissues but recent studies described the importance of the definition of the molecular markers gene status also in ctDNA from liquid biopsies, especially in plasma. The characterization of molecular markers on liquid biopsies has a wide range of advantages compared to tissue biopsies: they are not invasive (so analyses can be repeated serially in short times), they better reflect tumor heterogeneity and constant evolution of tumors; they are always possible, they are more cost-effective and ctDNA represents multiple metastatic sites simultaneously (Burrell RA and Swanton C, 2014; Strotman LN et al, 2016). In addition, liquid biopsies provide a good substitute to tissue biopsies when the tissue sample cannot be collected because of cancer localization or patient's unhealthy conditions.

Beside these advantages, liquid biopsies show two main disadvantages. Firstly, they cannot be used for the detection of molecular markers alterations in lesions located in brain or bones, since the blood-brain barrier and the low vascularisation of bone tissue cause a significant decrease or absence of ctDNA in plasma. Secondly, ctDNA is, generally, highly fragmented

and low concentrated (Burrell RA and Swanton C, 2014; Strotman LN et al, 2016), thus requiring methodologies characterized by high sensitivity.

The little number of disadvantages is completely negligible compared to the huge number of advantages, indeed the prospective of a better care for patients affected by cancer brought to the development of analyses on liquid biopsies. Consequently, we decided to develop and validate on plasmids and on *EGFR* mutant cell lines, in collaboration with PentaBase ApS, a real-time PCR assay for T790M analyses on liquid biopsies (SensiScreen® T790M liquid biopsy assay).

On the basis of the high aggressiveness of lung AC, we choose to start the development of the assay on the most important mutation of resistance of this histotype in order to be able to monitor the development of resistance to EGFR targeted therapies in short times.

First of all, sensitive assays on plasmids and cell lines reported that also this kit for liquid biopsies has LOD of 0.1%, corresponding to the detection of until 1 copy of mutant DNA in a wt background. After the validation we applied this assay, on plasma and serum samples of a second cohort. In addition, we evaluated plasma, serum and tissue samples of cohort two by the simplex and multiplex tissue kit specific for the different *EGFR* alterations described previously. In order to confirm SensiScreen® results, we applied also cobas® (Roche) methodology in some samples for which we had enough material and we compared the data. Interestingly, some patients were previously characterized by DS, IOT Oncomine cell-free nucleic acids assay (Thermo Fisher Scientific) or TheraScreen® (QIAGEN). In plasma, serum and tissue samples, SensiScreen® confirmed all the data obtained by the other methodologies and, compared to TheraScreen® (QIAGEN), it was able to evaluate four samples that were defined as not evaluable. In addition in two plasma samples it succeeded in the detection of new L858R mutations compared to TheraScreen® and cobas® (Roche) (corresponding to 14% of cases). These two patients, classified as L858R negative by cobas® (Roche) and TheraScreen® (QIAGEN) techniques on DNA from plasma, were subjected to a subsequent tissue biopsy for the molecular characterization, although the clinical conditions of the patients were not completely good. In tissue, cobas® (Roche), TheraScreen® (QIAGEN) and SensiScreen® assays demonstrated the presence of the L858R change. So, we can conclude that the application of SensiScreen® may prevent the use of tissue biopsies, which could be the source of side effects for the patients. Moreover, liquid biopsies testing influences the costs because they are cheaper and can avoid the repetition of tissue biopsies. The aforementioned data demonstrate that SensiScreen® is a good technique also for testing in liquid biopsies and can find a higher number of mutated cases than other methodologies. The

discovery of new mutated samples and the ability to characterize also the specimens that were not evaluable by other techniques can improve, influence and change the decision concerning the administration of targeted treatments and it is due to the higher sensitivity of SensiScreen®. Indeed the minor size of amplicons obtained by PentaBase assays define an LOD equal to 0.1-1%; whereas the sensitivities of DS, TheraScreen® (QIAGEN) and cobas® (Roche) are 9%, 1% and 2-5% respectively (Sholl LM et al, 2016; Sorber L et al, 2016). It is important to underline that, in all the methodologies, sensitivity differs on the basis of the mutation that is analysed and generally T790M mutation assays present the lowest sensitivity (Thress KS et al, 2016).

SensiScreen® should be preferable than Oncomine cell-free nucleic acids assay (Thermo Fisher Scientific), even if the LOD of IOT (Thermo Fisher Scientific), equal to up to 0.1%, is similar to the sensitivity of PentaBase assays (Sholl LM et al, 2016; Sorber L et al, 2016). Indeed, even if NGS approach gives information about multiple genes in one run, we have to consider that SensiScreen® data analyses and experiments are easier, faster (it takes only 90 minutes) and are less expensive in both tissue and liquid samples. The decision of the better approach to test liquid biopsies is complicated but we could define that NGS appears to be the most appropriate technique for use at the time of diagnosis. Subsequently, in the follow up period, we could concentrate our attention in only one gene and we could use a candidate-gene approach involving mutation-specific techniques (e.g. SensiScreen®, TheraScreen® or cobas®) in order to benefit from their advantages (i.e. low costs, simpler and faster workflow, easier data management). The determination of the methodology to apply is based on the sensitivity but depends also on the kind of mutation that we want to study. For example, the detection of C797S mutation is not possible with cobas® assay but is possible by IOT, but at the moment its clinical relevance is marginal because there are no drugs targeting such alteration.

Beside the aforementioned assays, applied in the second cohort of my project, many other tests can be considered for *EGFR* characterization in liquid biopsies. Some innovative assays are ddPCR, and BEAMing qPCR. These methodologies are defined by a LOD that is comparable to SensiScreen but they need trained personnel and they are more difficult, more expensive than PentaBase assays. Many papers report comparisons between NGS, ddPCR, BEAMing qPCR demonstrating that these assays characterize the same quantity of mutations. In literature these assays identify higher rates of mutated cases compared to cobas® (Roche) (Thress KS et al, 2015; Sholl LM et al, 2016; Sorber L et al, 2016; Feng Q et al, 2018). As a consequence we can hypothesize that a comparison between SensiScreen and ddPCR or

BEAMing qPCR could give a higher correspondence, for what concerns the mutated cases, than cobas (Roche) because these methodologies, like IOT (ThermoFisher), have a high LOD that is similar to the new PentaBase assays. Consequently, PentaBase results could find a higher correspondence with the data obtained by methods characterized by high sensitivity and specificity (e.g. IOT, ddPCR, BEAMing qPCR) than with cobas® (Roche) or TheraScreen® (QIAGEN).

The second aim of my research project was the investigation of a new molecular marker for targeted therapies. In particular, we decided to study ROR1, a transmembrane protein belonging to the receptor TK family whose endogenous ligand has not been discovered yet. Its expression is peculiar because it is found during embryo-development but it is not present in normal adult tissues (Rebagay G et al, 2012). However, recent studies report ROR1 expression in tumor tissues (e.g. B-CLL, B-ALL, MCL) and in breast cancer cell lines (Baskar S et al, 2008; DaneshManesh AH et al, 2008; Shabani M et al, 2008; Zhang S et al, 2012). ROR1 oncogenic activity is related to PI3K-AkT-mTOR pathway activation that brings to the enhancement of cellular migration, proliferation and survival (Rebagay G et al, 2012). Interestingly, one paper reports, in NSCLC, a significant association between ROR1 and TTF-1, the most important AC marker (Yamaguchi T et al, 2012). The ROR1 expression profile detected in cancers and its demonstrated association with TTF-1 led us to the investigation of its role as a new putative marker of response to targeted therapies in lung AC. This hypothesis arise from the fact that targeted therapies developed against ROR1 should be effective and specific only against tumour cells and consequently should not act against normal cells, because this receptor is found only in tumor cell lines and tissues (Baskar S et al, 2008; DaneshManesh AH et al, 2008; Shabani M et al, 2008; Rebagay G et al, 2012; Zhang S et al, 2012).

To determine whether ROR1 could be a promising new marker in lung AC, we evaluated its expression in the third cohort of this project. For this purpose, we selected a different cohort, characterized by 102 lung AC cases, because we need to test a group of patients with clear OS, PFS and treatment data. The patients were characterized for *EGFR*, *KRAS*, *BRAF*, *HER2*, *ALK* and *ROS1* molecular alterations by DS or FISH. These assays confirmed that, as previously described for cohort one, the third cohort is a representative population with rates of mutations similar to those described in the literature (Takeuchi T et al, 2006; Rikova K et al, 2007; Herbst RS et al, 2008; Michaloglou C et al, 2008; Gerber DE et al, 2014). In the third cohort TTF-1 expression was evaluated by IHC and 74.3% out of the evaluable cases

showed nuclear expression of this protein demonstrating the importance of this transcriptional factor as a fundamental marker for lung AC histotype (Rosell R et al, 2012).

The molecular and clinical-pathological data correlations were investigated through the application of the two-tailed Fisher's exact statistical test. Statistics failed to demonstrate significant correlations between the clinical-pathological characteristics and the molecular data, if we exclude the mutual exclusivity of *EGFR* and *KRAS* mutations.

Then, ROR1 expression was tested by TaqMan real-time PCR and the results were compared in normal and tumour tissues of the same patient by the application of the Livak method. The analyses were done only on the samples for which we had sufficient tissue material left for RNA extraction. ROR1 overexpression was detected in 28.6% out of the evaluable cases. The data obtained by ROR1 real-time demonstrated that this gene could be a new molecular marker for specific targeted therapies (now under evaluation in B-ALL, B-CLL, MCL and breast cancer cell lines and on animal models), because its mRNA is found in a not negligible number of lung AC samples, compared to or even higher than EGFR (Baskar S et al, 2008; DaneshManesh AH et al, 2008; Shabani M et al, 2008).

Even if our results are promising, we found an expression rate that is lower compared to the values reported in the literature: indeed, a single study, published in 2012, demonstrated that ROR1 (as evaluated by IHC) is expressed in 77% of lung cancers (Zhang S et al, 2012). This discrepancy can be explained by the different methodologies that we applied, or by the different ethnicity of the group analysed in literature. In fact, Zhang and colleagues evaluated ROR1 by IHC but we decided to estimate ROR1 expression by real-time PCR because the interpretation is more objective and because, to date, no commercial, standardized ROR1 antibodies have been developed (Zhang and colleagues used an home-made antibody). Furthermore, IHC and real-time PCR analyse ROR1 expression in two different molecular levels, protein and mRNA (respectively); as a consequence, expression data could be discordant between these two assays due to different mechanisms of regulation. Another fact that could explain the different rates of ROR1 expression between our data and those included in the paper of Zhang and colleagues is that this study investigated an Asian cohort whereas our experiments were conducted on a Caucasian cohort. Indeed, diversities between Asian and Caucasian populations exist in the mutational landscapes of lung AC: for example, in Asian patients the percentage of lung cancer ACs characterized by *EGFR* mutations is significantly higher compared to Caucasian patients (i.e. 70% in Asian patients compared to 20% in Caucasian population) and a similar discrepancy cannot be excluded also for ROR1 expression (Vargas AJ et al, 2016).

On the other hand, ROR1 expression is not correlated with any other molecular or clinical-pathological data, if we exclude a tendency of ROR1 overexpression in advanced stage cases, and no significant correlation was demonstrated between ROR1 and PFS, OS or treatment.
Noticeably, we found no correlation even between ROR1 and TTF-1. This could be due to the little number of patients for which we have assessed ROR1 and TTF-1 at first, or to the presence in the literature of only one study concerning the association between these two markers. In addition, also the ethnic differences could justify the discrepancy, indeed the values reported on the paper written by Yamaguchi T and colleagues were obtained in an Asian cohorts whereas we investigated Caucasian patients (Yamaguchi T et al, 2012).
To conclude, in the third cohort we decided to examine in depth the mechanism of ROR1 regulation in lung AC, starting from a recent study where the role of miRNA was evaluated with respect to ROR1 expression (Tan H et al, 2016). This study reports that a particular miRNA, miR-382, inhibits ROR1 through its hybridization with the mRNA of this tyrosine kinase receptor. In particular, the Authors highlighted that miR-382 binds ROR1 mRNA in the ROR1 3`-UTR region. In addition, they found that, in cell line models, the overexpression of miR-382 suppresses proliferation, migration and invasion in ovarian cancer cells (Tan H et al, 2016). These results are correlated to the role of tumor suppressor gene exploited by miR-382 through the inhibition of ROR1.
On these bases, we evaluated miR382 expression in our third cohort, and we found that in 48.1% out of the evaluable cases the miR-382 resulted over-represented in tumour cells if compared to normal tissue cells. This finding may justifies the low number of cases overexpressing ROR1 in our cohort, although statistical analyses (Sperman and Pearson correlation coefficient, two-tailed Fisher's exact test) did not show any significant correlation between ROR1 and miR-382 expression. This datum, however, deserves confirmation in a larger cohort, since we evaluated only 58 cases.
The relevance of miR-382 is related to its role in ROR1 regulation but may also play a role in the clinical setting, since miR-382 overexpressing patients should be less responsive to ROR1 therapies because this miRNA inhibits the oncogenic role of this tyrosine kinase receptor. In addition, miR-382 and ROR1 could be also prognostic markers because it has been highlighted that patients affected by cancer with high expression levels of ROR1 (i.e.: ovarian tumor) have a higher rate of relapse and a shorter median survival than patients expressing low levels of ROR1 (Zhang S et al, 2014).
In conclusion in my PhD project we developed a new more sensitive methodology for *EGFR* characterization in tissue and plasma samples and we studied ROR1 as a putative marker of

new molecular targeted therapies, which can enlarge the number of patients potentially treatable with targeted therapies. Our data suggest that the assays developed in collaboration with a Danish company may be introduced in clinical practice, with evident benefit for patients and for laboratories. As for ROR1, the characterization of this marker needs to be confirmed in larger cohorts, but our data are promising, especially if ROR-1 targeted therapies will be introduced into the clinical setting. Indeed a consistent fraction of lung AC (at least one fourth of Caucasian people, on the basis of our preliminary data) may be considered for the treatment with such therapies.

6. REFERENCES

- Anastas JN et al *"WNT signalling pathways as therapeutic targets in cancer"*, Nat Rev 13: 11-26 (2013).
- Ayeni D et al *"Emerging Agents and New Mutations in EGFR-Mutant Lung Cancer"*, Clin Cancer Res 21(17): 3818-3820 (2015).
- Bai H et al *"Epidermal growth factor receptor mutations in plasma DNA samples predict tumor response in Chinese patients with stages IIIB to IV non-small-cell lung cancer"*, J Clin Oncol 27(16): 2653-9 (2009).
- Banno E et al *"Sensitivities to various epidermal growth factor receptor-tyrosine kinase inhibitors of uncommon epidermal growth factor receptor mutations L861Q and S768I: What is the optimal epidermal growth factor receptor-tyrosine kinase inhibitor?"*, Cancer Sci 107(8): 1134-1140 (2016).
- Baskar S et al *"Unique cell surface expression of receptor tyrosine kinase ROR1 in human B-cell chronic lymphocytic leukemia"*, Clin Cancer Res 14(2): 396-404 (2008).
- Batson S et al *"Tyrosine kinase inhibitor combination therapy in first-line treatment of non-small-cell lung cancer: systematic review and network meta-analysis"*, Onco Targets and Therapies 10: 2473-2482 (2017).
- Bergethon K et al *"ROS1 rearrangements define a unique molecular class of lung cancers"*, J Clin Oncol 30(8): 863-870 (2012).
- Bettegowda C et al, *"Detection of circulating tumor DNA in early- and late-stage human malignancies"*, Sci Transl 6(224): 224ra24 (2014).
- Bicocca VT et al *"Crosstalk between ROR1 and the Pre-B cell receptor promotes survival of t(1;19) acute lymphoblastic leukemia"*, Cancer Cell 22: 656-667 (2012).
- Bidard FC et al *"Clinical validity of circulating tumor cells in patients with metastatic breast cancer: a pooled analysis of individual patient data"*, Lancet Oncol 15(4): 406-14 (2014).
- Bos M et al *"Activated RET and ROS: two new driver mutations in lung adenocarcinoma"* Transl Lung Cancer Res 2(2): 112-121 (2013).
- Burrell RA and Swanton C *"Tumour heterogeneity and the evolution of polyclonal drug resistance"* Mol Oncol 8(6): 1095-1111 (2016).
- Cappuzzo F et al *"Increased HER2 gene copy number is associated with response to gefitinib therapy in epidermal growth factor receptor-positive non-small-cell lung cancer patients"*, J Clin Oncol 23(22): 5007-5018 (2005).

- Chabon JJ et al *"Circulating tumor DNA profiling reveals heterogeneity of EGFR inhibitor resistance mechanisms in lung cancer patients"*, Nat Commun 7: 11815 (2016).
- Chen K et al *"Uncommon mutation types of epidermal growth factor receptor and response to EGFR tyrosine kinase inhibitors in Chinese non-small cell lung cancer patients"* Cancer Chemother Pharmacol 80(6): 1179-1187 (2017).
- Cheng F et al *"Circulating tumor DNA: a promising biomarker in the liquid biopsy of cancer"*, Oncotarget 7(30): 48832-48841 (2016).
- Chong CR and Janne PA "*The quest to overcome resistance to EGFR-targeted therapies in cancer"*, Nature medicine 19: 1389-1400 (2013).
- Christensen UB and Pedersen EB *"Intercalating nucleic acids containing insertions of 1-O (1-pyrenylmethyl)glycerol: stabilisation of dsDNA and discrimination of DNA over RNA"*, Nucleic Acids Res 30: 4918–4925 (2002).
- Christensen UB et al *"Intercalating nucleic acids: the influence of linker length and intercalator type on their duplex stabilities"*, Nucleosides Nucleotides and Nucleic Acids 23: 207–225 (2004).
- Ciuleanu T et al *"Efficacy and safety of erlotinib versus chemotherapy in second-line treatment of patients with advanced, non-small-cell lung cancer with poor prognosis (TITAN): a randomised multicentre, open-label, phase 3 study"*, Lancet Oncol 13(3): 300-308 (2012).
- Corrin B *"Pathology of the lung"*, 1st ed. Churchill Livingston (2000).
- Crowley E et al *"Liquid biopsy: monitoring cancer-genetics in the blood"*, Nat Rev Clin Oncol 10(8): 472-84 (2013).
- Cufer T et al "*Systemic therapy of advanced non-small-cell lung cancer: major-developments of the last 5-years"*, Eur J Cancer 49(6): 1216-1225 (2013).
- DaneshManesh AH et al *"ROR1, a cell surface receptor tyrosine kinase is expressed in chronic lymphocytic leukemia and may serve as a putative target for therapy"*, Int J Cancer 123: 1190-1195 (2008).
- Davare MA et al *"Foretinib is a potent inhibitor of oncogenic ROS1 fusion proteins"*, Proc Natl Acad Sci 110(48): 19519-19524 (2013).
- Dave H et al *"Restricted cell surface expression of receptor tyrosine kinase ROR1 in pediatric B-lineage acute lymphoblastic leukemia suggests targetability with therapeutic monoclonal antibodies"* Plos One 7(12): e52655 (2012).

- De la Bellacasa PR et al *"ALK and ROS1 as a joint target for the treatment of lung cancer: a review"*, Transl Lung Cancer Res 2(2): 72-86 (2013).
- Detterbeck FC *"The Eighth Edition Lung Cancer Stage Classification"*, CHEST 151(1): 193-203 (2017).
- Diaz LA and Bardelli A *"Liquid biopsies: genotyping circulating tumor DNA"*, J Clin Oncol 32(6): 579-586 (2014).
- Doebele RC et al *"Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small-cell lung cancer"*, Clin Cancer Res 18(5): 1472-1482 (2012).
- Douillard J et al *"First-line gefitinib in Caucasian EGFR mutation-positive NSCLC patients: a phase-IV, open-label, single-arm study"*, Br J Cancer 110(1): 55-62 (2014).
- Eberhard DA et al *"Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib"*, J Clin Oncol 23(25): 5900-5909 (2005).
- Feng Q et al *"A comparison of QuantStudio^{TM} 3D Digital PCR and ARMS-PCR for measuring plasma EGFR T790M mutations of NSCLC patients"*, Cancer Management and Research 10: 115-121 (2018).
- Forman D et al *"Cancer incidence in five continents"*, IARC Press, Vol X (2013).
- Francisci S et al *"Survival patterns in lung and pleural cancer in Europe 1999–2007: Results from the EUROCARE-5 study",* Eur J Cancer Vol 51, Issue 15: 2242-2253 (2015).
- Gainor JF and Shaw AT *"Emerging paradigms in the development of resistance to tyrosine kinase inhibitors in lung cancer"*, J Clin Oncol 31(31): 3987-3996 (2013).
- Garrido Castro AC and Felip E *"HER2 driven non-small-cell lung cancer (NSCLC): potential therapeutic approaches"*, Transl Lung Cancer Res 2(2): 122-127 (2013).
- Gautschi O et al *"A patient with BRAF V600E lung adenocarcinoma responding to vemurafenib"*, J Thorac Oncol 7(10): e23-24 (2012).
- Gautschi O et al *"Lung adenocarcinoma with BRAF G469L mutation refractory to vemurafenib"*, Lung Cancer 82(2): 365-367 (2013).
- Gautschi O et al *"Targeted therapy for patients with BRAF-mutant lung cancer: results from the European EURAF Cohort"*, J Thorac Oncol 10(10): 1451-1457 (2015).

- Gentile A et al *"ROR1 is a pseudokinase that is crucial for MET-driven tumorigenesis"*, Cancer Res 71(8): 3132-3141 (2011).
- Gerber DE et al, 2014 ASCO Educational Book: e353-e365 (2014).
- Goldstraw P et al *"The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer"*, J Thorac Oncol 11 (1): 39-51 (2015).
- Goto K et al *"Epidermal growth factor receptor mutation status in circulating free DNA in serum: from IPASS, a phase III study of gefitinib or carboplatin/paclitaxel in non-small-cell lung cancer"* J Thorac Oncol 7(1): 115-21(2012).
- Herbst RS and Lippman SM, *"Molecular signatures of lung cancer--toward personalized therapy"*, N Engl J Med 365(1): 76-78 (2007).
- Herbst RS et al *"Lung cancer"*, N Engl J Med 359(13): 1367-1380 (2008).
- Hurvitz SA et al *"Trastuzumab emtansine in the treatment of HER-2-positive metastatic breast cancer patients"*, J Clin Oncol 31(9), 1157–1163 (2013).
- Jänne PA et al *"Selumetinib plus docetaxel for KRAS-mutant advanced non-small-cell lung cancer: a randomised, multicentre, placebo-controlled, phase 2 study"*, Lancet Oncol 14: 38-47 (2013).
- Jänne PA et al *"AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer"*, N Engl J Med 372(18): 1689-1699 (2015).
- Jia Y et al *"Overcoming EGFR (T790M) and EGFR (C797S) resistance with mutant-selective allosteric inhibitors"*, Nature 534: 129–132 (2016).
- Jianjigian YY et al *"Dual inhibition of EGFR with afatinib and cetuximab in kinase inhibitor-resistant EGFR-mutant lung cancer with and without T790M mutations"*, Cancer Discov 4(9): 1036-1045 (2014).
- Joshi M et al *"Trametinib with or without vemurafenib in BRAF mutated non-small-cell lung cancer"*, Plos One 10(2): e0118210 (2015).
- Kaneda H et al *"Molecularly targeted approaches herald a new era of non-small-cell lung cancer treatment"*, Cancer Manag Res 5: 91-101 (2013).
- Kaneda T et al "*Possible differential EGFR-TKI efficacy among exon 19 deletional locations in EGFR-mutant non-small-cell lung cancer"*, Lung Cancer 86(2): 213-218 (2014).

- Karachaliou N et al *"ROR1 as a novel therapeutic target for EGFR-mutant non-small-cell lung cancer patients with the EGFR T790M mutation"*, Transl Lung Cancer Res 3(3): 122-130 (2014).
- Katayama R et al *"Mechanisms of acquired crizotinib resistance in ALK-rearranged lung Cancers"*, Sci Transl Med 4(120): 120ra17 (2012).
- Kim DW et al *"ARIAD's investigational medicine brigatinib demonstrates 54 percent confirmed objective response rate and 12.9-month median progression-free survival in ALTA study*", J Clin Oncol 30(suppl): abstr7533 (2012).
- Kim DW et al "*Activity and safety of ceritinib in patients with ALK-rearranged non-small-cell lung cancer (ASCEND-1): updated results from the multicentre, open-label, phase 1 trial"*, Lancet Oncol 17(4): 452-463 (2016).
- Kim TM et al *"Mechanisms of acquired resistance to AZD9291: a mutation-selective, irreversible EGFR inhibitor"*, J Thorac Oncol 10: 1736-1744 (2015).
- Kim SS et al *"Droplet digital PCR-based EGFR mutation detection with an internal quality control index to determine the quality of DNA"*, Sci Rep 8(1): 543 (2018).
- Klemm F et al *"β-catenin-indipendent WNT signaling in basal-like breast cancer and brain metastases"*, Carcinogenesis 32(3): 434-442 (2011).
- Ko B et al "EGFR T790M: revealing the secrets of a gatekeeper", Lung Cancer: Targets and Therapy 8: 147-159 (2017).
- Koudelakova V et al *"Non-small-cell lung cancer--genetic predictors"*, Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub 157(2): 125-136 (2013).
- Lee VH et al *"Association of exon 19 and 21 EGFR mutation patterns with treatment outcome after first-line tyrosine kinase inhibitor in metastatic non-small-cell lung cancer"*, J Thorac Oncol 8(9): 1148-1155 (2013).
- Leprieur EG et al *"Spotlight on crizotinib in the first-line treatment of ALK-positive advanced non-small-cell lung cancer: patients selection and perspectives"*, Lung Cancer: Targets and Therapies 7: 83-90 (2016).
- Livak KJ and Schmittgen TD *"Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method"*, Methods 25(4): 402-408 (2001).
- Malapelle U et al *"Next-generation sequencing techniques in liquid biopsy: focus on non-small-cell lung cancer patients"*, Transl Lung Cancer Res 5(5): 505-510 (2016).

- Mao C et al *"KRAS mutations and resistance to EGFR-TKIs treatment in patients with non-small-cell lung cancer: a meta-analysis of 22 studies"*, Lung Cancer 69(3): 272-278 (2010).
- Martin V et al *"An overview of the epidermal growth factor receptor fluorescence in situ hybridisation challenge in tumor pathology"*, Br J Cancer 108:668-675 (2013).
- Massarelli E et al *"KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer"*, Clin Cancer Res 13(10): 2890-2896 (2007).
- Mazières J et al "*Lung cancer patients with HER2 mutations treated with chemotherapy and HER2-targeted drugs: results from the European EUHER2 cohort"*, Annal Oncol 27(2): 281-286 (2015).
- Michaloglou C et al *"BRAF(E600) in benign and malignant human tumors"*, Oncogene 27(7): 877-895 (2008).
- Miller KD et al *"Cancer treatment and survivorship statistics, 2016"*, CA Cancer J Clin 66(4): 271-289 (2016).
- Minari R et al *"Third-generation epidermal growth factor receptor-tyrosine kinase inhibitors in T790M-positive non-small-cell lung cancer: review on emerged mechanisms of resistance",* Transl Lung Cancer 5(6): 695-708 (2016).
- Mitsudomi T and Yatabe Y *"Epidermal growth factor receptor in relation to tumor development: EGFR gene and cancer"*, FEBS J 277(2): 301-308 (2010).
- Mitsudomi T et al *"AZD9291 in pre-treated T790M positive advanced NSCLC: AURA2 phase II study"*, J Thorac Oncol 10(9): S320 (2015).
- Mok TS et al *"Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma"*, N Engl J Med 361(10): 947-957 (2009).
- Molinari F et al *"Increased detection sensitivity for KRAS mutations enhances the prediction of anti-EGFR monoclonal antibody resistance in metastatic colorectal cancer"*, Clin Cancer Res 17: 4901-4914 (2011).
- Moran DM et al *"KRAS mutation status is associated with enhanced dependency on folate metabolism pathways in non-small-cell lung cancer cells"*, Cancer Res 72: LB-499 (2014).
- Mukohara T et al *"Differential effects of gefitinib and cetuximab on non-small-cell lung cancers bearing epidermal growth factor receptor mutations"*, J Natl Cancer Inst 97: 1185-1194 (2005).

- Murray S et al *"Molecular predictors of response to tyrosine kinase inhibitors in patients with Non-Small-Cell Lung Cancer"*, J Exp Clin Cancer Res 31(77): 1-10 (2012).
- Nakamura H et al *"Accuracy of the cobas EGFR mutation assay in Non-small-cell lung cancer compared with three laboratory developed tests"*, Clin Lung Cancer S1525-7304(17)30309-1 (2017).
- Niederst MJ et al *"The allelic context of the C797S mutation acquired upon treatment with third generation EGFR inhibitors impacts sensitivity to subsequent treatment strategies"*, Clin Cancer Res 21(17): 3924-3933 (2015).
- Nokihara H et al *"Alectinib (ALC) versus crizotinib (CRZ) in ALK inhibitor naïve ALK-positive non-small cell lung cancer (ALK+ NSCLC): primary results from the J-ALEX study"*, J Clin Oncol 34(Suppl): abstr9008 (2016).
- Noronha V et al *"Epidermal growth factor receptor exon 20 mutation in lung cancer: types, incidence, clinical feature and impact on treatment",* Oncotargets and Therapy 10: 2903-2908 (2017).
- Novello S et al *"Metastatic non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up"*, Annal Oncol 27(5): v1-v27 (2016).
- Ortiz-Cuaran S et al *"Heterogeneous mechanisms of primary and acquired resistance to third generation EGFR inhibitors"*, Cancer Res 22: 4837-4847 (2016).
- Ou SH et al *"Alectinib in crizotinib-refractory ALK-rearranged non-small-cell lung cancer: a phase II global study"*, J Clin Oncol 34(7): 661-668 (2016).
- Oxnard GR et al *"Association between plasma genotyping and outcomes of treatment with osimertinib (AZD9291) in advanced non-small-cell lung cancer"*, J Clin Oncol 34(28): 3375-82 (2016).
- Paik PK et al *"Clinical characteristics of patients with lung adenocarcinomas harbouring BRAF mutations"*, J Clin Oncol 29(15): 2046-2051 (2011).
- Pao W et al *"Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain"*, PLoS Med 2(3): e73 (2005).
- Park JH et al *"Activation of the IGF1R pathway potentially mediates acquired resistance to mutant-selective 3rd-generation EGF receptor tyrosine kinase inhibitors in advanced non-small-cell lung cancer"*, Oncotarget 7(16): 22005-22015 (2016).

- Passaro A et al *"Personalized treatment in advanced ALK-positive non-small-cell lung cancer: from bench to clinical practice"*, Oncotargets and Therapy 9: 6361-6376 (2016).
- Pereira E et al *"Personalized circulating tumor DNA biomarkers dynamically predict treatment response and survival in gynecologic cancers"*, Plos One 10(12):e0145754 (2015).
- Piotrowska Z et al *"Heterogeneity underlies the emergence of EGFRT790 wild-type clones following treatment of T790M-positive cancers with a third-generation EGFR inhibitor"*, Cancer Discov 5 :713-722 (2015).
- Planchard D et al *"EGFR-independent mechanisms of acquired resistance to AZD9291 in EGFR T790M-positive NSCLC patients"*, Annal Oncol 26(10): 2073-2078 (2015).
- Rebagay G et al *"ROR1 and ROR2 in human malignancies: potentials for targeted therapy"*, Frontiers in Oncology Vol 2 Article 34 (2012).
- Ridge CA et al "Epidemiology of lung cancer", Semin Intervent Radiol 30: 93-98 (2013).
- Rikova K et al *"Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer"*, Cell 131: 1190-1203 (2007).
- Riva A et al *"SensiScreen® KRAS exon 2-sensitive simplex and multiplex real-time PCR-based assays for detection of KRAS exon 2 mutations"*, Plos One 12(6): e0178027 (2017).
- Rosell R et al *"Screening for epidermal growth factor receptor mutations in lung cancer",* N Engl J Med 361(10): 958-67 (2009).
- Rosell R et al *"Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial"*, Lancet Oncology 13(3): 239-246 (2012).
- Rosell R and Karachaliou N *"Lung cancer: using ctDNA to track EGFR and KRAS mutations in advanced-stage disease"*, Nat Rev 13: 401-402 (2016).
- Rossi A et al *"The role of cetuximab and other epidermal growth factor receptor monoclonal antibodies in the treatment of advanced non-small cell lung cancer"*, Rev Recent Clin Trials 3: 217-227 (2008).

- Saad N et al "*Epidermal growth factor receptor T790M mutation-positive metastatic non-small-cell lung cancer: focus on osimertinib (AZD9291)*", OncoTargets and Therapy 10: 1757-1766 (2017).
- Sacher AG et al *"Prospective validation of rapid plasma genotyping for the detection of EGFR and KRAS mutations in advanced lung cancer"*, JAMA Oncol 2(8): 1014-1022 (2016).
- Sandler A et al *"Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer"*, N Engl J Med 355(24): 2542-2550 (2006).
- Santarpia M et al *"Osimertinib in the treatment of non-small-cell lung cancer: design, development and place in therapy"*, Lung Cancer: Targets and Therapies 8: 109-125 (2017).
- Sequist LV et al *"Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations"*, J Clin Oncol 31(27): 3327-3334 (2013).
- Seto T et al *"Erlotinib alone or with bevacizumab as first-line therapy in patients with advanced non-squamous non-small-cell lung cancer harbouring EGFR mutations (JO25567): an open-label, randomised, multicentre, phase 2 study"*, Lancet Oncol 15(11): 1236-1244 (2014).
- Shabani M et al *"Expression profile of orphan receptor tyrosine kinase (ROR1) and Wilms' tumor gene 1 (WT1) in different subsets of B-cell acute lymphoblastic leukemia"*, Leuk Lymphoma 49: 1360-1367 (2008).
- Sahnane N et al *"EGFR and KRAS mutations in ALK-positive lung adenocarcinomas: biological and clinical effect",* Clin Lung Cancer 17(1): 56-61 (2016).
- Sharma VS et al *"Epidermal growth factor receptor mutations in lung cancer"*, Nat Rev 7(3): 169-181 (2007).
- Shaw AT et al *"Alectinib in ALK-positive, crizotinib-resistant, non-small-cell lung cancer: a single-group, multicentre, phase 2 trial"*, Lancet Oncol 17(2): 234-242 (2016).
- Shepherd FA et al *"Lung cancer in 2013: state of the art therapy for metastatic disease"*, ASCO Educational Book 2013: 339-346 (2013).
- Sholl LM et al *"Liquid biopsy in lung cancer"*, Arch Pathol Lab Med 140: 825-829 (2016).
- Siegel R et al *"Cancer statistics, 2015"*, Cancer J Clin 65: 5–29 (2015).

- Song H-N et al *"Acquired C797S mutation upon treatment with a T790M-specific third-generation EGFR inhibitor (HM61713) in non-small-cell lung cancer"*, J Thorac Oncol 11(4): e45-e47 (2016).
- Sorber L et al *"Circulating cell-free nucleic acids and platelets as a liquid biopsy in the provision of personalized therapy for lung cancer patients"*, Lung Cancer 107:100-107 (2016).
- Spagnolo F et al *"BRAF-mutant melanoma: treatment approaches, resistance mechanisms, and diagnostic strategies"*, Onco Targets Ther 8:157-68 (2015).
- Strotman LN et al *"Liquid Biopsies in Oncology and the Current Regulatory Landscape",* Mol Diagn Ther 20(5): 429-436 (2016).
- Sun JM et al *"Prognostic and predictive value of KRAS mutations in advanced non-small cell lung cancer"*, Plos One 8(5): e64816 (2013).
- Takeuchi T et al *"Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors"*, J Clin Oncol 24(11): 1679-1688 (2006).
- Takezawa K et al *"HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR mutant lung cancers that lack the second-site EGFR T790M mutation"*, Cancer Discov 2(10): 922-33 (2012).
- Tan H et al *"miR-382 inhibits migration and invasion by targeting ROR1 through regulating EMT in ovarian cancer"* Int J Oncol 48: 181-190 (2016).
- Thress KS et al *"EGFR mutation detection in ctDNA from NSCLC patient plasma: A cross-platform comparison of leading technologies to support the clinical development of AZD9291"*, Lung Cancer 90(3): 509-15 (2015).
- Toyokawa G and Seto T *"Updated evidence on the mechanisms of resistance to ALK inhibitors and strategies to overcome such resistance: clinical and preclinical data"*, Oncol Res Treat 38(6): 291-298 (2015).
- Travis WD et al, *"WHO/IASLC Classification of Tumor"*, Lyon: IARC Press (2004).
- Travis WD et al, *"WHO/IASLC Classification of Tumor"*, Lyon: IARC Press (2016).
- Tsao AS et al *"Scientific advances in lung cancer 2015"*, J Thorac Oncol 11(5): 613-638 (2016).
- Uchibori K et al *"Brigatinib combined with anti-EGFR antibody overcomes osimertinib resistance in EGFR-mutated non-small-cell lung cancer"*, Nat Commun 8 : 14768 (2017).

- Walter AO et al *"Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC"*, 3(12): 1404-1415 (2013).
- Weiler D et al *"Rapid response to trastuzumab emtansine in a patient with HER2-driven lung cancer"*, J Clin Oncol 10(4): p e16-e17 (2015).
- Wei P et al *"KRAS mutation is a weak, but valid predictor for poor prognosis and treatment outcomes in NSCLC: a meta-analysis of 41 studies"*, Oncotarget 7(7): 8373-8388 (2016).
- Wu Y-L et al *"Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC"*, Lancet Oncol 15(2): 213-222 (2014).
- Yam I et al *"EGFR array: uses in the detection of plasma EGFR mutations in non-small-cell lung cancer patients"*, J Thorac Oncol 7(7): 1131-40 (2012).
- Yamaguchi T et al *"NKX2-1/TITF1/TTF-1-induced ROR1 is required to sustain EGFR survival signaling in lung adenocarcinoma"*, Cancer Cell 21(3): 348-361 (2012).
- Yang JC et al *"Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: a combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6"*, Lancet Oncol 16(7): 830-838 (2015).
- Yun C-H et al *"The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP"*, Proc Natl Acad Sci USA 105(6): 2070-2075 (2008).
- Yoshida A et al *"Comprehensive analysis of RET and ROS1 rearrangement in lung adenocarcinoma"*, Modern Pathol: 1-10 (2013).
- Zarogoulidis K et al *"The role of second-line chemotherapy in small cell lung cancer: a retrospective analysis"*, Onco Targets Ther 6: 1493-500 (2013).
- Zhang H et al, *"ROR1 expression correlated with poor clinical outcome in human ovarian cancer"*, Sci Rep 4: 5811 (2014).
- Zhang S et al, "*The onco-embrionic antigen ROR1 is expressed by a variety of human cancers"*, Am J Pathol 181(6): 1903-10 (2012).
- Zhang S et al, "Ovarian cancer stem cells express ROR1, which can be targeted for anti-cancer-stem-cells-therapy", Proc Natl Acad Sci USA 111: 17266-17271 (2014).
- Zhao Z et al, "Response to crizotinib in a lung adenocarcinoma patient harbouring a novel SLC34A2-ROS1 fusion variant", Onco Targets and Therapy 10: 4129-4133 (2017).

- Zheng D et al, *"Plasma EGFR T790M ctDNA status is associated with clinical outcome in advanced NSCLC patients with acquired EGFR-TKI resistance"*, Sci Rep 6(20913): 1-9 (2016).
- Zhu Y et al, *"Estimation of cell-free circulating EGFR mutation concentration predicts outcomes in NSCLC patients treated with EGFR-TKIs"*, Oncotarget 8(8): 13195-13205 (2017).

- 8th edition of the TNM classification for lung cancer proposed by the IASLC
- Cancer.org. Cancer facts and figures 2016. American Cancer Society; 2016. Available from: http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2016/
- cobas® protocol. Available from: http://egfrmutationtestv2.roche.com/
- EMA website. Available from: http://www.ema.europa.eu/
- IARC website. Available from: https://ww.iarc.fr/
- MyCancerGenome: http://www.mycancergenome.org/content/disease/lungcancer/
- NCCN website. Available from: http://www.nccn.org/
- National Cancer Institute at the National Institutes of Health website. Available from: https://www.cancer.gov/about-cancer/
- The Human Protein Atlas. Available from: http://www.proteinatlas.org/
- Two-tailed Fisher's exact test. Available from: http://in-silico.net/statistics/

Acknowledgements

A big gratitude goes to Prof Luca Mazzucchelli for giving me the chance to carry on my PhD project at the Institute of Pathology in Locarno.
Many thanks to Dr Milo Frattini for all the teachings, for the time dedicated to the correction of my thesis and, first of all, for showing me how to work with dedication and interest.
I am grateful to Prof Renzo Boldorini for helping me during all the PhD period and for the correction of my thesis. Thank you very much to Dr Giulio Rossi and to Dr Umberto Malapelle for the evaluation of my PhD project.
Many thanks to all the people working in the laboratory of molecular pathology: Sandra, Francy, Vitto, Sara, Paola, Anto, Je, Giulia, Meli, Roby and Nicoló. In particular I would like to say *"thank you"* to Sandra for all the things that she explained to me, for all the support, for the funny moments and for demonstrating me how to work in lab with passion and enthusiasm. She will always be an example for me.
Thank you to my big strange family because they have always supported my choices.
I am grateful to Dany, my love, for filling every moment of my life with joy, for encouraging me in everything and for demonstrating me that life is unpredictable
A very special *"thank you"* goes to my crazy "little" sister Sophy, one of the main reasons of my life, for transforming every moment into a beautiful coloured rainbow and for demonstrating me how to be proud of myself.

I dedicate this thesis to my special mum and dad…thank you because you have always believed in my abilities…because you never give up…but, first of all, thank you for all your love…I owe it all to you.