Inês Filipa Paixão Mansinhos

Detection of New Actionable Mutations in Lung Cancer Precision Therapy

Dissertação de Mestrado em Bioquímica, orientada pelo Doutor Hugo Prazeres e pela Professora Doutora Paula Veríssimo, apresentada ao Departamento de Ciências e Tecnologia da Universidade de Coimbra

Agosto 2017

· U C ·

UNIVERSIDADE DE COIMBRA

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Detection of New Actionable Mutations in Lung Cancer Precision Therapy

Dissertação apresentada à Universidade de Coimbra para cumprimentos dos requisitos necessários à obtenção do grau de Mestre em Bioquímica, realizada sob a orientação científica do Doutor Hugo João Marques Prazeres (Laboratório Patologia Molecular, Serviço de Anatomia Patológica IPOCFG, EPE) e pela Professora Doutora Paula Cristina Veríssimo Pires (Departamento de Ciências da Vida, Faculdade de Ciência e Tecnologia, Universidade de Coimbra).

Inês Filipa Paixão Mansinhos

2017

Abstract

Lung cancer is the most common cause of cancer death around the world, in both sex. About 80%–85% of lung cancer cases are non-small-cell lung cancer (NSCLC) patients, the remaining 15%–20% are small-cell lung cancer (SCLC). NSCLC is divided into three categories called: adenocarcinoma, squamous-cell carcinoma and large cell carcinoma. Among them, adenocarcinoma cases account for around 40-50% of NSCLC patients. The prognosis for NSCLC is low with a five-year survival rate of less than 20%, and is even worse for SCLC with a five-year survival rate of less than 5%.

For a long time, the first-line treatments have been surgery, chemotherapy or radiotherapy. However, the discovery of several oncogenic driver mutations in patients with NSCLC, adenocarcinoma cases in particular, has allowed the development of personalized treatments based on these specific molecular alterations. Therefore, EGFR (Epidermal Growth Factor Receptor) mutations account for up to 15% of adenocarcinoma and primarily occurred in the tyrosine kinase (TK) domain of the gene. More than 80% of these mutations consist of in-frame deletions in exon 19 and the L858R point mutation in exon 21. Such mutations induced a constitutive activation of EGFR, making it a potential therapeutic target. Thus, EGFR-mutated patients can benefit from a specific first-line treatment specifically the TK inhibitors (TKI) that competitively inhibits fixation of adenosine triphosphate (ATP) in the catalytic binding site of TK domain. Other driver biomarkers in lung cancer have also been proposed and some of them might provide additional information for clinical decision-making.

In this way, the main goal of this project was to evaluate mutations in other potentially actionable targets – MET and ERBB2– in patients with adenocarcinoma by Sanger sequencing and to develop a Real Time PCR multiplex assay for rapid sensitive assessment of mutation profile in tissue and plasma. This assay, will also give the opportunity to monitor the evolution of mutational status in the plasma during the treatment for the prediction of relapse and control the appearance of clones with resistance mutations.

Of the total of 172 samples, 161 (88.9%) were classified as negative for alterations in exons 18, 19, 20 and 21 of EGFR, whereas 19 samples (11.1%) were classified as positive.

In total of the 19 alterations in EGFR, 73.7% were deletions in exon 19 and 21% was related to Leu858Arg mutation. A case of a T790M alteration was also founded in a patient. At a lower frequency, a case of a Leu861Gln was also reported. In MET gene, the same 172 samples were, also, analyzed. Of these, 9 samples (5.2%) harbored alterations in MET gene, including 2 intronic variants, 2 indel mutations and 5 pontual mutations in exon 14. ERBB2 alterations were analyzed in 69 samples and one case of an insertion of 12 bases in exon 20 were detected.

This work allowed us to conclude that an important proportion of cases harbors mutations in MET and ERBB2 and these patients could potentially be treated with approved drugs for these targets.

RESUMO

O cancro de pulmão é a causa mais comum de morte por cancro, em todo o mundo, em ambos os sexos. Cerca de 80% a 85% dos casos de cancro de pulmão são pacientes com cancro de pulmão de não pequenas células (CPNPC), sendo os restantes 15% -20% cancro de pulmão de pequenas células (CPPC). O CPNPC é dividido em três grupos: adenocarcinoma, carcinoma de células escamosas e carcinoma de células grandes. Entre eles, os casos de adenocarcinoma representam cerca de 40 a 50% dos pacientes com CPNPC. O prognóstico para CPNPC é baixo, com uma taxa de sobrevivência de cinco anos inferior a 20% sendo esta, ainda, pior para o CPPC, com uma taxa de sobrevivência de cinco anos inferior a 5%.

Durante muito tempo, os tratamentos de primeira linha foram a cirurgia, a quimioterapia ou a radioterapia. No entanto, a descoberta de várias mutações *drivers* da carcinogénese em pacientes com CPNPC, especialmente em casos de adenocarcinoma, permitiu o desenvolvimento de tratamentos personalizados com base nessas alterações moleculares específicas. Deste modo, as mutações no EGFR (Recetor do Fator de Crescimento Epidérmico) representam até 15% dos adenocarcinomas e ocorrem principalmente no domínio tirosina quinase (TK) do gene. Mais de 80% dessas mutações consistem em deleções *in-frame* no exão 19 e na mutação pontual L858R no exão 21. Tais mutações induzem uma ativação constitutiva do EGFR, tornando-se um potencial alvo terapêutico. Assim, os pacientes portadores de mutações no EGFR podem beneficiar de um tratamento específico de primeira linha, mais especificamente, de inibidores de TK (TKI) que, de forma competitiva, inibem a fixação da adenosina trifosfato (ATP) ao local de ligação catalítica do domínio TK. Foram, também, propostos outros *driver* biomarcadores em cancro de pulmão podendo, alguns deles, fornecer informações adicionais para a tomada de decisões clínicas.

Desta forma, o principal objetivo deste projeto foi avaliar mutações noutros alvos potencialmente acionáveis - MET e ERBB2 - em pacientes com adenocarcinoma, através da sequenciação de Sanger e desenvolver um ensaio multiplex de PCR em tempo real, para uma rápida e sensível avaliação do estado mutacional em tecido e em plasma. Este ensaio também dará a oportunidade de monitorizar a evolução do estado mutacional

no plasma durante o tratamento, para a predição de recidiva e controlo do aparecimento de clones com mutações de resistência.

Do total de 172 amostras, 161 (88,9%) foram classificadas como negativas para alterações nos exões 18, 19, 20 e 21 do EGFR, enquanto 19 (11,1%) foram classificadas como positivas. No total das 19 alterações encontradas no EGFR, 73,7% foram deleções no exão 19 e 21% relataram a mutação Leu858Arg. Um caso de uma alteração T790M foi, também, encontrado num paciente. Numa frequência mais baixa, um caso Leu861Gln também foi relatado. No gene MET, as mesmas 172 amostras foram, igualmente, analisadas. Destas, 9 amostras (5,2%) apresentaram alterações no gene, incluindo 2 variantes intrónicas, 2 mutações indel e 5 mutações pontuais, no exão 14. As alterações no ERBB2 foram analisadas em 69 amostras, tendo sido detetado um caso de inserção de 12 bases no exão 20.

Este trabalho permitiu concluir que uma proporção importante de casos apresenta mutações no MET e ERBB2, sendo que tais pacientes poderiam ser tratados com fármacos aprovados para esses alvos.

LIST OF FIGURES

Figure 1. Distribution of oncogenic driver mutations in lung adenocarcinoma……29

Figure 2. Pathways activated by EGFR binding……30

Figure 3. Location of EGFR mutations……33

Figure 4. Timeline of discovery in lung cancers harboring alterations of the MET pathway……36

Figure 5. The MET receptor and selected MET pathway-directed targeted therapies……36

Figure 6. Activating MET splice site mutation……38

Figure 7. Members of the ERBB/HER superfamily of cell-surface receptor tyrosine kinase……42

Figure 8. Location of the major mutation site and the ATP-binding pocket of the ERBB2 kinase domain in the modeled ERBB2–Afatinib complex structure……46

Figure 9. Diagram of the most frequent alterations of MET exon 14……67

Figure 10. Electropherograms of two cases of intronic variants in MET gene……69

Figure 11. Electropherograms of two cases of silent mutations in codon 3024 of exon 14 MET gene……70

Figure 12. Electropherograms of one case of a silent mutation in codon 3015 of 14 MET gene……71

Figure 13. Electropherograms of one case of a nonsense mutation in codon 3009 of 14 MET gene……72

Figure 14. Electropherograms of one case of a missense mutation in codon 3013 of 14 MET gene……74

Figure 15. Electropherograms of two cases of indel mutations in exon 14 of MET gene……75

Figure 16 - Diagram of the most frequent alterations of ERBB2 exon 20……76

Figure 17. Electropherograms of one case of an insertion in exon 20 of ERBB2……77

Figure 18. Flowchart summarizing the procedure used for designing the assay and the primers and probes for the genotyping real-time PCR protocol……78

Figure 19. TaqMan assay……82

LIST OF TABLES

Table 1. Summary of case reports of patients with MET exon 14 and responses to MET treatment……..40

Table 2. Targeted drug therapies, their mechanisms, and available clinical trials in HER2 positive NSCLC……..45

Table 3. List of primers used in PCR……59

Table 4. EGFR mutations identified in 19 patients with routine genetic testing……66

Table 5. List of primers designed for RT-PCR……81

Table 6. List of probes designed for RT-PCR……81

Table 7. Fluorophores used for RT-PCR assay……83

CONTENT

ABSTRACT 5

RESUMO 7

LIST OF FIGURES 9

LIST OF TABLES 11

LIST OF ABBREVIATIONS 15

1. INTRODUCTION 21

1.1 Cancer as a leading cause of death 21

1.1.1 Lung cancer 21

1.1.1.1 Epidemiology 21

1.1.1.2 Histology 23

1.1.1.2.1 Small cell lung cancer 23

1.1.1.2.2 Non-small cell lung cancer 24

1.1.1.2.2.1 Adenocarcinoma 24

1.1.1.2.2.2 Squamous Cell Carcinoma 25

1.1.1.2.2.3 Large Cell Carcinoma 25

1.2 Molecular alterations in NSCLC 26

1.2.1 Molecular alterations in Adenocarcinoma 28

1.2.1.1. EGFR Mutations as a paradigm of personalized treatment in Adenocarcinoma 29

1.3 New Targetable Oncogenes in Non–Small-Cell Lung Cancer 34

1.3.1 MET 35

1.3.2 ERBB2 41

2. AIMS 51

3. MATERIALS AND METHODS 55

3.1 Samples 55

3.2 Methods 55

3.2.1 DNA extraction 55

3.2.1.1 DNA extraction from Paraffin-embedded tissues 56

3.2.1.2 DNA extraction from Bronchoalveolar lavage 57

3.2.1.3 DNA extraction from Plasma 59

3.2.2 DNA concentration and Purity 59

3.2.3 PCR Amplification 60

3.2.4 DNA integrity 61
3.2.5 Sanger sequencing 62
4. RESULTS AND DISCUSSION 65
4.1 MET Mutation Analysis 67
4.2 ERBB2 Mutation Analysis 76
4.3 Design of a multiplex RT-PCR assay 77
4.3.1 Mapping of the target sequences 79
4.3.2 Primer & Probe design 79
4.3.3 Genotyping Assay using TaqMan probes 82
4.3.4 Multiplexing and use of Fluorophores 83
5. CONCLUSION AND FUTURE PERSPECTIVES 87
6. BIBLIOGRAPHY 89
Appendix I

LIST OF ABBREVIATIONS

A	Adenine
A	Absorbance
ACS	American Cancer Society
Ala	Alanine
ALK	Anaplastic lymphoma kinase
AJCC	*American Joint Committee on Cancer*
AKT1	AKT serine/threonine kinase 1
Arg	*Arginine*
ARAF	Serine/threonine-protein kinase A-Raf
BDGP	Berkely Drosophila Genome Project
BRAF	Serine/threonine-protein kinase B-Raf
C	Cytosine
C	Cysteine
CDK4	Cyclindependent kinase 4 gene
CLCGP	*Clinical Lung Cancer Genome Project*
CVD	Cardiovascular diseases
D	Aspartate
Del	Deletion
ddNTPs	Dideoxynucleotides
DDR2	Discoidin Domain Receptor Tyrosine Kinase 2
E	Glutamate
EB	Elution buffer
EDTA	Ácido etilenodiamino tetra-acético
EGF	Hepatocyte growth factor
EGFR	Epidermal growth factor receptor
ERBB	Erb-B1 Receptor Tyrosine Kinase
Ex	Exon
FDA	Food and Drug Administration
FGFR	Fibroblast Growth Factor Receptor 1

FFPE	Formalin-fixed, paraffin-embedded tissues
G	Glycine
G	Guanine
Gln	Glutamine
HER	Human Epidermal Growth Factor Receptor
HRAS	HRas proto-oncogene, GTPase
IARC	International Agency For Research On Cancer
Ivs	InterVening Sequence
Ins	insertion
IPT	Immunoglobulin-plexin transcription
KRAS	Kirsten rat sarcoma viral oncogene homolog
L	Leucine
Leu	Leucine
M	Methionine
mAb	Monoclonal antibody
MDM2	Murine double minute gene
MEK-1	Mitogen-activated protein kinase kinase 1
MET	Mesenchymal-epithelial transition factor
mTOR	Mammalian target of rapamycin
NGM	*Network Genomic Medicin*e
NRAS	NRAS proto-oncogene, GTPase
NRG1	Neuregulin 1
NSCLC	Non-small cell lung cancer
OD	Optical density
P	Proline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase Chain Reaction
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
Pro	Proline
PSI	Plexin semaphoring integrin domain

PTB	Phosphotyrosine-binding
PTEN	Phosphatase and tensin homolog
R	Arginine
RET	Proto-oncogene tyrosine-protein kinase receptor
RIT1	Ras like without CAAX 1
ROS	ROS proto-oncogene 1, receptor tyrosine kinase
RT-PCR	Real-time Polymerase Chain Reaction
S	Serine
SCC	Squamous cell carcinoma
SCLC	Small cell lung cancer
Ser	Serine
SH2	Src homology 2
T	Thymine
TGFα	Transforming growth factor alpha
TK	Tyrosine kinase
TKI	Tyrosine kinase inhibitor
TNM	Tumour, node and metastasis
TP53	Tumor protein p53
Tyr	Tyrosine
UICC	*Union Internationale Contre le Cancer*
UVs	Unclassified variant
V	Valine
Val	Valine
X	Codon Stop
Y	Tyrosine
WHO	World Health Organization

Chapter I

1. INTRODUCTION

1.1 Cancer as a leading cause of death

Cancer is the result of genetic variations induced by multiple carcinogenic factors and can be generally described as a disease caused by uncontrolled cell growth, which can invade and spread to distant sites of the body (Hanahan & Weinberg, 2011). It is subdivided in several types that are characterized by their localization and tissue of origin (Zhang *et al.*, 2016). Lung, prostate, colorectal, stomach, and liver cancer are the most common types of cancer in men, while breast, colorectal, lung, uterine cervix, and stomach cancer are the most common among women (WHO, 2011). Despite cardiovascular diseases (CVD) are still the deadliest diseases worldwide causing 17.3 million deaths in 2008 (cancer: 7.3 million (2008)), cancer is the second leading cause of death, with around 8.2 million cases of death in 2012 (IARC, 2012). Nowadays there are more cancer-related than CVD-related deaths among men in the 10 European countries. It is predicted that the total numbers of new cancer cases will increase to 21.4 million by 2030 (WHO, 2011).

1.1.1 Lung cancer

1.1.1.1 Epidemiology

Until the 1930s, primary lung cancer was a rare condition. Today, worldwide, lung cancer is the most frequent and deadliest type of cancer accounting for 1.8 million new cases (12.9% of all new cancer cases) and leading to 1.59 million deaths (19.3%) in 2012 (IARC, 2012). The highest incidence occurs in Europe and North America, with parts of Africa having the lowest incidence (Teh & Belcher, 2014).

Between 85% and 90% of patients with lung cancer have a positive smoking history, although only 15% of smokers go on to develop lung cancer (Vineis *et al.*, 2005). The risk

of developing a smoking-related lung cancer is proportional to the number of cigarettes smoked and length of smoking history (Dela Cruz, Tanoue & Matthay, 2011). Other risk factors that may increase the risk of developing lung cancer are passive smoking, radon exposure, occupational exposure to asbestos, silica and uranium, previous radiotherapy to the lungs, decreased fruit and vegetable consumption, extreme air pollution such as intense indoor exposure to smoky coal and genetic and familial factors (Teh & Belcher, 2014; American Cancer Society, 2016).

The outcome for lung cancer patients is very poor - the five-year survival rate is only about 16.8% for all patients (Cancer Screening, 2016).

In general patients do not feel any symptoms of early stage lung cancers, therefore about 70% of patients are diagnosed at late stage disease (stage III or IV). Unfortunately a diagnosis at late stage goes often in line with poor survival, which can be as low as 9.5 months of median survival time for stage IV patients (*Clinical Lung Cancer Genome Project* (CLCGP) *Network Genomic Medicin*e (NGM), 2013). Despite advances in clinical therapies, the 1 and 5-year survival rates for lung cancer patients remain at a discouraging 44% and 17%, respectively (American Cancer Society, 2016).

As for the classification of lung cancer, the *Union Internationale Contre le Cancer* (UICC) and *the American Joint Committee on Cancer* (AJCC) have published the 7th edition of the international standardized TNM staging system. The TNM staging system describes: (1) the size and growth stage of the tumor (T), (2) if the tumor already spread to the lymph nodes (N) and (3) if the tumor has already metastasized (M). These criteria are used to judge a patient's tumor stage and are updated regularly (Detterbeck, Boffa & Tanoue, 2009).

1.1.1.2 Histology

The classification of lung cancer is assessed according to histological stainings. It can be coarsely divided, according to clinical, histological, and neuroendocrine features, into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), being NSCLC histologically a wide classification, when compared to SCLC (Petersen, 2011; Travis, Brambilla, & Riely, 2013).

Until late in the first decade of the 21st century, histology in the NSCLC group did not affect the choice of treatment (Travis, Brambilla & Riely, 2013). Standard treatment consisting of a platinum drug in combination with a third generation cytotoxic drug was used to treat all kinds of NSCLC. In 2008, Scagliotti *et al*. published the results of a Phase III study, in which a pre-planned subgroup analysis showed a marked difference in survival between squamous cell carcinomas and non-squamous cell carcinomas (e.g. adenocarcinoma). After that study, the focus on histology began to increase. Today, tumor histology is vital when choosing an individualized treatment (Reck *et al*., 2013).

With the increasing amount of genomics data, there is a future trend to diagnose and classify lung cancers based on genomics data, as every subtype has its own molecular profile. Therefore, molecular pathology can lead to a completely unbiased diagnosis and classification system in the near future and could already eliminate and further subdivide the group of large cell carcinomas (CLCGP & NGM, 2013).

1.1.1.2.1 Small cell lung cancer

Small cell lung cancer accounts for around 15% of all diagnosed lung cancer cases worldwide (Wahbah *et al*., 2007) and is highly associated with smoking. More than 90% of SCLC patients are or have been heavy smokers (Rosell & Wannesson, 2012; Meerbeeck, Fennell & Ruysscher, 2011). SCLC patients typically respond very well to initial standard chemotherapy with platinum and etoposide but show very soon a resistance phenotype and tumor relapse. The tumor is characterized by early metastasis and patients show a very poor 2-year survival, which is of about 5% for high stage or

~15% for low stage disease (Pleasance *et al*., 2010). Up to date, no gene could be identified as therapeutical target in SCLC. Only the tumor suppressor genes tumor protein p53 (TP53) and retinoblastoma (RB1) are consistently inactivated in SCLC (Peifer *et al*., 2012; Rudin *et al*., 2012).

1.1.1.2.2 Non-small cell lung cancer

NSCLC, accounts for at least 80% of lung cancers, and can be divided into three main groups: adenocarcinoma, large cell carcinoma and squamous cell carcinoma (Pao & Hutchinson, 2012). Other NSCLC subtypes, such as adenosquamous carcinoma and sarcomatoid carcinoma, are very rare (American Cancer Society (ACS), 2016).

1.1.1.2.2.1 Adenocarcinoma

Adenocarcinoma accounts for ~50% of all lung cancer cases and is the most frequent subtype of lung cancer. It is the leading cause for cancer-related mortality with more than one million deaths worldwide. This type of lung cancer occurs mainly in current or former smokers, but it is also the most common type of lung cancer seen in non-smokers. It is more common in women than in men, and it is more likely to occur in younger people than other types of lung cancer.

Adenocarcinoma histology represents the large majority among the NSCLC subgroup. It is usually found in outer parts of the lung and it is defined as malignant epithelial neoplasms characterized by gland formation. Although this type of tumor tends to grow slower and is more likely to be found before it has spread than other types of lung cancer, this may vary from patient to patient (ACS, 2016).

1.1.1.2.2.2 Squamous Cell Carcinoma

Squamous cell carcinoma (SCC) accounts for 25-30% of NSCLC cases. These cancers start in precursors of squamous cells, which are flat cells that line the inside of the airways in the lungs. They are often linked to a history of smoking and tend to be found in the central part of the lungs, arising in large central airways that are proximal to segmental bronchi (ACS, 2016).

1.1.1.2.2.3 Large Cell Carcinoma

Large cell carcinoma is a very heterogeneous group within lung cancer. It is poorly differentiated and accounts for about 10% of lung cancer cases. It can occur in any part of the lung. It tends to grow and spread quickly, which can make it harder to treat. A subtype of large cell carcinoma, known as large cell neuroendocrine carcinoma, is a fast-growing cancer that is very similar to small cell lung cancer (ASC, 2016; Teh & Belcher, 2014). Due to its poor differentiation and molecular similarities to other lung cancer subtypes, some hypothesize that large cell carcinoma are only poorly differentiated variants of lung adenocarcinoma or squamous cell carcinoma and, therefore, this subtype might be a disappearing diagnostic entity (CLCGP & NGM, 2013).

It is commonly known that about 50% of NSCLC is diagnosed in advanced stage and, for the majority of these patients, even if encouraging data regarding immunotherapy has been emerging, to date prognosis remains poor and chemotherapy still represents the mainstay of treatment. However, approximately 15-20% of advanced NSCLC presents a targetable driver mutation, a condition that dramatically changes therapeutic perspectives and patient outcome (Bordi *et al*., 2015).

1.2 Molecular alterations in NSCLC

The molecular basis of lung cancer is complex and heterogeneous (Larsen & Minna 2011). Lung cancer develops through a multistep process involving development of multiple genetic and epigenetic alterations, particularly activation of growth promoting pathways and inhibition of tumor suppressor pathways. Greater understanding of the multiple biochemical pathways involved in the molecular pathogenesis of lung cancer is crucial to the development of treatment strategies that can target molecular aberrations and their downstream activated pathways (Larsen & Minna, 2011).

Oncogenic driver mutations refer to mutations that are responsible for both the initiation and maintenance of the cancer (Suda, Tomizawa & Mitsudomi, 2010). These mutations are often found in genes that encode for signalling proteins that are critical for maintaining normal cellular proliferation and survival. The presence of mutations on these genes will confer growth advantage on cancer cells, favoring their being selected during tumor progression (Bronte *et al*. 2010). Recently, treatment paradigms for non–small-cell lung cancer (NSCLC), which accounts for at least 80% of lung cancers, have shifted from one based only on histology (for adenocarcinoma, squamous-cell carcinoma and large-cell carcinoma) to one that incorporates molecular subtypes involving particular genetic alterations that drive and maintain tumorigenesis (Pao & Hutchinson, 2012). Consequently, NSCLC, especially lung adenocarcinomas, can be further sub-classified by their genetic mutation profiles, making personalized treatment strategies based on the identification of oncogenic driver mutations feasible (Shames & Wistuba, 2014).

Lung tumorigenesis frequently targets the activation of growth factor signaling proteins (e.g. EGFR, KRAS, BRAF, MEK-1, HER2, MET, ALK and RET) as well as inactivation of tumor suppressor genes (e.g. P53, PTEN) (Larsen & Minna, 2011). Activation of growth factor signaling proteins can occur by gene amplification or other genetic alterations, including point mutations and structural rearrangements, all leading to uncontrolled signalling through oncogenic pathways.

So called “Oncogene addiction”, results when cell survival depends on persistent activation of the aberrant signalling (Lynch *et al*. 2004, Weinstein, 2002) making these

oncogenic mutations ideal candidates for targeted therapies. Indeed, while the mechanisms underlying oncogene addiction are not fully understood, this principle provides a basis for pharmacologically targeting the required oncogenes in these tumor cells and thereby achieving a sort of cancer-cell selectivity (Sharma & Settleman, 2007).

The concept of oncogenic mutations has further evolved to distinguish between "driver" as well as "passenger" mutations. Oncogenic driver mutations are mutations that are causally implicated in oncogenesis; conferring growth advantage on cancer cells and are considered being positively selected in the tumor microenvironment. In contrast, passenger mutations are present and found to be mutated in tumors, but have not been selected and are not believed to confer clonal growth advantage, and therefore do not contribute directly to tumor development. The occurrence of passenger mutations is believed to represent somatic mutations without functional consequences during cell divisions (Stratton, Campbell & Futreal, 2009).

Oncogenic driver mutations are usually identified from analysis of large genomic datasets with high throughput technology, like massive parallel sequencing screening, in which it is possible to show significant difference in small region abnormalities of the genome that are present at more than the normal expectation rate of somatic mutations. These small genetic abnormalities may range from single base substitution to whole gene deletions (Nakagawa *et al*., 2015). To qualify as oncogenic driver mutations, these genetic abnormalities need to be tested and validated with *in vitro* experiments and in independent tumor sets or patient cohorts, (Stratton, Campbell & Futreal, 2009).

Oncogenic driver mutations have been identified in more than 40% of lung adenocarcinomas and they include mutations in genes such as EGFR, ALK, HER2, KRAS2, RET, ROS1, AKT1, BRAF, MEK1, MET, NRAS and PIK3CA (Pao & Hutchinson, 2012). Their individual presence is almost always exclusive of other driver mutations in the same tumor clone (Sequist *et al*., 2011; Yip *et al*., 2013; Pao & Hutchinson, 2012), although other mutations may be present at lower rate in subclonal populations, and expand as a result of therapy. The presence of some of these mutations could be associated with gender, smoking status or ethnic differences, and this may provide insight into the patient factors in lung cancer. Identification and study into the roles these oncogenic

mutations would have the potential to guide development of further therapeutic targeting opportunities (Larsen & Minna, 2011).

Signalling pathways regulated by oncogenes and tumor suppressor genes are often interconnected with cross-talks between carcinogenetic pathways (Azad, Lawen & Keith, 2015). Adding to this, another level of complexity arises from the occurrence of mutational evolution of tumors over time and during the natural course of disease progression or in response to selection pressure exerted by anti-cancer therapy (Larsen & Minna, 2011).

Overall, tobacco smoking-related mutations appear to have a stronger role in tumor initiation, while mutations associated with endogenous processes are more prominent at a later stage of tumor development and are associated with tumor progression (Shi *et al*., 2016).

1.2.1 Molecular alterations in Adenocarcinoma

The most common mutated driver oncogenes are kirsten rat sarcoma viral oncogene homolog (KRAS) and mutations in the gene encoding for the epidermal growth factor receptor (EGFR) represent the second driver mutations identified in NSCLC (Bordi *et al.,* 2015). Both these genes are mutated in about 30-45% of all lung adenocarcinomas and, in 25-50% of lung adenocarcinomas no oncogenic driver mutation is detected (Figure 1) (Cancer Genome Atlas Research Network, 2014; Pao & Hutchinson, 2012).

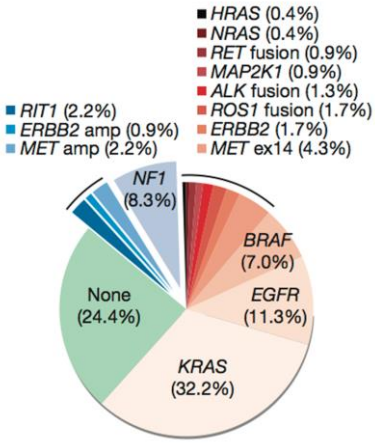

Figure 1: Distribution of oncogenic driver mutations in lung adenocarcinoma (Cancer Genome Atlas Research Network, 2014).

1.2.1.1. EGFR Mutations as a paradigm of personalized treatment in Adenocarcinoma

The Epidermal Growth Factor Receptor (also known as EGFR/ERBB1/HER1) belongs to a ErbB family of tyrosine kinase receptors, that includes EGFR, HER2/ ERBB2, HER3/ERBB3 and HER4/ERBB4, is located on the short arm of chromosome 7 (7p11.2) and encodes a 170-kDa type I transmembrane growth factor receptor. EGFR encodes an extracellular ligand-binding domain, a transmembrane domain, and an intracellular region consisting of a highly conserved tyrosine-kinase domain (Gazdar, 2014). Ligand binding induces homodimerization (if dimerization is established between two EGFR receptors) or heterodimerization of the receptor with other members of the EGFR family, resulting in autophosphorylation of the tyrosine residues (Pakkala & Ramalingam, 2017; Quintanal-Villalonga *et al*., 2016). These phosphorylated tyrosine residues provide a docking site for proteins containing Src homology 2 (SH2) or phosphotyrosine-binding (PTB) domains, leading to the activation of intracellular signaling pathways (Carcereny *et al*; 2015). Overexpression of EGFR and its ligand, the transforming growth factor alpha (TGFα), or activating mutations in EGFR, lead to activation of signalling pathways including the RAS-RAF-MEK-ERK pathway, the PI3K-Akt-mTOR pathway, and the stress

activated pathway involving JAK and STAT (Tebbutt, Pedersen & Johns, 2013). These signal cascades lead to profuse cellular responses, which include increased cell proliferation, motility, resistance to apoptosis, differentiation, invasion and metastasis (Figure 2). Independently from the role of EGFR in the membrane, where it activates its associated signalling pathways, EGFR is internalized to the nucleus (Pakkala & Ramalingam, 2017). Once in the nucleus, EGFR is capable of acting as a coactivator for several oncogenes as Cyclin D1, nitric oxide synthase, Aurora Kinase A, c-Myc, and B-Myb. Furthermore, nuclear EGFR promotes DNA replication and repair through its association to proliferating cell nuclear antigen (PCNA) and DNA dependent protein kinase (Quintanal-Villalonga *et al*., 2016). The presence of mutations in EGFR implicates a receptor constitutively activated that continuously gives the cell input favoring proliferation (Bordi *et al.,* 2015).

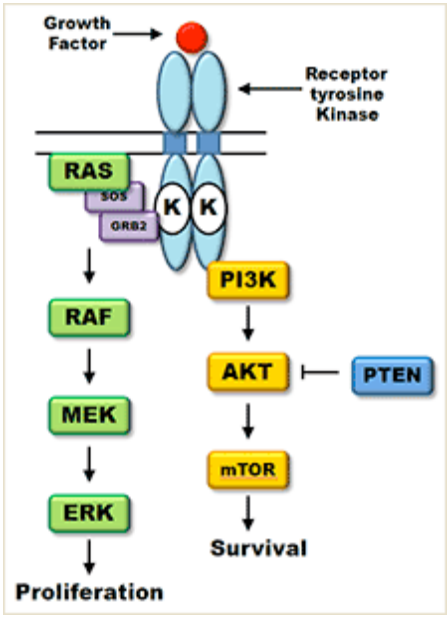

Figure 2: Pathways activated by EGFR binding. Scheme of RAS-RAF-MEK-MAPK and PI3K-AKT-mTOR pathways activated by binding to EGFR receptor. "K" denotes the tyrosine kinase domain (Lovly, Horn & Pao, 2015).

EGFR mutations are located in exons 18 to 21, which encode the ATP binding site of the tyrosine kinase domain (Pakkala & Ramalingam, 2016) (Figure 3). Although multiple mutations have been identified, the vast majority of EGFR mutations (80-90%) are either exon 19 mutations (deletion of 4 highly conserved amino acids (E746_A750 del)) or exon 21 mutations (point mutation (CTG to CGG) that results in substitution of leucine by arginine at codon 858 (L858R)), 45% and 41% respectively (Quintanal-Villalonga *et al*., 2016). There are also a range of less common mutations including in frame duplications or insertions in exon 20 (~ 5% - 10%), of which there are many variants that are often associated with resistance to EGFR tyrosine kinase inhibitors (TKIs) (Su *et al*., 2012).

It has been demonstrated that EGFR activating gene alterations are more common in patients with specific clinic-pathological characteristics, such as female, never smoker, Asiatic origin and adenocarcinoma histological subtype (Bordi *et al.,* 2015). In fact, activating mutations of EGFR have been reported in 10 - 15% of unselected Western patients and 30-50% of Asian populations (Kono *et al*., 2009). Differences in the reported prevalence rates of various mutations may in part relate to different patient populations but also depends on the sensitivity of mutation analysis techniques utilized in different studies (Gazdar, 2014).

Platinum-based chemotherapy has been found to provide a survival benefit for patients with advanced lung cancer; however, most patients do not survive longer than 1 year (Jackman *et al*., 2007). In the last decade, the discovery of EGFR mutations and subsequent therapies targeting this receptor has changed the treatment patterns and outcomes of non-small cell lung cancer (NSCLC) (Xu *et al.,* 2016). In fact, EGFR mutations represent the most important factor for prediction of response to EGFR tyrosine kinase inhibitors (Jackman *et al*., 2007). First-generation EGFR TKIs, such as Erlotinib (approved in 2003 by FDA) and Gefitinib (approved in 2004), are reversible inhibitors that compete for kinase domain binding with endogenous ATP, thus preventing its tyrosine-phosphorylating activity and blocking downstream signaling (Ogunleye *et al*., 2015). After the development of Imatinib and the uncover of differential response according to c-Myc mutations, these agents are one of the the “poster child” for targeted therapy for solid tumors as a subset of patients with non-small-cell lung cancer (NSCLC) that

contains mutations in the kinase domain of EGFR will respond better to EGFR TKIs, despite EGFR-targeted TKIs have generally performed poorly as single agents (Tebbutt, Pedersen & Johns, 2013). It is of notice that not all mutations seem to confer the same sensitivity to TKIs. Patients with exon 19 deletion treated with EGFR TKIs seem to have a better outcome compared with those with exon 21 mutation (Pakkala & Ramalingam, 2016). Afatinib, approved by the FDA in July 2013, is a second-generation EGFR TKI that binds irreversibly to the free cysteine in the kinase domain of EGFR, forming a direct chemical covalent bond with the EGFR receptor (Ogunleye *et al*., 2015). In addition, Afatinib binds all the ErbB receptors, not just EGFR (Morgensztern *et al*., 2014). Although Afatinib has proved slightly more effective, it also has the most side effects, and Gefitinib is generally the best tolerated drug that is approved in the first-line setting (Pakkala & Ramalingam, 2016).

Indeed, TKIs are associated with significant increase in response rate (approximately 70%) and improvement in progression free and overall survival (OS) (Bordi *et al.,* 2015). All three drugs , that can be used alone (without chemo) as the first treatment for advanced NSCLCs, have been approved on the basis of randomized trials showing superior progression-free survival, objective responses, and more favorable safety profiles when compared with standard first-line platinum-based doublet chemotherapy in patients with EGFR-mutant NSCLC (Park *et al.,* 2016). However, for drug prescription purpose, the presence of EGFR mutation needs to be demonstrated and therefore neoplastic tissue sample is always required (Bordi *et al.,* 2015).

EGFR inhibitors mentioned above can often shrink tumors for several months or more but eventually these drugs will stop working for most people, usually because the cancer cells develop secondary mutations in the EGFR gene (in most patients within a year), such T790M (in exon 20) (Pakkala & Ramalingam, 2016). T790M is found in about 50% of patients that develop acquired TKI resistance, due to molecular mechanisms like bypass signalling. This mechanism involves the reactivation of downstream signalling pathways via amplification of other TKRs (like MET or HER2) and mutations of downstream members of EGFR-signalling pathway (such as PIK3CA, KRAS, and BRAF) and even through ALK gene rearrangement (Quintanal-Villalonga *et al*., 2016). Recently,

a third generation EGFR TKIs – Osimertinib (AZD9291) – has been approved and showed to provide patients with a second-line option against T790M mutated lung cancer (Pakkala & Ramalingam, 2016).

Inevitably, cancer cells will continue to evolve and become resistant to the 3rd generation EGFR-TKIs in the clinic, resulting in disease progression (Govindan, 2015). Therefore, there is a great challenge and urgent need for a comprehensive understanding of how this resistance develops and how to develop effective strategies to delay or overcome resistance. Recent studies have identified a novel acquired EGFR C797S mutation in cultured cell lines and clinical tumors resistant to 3rd generation EGFR-TKIs (Thress *et al*., 2015; Ercan *et al*., 2015; Niederst *et al*., 2015). However, C797S mutation was detected only in some NSCLCs with T790M mutation (36%), suggesting the existence of other resistance mechanisms.

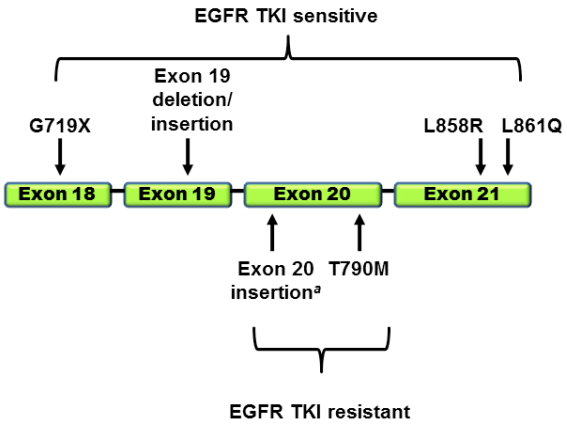

Figure 3: Location of EGFR mutations. This schematic represents the mutations found within exons 18-21. It also indicates if the identified mutations are sensitized to first-line TKI therapies (Lovly, Horn & Pao, 2015).

Apart from second- and third-generation EGFR-TKIs, other treatment strategies are being developed to overcome acquired resistance. The switching to chemotherapy after resistance has appeared to be the most accepted approach, although there are several retrospective studies with inconsistent results to this respect (Xu *et al*., 2015). Another

alternative therapy which is currently under clinical assessment is the combination of EGFR-TKIs and chemotherapy. Up to date, the results on the effectiveness of this combination therapy are not conclusive, but ongoing clinical trials on this issue could clarify if this approach could be beneficial for patients with EGFR-TKI acquired resistance (Janjigian *et al*., 2011). Thanks to the identification of the molecular mechanisms leading to acquired resistance to TKIs, approaches with a more targeted design are being designed. Many current research works bet on the combination of an EGFR-TKI with another molecularly targeted agent, for therapeutic tumor resensitization to anti-EGFR-therapy, with interesting preclinical results (Pakkala & Ramalingam, 2016).

Nowadays, using the molecular platforms that have been developed, detailed information about the presence or absence of a very high number of different molecular alterations can be acquired simultaneously and in a very short time. Thanks to the availability of targeted drugs, much of this information is not only useful to increase our knowledge about molecular characteristics of different tumors, but it can be also useful to select patients that are candidates for the treatment with specific agents.

1.3 New Targetable Oncogenes in Non–Small-Cell Lung Cancer

The identification of oncogenic driver mutations underlying sensitivity to epidermal growth factor receptor (EGFR) and anaplastic lymphoma kinase (ALK) tyrosine kinase inhibitors has led to a surge of interest in identifying additional targetable oncogenes in non–small-cell lung cancer.

Many studies have been carried out to identify novel “oncodrivers” in adenocarcinoma with the help of high throughput technologies. In a collaborative work 188 human lung adenocarcinomas DNA samples were sequenced for 623 genes with a potential role in cancer. This analysis revealed more than 1000 somatic mutations which occurred preferably in 26 genes, 30% of which were TKRs (Quintanal-Villalonga *et al*., 2016). Among these new potentially oncogenic gene alterations have been characterized are

included MET mutations/ deletions, BRAF mutations, HER2 insertions, PIK3CA mutations, FGFR1 amplifications, DDR2 mutations, ROS1 rearrangements and RET rearrangements (Park *et al*., 2015). Furthermore, efforts in next generation sequencing studies have revealed rarer driver mutations including NRG1, ERBB4, ARAF, RIT1, HRAS, and NRAS mutations that may serve as potential therapeutic targets. In this way, other targeting therapies against these genes have being actively investigated in lung cancer with the goal to discover new potential therapeutic targets The number of predictive biomarkers with their matched targeted drugs entering clinical trials is expected to increase dramatically. (Yeung *et al*., 2015).

1.3.1 MET

MET (c-MET or mesenchymal-epithelial transition factor) is a receptor tyrosine kinase that was first characterized as a proto-oncogene in 1984, in a chemically transformed osteosarcoma cell line (Figure 4) (Jenkins *et al*., 2015). The MET gene, located on chromosome 7q21-q31, is approximately 125 kilobases long, with 21 exons. The extracellular region of MET contains semaphorin, cysteine-rich, and immunoglobulin domains; the intracellular region consists of a juxtamembrane domain, a tyrosine kinase catalytic domain, and a carboxyterminal docking site (Figure 5) (Drilon *et al*., 2016). Binding of hepatocyte growth factor (HGF) to MET induces phosphorylation of the docking site and stimulates downstream signal pathways RAS/ERK/MAPK, PI3K-AKT, Wnt/β-catenin, and STAT signaling pathways. (Park *et al*., 2015; Drilon *et al*., 2016). These pathways are known to regulate various cellular processes, including cell proliferation, cell motility, cell scattering, cell invasion, metastasis, angiogenesis, epithelial to mesenchymal transition and survival (Feng, Thiagarajan & Ma 2012; Park *et al*., 2015).

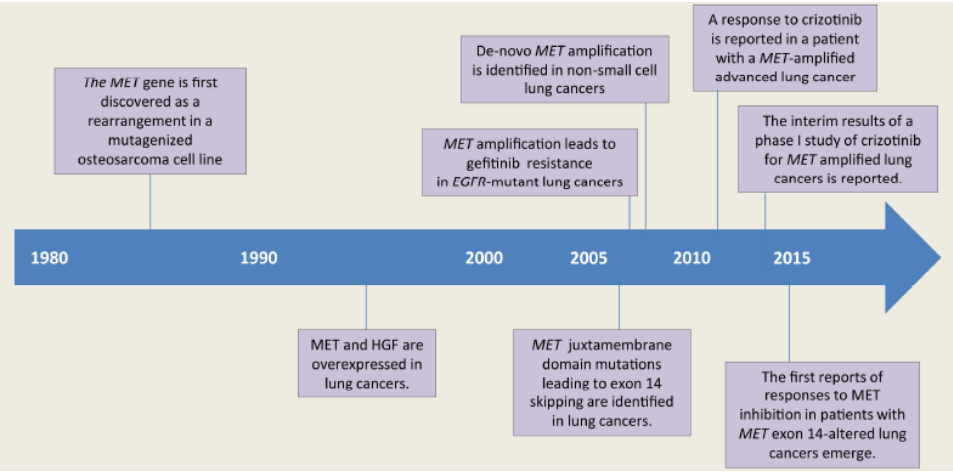

Figure 4: Timeline of discovery in lung cancers harboring alterations of the MET pathway (Drilon *et al*., 2016).

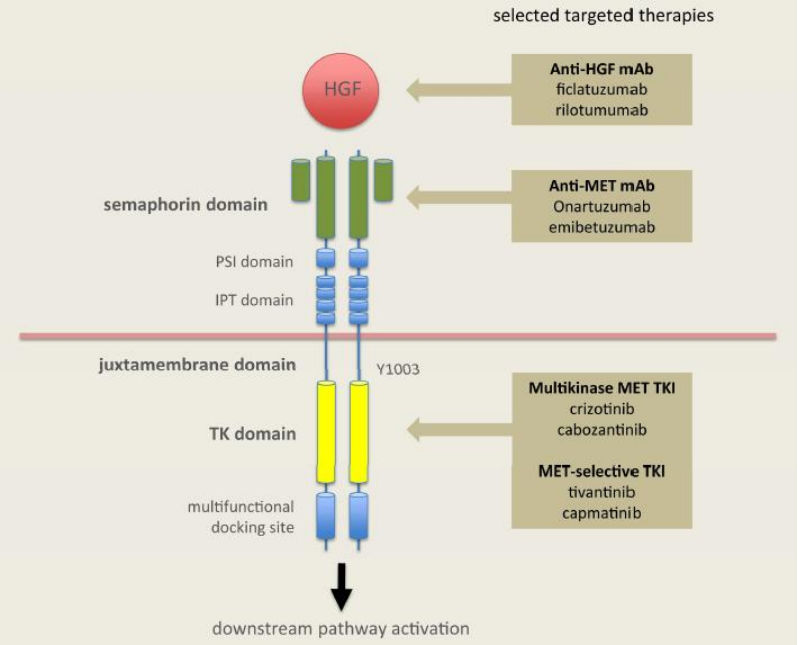

Figure 5: The MET receptor and selected MET pathway-directed targeted therapies (HGF, hepatocyte growth factor; IPT, immunoglobulin-plexin transcription; mAb, monoclonal antibody; MET, mesenchymal epithelial transition receptor; PSI, plexin semaphoring integrin domain; TK, tyrosine kinase; TKI, tyrosine kinase inhibitor) (Drilon *et al*., 2016).

MET protein is expressed in 22.2–74.6% of NSCLC, and it has been associated with poor prognosis in several studies. Some of those studies also reported that MET expression is more common in adenocarcinoma than in other histologic types, although it is also common in squamous cell carcinoma (Park *et al*., 2015; Schrock *et al*., 2016). Indeed, MET has proved to be of relevance in NSCLC after the large scale molecular profiling work by The Cancer Genome Atlas (2014) in lung adenocarcinoma. MET alterations were found in 7% of tumors and were mutually exclusive with other known oncogenes, supporting the role of MET as an oncogene. The most common alterations for this gene are overexpression, amplification, and exon 14 skipping (Quintanal-Villalonga *et al*., 2016). These alterations result in aberrant MET activation which can be mediated through HGF-dependent or HGF-independent mechanisms, which causes MET pathway dysregulation (Park *et al*., 2015).

While tumors such as sporadic and hereditary renal cell carcinomas harbor activating mutations of the MET kinase domain, lung cancers commonly harbor mutations in the extracellular and juxtamembrane domains. The juxtamembrane domain are encoded in part by MET exon 14 and contains critical regulatory elements, including tyrosine Y1003 in a DpYR motif, the direct binding site for Cbl, an E3 ubiquitin ligase that causes ubiquitination, receptor endocytosis, and degradation of MET (Onozato *et al*., 2009; Awad *et al*., 2016) (Figure 5). MET can have splice mutations in the juxtamembrane region, which can lead to exon 14 deletion. These somatic mutations are associated with ligand-mediated proliferation and tumor growth by decreased ubiquitination and delayed down-regulation of receptors, and are known to be important activating mechanisms of the MET pathway (Figure 6) (Park *et al*., 2015; Onozato et al., 2009). In fact, the extracellular semaphoring domain is thought to be required for receptor activation and dimerization, however, the relevance of mutations in this domain remains unclear. In contrast, juxtamembrane domain mutations often result in MET exon 14 alterations (Drilon *et al*., 2016).

MET exon 14 alterations are extremely diverse. Base substitutions or indels disrupt several gene positions important for splicing out introns flanking METex14, including the branch point, polypyrimidine tract, 3’ splice site of intron 13, and the 5’ splice site of

intron 14 (Drilon *et al*., 2016). In this way, deletion of exon 14 plays has an important role in the development of lung adenocarcinomas, conferring various advantages to the tumor cells and that MET splice mutations are ‘‘driver mutations’’ and not just ‘‘passenger mutations” (Onozato *et al*., 2009).

Schrock and colleagues (2016), have shown that concurrent, murine double minute gene (MDM2) amplification, cyclindependent kinase 4 gene (CDK4) amplification, and EGFR amplification, and KRAS mutation were observed in 35%, 21%, 6.4%, 3% of patients with MET exon 14, respectively.

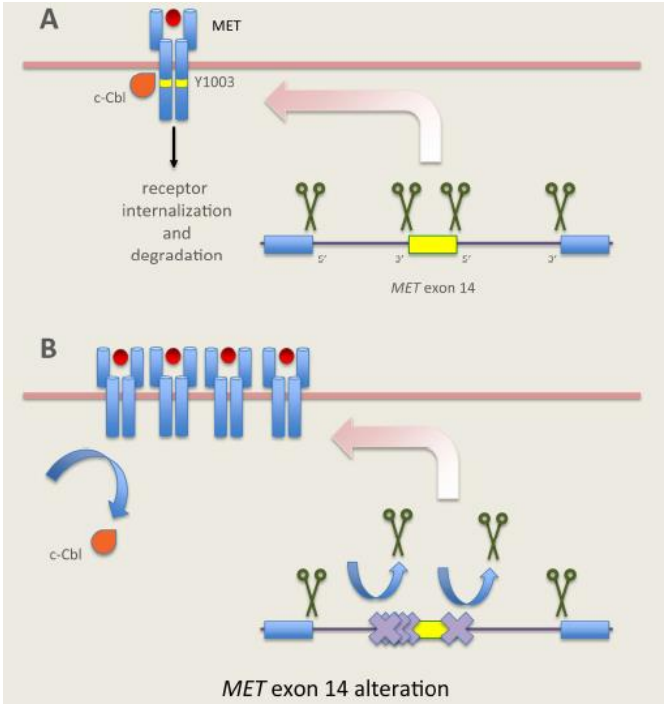

Figure 6: Activating MET splice site mutation. A – WT. Normally, introns flanking METex14 in pre-mRNA are spliced out, resulting in mRNA containing METex14 that is translated into a functional MET receptor (METex14 encodes part of the juxtamembrane domain containing Y1003 (the c-Cbl E3 ubiquitin ligase binding site) and ubiquitination tags the MET receptor for degradation.); B – Mutant. Juxtamembrane domain mutations that disrupt splice sites flanking METex14 result in aberrant splicing. These mutations result in METex14 skipping, producing a truncated MET receptor lacking the Y1003 c-Cbl binding site. Losing this binding site results in decreased ubiquitination and degradation of the MET protein, sustained MET activation, and oncogenesis (Drilon *et al*., 2016).

Other mechanism of MET is copy-number gains. MET copy-number gains arise from two distinct processes: polysomy and amplification. High polysomy occurs when there are multiple copies of chromosome 7 in tumor cells, secondary to factors such as chromosomal duplication. True amplification occurs in the setting of focal or regional gene duplication, via processes such as breakage-fusion-bridge mechanisms. As opposed to polysomy, it is believed that amplification represents a state of true biologic selection for MET-activation as an oncogenic driver and it is thought to dysregulate MET pathway signaling via protein overexpression and constitutive kinase activation (Drilon *et al*., 2016). Lutterbach *et al*., in an *in vitro* study, demonstrated that the level of tyrosine phosphorylation was greater in a MET-amplified cell line than in a non-amplified one, and the knockdown of MET in the amplified cell line caused growth inhibition, cell cycle arrest, and apoptosis. In addition, MET amplification has been identified as the mechanism resulting in EGFR-TKI resistance in about 20% of resistant tumors (Park *et al*., 2015; Bonanno, Jirillo & Favaretto, 2011).

Clinical trials of MET-directed therapies have taken two approaches: monoclonal antibody therapy directed against the receptor or HGF ligand; and tyrosine kinase inhibition (small molecule inhibitors). More than twenty agents have undergone preclinical and clinical study (Paik *et al*., 2015). The small molecule TKIs are further subdivided into multikinase and selective MET inhibitors. Examples of multikinase MET inhibitors include Crizotinib, Cabozantinib, MGCD265, AMG208, Altiratinib, and Golvatinib. Selective MET inhibitors include the ATP-competitive agents Capmatinib and Tepotinib (MSC2156119J), and the ATP-non-competitive agent Tivantinib. Monoclonal antibody therapy is divided into anti-MET antibodies (e.g. Onartuzumab and Emibetuzumab [LY2875358]), and anti-HGF antibodies (e.g. Ficlatuzumab [AV-299] and Rilotumumab [AMG 102]) (Figure 5) (Drilon *et al*., 2016).

Although some of these drugs have demonstrated high efficacy *in vitro*, clinical trials results have been disappointing (Quintanal-Villalonga *et al*., 2016). The findings have ranged from impressively large responses in molecularly pre-selected subtypes of NSCLC in single-arm trials to the prominent failure of large phase III studies in different trial

populations (Drilon *et al*., 2016). Nonetheless, in small studies it has been shown that Crizotinib, an ATP competitive and selective potent inhibitor approved by FDA for the therapy of lung adenocarcinomas harboring ALK or ROS1 fusions, was clinically active in tumors with high level MET amplification (Schwab *et al*., 2014; Jenkins *et al*., 2015). Further, MET exon 14 alterations were recently identified as a molecular mechanism potentially associated with Crizotinib and Cabozantinib sensitivity in solid tumors (Mahjoubi *et al*., 2016; Jenkins *et al*., 2015; Paik *et al*., 2015; Schrock *et al*., 2016).

Table 1: Summary of case reports of patients with MET exon 14 and responses to MET treatment (Heist *et al*., 2016).

MET exon 14	MET amplification	Selected other alterations	Drug	Age (yr)	Sex	Smoking	Histology	Reported estimated PFS (mo)	Reference
c.3028G>C	*MET* amp	*MDM2* amp and multiple others	Cabozantinib	80	Female	Never	Adenocarcinoma	5.1+	22
c.3024_3028delAGAA GGTATATT	No amp IHC H Score 300	Multiple	Crizotinib	78	Male	Former	Adenocarcinoma	3.6	22
c.3028+1G>T	*MET* amp IHC NA	Multiple	Crizotinib	65	Male	Former	Adenocarcinoma	4.6+	22
c.3028G>C	No amp IHC H score 300	*CDK4* amp, *MDM2* amp	Crizotinib	90	Female	Never	Adenocarcinoma	3.1+	22
c.2888-5_2944del62	NA	*TPS3 ZMYM3*	Crizotinib	84	Female	Never	Histiocytic sarcoma	11	18
c.3028G>C	FISH not done 3+IHC	*TPS3*	INC280	82	Female	Former	Large cell	5+	18
c.3028+1G>T	Copy number 4; *MET:CEP7* 2.3	None reported	INC280	66	Female	Former	Squamous	13	18
c.3028 G>A	Borderline *MET:CEP7* 2.2	SNaPshot wt	Crizotinib	73	Male	Former	Squamous	6+	This study
c.3028G>C	BA (presume negative, Foundation)	*MDM2* amp	Crizotinib	76	Female	Former	Squamous	NA	25
Chr7:g.116412043G>C	No amp *MET/CEP7* 0.96	NA	Crizotinib	71	Male	Former	Adenocarcinoma	6+	26
c.2887-18>2887-7del12	NA	*CDKN2A/B* loss, *CDK4* amp, *MDM2* amp	Crizotinib	86	Male	Never	Adenocarcinoma	Crizotinib discontinued because of pneumonitis	27
Intron 14 +3 A>G	9 copies	None reported	Crizotinib	74	Female	Former	Sarcomatoid	NA	21

Abbreviations: amp, amplification; FISH, fluorescent in situ hybridization; IHC, immunohistochemistry; NA, not available; PFS, progression-free survival.

In the presence of an active MET-inhibitor, precedent from other driver states suggests monotherapy against MET should display clear evidence of anti-cancer activity. To date, two partially overlapping MET-related states in NSCLC have shown promise: MET exon 14 (METex14) alterations and MET gene amplification (Drilon *et al*., 2016).

While research into the MET pathway as a driver of oncogenesis has stretched well over three decades, advances in technology and appropriate patient selection have been reinvigorated by the initial data of effective therapeutic effect for lung cancers harboring METex14 alterations and/or MET amplification as their primary oncogenic driver. Attempts to define the criteria for optimal use of a MET-based precision therapy continue and future clinical trials with a strong focus on molecular enrichment are likely

to succeed in this arena. Both patients and health providers look forward to regulatory approval.

I.3.2 ERBB2

Human epidermal growth factor 2 (also known as ERBB2 or HER2), located on 17q12 chromosome, is a tyrosine kinase receptor that belongs to the super-family of cell-surface receptor tyrosine kinase ERBB/HER. This super-family consists of four different types of tyrosine kinase receptors that include ERBB1/HER1 (also known as EGFR) ERBB2/HER2, ERBB3/HER3 and ERBB4/HER4 (Kovacs *et al*., 2015). Each receptor is composed of an extracellular domain for ligand binding, an α-helical transmembrane segment and an intracellular tyrosine kinase domain. ERBB receptors are activated by binding to growth factors of the EGF family. Ligand binding to the extracellular domain initiates a conformational rearrangement that allows dimerization. Receptor dimerization is essential for ERBB function and can occur between two different ERBB receptors (heterodimerization) or between two molecules of the same receptor (homodimerization). Ligand binding promotes receptor dimerization and self-phosphorylation of specific tyrosine residues within the cytoplasmic tail, that serves as docking sites for various adaptor proteins containing SH2 and PTB domains, allowing activation of intracellular signaling cascades, including MAPK and PI3K/Akt pathways (Peters & Zimmermann, 2014). Although all four ERBB receptors possess the same essential domains, the functional activity of each domain varies. In fact, ERBB3 can bind to several ligands but lacks intrinsic tyrosine kinase activity and can only heterodimerize with other ERBB receptors. Instead, ERBB2 has an active tyrosine kinase domain but lacks a specific ligand and is the favored heterodimerization partner of the other ERBB receptors (Figure 7) (Pao & Chmielecki, 2010). ERBB2 activates signal transduction primarily via a hetero-dimerization with HER1 and HER3 rather than homo-dimerization (Roskoski, 2014; Kovacs *et al*., 2015). Additionally, specific protein-protein and protein-lipid interactions of single-span HER transmembrane domains are important for proper receptor activation and mechanism(s) that reduce or enhance such interactions (e.g., by

means of mutations), can affect downstream activity independently of kinase domains mutations (Bocharov *et al*., 2016).

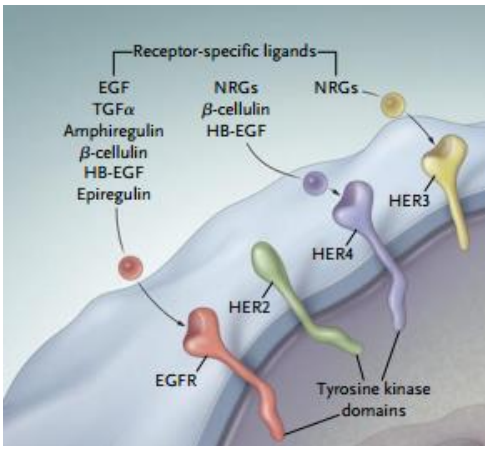

Figure 7: Members of the ERBB/HER superfamily of cell-surface receptor tyrosine kinase. Discrimination of all four members and their receptor specific ligands. The structural differences of the receptors are perceptible in: EGFR (ERBB1/HER1), ERBB3/HER3 and ERBB4/HER4, each presenting a ligand-binding domain where the receptor-specific ligands can bind; and EGFR (ERBB1/HER1), ERBB2/HER2 and ERBB4/HER4, each presenting a tyrosine kinase domain (adapted from: Ciardiello & Tortora, 2008).

ERBB2 has been a recognized proto-oncogene in human cancers since it was found to be amplified in breast cancers and gastric cancer more than two decades ago (Yoshizawa *et al*., 2014). Additionally, many preclinical studies have shown that overexpression of ERBB2 or mutations of the ERBB2 kinase domain play an important role in oncogenic transformation and tumorigenesis. Although alterations in this gene have been identified as oncogenic drivers and potential therapeutic targets in lung cancers, very little is known about the benefit of ERBB2-targeted therapy for NSCLCs harboring ERBB2 alterations (Suzawa *et al*., 2016). Three principal mechanisms of oncogenic activation of ERBB2 have been described: ERBB2 gene amplification, gene mutation resulting in molecular alterations of the receptor or ERBB2 protein overexpression (Peters & Zimmermann, 2014). In NSCLC patients, ERBB2 overexpression, ERBB2 amplification and ERRB2 mutations were reported in 11–32%, 2–23% and 1–5% respectively (Suzawa *et al*., 2016; Suzuki *et al*., 2015). In lung adenocarcinomas, ERBB2 mutations have been

reported in approximately 1–4% and ERBB2 amplifications in 2–5%. (Li *et al*., 2015; Kris *et al*., 2015; Ou *et al*., 2016)

Like EGFR, ERBB2 mutations are usually mutually exclusive with other driver mutations and it seems to be more frequent in never-smokers, Oriental ethnicity, female gender and adenocarcinomas, which suggest that similar genetic factors and possibly carcinogen(s) or other environmental factor(s) affect the occurrence of mutations in both genes (Shigematsu *et al*., 2005; Suzuki *et al*., 2015). However, these mutations have also been described in male heavy smokers with adenocarcinoma histology, suggesting that ERBB2 testing should be guided by the NSCLC subtype rather than features of a particular patient subgroup (Mar, Vredenburgh & Wassera, 2014).

ERBB2-activating mutations in NSCLCs were first described in 2004, by Stephens *et al*. These mutations occur in the first four exons of the tyrosine kinase domain (exons 18–21), including the most frequently observed alteration – a 12-bp duplication/insertion of the amino acid sequence YVMA in exon 20 at codon 776 (Suzuki *et al*., 2015). In fact, the most of mutations found in this gene in adenocarcinoma refers to exon 20 insertions, which have been found to induce constitutive activation of the ERBB2 kinase in a ligand-independent fashion, similar to the effect of EGFR mutations (parallel exon 20 insertion mutations in the EGFR gene can be identified with a similar frequency in NSCLC) (Oxnard, Binder & Janne, 2013).

From the initial report of Stephens *et al*., other groups have progressively demonstrated the presence of ERBB2 mutations in several human malignancies, including breast and colorectal cancer and urothelial carcinoma. From these studies, it has emerged that the type and location of ERBB2 mutations in nonlung tumors are different from those observed in NSCLC, with a higher prevalence of missense mutations over deletions/insertions and the involvement of exons other than 20, suggesting that functional mechanisms of the HER2 mutations in other tumor types might be different from that of lung cancer (Ricciardi *et al*., 2014).

Although oncogenic tyrosine kinase mutations most frequently alter the intracellular ATP- binding pocket, as EGFR exon 19 and 21 as well as in HER2 exon 19 or 20 mutations, rare mutations affecting the extracellular domain have recently been described, resulting in constitutively dimerized and activated HER2 (Greulich *et al*., 2012). Mutations in the transmembrane domain of HER2 (G660D and V659E) have also been described in familial lung adenocarcinomas (Yamamoto *et al*., 2014).

ERBB2 amplification seems far less common in NSCLC compared with breast cancer and is more frequently a consequence of chromosome 17 polysomy rather than gene amplification (Ricciardi *et al*., 2014). In addition, ERBB2 amplification has been proposed as a mechanism of acquired resistance to the anti-EGFR monoclonal antibody Cetuximab in colorectal cancer and activation of HER family signaling has been associated with acquired resistance to ALK inhibitors, suggesting that activation of alternative receptor tyrosine kinase signaling (bypass tracks) is a common mechanism of escape to targeted agents. Furthermore, Takezawa *et al*., reported ERBB2 amplification as a novel mechanism of resistance to EGFR TKIs in NSCLC patients.

According to Li and colleagues, ERBB2 mutations are not associated with ERBB2 amplification, thus suggesting a distinct entity and therapeutic target. In this way, ERBB2 -positive lung cancer may not be an adequate term, whereby the specific type of ERBB2 mutation, presence and degree of ERBB2 amplification, and ERBB2 protein expression should be precisely defined for each patient in future studies of ERBB2 -targeted agents. In 2015, Suzuki *et al*., referred that only ERBB2 expression and mutation are associated with ERBB2 phosphorylation at Tyr1221/1222. Furthermore, amongst the patients harboring ERBB2 mutations, analysis revealed that ERBB2 amplification was an unfavorable prognostic factor, while ERBB2 phosphorylation at Tyr1221/1222 was a favorable prognostic factor.

In the landscape of lung cancer biomarkers-based precision medicine, ERBB2 as a target remains poorly described. While in breast cancer ERBB2 overexpression or gene amplification is widely known to be associated with sensitivity to ERBB2-targeting drugs like Trastuzumab, Lapatinib, Pertuzumab, and Trastuzumab-emtansine, clinical research

in lung cancer has been slowed down after the first negative clinical trials of Trastuzumab added to chemotherapy in advanced NSCLC (Peters & Zimmermann, 2014). Nevertheless, recent studies have again begun to investigate the use of novel anti-HER2 drug in lung cancer patients (Table 2) (Yoshizawa *et al*., 2014). The most promising compounds are irreversible TKIs targeting ERBB2 and EGFR, such as Neratinib (HKI-272), Dacomitinib (PF-00299804), and Afatinib (BIBW- 2992) (Oxnard, Binder & Janne, 2013). In fact, phase I and II trial data suggest that these compounds have some activity: Neratinib and Dacomitinib have both been found to effectively inhibit the growth of ERBB2-mutant lung cancer cell lines as well as cell lines transformed by the introduction of ERBB2 (Oxnard, Binder & Janne, 2013; Ricciardi *et al*., 2014); Afatinib was found to induce modest regressions when introduced into transgenic mouse model, and this effect was potentiated by the addition of rapamycin, a mammalian target of rapamycin (mTOR) inhibitor, suggesting a particular dependence on the Akt/mTOR pathway in ERBB2-mutant lung cancer (Oxnard, Binder & Janne, 2013; Peters & Zimmermann, 2014). Further, according to De Grève *et al*., Afatanib shows objective response in three patients, even after failure of other EGFR- and/or ERBB2-targeted treatments. Enrollment of patients in clinical trials with novel agents targeting ERBB2 or downstream components of the PI3K/AKT/mTOR and MEK/ERK pathways is another clinical option for treating ERBB2 positive NSCLC (Table 2).

Table 2: Targeted drug therapies, their mechanisms, and available clinical trials in HER2 positive NSCLC (Mar, Vredenburgh & Wassera, 2014).

Drug class	Representative drug	Mechanism	Available data in NSCLC
Antibodies targeting HER2	*Trastuzumab*	Targets extracellular domain IV of HER2 receptor, prevents dimerization	[43–46]
	Pertuzumab	Targets extracellular domain II of HER2 receptor, prevents dimerization	None
	Trastuzumab-emtansine (TDM1)	A cytotoxic microtubule inhibitor, DM-1, is conjugated to trastuzumab, which delivers it to HER2 labeled tumor cells	Ongoing (NCT02289833)
HER-family tyrosine kinase inhibitors	*Lapatinib*	Inhibits EGFR/HER1 and HER2	None
	Afatinib	Inhibits EGFR/HER1, HER2, and HER4	[50]
	Neratinib	Inhibits EGFR/HER1 and HER2	[57]
	Dacomitinib	Inhibits EGFR/HER1, HER2, and HER4	[56]
Mammalian target of rapamycin (mTOR) inhibitors	*Temsirolimus*	Binds to FKBP-12 protein with the resulting complex inhibiting mTOR, causes G1 growth arrest of tumor cells	Ongoing (NCT01827267)
Phosphoinositide-3 kinase (PI3K) inhibitors	*Buparlisib*	Inhibits four isosomes of class I PI3K (α, β, γ, δ)	None
Heat shock protein 90 (HSP90) inhibitors	*Ganetespib*	Inhibits hsp90 molecular chaperone, leads to simultaneous degradation of critical oncoproteins including HER2	None
Insulin growth factor 1 receptor (IGF-1R) Inhibitors	*Cixutumumab*	Prevents natural ligand binding to IGF-1R, prevents activation of PI3K/AKT signaling pathway	None
Fc-modified chimeric monoclonal antibody	*Margetuximab*	Binds the HER2 receptor, mediates antibody-dependent cellular cytotoxicity via recruitment of immune cells	Ongoing (NCT01148849)

In a recent study, Suzawa *et al*., reported again efficacy of Afatinib, in both ERBB2-amplified and ERBB2-mutant NSCLC cell lines, which downregulated the phosphorylation of ERBB2 as well as their downstream signaling, and induced an antiproliferative effect, inducing cell cycle arrest and apoptosis in ERBB2-dependent cells. In addition, these effects were confirmed *in vivo* by using a xenograft mouse model of ERBB2-altered lung cancer cells. According to the COSMIC database (http://cancer.sanger.ac.uk/cosmic), A775insYVMA, G776VC, G776LC, and P780insGSP mutations account for 74% of all HER2 mutations in lung cancers. In addition, mutations in the αC–β4 loop (M774–R784) of the ERBB2 kinase domain accounted for 81% of ERBB2 mutations. As we can see, the 3-D structure of the modeled ERBB2–Afatinib complex (Figure 8) shows that the major mutations of ERBB2 are concentrated in an exposed area (A775–P780) of the αC–β4 loop, which is located on the back side of the ATP-binding pocket. Thus, these mutations are considered not to directly inhibit the binding of Afatinib to the ERBB2 kinase domain, which support the assertion that Afatinib is effective for the majority of ERBB2 mutations. This results and the aforementioned suggest that Afatinib can be a therapeutic option as an ERBB2-targeted therapy for NSCLC harboring HER2 amplification or mutations.

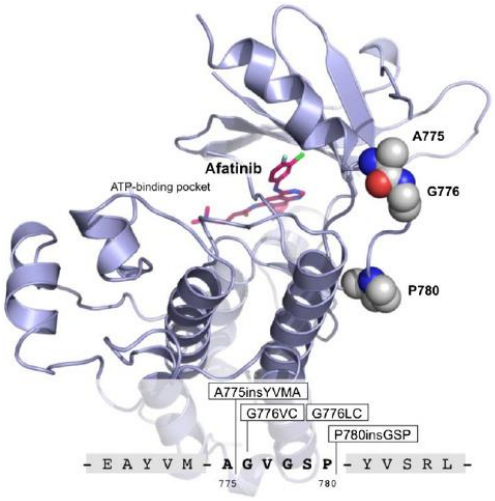

Figure 8: Location of the major mutation site and the ATP-binding pocket of the ERBB2 kinase domain in the modeled ERBB2–Afatinib complex structure. Afatinib is shown as a stick. Residues at the mutation site of the ERBB2 kinase domains (A775, G776, and P780) are shown as spheres. These residues are concentrated in an exposed area on the protein surface (A775–P780) of the αC–β4 loop (M774–R784), which is located on the back side of the ATP-binding pocket (Suzawa *et al*., 2016).

Dysregulation of the ERBB2 pathway, in particular ERBB2 mutations, may represent a possible novel therapeutic target in NSCLC, paving the way for a new generation of targeted agents in NSCLC (Ricciardi *et al*., 2014). In this way, testing for activating ERBB2 kinase domain aberrations, both point mutations and exon 20 insertions, should therefore be incorporated into standard multiplex molecular screening in lung adenocarcinoma. Although several clinical approaches for inclusion of ERBB2 targeted agents into patient treatment plans exist, the optimal choice of molecular therapy and timing of administration within the course of disease remain poorly defined (Mar, Vredenburgh & Wassera, 2014).

Chapter II

2. AIMS

A consecutive series of DNA samples from NSCLC patients previously analyzed for EGFR mutations was studied. The major goals of this work were:

- To investigate the frequency of mutations of MET exon 14 and ERBB2 exon 20 mutations by Sanger sequencing;
- To characterize the frequency and pattern of ERBB2 exon 20 mutations;
- To determine the frequency of MET exon 14 mutations;
- To evaluate the mutual exclusivity hypothesis of mutations involving the EGFR,
- MET and ERBB2 in NSCLC carcinogenesis;
- Development of Real Time PCR multiplex assay, to MET and ERBB2 mutations:
    - Rapid and sensitive assessment of mutation in tissue and plasma;
    - Potentially monitor the evolution of mutational status in the plasma during the treatment for the prediction of relapse and control the appearance of clones with resistance mutations.

Chapter III

3. MATERIALS AND METHODS

3.1 Samples

A consecutive series of DNA samples from 300 patients with NSCLC (formalin-fixed, paraffin-embedded tissues sections, bronchoalveolar lavages and plasma sent for routine analysis) were retrospectively retrieved from the archives of the Molecular Pathology Laboratory. These patients were referred to the IPO – Coimbra, Molecular Pathology Department, between January 2012 and July 2017, for EGFR exons 18, 19, 20 and 21 mutation analysis by PCR amplification, qPCR, followed by DNA sequencing.

Of the total of 300 DNA samples, only 172 samples, mostly of adenocarcinomas were analyzed since in several samples the amount of DNA in the remaining material was not sufficient. Some samples could not be analyzed due to failed amplification, which could presumable be attributed to lack of purity of the initial DNA sample (more frequent in formalin-fixed, paraffin-embedded tissues).

3.2 Methods

3.2.1 DNA extraction

In general, the DNA extraction procedure consists of two parts: a technique to lyse the cells gently and solubilize the DNA; and an enzymatic or chemical method to remove contaminating proteins, RNA, or macromolecules.

Firstly, the cells in a sample are separated from each other, often by a physical means such as grinding or vortexing, and put into a solution containing salt. The positively charged sodium ions in the salt help protect the negatively charged phosphate groups that run along the backbone of the DNA. A detergent is then added, which breaks down the lipids in the cell membrane and nuclei. DNA is released as these membranes are

disrupted. To get a clean sample of DNA, it is necessary to remove as much of the cellular debris as possible. This can be done by a variety of methods. Often a protease is added to degrade DNA-associated proteins and other cellular proteins. For DNA precipitation, initially, is add phenol/chloroform extraction to get rid of proteins (the addition of phenol and chloroform leads to the appearance of two phases: an upper aqueous phase containing the nucleic acids and an organic phase containing proteins solubilized in phenol and lipids dissolved in chloroform). Secondly, ethanol precipitation is a commonly used technique for concentrating and de-salting nucleic acids preparations in aqueous solution. The basic procedure is that salt and ethanol are added to the aqueous solution, which forces the precipitation of nucleic acid nucleic acids out of solution and removes salts and small organic molecules that may still be present in the sample. After precipitation the nucleic acids can then be separated from the rest of the solution by centrifugation. The pellet is washed in cold ethanol then after a further centrifugation step the ethanol is removed, and the nucleic acid pellet is allowed to dry before being resuspended in clean water.

3.2.1.1 DNA extraction from Paraffin-embedded tissues

Tissue sections, with 5-10µm of thickness, were obtained from formalin-fixed, paraffin-embedded tissues (FFPE) and were placed on 0.1% Poli-L-lysine coated slides and dried at 37°C overnight.

DNA extraction from paraffin-embedded tissues began with the slides being soaked on 500 µl lysis solution [400 ul 1x Lysis (10mM TrisHCl, (pH=8.0) (stock=0.5 M); 0.1M EDTA (pH=8.0) (stock=0.5 M); 0,5% SDS (stock10%)); 50 µl SDS 5%; 100 µl Proteinase K] and incubate for 3-4 hours at 56°C or overnight.

After incubation at 56°C, the digested tissue was directly centrifuged for 1 min at 10,000 g. After centrifugation, paraffin forms a solid layer on the surface of the solution and at the bottom stays the undigested tissue pellet. With a pippete tip, "drill" paraffin layer and with another tip, recover the bottom solution (phase containing DNA) to the new 1.5ml tube.

The bottom phase containing the DNA was removed and transferred into a new labelled 1.5mL microcentrifuge tube along with an equal volume (approximately 500mL) of phenol-chloroform (1:1) and shake vigorously by hand, for 1-5 min, to allow the phenol-chloroform to mix with the specimen, until an emulsion was formed. After which, the tube was put on the microcentrifuge, for 1 min at 10.000g, to continue mixing. The top phase containing the DNA was transferred to new 1.5mLmicrocentrifuge tube. Repeated the extraction with phenol-chloroform 2-3x, until the interface is clear, so that there is no precipitated protein (white interface). A final separation was performed with 500 µl of chloroform 100% and shake vigorously by hand, for 1-5 min. Centrifuged at 10,000g for 1min and the upper aqueous phase was removed to a new 1.5mL tube.

For DNA precipitation, added 1/10 vol (50 µl) of sodium acetate (3M, pH 5.2) and mixed. After that, 2 volumes (1000 µl) of cold ethanol 99% were added to the solution and shake gently. Subsequently, there are two options: a) If the DNA is fragmented: DNA is precipitated for 3 hours or preferably overnight; after precipitation, DNA is centrifuged at 10,000 g for 30 min at 4°C; remove the supernatant and wash for 5 min (1 ml of 70% ethanol); new centrifugation and repeat the wash with 70% ethanol; the ethanol is removed and dry, but not excessively; b) If the DNA forms a novel: "catch" the DNA with a tip p100 for a new tube; wash 2x the DNA novel for 5 min (1ml of 70% ethanol); remove the novel to a new eppendorf and dry, but not excessively.

Dissolve the recovered DNA in 20-100 µl of ultrapure water (depending on the pellet size of precipitated DNA). Place the DNA at 4ºC.

3.2.1.2 DNA extraction from Bronchoalveolar lavage

20 ml of bronchoalveolar lavage was centrifuged for 5 minutes (10 500 rpm) on a Centrifuge 5810 R (Eppendorf). The supernatant was removed and the pellet was resuspended in 500 µl lysis solution [400 ul 1x Lysis (10mM TrisHCl, (pH=8.0) (stock=0.5 M); 0.1M EDTA (pH=8.0) (stock=0.5 M); 0,5% SDS (stock10%)); 50 µl SDS 5%; 100 µl

Proteinase K] and incubate for 3-4 hours at 56°C or overnight. After incubation at 56°C, the digested tissue was directly centrifuged for 1 min at 10,000 g.

The bottom phase containing the DNA was removed and transferred into a new labelled 1.5mL microcentrifuge tube along with an equal volume (approximately 500mL) of phenol-chloroform (1:1) and shake vigorously by hand, for 1-5 min, to allow the phenol-chloroform to mix with the specimen, until an emulsion was formed. After which, the tube was put on the microcentrifuge, for 1 min at 10.000g, to continue mixing. The top phase containing the DNA was transferred to new 1.5mLmicrocentrifuge tube. Repeated the extraction with phenol-chloroform 2-3x, until the interface is clear, so that there is no precipitated protein (white interface). A final separation was performed with 500 µl of chloroform 100% and shake vigorously by hand, for 1-5 min. Centrifuged at 10,000g for 1min and the upper aqueous phase was removed to a new 1.5mL tube.

For DNA precipitation, 1/10 vol (50 µl) of sodium acetate (3M, pH 5.2) was added and mixed. After that, 2 volumes (1000 µl) of cold ethanol 99% were added to the solution and shaked gently. Subsequently, there are two options: a) If the DNA is fragmented: DNA is precipitated for 3 hours or preferably overnight; after precipitation, DNA is centrifuged at 10,000 g for 30 min at 4°C; remove the supernatant and wash for 5 min (1 ml of 70% ethanol); new centrifugation and repeat the wash with 70% ethanol; the ethanol is removed and dry, but not excessively; b) If the DNA forms a novel: "catch" the DNA with a tip p100 for a new tube; wash 2x the DNA novel for 5 min (1ml of 70% ethanol); remove the novel to a new eppendorf and dry, but not excessively.

Dissolve the recovered DNA in 20-100 µl of ultrapure water (depending on the pellet size of precipitated DNA). Place the DNA at 4ºC.

3.2.1.3 DNA extraction from Plasma

For each testing 2 mL of the plasma was used. In the Cobas® DNA Sample Preparation kit, Proteinase K, WBI (wash buffer I) and WBI (wash buffer II) were prepared according to the manufacturer's instructions. The plasma was mixed with 250 µL Proteinase K and 2 mL DNA PBB (binding buffer) and incubated at room temperature for 30 minutes. Then 500 µL isopropanol was mixed with the lysate and transferred into the High Pure Extender Assembly. The High Pure Extender Assemblies were centrifuged at 4000g for 1 min. The extenders were removed from the filters; the filters were placed in new collection tubes and washed with WBI and WBII according the manufacturer's instructions. The DNA was eluted in 100 µL DNA EB (elution buffer).

3.2.2 DNA concentration and Purity

A quantitative spectrophotometric assay of DNA was performed using an Infinite M200 spectrophotometer (Tecan Group Ltd, Mannedorf, Switzerland). Absorbance was measured at wavelengths of 260 and 280 (A260 and A280, respectively) nm. The concentration of nucleic acids was measured by absorbance at 260nm and given in ng/µl, while the ratio of 260/280nm indicates the purity of the sample. The absorbance quotient (OD260/OD280) provides an estimate of DNA purity. An absorbance quotient value of 1.6 < ratio < 2.0 was considered to be good, purified DNA. A ratio of >2.0 indicates RNA contamination where as a ratio of <1.6 is indicative of protein or phenols contamination. In case of contamination with proteins it will be necessary to carry out an additional treatment to remove the proteins from the DNA solution (for example, addition of proteinase K). If contamination is due to the presence of phenols, the sample needs to be cleaned with chloroform, isoamyl alcohol and ethanol.

The absorbance ratio A260/A230 is used as an additional measure to determine DNA purity, as at 230 nm the maximum absorbance of salts, carbohydrates or other contaminants present in the solution is detected. DNA is generally considered to be pure

when the A260/A230 ratio is around 1.5-2.2. A ratio of less than 1.5 may be indicative of the presence of contaminants in the sample. However, it must be kept in mind that the information provided by this measure is not as accurate as the A260/A280 ratio, and that it can be distorted by a low concentration of DNA in the sample since one would be overestimating the concentration of salts in the resuspension buffer.

After the specimens were analysed in i-control™ software, the concentration of genomic DNA was diluted to approximately 200ng/µL for use in PCR.

3.2.3 PCR Amplification

PCR (Polymerase Chain Reaction) is a revolutionary method developed by Kary Mullis in the 1980s. PCR is based on using the ability of DNA polymerase to synthesize new strand of DNA complementary to the offered template strand. Because DNA polymerase can add a nucleotide only onto a preexisting 3'-OH group, it needs a primer to which it can add the first nucleotide. This requirement makes it possible to delineate a specific region of template sequence that the researcher wants to amplify. At the end of the PCR reaction, the specific sequence will be accumulated in billions of copies (amplicons).

The adequacy of DNA extracts for the PCR-based assays was amplified by PCR using specific primers for each gene (table 3). PCR (Bio-rad-icycler-582BR-thermocycler) was carried out in 25 µl total reaction volumes, each containing 200 ng/ µl template DNA, 0.3 µl (10 p.moles) of each primer, 2.5 µl 1× Taq buffer (Nzytech), 1 µl dNTPs (200 uM), 1 µl MgCl2 (2 mM), and 2 U Taq DNA polymerase. The reaction mixture was heated to 94°C for 4 min, followed by 45 cycles, each consisting of 0.30 min denaturation at 94°C, 0.35 min annealing at 62.5°C, 0.40 min extension at 72°C, and a final 10-min extension at 72°C.

Table 3: List of primers used in PCR. ERBB2=Human epidermal growth factor 2; MET= c-MET or mesenchymal-epithelial transition factor.

Gene	Mutation locus	Forward	Reverse
ERBB2	p.A775_G776insYVMA p.G776>VC p.P780_Y781insGSP p.V777L p.G776>LC (ins ttgt) p.G776>LC (ins cttt) p.G776>LC (ins ttat)	5'-CCCTCTCAGCGTACCCTTGT-3'	5'-CCCGGACATGGTCTAAGAGG-3'
MET	c.2887-18_2887-7del	5'-ACCCATGAGTTCTGGGCACT-3'	5'-CACTTCGGGCACTTACAAGC-3'
MET	c.3009C>G c.3028G>A c.3028G>C c.3028G>T c.3028+1G>T c.3001_3021del c.3024_3028+7del	5'-AGGCTTGTAAGTGCCCGAAG-3'	5'-CAACAATGTCACAACCCACTGA-3'

3.2.4 DNA integrity

The PCR amplification products (5 µl) were subjected to electrophoresis (OWL Separation Systems, Inc. D3) on 2% (w/v) agarose gel in 1× TAE buffer at 80 V for 30 min and stained with Gel Star (Lonza), to test the integrity of genomic DNA. Each DNA sample was graded, according to the electrophoretic migration of sample DNA compared with a known molecular weight marker (Thermo Scientific), and viewed under UV light (TFX-40M, Vilber Lourmat).

Agarose gel electrophoresis can be used to assess the intactness of purified DNA. A DNA sample of high integrity shows a single, perfectly defined band at the top of the agarose gel. A degraded DNA sample will show a smear in the gel, which will be more pronounced when the degradation of the sample is higher.

3.2.5 Sanger sequencing

The DNA sample to be sequenced is combined in a tube with primer, DNA polymerase, and DNA nucleotides (dATP, dTTP, dGTP, and dCTP). The four dye-labeled, chain-terminating dideoxynucleotides (ddNTPs) are added as well, but in much smaller amounts than the ordinary nucleotides. ddNTPs are similar to regular, or deoxy, nucleotides, but with one key difference: they lack a hydroxyl group on the 3' carbon of the sugar ring. Once a ddNTP has been added to the chain, there is no hydroxyl available and no further nucleotides can be added. The chain ends with the ddNTP, which is marked with a particular color of dye depending on the base (A, T, C or G) that it carries.

Firstly, the mixture is first heated to denature the template DNA, then cooled so that the primer can bind to the single-stranded template. Once the primer has bound, the temperature is raised again, allowing DNA polymerase to synthesize new DNA starting from the primer. DNA polymerase will continue adding nucleotides to the chain until it happens to add a ddNTP instead of a normal one. At that point, no further nucleotides can be added, so the strand will end with the ddNTP.

This process is repeated in a number of cycles. By the time the cycling is complete, it's virtually guaranteed that a ddNTP will have been incorporated at every single position of the target DNA in at least one reaction. That is, the tube will contain fragments of different lengths, ending at each of the nucleotide positions in the original DNA. The ends of the fragments will be labeled with dyes that indicate their final nucleotide.

For sequencing, to the remaining 20 µl PCR products was added 1µl Exo / AP mix (Thermo Scientific), followed by incubation at 37°C for 30 minutes and 80°C for 20 minutes to remove residual primers and unreacted dNTPs. Thermocycling was performed on a Bio-rad-icycler-582BR-thermocycler. Sequencing was performed using Sanger sequencing, in Stab Vida Laboratory (Lisbon). Chromatograms were visualized using Chromas 2.6.2 software (Gene Codes Corporation, MI, USA, version 5.0.1).

Chapter IV

4. RESULTS AND DISCUSSION

The NSCLC tumor sample series studied at IPO – Coimbra, Molecular Pathology Department includes a significant proportion of biopsies or cytology specimens, but also formalin-fixed, paraffin-embedded (FFPE) tumor samples, fresh tissues, bronchoalveolar lavages and peripheral blood (leucocytes and plasma) which allow a smaller cellular representation of the tumor. In this study, FFPE, bronchoalveolar lavages and plasma were the main materials used. Out of the 300 samples initially identified, the remaining DNA quantity and/or quality was not enough in 128 cases, sufficient DNA for the assays or and purity of the samples allowed the analysis of 172 cases. This sample series mostly includes adenocarcinomas, since this NSCLC histological type is preferentially referred for detection of *EFGR* gene mutations as predictive markers for TKI therapy. Notwithstanding these caveats, this study allowed the characterization of the mutational pattern of *MET and ERBB2* in a significant series of NSCLCs.

The standard methodology in use at the IPO – Coimbra, Molecular Pathology Department to detect mutations in the *EGFR* gene relies on a Cobas® EGFR Mutation Test v2 (CE-IVD), that consists in a real-time PCR test that identifies 42 mutations in exons 18-21 of the epidermal growth factor receptor (EGFR) gene, including L858R, L861Q, exon 19 deletions, and the TKI-resistance mutation, T790M. It is designed to enable testing of both tissue and plasma specimens with one kit, and allows labs to mix-batch tissue and plasma on the same plate. One of the methodologies designed in this study consists also of a real-time PCR assay, for the purpose of detecting a large number of alterations in MET and ERBB2, both the ones previously descripted in the literature as well as the ones detected by DNA Sanger sequencing of the studied cases.

As such, one of the main goals of this study was to characterize the frequency and pattern of alterations in MET and ERBB2 genes and to evaluate the mutual exclusivity of EGFR, MET, and ERBB2 mutations.

According to Ravi Salgia, sensitizing *EGFR* mutations are found in around 10% of Caucasian patients and our results comply with these numbers. Of the total of 172 samples, 161 (88.9%) were classified as negative for alterations in exons 18, 19, 20 and

21 of EGFR, whereas 19 samples (11.1%) were classified as positive. In total of the 19 alterations in EGFR, 73.7% were deletions in exon 19 and 21% was related to a point mutation (CTG to CGG) that results in substitution of leucine by arginine at codon 858 (Leu858Arg). A case of a resistance mutation – T790M – was also reported. At a lower frequency, a case of a point mutation was also found in a patient (Leu861Gln).

These results are in agreement with those described by Quintanal-Villalonga *et al*, who reports that 80-90% of EGFR alterations are deletions in exon 19. All of these patients, according with described by literature, are sensitive to EGFR–TKI therapy.

Table 4: EGFR mutations identified in 19 patients with routine genetic testing.

EGFR Exon	HGVS Protein	HGVS DNA	No. Samples
19	p.Lys745_Glu749del	c.2233_2247del	10
	p.Glu746_Ala750del	c.2235_2249del	2
	p.Leu747_Ala750del	c.2239_2250del	1
20	p.Thr790Met	c.2369C>T	1
21	p.Leu858Arg	c.2573T>G	4
	p.Leu861Gln	c.2582T>A	1

DNA mutational analysis via direct sequencing – Sanger – is considered the gold standard for characterizing mutations and is generally performed on PCR products using sequencing primers spanning the DNA region of interest (Khoo *et al*., 2015). Our primers were designed to detect the most frequent alterations in our genes of interest.

4.1 MET Mutation Analysis

MET is an RTK in the MET/RON family whose only known ligand is hepatocyte growth factor. Two main mechanisms of aberrant MET activation have been reported in lung cancer, including MET gene amplification and MET exon 14 skip mutations. Over 100 mutations in MET-mutated cancers resulting in exon 14 skipping have been described (Frampton *et al*. 2015). Figure 9 shows the most frequent mutations in MET exon 14. Among these known alterations, there are predominantly deletions and point mutations.

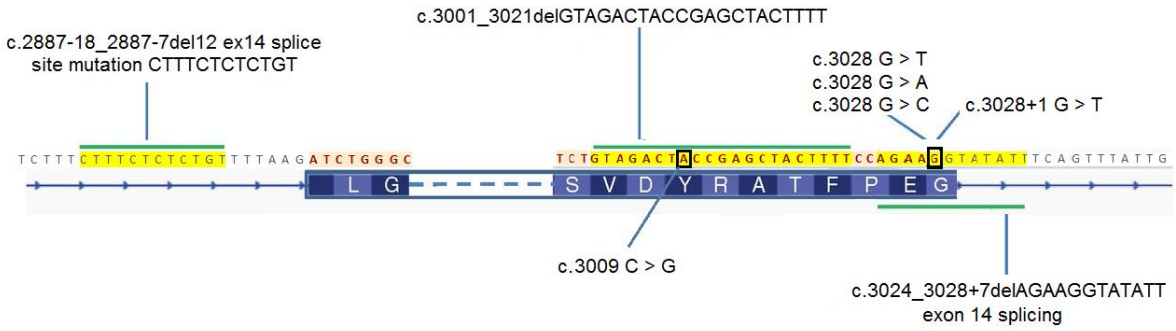

Figure 9: Diagram of the most frequent alterations of MET exon 14.

All these alterations were reported in previous studies (Jenkins *et al*. 2015; Paik *et a*l. 2015; Frampton *et al*. 2015; Kollmannsberger *et al*. 2015 and Awad *et al*. 2016), and are considered actionable mutations, since the patients harboring these mutations had partial or total therapeutic response to TKIs. Among the tested drugs, Crizotinib was the one that demonstrated most cases with a complete response (Schrock *et al*., 2016).

MET mutational analysis by DNA sequencing was achieved in the total of the 172 samples previously analyzed for EGFR mutations. The results indicated that 9 samples

(5.2%) harbored alterations in MET gene, and the remaining 162 (94.8%) were classified as wild-type. In total of altered samples, we found 2 intronic variants (cases 128 and 136), 2 indel mutations (cases 131 and 135) and 5 pontual mutations in exon 14 (cases 24, 82, 112, 128 and 131). Among these results two samples showed two MET mutations (case 128 and 131).

MET exon 14 alterations include a heterogeneous group of mutations (affecting MET exon 14 but also its adjacent intronic regions, some of them altering the process of splicing producing a Met variant that lacks the exon 14. The alternative splicing represents a physiological process that leads to the production of multiple protein isoforms from the same genetic information codified by a single gene (Drilon et al., 2016). Although carefully regulated, a series of pathological mechanisms (as gene fusions, splice site mutations or mutations in genes encoding splicing factors) can trigger the production of alternative RNA transcripts sustaining the development of different types of disease, such as cancer (Srebrow & Kornblihtt, 2006). The two intronic variants reported in this study - ivs14+38A>G - are both pontual mutations at the nucleotide position 3056 of intron 14, in which an Adenine is replaced by a Guanine (c.3056 A>G) (Figure 10). Transcript variants within an intron have many possibilities for regulating genes. Intronic variants may affect alternative splicing of the mRNA, affecting the splice donor or acceptor sites and secondary structure. For the purpose of assessing the possible alternative splicing induced by this change, a splice-site analysis tool developed by the Berkely Drosophila Genome Project (BDGP) (www.fruitfly.org/seq_tools/splice.html) was used. According to our *in silico* analysis, this alteration does not give rise to a new splicing, and therefore it should not be assumed as pathogenic. In this way, since this alteration has no established clinical significance, it is classified as unclassified variant (UVs).

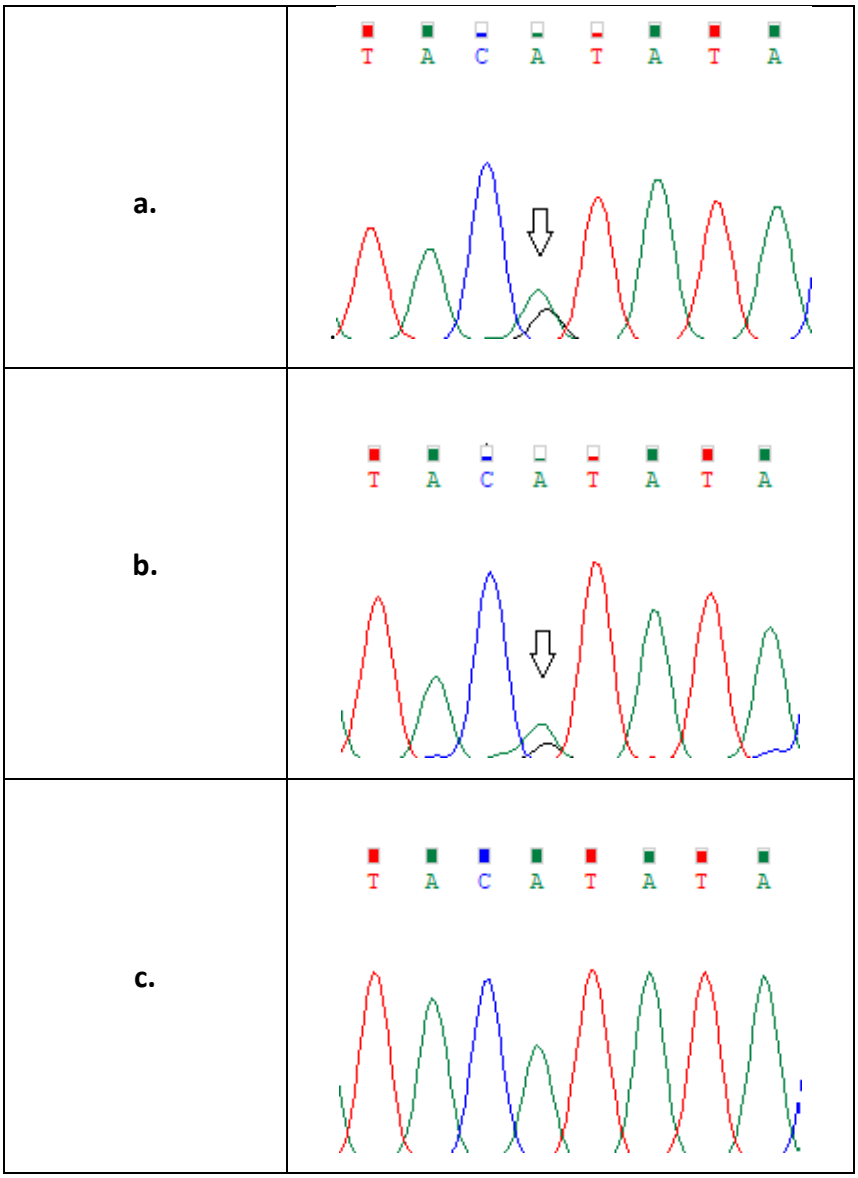

Figure 10: Electropherograms of two cases of intronic variants in MET gene (↙). a) ivs14+38A>G (case 128); b) ivs14+38A>G (case 136); c) Wild-type sample.

In the figure below, we show a silent mutation in two cases (128 and 131). Adenine is substituted by a guanine, at nucleotide position 3028 (CCG>CCA), which originates the same coded amino acid (Proline) (Figure 11). Since silent mutations do not alter amino acids, they are generally considered nonfunctional in cancer. However, Supek *et al*, in 2014, presented a compelling analysis suggesting that such silent mutations can be oncogenic by altering transcript splicing and thereby affecting protein function. We

performed an *in silico* analysis using the above mentioned splice site prediction tool and no alternative splicing was predicted. As such, this alteration should not be assumed as pathogenic because it does not give rise to a new splicing. The same assumptions can be made about case 82, which reports another silent mutation without predicted splicing changes, in which a Thymine is substituted by an Adenine, giving rise to the same amino acid –Alanine (Ala1005Ala) (Figure 12).

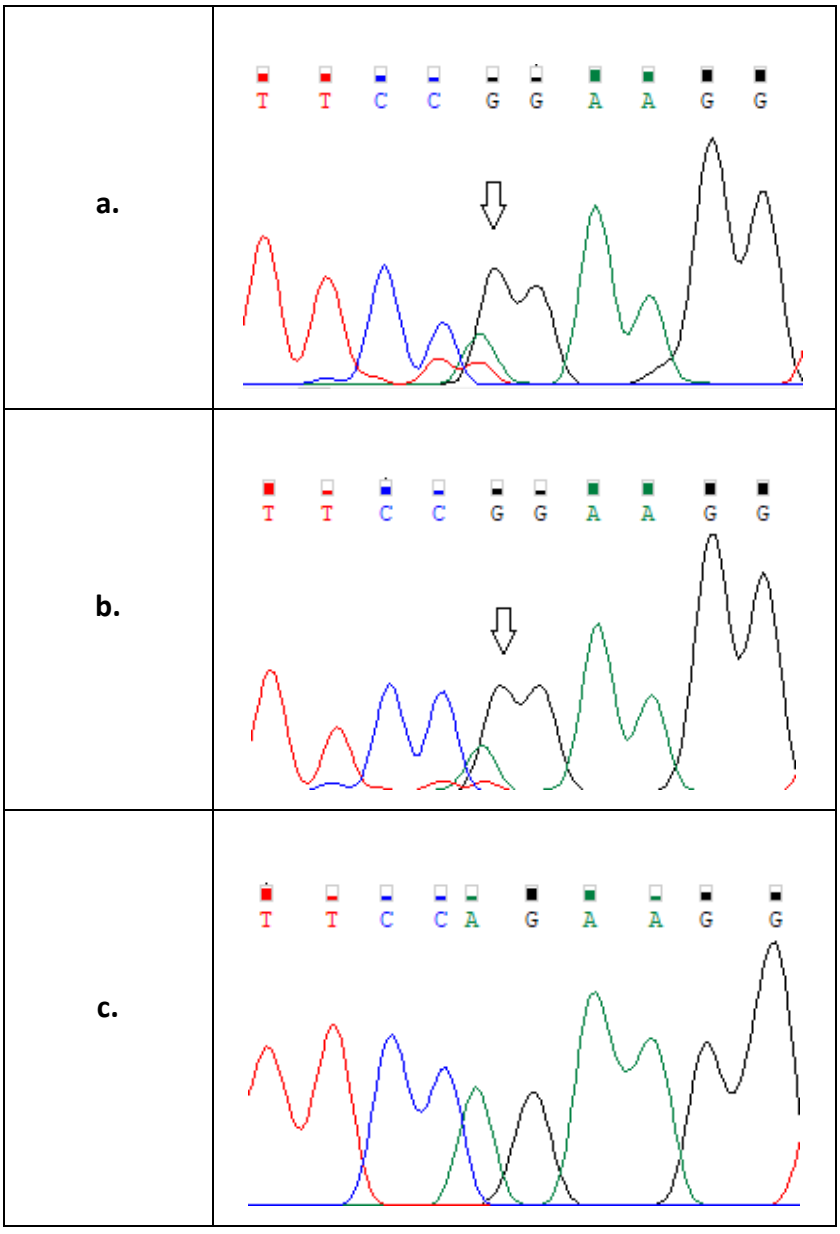

Figure 11: Electropherograms of two cases of silent mutations in codon 3024 of exon 14 MET gene (↙). a) c.3024A>G; Pro1008Pro (case 131); b) c.3024A>G; Pro1008Pro (case 128); c) Wild-type sample.

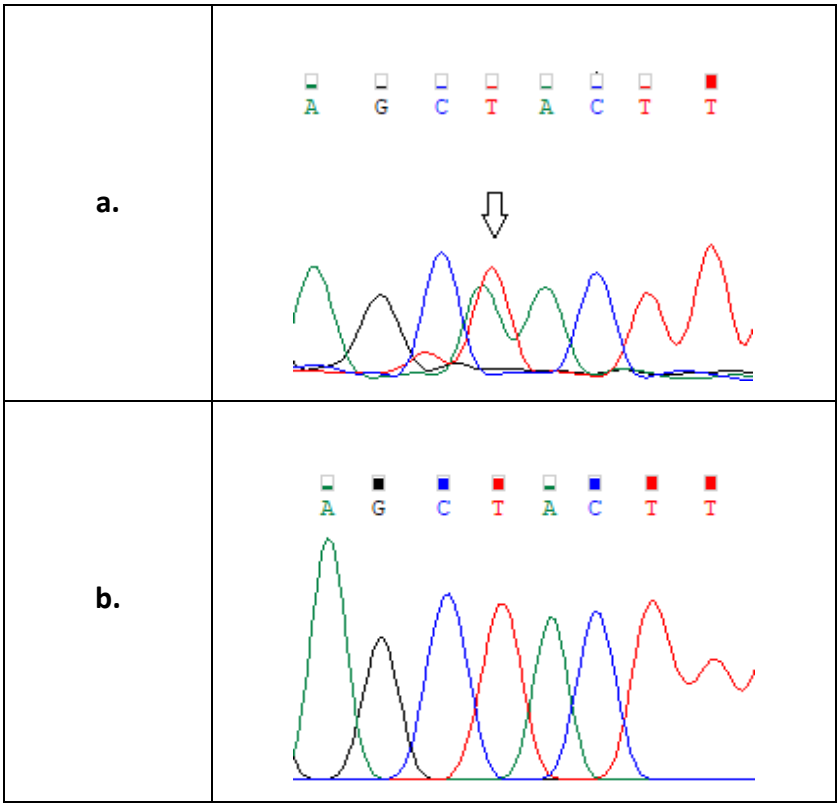

Figure 12: Electropherograms of one case of a silent mutation in codon 3015 of 14 MET gene (↙). a) c.3015T>A; Ala1005Ala (case 82); b) Wild-type sample.

In one case (case 24) we identified a mutation in tyrosine residue 1003, in which the substitution of a Cytosine for a Guanine, led to a stop codon (UAG) in the reading frame (Figure 13). This nonsense mutation – Y1003X – was reported in previous studies. According to the *The Cancer Genome Atlas Research Network*, MET Y1003X results in the formation of a stop codon in the juxtamembrane domain of the Met protein which disrupts several splice enhancer sites. Consequently, Y1003X results in exon skipping deletion of exon 14, increased Met kinase activity, and downstream Akt and Erk signaling. Schrock and colleagues, in the last year, identified 6 patients with this alteration and, according to them, 1003 mutations are functionally analogous with the

clinical sequelae of METex14, although responses in patients with Y1003X mutations to MET-targeted therapies have not yet been reported.

Physiological Met signalling is tightly regulated, with activation of Met being directly and acutely coupled to its degradation. According to Awad *et al*., (2016), activated Met is rapidly internalized and delivered to the sorting endosome, from which a proportion is recycled back to the membrane, while the rest is directed to the multivesicular body (MVB) and then undergoes degradation in the lysosome. Ubiquitination of Met is required for efficient sorting by the endosome, and is dependent on phosphorylation of Tyr-1003 in the Met juxtamembrane domain, which leads to binding of the CBL tyrosine kinase binding domain and CBL activation (Park *et a*l., 2015). When Met receptor in which Tyr-1003 is missing or mutated, is not directed to the MVB, but is instead trafficked back to the cell surface (Awad *et al*., 2016).

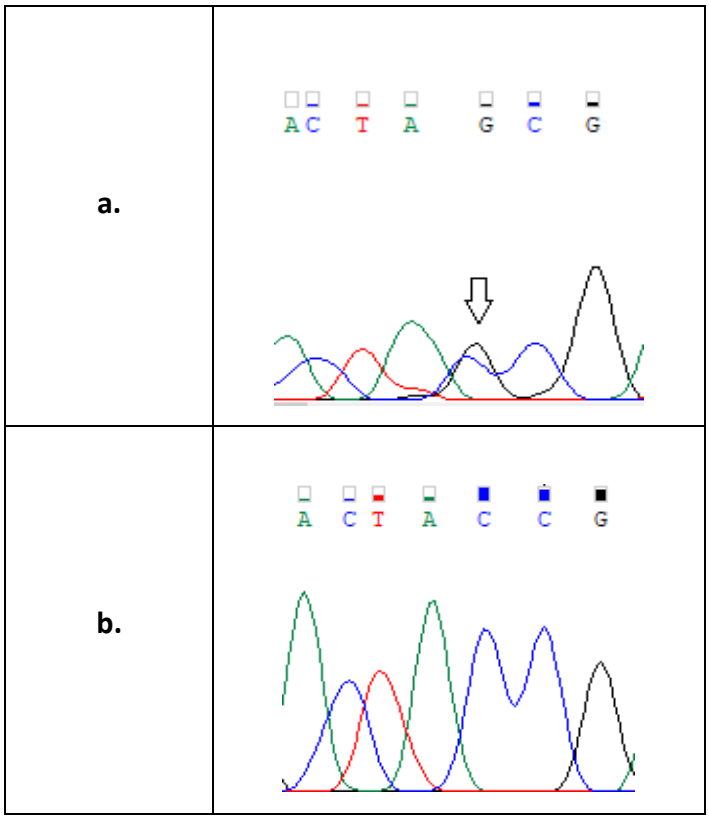

Figure 13: Electropherograms of one case of a nonsense mutation in codon 3009 of 14 MET gene (↙). a) c.3009C>G; Tyr1003X (case 24); b) Wild-type sample.

Like EGFR, MET mutations are usually mutually exclusive with other driver mutations (Quintanal-Villalonga *et al*., 2016). All the mutations found in this project are in agreement with such evidence, except for the nonsense mutation described above. The case number 24 has a concomitant EGFR/MET mutations presents a EGFR p.Leu858Arg exon 21 mutation that is, not only a class II activating mutation, but also a TKI sensitizing mutation (Xu *et al*., 2015). On the other hand, the concomitant MET mutation (p.Tyr1003X) decreases ubiquitination and increases stability of the MET protein. In this sample, two different scenarios could be considered: each mutation took place in different tumor clones, or both mutations occur in the same clone at different time points. The first scenario, observed in the study Benesova *et al*. (2010), but for co-mutations between EGFR and KRAS, would indicate the polyclonal nature of this lung tumor through the existence of a subpopulation with a TKI sensitizing EGFR mutation (p.Leu858Arg), as well as another subpopulation with the other mutation. TKI treatment in such cases could result in a partial response until the TKI-sensitive subpopulation (p.Leu858Arg) eventually disappears and only the TKI-resistant subpopulation (the other mutation) remains. The second scenario postulates the p.Leu858Arg EGFR mutation as the primary event and the occurrence of the p.Tyr1003X MET mutation as a mechanism of disease progression leading to a metastatic NSCLC resistant to TKI therapy.

Another mutation identified in our series in case number 112, occurred more precisely at nucleotide position 3013 of the exon 14 (Figure 14). It is a missense mutation that results in the smaller hydrophobic amino acid alanine being replaced by the smaller polar amino acid serine – Ala1005Ser – which may lead to a possible structural alteration in the protein. Perschard and colleagues, have demonstrated that the Met receptor is predicted to form a salt bridge between 1002Asp and 1004Arg in the DpYR motif to stabilize the peptide conformation most favorable to expose 1003Tyr of Met toward the phosphotyrosine binding pocket of c-Cbl. According to them, the substitution of either 1002Asp or 1004Arg in the DpYR motif would result in a loss of the salt bridge, which would be expected to alter the projected orientation of 1003Tyr toward the

phosphotyrosine binding pocket of c-Cbl, decreasing the binding affinity of the TKB domain of c-Cbl to the Met receptor. In our case, alanine and serine are not charged amino acids, so their side chains do not make salt bridges.

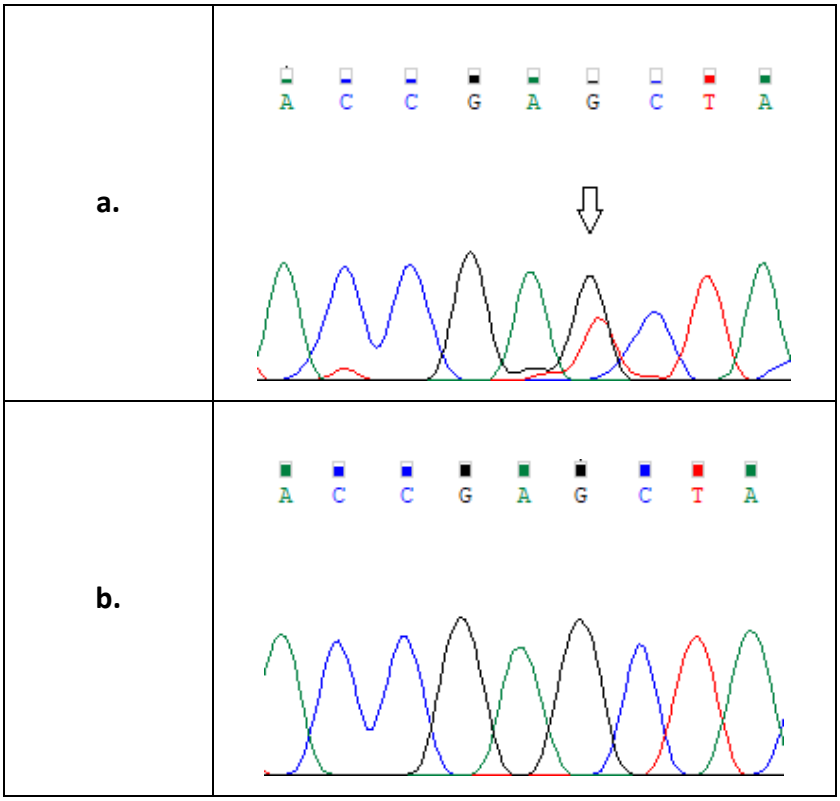

Figure 14: Electropherograms of one case of a missense mutation in codon 3013 of 14 MET gene (↙). a) c.3013G>T; Ala1005Ser (case 112); b) Wild-type sample.

The remaining alterations founded in MET gene are indel mutations (cases 131 and 135). It is notable a length difference between two alleles but it is unknowable if the difference was originally caused by a sequence insertion or by a sequence deletion. Both alterations should be in start of de sequencing, so it is not easily detected by electropherograms their exact location. The last base in the wild-type samples is a guanine (position 119/120 of electropherogram) and in both cases we have more bases amplified, showing an existence of an insertion or and deletion. In case number 131 it is noteworthy a difference of 16 bases and in case number 135 there are a discrepancy of 7 bases (Figure 15). Dozens of distinct MET exon 14 sequence variants have been

described. Moreover, they include base substitutions, deletions, insertions, or complex indels that can be located at splice acceptor or donor sites. Schrock, in 2016, analyzed 298 cases with METex14 alterations, which have detected 165 different variants predicted to affect MET exon 14. These included 157 base substitutions and 145 indels. This alterations affected the splice acceptor site in 104 cases, the splice donor site in 191 cases, and the approximately 25–base pair intronic noncoding region immediately adjacent to the splice acceptor site in seven cases. Jenkins and colleagues, have founded an intronic deletion, c. 2887-18_2887-7del12, that resides in the polypyrimidine tract just upstream of the exon 14 5' splice acceptor site. Mutations affecting this tract or the 3' splice donor site cause in-frame skipping of exon 14. In our results, the information of the electropherograms don't give us the exactly location of the alterations, so we cannot conclude if splice acceptor site or splice donor site are affected, although, according to the described MET mutation profile, this is very likely.

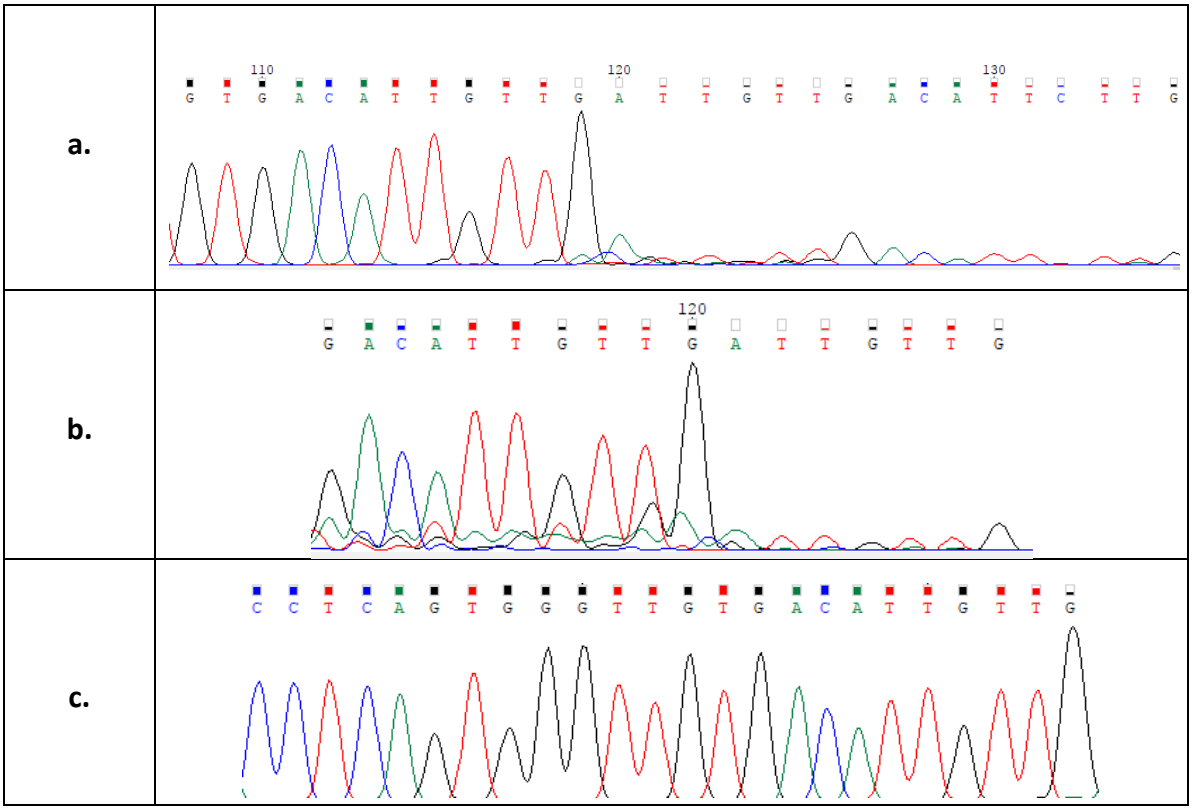

Figure 15: Electropherograms of two cases of indel mutations in exon 14 of MET gene. a) Case 131 – indel mutation of 16 bases; b) Case 135 – indel mutation of 7 bases; c) Wild-type sample.

4.2 ERBB2 Mutation Analysis

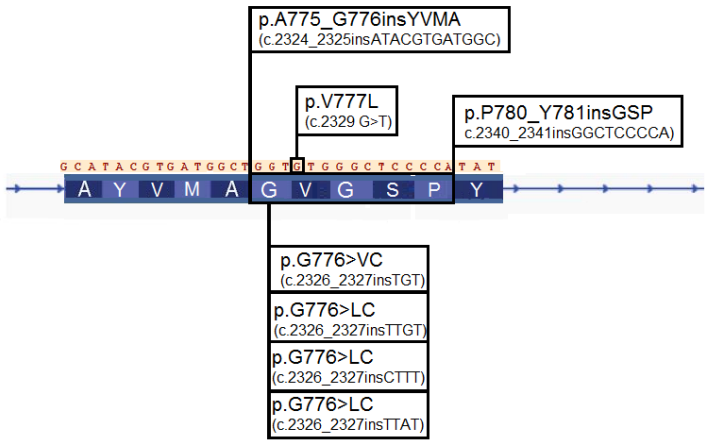

Figure 16: Diagram of the most frequent alterations of ERBB2 exon 20.

Within the duration scheduled for this work, ERBB2 mutational analysis by DNA sequencing was achieved in 69 samples out of the 172 that were previously analyzed for EGFR mutations and for MET. The results indicated that a single sample (1.5%) harbored an alteration in ERBB2 gene, and the remaining 68 (98.5%) were wild-type.

This alteration was an in-frame insertion of 12 nucleotides. This is predicted to result in a change of only a few amino acids, and, therefore it may still be possible for the protein to function, even though its sequence may be slightly different.

The great majority of ERBB2 mutations found in lung cancer usually involve small in-frame insertions and point mutations in exon 20 which have been observed mostly in adenocarcinomas. Most of the exon 20 insertions occur between amino acids 775 and 781, with p.A775_G776insYVMA as the most frequently observed alteration (50-80% of cases). These insertions induce a conformational change of the autoinhibitory αC-β4 loop, thus narrowing the ATP-binding cleft and promoting enhanced kinase activity (Suzawa *et al*., 2016). According to Suzuki *et al*, p.A775_G776insYVMA, p.G776VC, p.G776LC, and p.P780insGSP mutations account for 74% of all ERBB2 mutations in lung cancers (Figure 16).

The exact location of the alteration detected in ERBB2 could not easily be mapped through analysis of the obtained electropherogram, so we cannot predict the exact consequences of this mutation although, according to the described mutational profile of ERBB2, this is very likely to consist of an insertion affecting the ATP-binding cleft and promoting enhanced kinase activity, similar to the remaining described mutations.

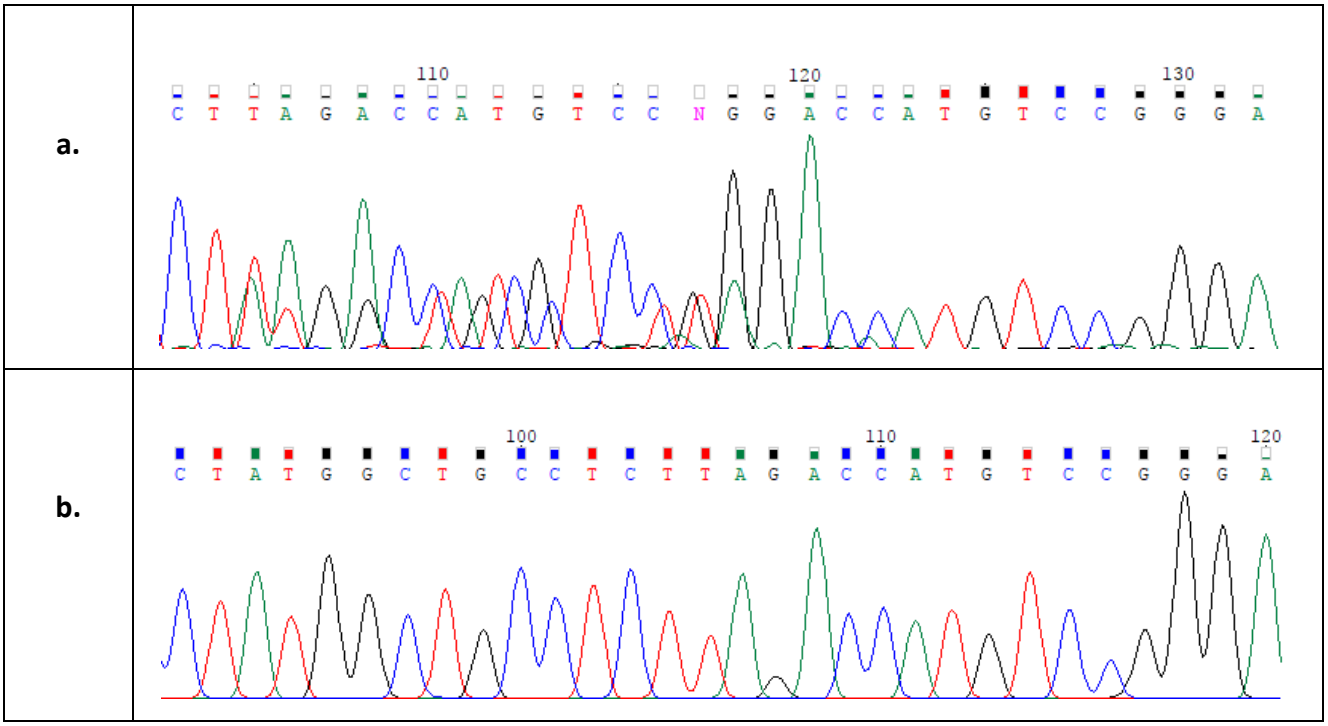

Figure 17: Electropherograms of one case of an insertion in exon 20 of ERBB2 a) insertion of 12 bases (case 49); b) Wild-type sample.

4.3 Design of a multiplex RT-PCR assay

Quantitative real-time PCR (qPCR), also known as Real-Time PCR, is widely and successfully used in clinical and biological fields for quantification of nucleic acid sequences. This is a sensitive and specific technique in which the DNA amount is monitored during the reaction by using fluorescent dyes that are incorporated into the PCR product. This technique became possible after introduction of an oligonucleotide

probe which was designed to hybridize within the target sequence. Cleavage of the probe during PCR because of the 5' nuclease activity of Taq polymerase can be used to detect amplification of the target-specific product. The increase in the fluorescent signal is directly proportional to the number of PCR product molecules generated.

As an output of our work, we designed a RT-PCR assay to detect new actionable mutations in MET and ERBB2 according to the procedure outlined in the flowchart below (Figure 18).

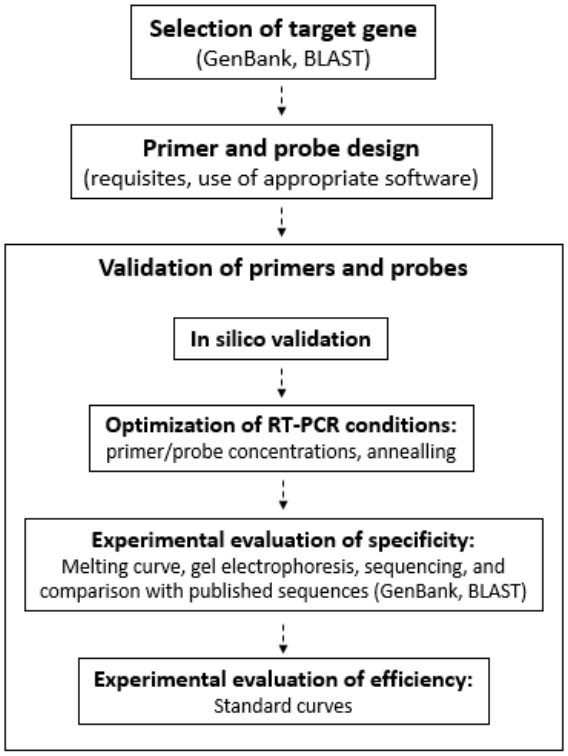

Figure 18: Flowchart summarizing the procedure used for designing the assay and the primers and probes for the genotyping real-time PCR protocol.

4.3.1 Mapping of the target sequences

We have assembled the mutations described in the literature and the alterations detected in this study, and we have mapped their location in the gene sequence (e.g., Ensemble) in order to define the target sequences of the assay. Then, we designed primers to amplify these regions and we designed probes to detect the respective mutations.

4.3.2 Primer & Probe design

A preliminary and key step when designing primers and probes for qPCR is the selection of target sequences of nucleic acid where they can hybridize. Once the selection of the target sequences has been done, the next step is to find potential primers or probes targeting regions on the corresponding gene sequences. Indeed, design of primers and probes is one of the most crucial factors affecting the success and quality of quantitative real-time PCR (qPCR) analyses, since an accurate and reliable amplification depends on using efficient primers and probes.

This can be done manually using sequence alignment program or automatically using primer design software. In both techniques it should be checked that the suggested primer and probe set achieves the following criteria, for avoid or reduce the formation of primer-dimers and other non-specific products: amplicon length, melting temperature (Tm), primer and probe length, GC content, self-complementary, primer-dimer and hairpin formation, degree of degeneracy, 5’ end stability, and 3’ end specificity (Integrated DNA Techologies, Inc [US]).

Primers were designed using the Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/primer3/). The list of primers is represented in Table 2. The factors to take into account when designing the optimal primers for the gene of interest are:

- Primers sequence length is within the optimal range of 18-30 base pairs.
- The G/C content is between 35% and 65%.
- The primers have no Primers should not have 4 or + G consecutive regions.
- The melting temperature should be 60°C - 65°C (optimum=62°C) and pairs should not differ by more than 1-2°C.
- The primers result in an amplicon length that is between the 100-200 bp optimal range for qPCR.

Likewise, there are some tips to consider when designing probes:

- The probe should be in close proximity to the forward or reverse primer, but should not overlap with a primer-binding site on the same strand.
- Probes should have a Tm 6–8°C higher than the primers.
- GC content should be 35−65% and avoid a G at the 5’ end to prevent quenching of the 5’ fluorophore.

The list of designed primers and probes is represented the tables 5 and 6.

Table 5: List of primers designed for RT-PCR. All probes were designed to detect the most frequents alterations in each gene, according to the literature. ERBB2=Human epidermal growth factor 2; MET= c-MET or mesenchymal-epithelial transition factor.

Gene	Mutation locus	Forward	Reverse
ERBB2	p.A775_G776insYVMA p.G776>VC p.P780_Y781insGSP p.V777L p.G776>LC (ins ttgt) p.G776>LC (ins cttt) p.G776>LC (ins ttat)	5'-CCCTCTCAGCGTACCCTTGT-3'	5'-CCCGGACATGGTCTAAGAGG-3'
MET	c.2887-18_2887-7del	5'-ACCCATGAGTTCTGGGCACT-3'	5'-CACTTCGGGCACTTACAAGC-3'
MET	c.3009C>G c.3028G>A c.3028G>C c.3028G>T c.3028+1G>T c.3001_3021del c.3024_3028+7del	5'-AGGCTTGTAAGTGCCCGAAG-3'	5'-CAACAATGTCACAACCCACTGA-3'

Table 6: List of probes designed for RT-PCR. All probes were designed to detect the most frequents alterations in each gene, according to the literature. ERBB2=Human epidermal growth factor 2; MET= c-MET or mesenchymal-epithelial transition factor.

Gene	Mutation	Probe name	Probe sequence
ERBB2	p.A775_G776insYVMA	ERBB2_1	ACACCGCCATCACGTATAGCCATCA
ERBB2	p.G776>VC	ERBB2_2	GAGCCCACACACAGCCATCA
ERBB2	p.P780_Y781insGSP	ERBB2_3	AGACATATGGGGAGCCTGGGGAG
ERBB2	p.V777L	ERBB2_4	GGGGAGCCCAAACCAGCCAT
ERBB2	p.G776>LC (ins ttgt)	ERBB2_5	GAGCCCACACACAACAGCCATCA
ERBB2	p.G776>LC (ins cttt)	ERBB2_6	AGCCCACACAAAGCAGCCATCA
ERBB2	p.G776>LC (ins ttat)	ERBB2_7	GGAGCCCACACATAACAGCCATCA
ERBB2	WT	ERBB2_WT	GAGCCCACACCAGCCATCACGTA
MET	c.3028G>A	MET_1	ACTGAAATATACTTTCTGGAAAAGTAGCTCGG
MET	c.3028G>C	MET_2	TGAAATATACGTTCTGGAAAAGTAGCTCGG
MET	c.3028G>T	MET_3	ACTGAAATATACATTCTGGAAAAGTAGCTCGG
MET	c.3028+1G>T	MET_4	CTGAAATATAACTTCTGGAAAAGTAGCTCGGT
MET	c.3001_3021del	MET_5	GAAATATACCTTCTGGAGATTCATTTGAAACCA
MET	c.2887-18_2887-7del	MET_6	CACTGCCCAGATCTTAAAAAAGAGCTTGT
MET	c.3024_3028+7del	MET_7	AGAACAATAAACTGGGAAAAGTAGCTCGG
MET	c.3009C>G	MET_8	CTTCTGGAAAAGTAGCTCGCTAGTCTACAGATTCA
MET	WT	MET_WT	ACTGAAATATACCTTCTGGAAAAGTAGCTCGG

4.3.3 Genotyping Assay using TaqMan probes

TaqMan® probes were designed to bind to a specific sequence (mutation or wild-type) in the target sequence.

Assays were designed to include an oligonucleotide probe containing a fluorescent reporter dye at the 5' end and quencher at the 3' end. This oligonucleotide was designed in such a way that it will bind to downstream of the primer binding site in the target DNA molecule and two TaqMan probes, one complementary to the wild allele and the other one to the mutated allele. While the probes are inactive, the quenchers interact with the dye through FRET mechanism and disable fluorescent activity. During annealing stage of PCR, TaqMan probes hybridize with corresponding DNA molecules (either wild-type or mutated). During the extension stage of the PCR, the hybridized probe(s) are cleaved due to 5`-nuclease activity of Taq polymerase. This leads to increasing of the dye fluorescence intensity. In this way, the genotype of the sample is determined by the ratio of fluorescence intensity of two different dyes present on distinct probes for wild-type and mutated alleles (Figure 19).

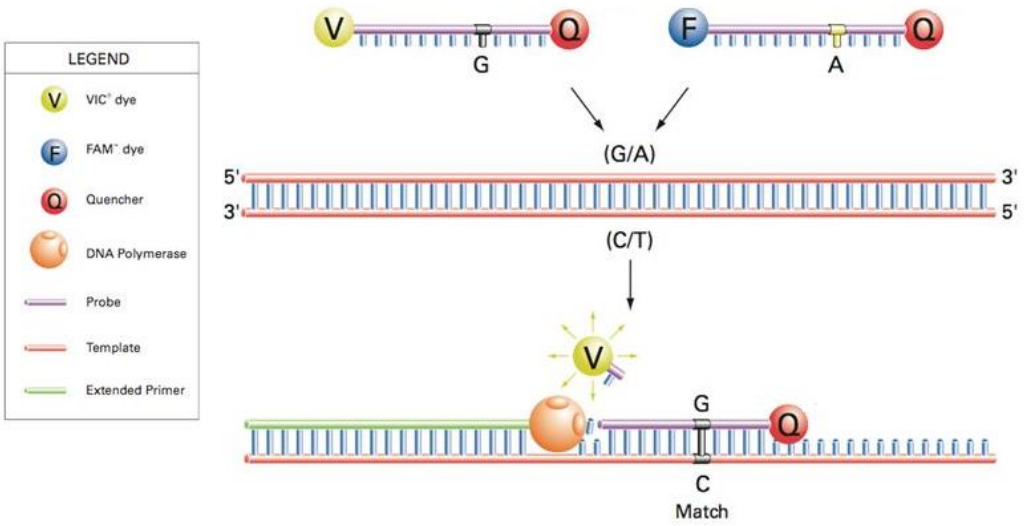

Figure 19: TaqMan assay. (From: *www.testgen.ru/en/technology/genotyping.html*)

4.3.4 Multiplexing and use of Fluorophores

All probes were TaqMan® probes with ABI dyes - FAM™, HEX® or Cy5™ - and BBQ as the quencher. These dyes were chosen as they are compatible with the Real-Time PCR Cobas z 4800, which is used at Molecular Pathology Laboratory of IPO Coimbra FG, EPE.

The assay was multiplexed so that each well of the RT-PCR plate (for each gene) contains the 3 fluorophores - FAM™, HEX® and Cy5™. FAM™, HEX® dyes will detect the mutations descripted in table x. The Cy5 ™ dye was chosen for the wild-type alleles.

Table 7: Fluorophores used for RT-PCR assay. ERBB2=Human epidermal growth factor 2; MET= c-MET or mesenchymal-epithelial transition factor.

	FAM™	HEX®	Cy5™
MET	c.3009C>G c.3028G>A c.3028G>C c.3028G>T c.3028+1G>T	c.2887-18_2887-7del c.3001_3021del c.3024_3028+7del	WT
ERBB2	p.A775_G776insYVMA p.G776>LC (ins ttgt) p.G776>LC (ins cttt) p.G776>LC (ins ttat) p.G776>VC	p.P780_Y781insGSP p.V777L	WT

Chapter IV

5. CONCLUSION AND FUTURE PERSPECTIVES

Taking into account the results obtained in this study, we take the following conclusions:

- The discovery and success of targeted therapies has launched a new era of lung cancer research focused on the detection and treatment of targetable alterations.
- The frequency of EGFR mutations in NSCLC was 11.1%, being the majority alterations in exon 19, which is in accordance to the literature.
- Were found 5.2% of cases that harbored alterations in the MET gene and 1.5% in ERBB2, both of which are targeted by drugs approved in different settings (Crizotinib and Transtuzumab).
- This demonstrates that a proportion of NSCLCs could potentially be treated with approved drugs for these targets, with evidence accumulating to show that a clinical benefit is achievable.
- EGFR and MET concomitant mutations may co-exist (presumably in different clonal populations) as a sign of tumor heterogeneity or in response to “kinase switching” in response to treatment, and this occurred in about 1% of the samples analyzed. Nonetheless, mutual exclusivity of primary genetic changes in lung carcinogenesis is the general rule.
- The alterations in MET and ERBB2 genes are found in the frequent altered hotspots involved in auto-inhibitory domains of these receptors (exon 14 for MET gene and exon 20 for ERBB2), establishing that apparently deleterious mutations can occur in oncogenes and still result in their activation.
- The strategy of Sanger sequencing analysis to identify cases of NSCLC with MET and ERBB2 allows identification of a broad spectrum of genetic changes and thus can contribute to expand the mutation profile of these targets.
- Future perspectives: Validate the RT-PCR assay with the samples already detected through Sanger sequencing.

6. BIBLIOGRAPHY

American Cancer Society. Lung Cancer (Non-Small Cell). Atlanta, Ga: American Cancer Society; 2016.

Arcila M, Chaft J, Nafa K, et al. Prevalence, clinicopathologic associations, and molecular spectrum of ERBB2 (HER2) tyrosine kinase mutations in lung adenocarcinomas. Clin Cancer Res. 2012; 18: 4910–4918.

Awad M, Oxnard G, Jackman D. MET Exon 14Mutations in Non–Small-Cell Lung Cancer Are Associated With Advanced Age and Stage-Dependent MET Genomic Amplification and c-Met Overexpression. J Clin Oncol. 2016 Mar 1;34(7):721-30.

Azad A, Lawen A, Keith J. Prediction of signaling cross-talks contributing to acquired drug resistance in breast cancer cells by Bayesian statistical modeling. BMC Systems Biology. 2015; 9:2.

Benesova L, Minarik M, Jancarikova D, et al. Multiplicity of EGFR and KRAS mutations in non-small cell lung cancer (NSCLC) patients treated with tyrosine kinase inhibitors. Anticancer Research. 2010; 1667-1672.

Bocharov E, Lesovoy D, Pavlov K, et al. Alternative packing of EGFR transmembrane domain suggests that protein-lipid interactions underlie signal conduction across membrane. Biochim Biophys Acta. 2016; 1858:1254–61.

Bonanno L, Jirillo A and Favaretto A. Mechanisms of acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors and new therapeutic perspectives in non small cell lung cancer. Curr Drug Targets. 2011; 12(6): p. 922-33.

Bordi P, Marzia D, Bordi P and Danesi R. Circulating DNA in diagnosis and monitoring EGFR gene mutations in advanced non-small cell lung cancer. Transl Lung Cancer Res. 2015; 4(5): 584–597.

Bronte G, Rizzo S, La Paglia L, et al. Driver mutations and differential sensitivity to targeted therapies: a new approach to the treatment of lung adenocarcinoma. Cancer Treatment Reviews 2010; 36S3: S21-S29.

Carcereny E, Morán T, Capdevila L, et al. The epidermal growth factor receptor (EGRF) in lung cancer. Translational Respiratory Medicine. 2015; 3:1.

Ciardiello F and Tortora G. EGFR antagonists in cancer treatment. N Engl J Med. 2008; 1160-1174.

Clinical Lung Cancer Genome Project (CLCGP), Network Genomic Medicine (NGM). A genomics-based classification of human lung tumors. Science Translational Medicine. 2013; 5(209), 209ra153–209ra153.

Cooper W, Lam D, O'Toole S and Minna J. Molecular biology of lung cancer. Journal of Thoracic Disease. 2013; 5(Suppl 5):S479-S490.

De Grève J, Teugels E, Geers C, et al. Clinical activity of afatinib (BIBW 2992) in patients with lung adenocarcinoma with mutations in the kinase domain of HER2/neu. Lung Cancer. 2012; 76(1):123-7.

De Sousa E Melo F, Vermeulen L, Fessler E and Medema JP. Cancer heterogeneity-a multifaceted view. EMBO reports. 2013; 686-95.

Dela Cruz C, Tanoue L and Matthay R. Lung Cancer: Epidemiology, Etiology, and Prevention. Clinics in chest medicine. 2011; 32(4).

Detterbeck F, Boffa D and Tanoue L. The New Lung Cancer Staging System. Chest. 2009; 260-271.

Dietel M, Jöhrens K, Laffert M, et al. A 2015 update on predictive molecular pathology and its role in targeted cancer therapy: a review focussing on clinical relevance. Cancer Gene Ther. 2015; 22(9):417-30

Drilon A, Cappuzzo F, Ou S, et al. Targeting MET in Lung Cancer: Will Expectations Finally Be MET?. J Thorac Oncol. 2016; 15-26.

Ercan D, Choi H, Yun C, et al. EGFR mutations and resistance to irreversible pyrimidine based EGFR inhibitors, Clin. Cancer Res. 2015; 3913–3923.

Feng Y, Thiagarajan P and Ma P. MET signaling: novel targeted inhibition and its clinical development in lung cancer. J Thorac Oncol. 2012; 7(2):459-67.

Frampton G, Ali S, Rosenzweig M, et al. Activation of MET via Diverse Exon 14 Splicing Alterations Occurs in Multiple Tumor Types and Confers Clinical Sensitivity to MET Inhibitors. Cancer Discov. 2015; 5(8):850-9.

Gazdar A. EGFR Mutations in Lung Cancer: Different Frequencies for Different Folks. J Thorac Oncol. 2014; (2):139-40.

Gazdar A. Epidermal growth factor receptor inhibition in lung cancer: the evolving role of individualized therapy. Cancer Metastasis Rev. 2010; 29(1):37-48.

Govindan R. Overcoming resistance to targeted therapy for lung cancer, N. Engl. J. Med. 2015; 1760–1761.

Greulich H, Kaplan B, Mertins P, et al. Functional analysis of receptor tyrosine kinase mutations in lung cancer identifies oncogenic extracellular domain mutations of ERBB2. Proc Natl Acad Sci U S A. 2012; 109:14476-81.

Haber D and Velculescu V. Blood-based analyses of cancer: circulating tumor cells and circulating tumor DNA. Cancer Discov. 2014; 4: 650–661.

Hanahan D and Weinberg R. Hallmarks of cancer: the next generation. Cell. 2011; 144:646–674.

Hecht S. Tobacco smoke carcinogens and lung cancer. Journal of the National Cancer Institute. 1999; 1194–1210.

Hecht S. Cigarette smoking and lung cancer: chemical mechanisms and approaches to prevention. The Lancet Oncology. 2002; 3(8), 461–469.

Heist S, Shim S, Gingipally S, et al. MET exon 14 skipping in non-small cell lung cancer. Oncologist. 2016; 21:481–486.

Khoo C, Rogers T-M, Fellowes A, et al. Molecular methods for somatic mutation testing in lung adenocarcinoma: EGFR and beyond. Translational Lung Cancer Research. 2015; 4(2):126-141.

Kris M, Camidge D, Giaccone G, et al. Targeting HER2 aberrations as actionable drivers in lung cancers: phase II trial of the pan-HER tyrosine kinase inhibitor dacomitinib in patients with HER2-mutant or amplified tumors. Ann Oncol. 2015; 26(7):1421-7.

IARC. (2012). GLOBOCAN 2012. Retrieved October 27, 2016, from www.globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

Jackman D, Yeap B, Linderman N, et al. Phase II clinical trial of chemotherapy-naive patients > or = 70 years of age treated with erlotinib for advanced non-small-cell lung cancer. J Clin Oncol. 2007; 25(7): 760-6.

Janjigian Y, Azzoli C, Krug L, et al. Phase I/II trial of cetuximab and erlotinib in patients with lung adenocarcinoma and acquired resistance to erlotinib. Clin Cancer Res. 2011; 17(8):2521-7.

Janne P, Smith I, McWalter G, et al. Impact of KRAS codon subtypes from a randomised phase II trial of selumetinib plus docetaxel in KRAS mutant advanced nonsmall- cell lung cancer. British Journal of Cancer. 2015; 199-203.

Jenkins R, Oxnard G, Elkin S, et al. Response to crizotinib in a patient with lung adenocarcinoma harboring a MET splice site mutation. Clin Lung Cancer. 2015; 16(5):e101-4.

Larsen J and Minna J. Molecular biology of lung cancer: clinical implications. Clin Chest Med. 2011; 32(4):703-40.

Keating G. Afatinib: a review of its use in the treatment of advanced nonsmall cell lung cancer. Drugs. 2014; 207-21.

Kong-Beltran M, Seshagiri S, Zha J, et al. Somatic mutations lead to an oncogenic deletion of met in lung cancer. Cancer Res 2006; 66:283-9.

Kono S, Marshall M, Ware K and Heasley L. The Fibroblast Growth Factor Receptor Signaling Pathway as a Mediator of Intrinsic Resistance to EGFR-specific Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. Drug Resist Updat. 2009; 12(4-5):95-102.

Kovacs E, Zorn J, Huang Y, et al. A structural perspective on the regulation of the epidermal growth factor receptor. Annu Rev Biochem. 2015; 84:739–64.

Lawrence M, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013; 499:214–218.

Lin L and Bivona T. Mechanisms of Resistance to Epidermal Growth Factor Receptor Inhibitors and Novel Therapeutic Strategies to Overcome Resistance in NSCLC Patients. Chemother Res Pract. 2012; 2012:817297.

Lovly C, Horn L and Pao W. EGFR. My Cancer Genome. 2013. http://www.mycancergenome.org/content/disease/lungcancer/egfr/?tab=0 (Updated October 16)

Lovly C, Horn L and Pao W. KRAS in Non-Small Cell Lung Cancer (NSCLC). My Cancer Genome. 2015.
https://www.mycancergenome.org/content/disease/lung-cancer/kras/ (Updated June 18).

Lutterbach B, Zeng Q, Davis L, et al. Lung cancer cell lines harboring MET gene amplification are dependent on Met for growth and survival. Cancer Res. 2007; 2081–2088.

Mahjoubi L, Gazzah A, Besse B, et al. A never-smoker lung adenocarcinoma patient with aMETexon 14 mutation (D1028N) and a rapid partial response after crizotinib. Invest New Drugs. 2016; 34(3):397-8.

Mar N, Vredenburgh J and Wasser J. Targeting HER2 in the treatment of non-small cell lung cancer. Lung Cancer. 2015; 87(3):220-5.

Mazières J, Peters S, Lepage B, et al., Lung cancer that harbors an HER2 mutation: epidemiologic characteristics and therapeutic perspectives. J Clin Oncol, 2013. 31(16): p. 1997-2003.

Morgensztern D, Campo M, Dahlberg S et al. Molecularly Targeted Therapies in Non–Small-Cell Lung Cancer Annual Update 2014. J Thorac Oncol. 2015; 10:S1–S63.

Nakagawa H, Wardell C, Furuta M, et al. Cancer whole-genome sequencing: present and future. Oncogene. 2015; 34(49):5943-50.

Niederst M, Hu H, Mulvey H, 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. 2015; 3924–3933.

Ogunleye F, Ibrahim M, Stender M et al. Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Advanced Non-Small Cell Lung Cancer - A paradigm shift in stage IV non-small cell lung cancer treatment. Am J Hematol Oncol. 2015; 11(1):16-25.

Onozato R, Kosaka T, Kuwano H, et al. Activation of MET by Gene Amplification or by Splice Mutations Deleting the Juxtamembrane Domain in Primary Resected Lung Cancers. J Thorac Oncol. 2009; 4:5–11.

Ou SH, Schrock A, Bocharov E, et al. HER2 transmembrane mutations (V659/G660) that stabilize homo- and heterodimerization are rare oncogenic drivers in lung adenocarcinoma that respond to
Afatinib. Journal of Thoracic Oncology. 2017; 446-457.

Oxnard G, Binder A and Janne P. New targetable oncogenes in non-small-cell lung cancer. J Clin Oncol. 2013; 31:1097–104.

Paik P, Drilon A, Fan P, et al. Response to MET inhibitors in patients with stage IV lung adenocarcinomas harboring MET mutations causing exon 14 skipping. Cancer Discov. 2015; 5:842–849.

Pakkala S and Ramalingam S. Epidermal Growth Factor Receptor Mutated Advanced Non–Small Cell Lung Cancer: A Changing Treatment Paradigm. Hematology/oncology clinics of North America. 2017; 31:83-99.

Pao W and Chmielecki J. Rational, biologically based treatment of EGFR mutant non-small-cell lung cancer. Nature Reviews Cancer. 2010; 760-774.

Pao W and Girard N. New driver mutations in non-small-cell lung cancer. Lancet Oncol. 2011; 12:175-80.

Pao W, Miller V, Zakowski M, et al. EGFR receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101:13306–13311.

Park S, More S, Murtuza A, et al. New Targets in Non–Small Cell Lung Cancer. Hematol Oncol Clin. 2017; 113–129.

Park S, Koh J, Kim DW, et al. MET amplification, protein expression, and mutations in pulmonary adenocarcinoma. Lung Cancer. 2015; 90(3):381-7.

Park K, Tan E, O'Byrne K, et al. Afatinib versus gefitinib as first-line treatment of patients with EGFR mutation-positive non-small-cell lung cancer (LUX-Lung 7): a phase 2B, open-label, randomized controlled trial. Lancet Oncol. 2016; 17(5):577-89.

Peschard P, Ishiyama N, Lin T, et al. A Conserved DpYR Motif in the Juxtamembrane Domain of the Met Receptor Family Forms an Atypical c-Cbl/Cbl-b Tyrosine Kinase Binding Domain Binding Site Required for Suppression of Oncogenic Activation. J Biol Chem. 2004; 279(28):29565-71.

Peters S and Zimmermann S. Targeted therapy in NSCLC driven by HER2 insertions. Transl Lung Cancer Res. 2014; 3(2):84-8.

Quintanal-Villalonga A, Paz-Ares L, Ferrer I and Molina-Pinelo S. Tyrosine Kinase Receptor Landscape in Lung Cancer: Therapeutical Implications. Disease Markers. 2016; 2016:9214056.

Reck M, Heigener D, Mok T, et al. Management of non-small-cell lung cancer: recent developments. Lancet. 2013; 382(9893): p. 709-19.

Roskoski R. ErbB/HER protein-tyrosine kinases: Structures and small molecule inhibitors. Pharmacol Res. 2014; 87:42–59.

Rosell R, Carcereny E, Gervais R, et al, Erlotinib versus chemotherapy (CT) in advanced non-small cell lung cancer (NSCLC) patients (p) with epidermal growth factor receptor (EGFR) mutations: Interim results of the European Erlotinib Versus Chemotherapy (EURTAC) phase III randomized trial (2011 ASCO Annual Meeting). J Clin Oncol 29. 2011; (suppl; abstr 7503).

Salgia R. Mutation testing for directing upfront targeted therapy and post-progression combination therapy strategies in lung adenocarcinoma. Expert Review of Molecular Diagnostics. 2016; 16(7):737-749.

Sangwan V and Park M. Receptor tyrosine kinases: role in cancer progression. Curr. Oncol. 2006; 13, 191–193.

Scagliotti G, Parikh P, von Pawel J, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced stage non-small-cell lung cancer. J Clin Oncol. 2008; 26(21):3543-51.

Schrock A, Frampton G, Suh J, et al. Characterization of 298 Patients with Lung Cancer Harboring MET Exon 14 Skipping Alterations. J Thorac Oncol. 2016;11(9):1493-502.

Shi J, Hua X, Zhu B, et al. Somatic Genomics and Clinical Features of Lung Adenocarcinoma: A Retrospective Study. PLoS Med. 2016;13(12):e1002162.

Shigematsu H, Takahashi T, Nomura M, et al. Somatic Mutations of the HER2 Kinase Domain in Lung Adenocarcinomas. Cancer Res. 2005 Mar 1;65(5):1642-6.

Schwab R, Petak I, Kollar M, et al. Major partial response to crizotinib, a dual MET/ALK inhibitor, in a squamous cell lung (SCC) carcinoma patient with de novo c-MET amplification in the absence of ALK rearrangement. Lung Cancer, 2014; 83(1): p. 109-11.

Sharma S and Settleman J. Oncogene addiction: setting the stage for molecularly targeted cancer therapy. Genes Dev. 2007; 21(24):3214-31.

Srebrow A and Kornblihtt A. The connection between splicing and cancer. J Cell Sci. 2006; 119:2635-41. Stephens et al., 2004 Lung cancer: Intragenic ERBB2 kinase mutations in tumours

Su K, Chen H, Li K, et al. Pretreatment epidermal growth factor receptor (EGFR) T790M mutation predicts shorter EGFR tyrosine. J Clin Oncol. 2012; 30(4):433-40.

Supek F, Miñana J, Valcárcel T, et al. Synonymous mutations frequently act as driver mutations in human cancers. Cell. 2014; 156(6):1324-35.

Suzuki M, Shiraishi K, Yoshida A, et al. HER2 gene mutations in non-small cell lung carcinomas: concurrence with Her2 gene amplification and Her2 protein expression and phosphorylation. Lung Cancer. 2015; 87(1): p. 14-22.

Takezawa K, Pirazzoli V, Arcila M, et al. HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR mutant lung cancers that lack the second-site EGFRT790M mutation. Cancer Discov. 2012; 2:922–933.

Tebbutt N, Pedersen M and Johns TG. Targeting the ERBB family in cancer: couples therapy. Nat Rev Cancer. 2013; 13:663–673.

The E and Belcher E. Lung cancer: diagnosis, staging and treatment. Cardiothoracic surgery II. 2014; 242-248.

The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014; 511:543–550.

Travis W, Brambilla E and Riely G. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. J Clin Oncol. 2013; 31(8): p. 992-1001.

Thress K, Paweletz C, Felip E, et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M,.Nat. Med; 2015. 560–562.

Vineis P, Airoldi L, Veglia F, et al. Environmental tobacco smoke and risk of respiratory cancer and chronic obstructive pulmonary disease in former smokers and never smokers in the EPIC prospective study. BMJ. 2005; 330:277.

WHO. (2011). Global status report on noncommunicable diseases 2010 (pp. 1–176)

Wu D, Wu T, Chen C, and Lee H. PAK1 is a novel therapeutic target in tyrosine kinase inhibitor-resistant lung adenocarcinoma activated by the PI3K/AKT signaling regardless of EGFR mutation. Clinical Cancer Research. 2016; 1;22(21):5370-5382.

Xu M, Xie Y, Ni S and Liu H. The latest therapeutic strategies after resistance to first generation epidermal growth factor receptor tyrosine kinase inhibitors (EGFR TKIs) in patients with non-small cell lung cancer (NSCLC). Annals of Translational Medicine. 2015; 3(7):96.

Yamamoto H, Higasa K, Sakaguchi M, et al. Novel germline mutation in the transmembrane domain of HER2 in familial lung adenocarcinomas. J Natl Cancer Inst. 2014; 106:djt338.

Yeung S, Tong J, Law P, et al. Profiling of Oncogenic Driver Events in Lung Adenocarcinoma Revealed MET Mutation as Independent Prognostic Factor. J Thorac Oncol. 2015; 10(9):1292-300.

Yoshizawa A, Sumiyoshi S, Sonobe M, et al. HER2 status in lung adenocarcinoma: A comparison of immunohistochemistry, fluorescence in situ hybridization (FISH),dual-ISH, and gene mutations. Lung Cancer. 2014; 85(3):373-8.

Zhang Y, Wang D, Shi L, et al. Genome analyses identify the genetic modification of lung cancer subtypes. Seminars in Cancer Biology. 2016; 20-30.

APPENDIX I

Discrimination of the mutational status of *EGFR*, *MET* and *ERBB2* in 172 NSCLC samples

ND – Not done

Wt – Wild type

Case number	EGFR alterations	MET alterations	ERBB2 alterations
1	Wt	wt	wt
2	Wt	wt	wt
3	Wt	wt	wt
4	Wt	wt	wt
5	Wt	wt	wt
6	p.Leu858Arg	wt	ND
7	Wt	wt	ND
8	Wt	wt	ND
9	Wt	wt	wt
10	Wt	wt	wt
11	Wt	wt	ND
12	p.Lys745_Glu749del	wt	wt
13	Wt	wt	ND
14	p.Lys745_Glu749del + p.Thr790Met	wt	wt
15	p.Lys745_Glu749del	wt	wt
16	wt	wt	wt
17	wt	wt	wt
18	wt	wt	ND
19	wt	wt	ND
20	wt	wt	wt
21	wt	wt	wt

22	wt	wt	wt
23	wt	wt	ND
24	p.Leu858Arg	Tyr1003X (c.3009C>G)	wt
25	wt	wt	wt
26	wt	wt	ND
27	wt	wt	ND
28	wt	wt	wt
29	wt	wt	wt
30	wt	wt	wt
31	wt	wt	wt
32	wt	wt	ND
33	wt	wt	ND
34	wt	wt	ND
35	wt	wt	wt
36	wt	wt	wt
37	wt	wt	wt
38	wt	wt	wt
39	wt	wt	wt
40	wt	wt	wt
41	wt	wt	wt
42	wt	wt	wt
43	wt	wt	wt
44	wt	wt	wt
45	wt	wt	wt
46	wt	wt	wt
47	wt	wt	wt
48	wt	wt	wt
49	wt	wt	Insertion (12 bases)
50	wt	wt	wt
51	wt	wt	wt

52	wt	wt	wt
53	wt	wt	wt
54	wt	wt	wt
55	wt	wt	wt
56	wt	wt	wt
57	p.Leu747_Ala750del	wt	wt
58	wt	wt	wt
59	p.Leu861Gln	wt	wt
60	wt	wt	ND
61	wt	wt	ND
62	p.Glu746_Ala750del	wt	wt
63	p.Lys745_Glu749del	wt	wt
64	wt	wt	ND
65	wt	wt	ND
66	p.Leu858Arg	wt	ND
67	p.Leu858Arg	wt	wt
68	wt	wt	ND
69	wt	wt	ND
70	wt	wt	ND
71	wt	wt	ND
72	wt	wt	ND
73	wt	wt	wt
74	wt	wt	ND
75	wt	wt	ND
76	wt	wt	ND
77	wt	wt	wt
78	p.Lys745_Glu749del	wt	wt
79	wt	wt	ND
80	wt	wt	ND
81	wt	wt	ND
82	wt	Ala1005Ala (c.3015T>A)	wt

83	wt	wt	ND
84	wt	wt	ND
85	wt	wt	ND
86	wt	wt	wt
87	wt	wt	ND
88	wt	wt	ND
89	wt	wt	ND
90	wt	wt	ND
91	wt	wt	ND
92	wt	wt	wt
93	wt	wt	ND
94	wt	wt	wt
95	wt	wt	ND
96	wt	wt	wt
97	wt	wt	wt
98	p.Lys745_Glu749del	wt	wt
99	wt	wt	ND
100	wt	wt	ND
101	wt	wt	ND
102	wt	wt	ND
103	wt	wt	ND
104	p.Lys745_Glu749del	wt	wt
105	wt	wt	ND
106	wt	wt	ND
107	p.Lys745_Glu749del	wt	ND
108	wt	wt	ND
109	wt	wt	ND
110	p.Glu746_Ala750del	wt	wt
111	wt	wt	ND
112	wt	Ala1005Ser (c.3013G>T)	wt
113	wt	wt	wt

114	wt	wt	ND
115	wt	wt	ND
116	wt	wt	ND
117	wt	wt	ND
118	p.Lys745_Glu749del	wt	wt
119	wt	wt	wt
120	wt	wt	ND
121	wt	wt	ND
122	wt	wt	ND
123	wt	wt	ND
124	wt	wt	ND
125	wt	wt	ND
126	wt	wt	ND
127	wt	wt	ND
128	wt	ivs14+38A>G + Pro1008Pro (c.3024A>G)	wt
129	wt	wt	ND
130	wt	wt	ND
131	wt	Pro1008Pro (c.3024A>G) + Indel mutation (16 bases)	wt
132	wt	wt	ND
133	wt	wt	ND
134	wt	wt	ND
135	wt	Indel mutation (7 bases)	wt
136	wt	ivs14+38A>G	wt
137	wt	wt	wt
138	wt	wt	ND
139	wt	wt	ND
140	wt	wt	ND
141	wt	wt	ND
142	wt	wt	ND

143	wt	wt	ND
144	wt	wt	ND
145	wt	wt	ND
146	wt	wt	ND
147	wt	wt	ND
148	wt	wt	ND
149	wt	wt	ND
150	wt	wt	ND
151	wt	wt	ND
152	wt	wt	ND
153	wt	wt	ND
154	wt	wt	ND
155	wt	wt	ND
156	wt	wt	ND
157	wt	wt	ND
158	wt	wt	ND
159	p.Lys745_Glu749del	wt	ND
160	wt	wt	ND
161	wt	wt	ND
162	wt	wt	ND
163	wt	wt	ND
164	wt	wt	ND
165	wt	wt	ND
166	wt	wt	ND
167	wt	wt	ND
168	wt	wt	ND
169	wt	wt	ND
170	wt	wt	ND
171	wt	wt	ND
172	wt	wt	ND