Beatriz de Almeida Gonçalves Carrilho Mestre

DETECTION OF GENETIC BIOMARKERS OF LUNG CANCER USING EXHALED BREATH CONDENSATE

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

Agosto 2017

UNIVERSIDADE DE COIMBRA

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

DETECTION OF GENETIC BIOMARKERS OF LUNG CANCER USING EXHALED BREATH CONDENSATE

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 de Patologia Molecular, Serviço de Anatomia Patológica IPOCFG,EPE) e da Professora Doutora Paula Cristina Veríssimo Pires (Departamento de Ciências da Vida, Faculdade de Ciência e Tecnologia)

Beatriz de Almeida Gonçalves Carrilho Mestre

2017

ABSTRACT

Lung cancer, the second most diagnosed type of cancer, is one of the primary oncological causes of death in the world. This pathology has a "silent" course, without characteristic symptoms that would allow its precocious detection being absent. This leads to a late detection that often occurs in a non-operable phase of the cancer, contributing to a 5 year survival rate of only 18,7 %. In the rare cases that the cancer is detected in an initial state, the survival rate augments significantly to 50-70 %. Therefore, controlling lung cancer-derived mortality relays on the development of specific methods for precocious diagnosis that are employable to the population basis as a secondary prevention measure.

In the last four decades early diagnosis methods have been widely investigated as an option to reduce lung cancer mortality. In the 70's and the 80's several studies of screening were conducted with the use of thorax radiography and expectoration cytology methods in risk populations. However, these studies failed to show a reduction in mortality rate and, therefore, are not recommended as screening methods. Recently, *National Lung Screening Trial* (NLST) demonstrated that low-dose computerized tomography lead to a 20% reduction in the mortality rate in one risk population. The success of this method lead to its implementation in the USA. Nevertheless, European screening tests have yet to proof a clear advantage in the implementation of the latter method in high risk populations. Therefore, in the current state of the art there is no trial consensually recommended to be a part of a screening program in asymptomatic populations.

With the rise in cancer biology knowledge it is now known that several genetic alterations are associated with the development of this disease, particularly: mutations in oncogenes, methilation in CpG islands of tumor suppressor genes, loss of heterozygoty and microsatellite instability. These genetic and epigenetic alterations have not shown enough sensibility and specificity alone but a panel of molecular markers that reaches a desirable sensitivity and specificity for early diagnosis is very promising.

It was found that in cancer there is a higher release of DNA by tumor tissue into body fluids. In what concerns the lungs, the inferior airways are covered by a protector

fluid and studies have demonstrated that it is possible to obtain DNA from exhaled breath that it is thought to come from this fluid. Also, it was possible to detect molecular alterations characteristic of lung cancer in these types of samples, what puts into perspective its application as biological material for analysis of cancer molecular markers. In this sense, exhaled breath condensate has opened frontiers in the development of methodologies for early diagnosis in the lung cancer field. Based on these studies, the startup *Infogene* developed and patented a portable and self-collection blowing device denominated *Oncosopro,* which allows the recovery of biomolecules from exhaled breath condensate through a small filter.

The current thesis has as a main goal the optimization of the laboratory protocol for sample recovery by the patented blowing device, in order to obtain the major possible quantity of biological material for posterior molecular analysis. The efficiency of several protocols was evaluated based on two criteria: quantity and quality of the DNA, measured through espectrophotometry and another fluorometric method; and the integrity and quality of the DNA obtained for further amplification through PCR and electrophoresis. Through the modification of several parameters in the protocol, a higher quantity of DNA was obtained, meeting the stated goal. A posterior population trial in control groups formed by healthy people, smokers and ex-smokers over 50 years old showed this optimized extraction method is effective for recovery and amplification of the exhaled breath DNA.

A future phase will rely on the identification of a panel of genetic molecular markers sensible and specific for early detection of lung cancer as well as a method of prediction of the respective therapeutic answer using exhaled breath. Clinical utility can be validated through determination of several parameters, such as specificity, sensibility, positive predictive value and negative predictive value of this DNA recovering method. Upon the clinical validation, this method might be employed into screening programs in order to reduce the mortality rate in this oncological disease.

RESUMO

O cancro de pulmão é uma das principais causas oncológicas de morte em todo o mundo, constituindo o segundo tipo de cancro mais frequentemente diagnosticado. Esta patologia tem um decurso "silencioso", não produzindo sintomas característicos que permitam a sua detecção precoce, o que faz com que normalmente seja detectada tardiamente, já numa fase inoperável. Este diagnóstico tardio contribui para uma taxa de sobrevivência a 5 anos de apenas 18.7 %. No entanto, a sobrevida é significativamente melhor quando o cancro do pulmão é detectado em estágio inicial, com a taxa de sobrevivência a 5 anos a aumentar para 50-70%,. Assim sendo, o controlo da mortalidade por cancro do pulmão está dependente do desenvolvimento de métodos específicos de diagnóstico precoce que possam ser empregues no rastreio de base populacional como forma de prevenção secundária.

Nas últimas quatro décadas, o diagnóstico precoce tem sido amplamente investigado como uma opção para reduzir a mortalidade do cancro do pulmão. Nos anos 70 e 80 do século XX, foram praticados vários estudos de rastreio com recurso à radiografia ao tórax e citologia da expectoração em população de risco. No entanto, estes estudos não geraram qualquer redução na taxa de mortalidade e, por isso, não são recomendados como método de rastreio. Recentemente, a *National Lung Screening Trial* (NLST) demostrou que a tomografia computorizada de baixa dose levou a uma redução de 20% na taxa de mortalidade numa população de alto risco. O sucesso deste levou a que este método de rastreio fosse implementado nos Estados Unidos da América. Contudo, ensaios Europeus de rastreio do cancro do pulmão ainda não relataram uma vantagem clara na implementação deste método numa população de alto risco. Deste modo, no actual estado da arte não existe nenhum teste que seja consensualmente recomendado para fazer parte de um programa de rastreio em população assintomática.

Com o aumento do conhecimento da biologia do cancro do pulmão é sabido que diversas alterações genéticas estão associadas ao desenvolvimento desta doença, particularmente: mutações em oncogenes, metilação nas ilhas CpG de genes supressores de tumores, perda de heterozigotia e instabilidade de microssatélites. Estas alterações genéticas e epigenéticas não têm apresentado sensibilidade e especificidade aceitável por si só, pelo que a utilização de um painel de marcadores moleculares que atinja uma

sensibilidade e especificidade credível no diagnósticos precoce do cancro do pulmão é muito promissor.

Constatou-se que no cancro há uma maior libertação de DNA por parte do tecido tumoral para os fluidos corporais. No que toca ao pulmão, as vias aéreas inferiores são revestidas por um fluido protector e estudos demonstraram que é possível obter DNA do ar exalado que se pensa provir deste fluido. Além disto, foi possível detectar alterações moleculares características do cancro do pulmão neste tipo de amostras, o que perspectiva a sua aplicação como material biológico de partida para a análise de marcadores moleculares do cancro. Neste sentido, o ar exalado condensado tem aberto fronteiras no desenvolvimento de metodologias para diagnóstico precoce na área do cancro do pulmão. Com base nestes estudos, a Starup *Infogene* desenvolveu e patenteou um dispositivo de sopro, designado *Oncosopro*, portátil e de auto-colheita, que permite a recuperação de biomoléculas do ar exalado condensado através da sua captura num filtro.

A presente tese tem como principal objectivo optimizar o protocolo laboratorial de recuperação da amostra biológico a partir do dispositivo de sopro, de modo a obter a maior quantidade possível de material biológico para análises moleculares posteriores. A eficiência dos diversos protocolos foi avaliada com base em 2 critérios: quantidade e qualidade de DNA, pela medição da absorvância a 260, 280 e 230 nm por espectrofotometria e por um método fluorométrico e; a integridade e qualidade do DNA obtido para amplificação, através de PCR e electroforese em gel de agarose. Através da modificação de diversos parâmetros no protocolo, obteve-se uma maior quantidade de DNA, cumprindo-se o objectivo estipulado. O cumprimento deste levou à posterior realização de um estudo populacional num grupo de controlos, constituído por pessoas saudáveis, fumadoras e ex-fumadoras, com idade superior a 50 anos. Neste estudo foi possível obter e amplificar o DNA proveniente do ar exalado em todas as amostras da população, revelando a eficácia do método de extração otimizado.

Uma vez alcançada esta meta com sucesso, uma etapa futura passará pela identificação de um painel de marcadores moleculares genéticos sensíveis e específicos para a detecção precoce e previsão de resposta terapêutica do cancro do pulmão pelo ar exalado, através de um estudo clínico. Uma vez estabelecido o painel de marcadores será necessário proceder-se á validação clinica para determinar a especificidade,

sensibilidade, valor preditivo positivo e valor preditivo-negativo deste teste baseado no sopro. Caso a utilidade clinica seja comprovada, este método pode ser empregue em programas de rastreio, de modo a reduzir a taxa de mortalidade desta doença oncológica.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADC	Adenocarcinoma
ALF	Airway Lining Fluid
ALK	Anaplastic Lymphoma Kinase
AT2	Alveolar Epithelial 2
BAL	Bronchoalveolar Lavage
BASCs	Bronchioloalveolar Stem Cells
BRAF	B-Ras proto-oncogene, serine/threonine kinase
CCND1	Cyclin D1 protein
CK 7	Cytoqueratin 7
CT	Computed Tomography
CXR	Chest X-ray
DAPK	Death Aassociated Protein Kinase
DCC	Deleted in Colorectal Cancer
dNTP	Deoxyribonucleotide Triphosphates
EDTA	Ethylene Diamine Tetracetic Acid
EGFR	Epidermal Growth Factor Receptor
EML4	Microtubule-associated Protein of Echinoderm 4 Gene
ERBB2	Human Epidermal growth factor Receptor 2
ERCC1/2	Excision repair cross-complementation group ½
FGFR1	Fibroblast growth factor receptor
F-qMSP	Fluorescent Quantitative Methylation-Specific Polymerase Chain Reaction
FTA	Flinders Technology Associates
IASLC / ATS / ERS	International Association for the Study of Lung Cancer / American Thoracic Society / European Respiratory Society
IMH	Immunohistochemical
KRAS	Kirsten Rat Sarcoma
KRT 5	Keratin 5

LCC	Large Cell Carcinoma
LOH	Loss of Heterozygosity
LDCT	Low-Dose Computed Tomography
MA	Microsatellite Alteration
MET	Mesenchymal–Epithelial Transition
mtDNA	Mitochondrial DNA
miRNA	Micro RNA
MGMT	O-6-methylguanine-DNA Methytransferase
MMR	Mismatch repair
MSI	microsatellite instability
MYCL1	L-myc-1 proto-oncogene protein
NLST	National Lung Screening Trial
NSCLC	Non-Small Cell Cancer
NSCLC-NOS	Non- Small Cell Carcinoma-Not Otherwise Specified
NO	Nitric oxide
OSCC	oral squamous cell carcinomas
PCR	Polymerase Chain Reaction
PDCD4	Programmed cell death 4
PI3K	Phosphatidylinositide 3-Kinase
PIK3CA	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform
PTEN	Phosphatase and Tensin Homolog
PTFE	Polytetrafluoroethylene
RASSF1A	Ras Association Domain-Containing Protein 1
RB1	Retinoblastoma-associated protein
RIT 1	Ras-Like Protein Expressed In Many Tissues
ROS	Reactive Oxygen Species
ROS1	ROS proto-oncogene 1,receptor tyrosine kinase
rRNA	Ribosomal ribonucleic acid
SCC	Squamous Cell Carcinoma
SCLC	Small Cell Cncer
Scgb1a1	Secretoglobin Family 1a Member 1
SDS	Sodium Dodecyl Sulphate

SI	Spuctum Induction
SP-B	Surfactant Protein-B
SOX2	SRY-box2
STK11	Serine/threonine kinase 11
tBGS	Tag-adaptade Bisulfite DNA Sequencing
TE	Tris-HCL EDTA
TERT	Telomerase Reverse Transcriptase
TKIs	Tyrosine Kinase Inhibitors
TNM	Tumor Nodule Metasis
TPM1	Tropomyosin-1
TP 53/73	Tumor protein 53/73
TRIS	Tris-Hyroxymethul Methane
Trp63	Transcription Factor Transformation-Related Protein 63
tRNA	Transfer ribonucleic acid
TTF-1	Thyroid Transcription Factor 1
TURBO-DECCS	Transportable Unit for Research on Biomarkers Obtained from Disposable Exhaled Condensate Collection Systems
UTR	Untranslated Region
UV	Ultraviolet

CHAPTER I

1. INTRODUCTION

1.1 Epidemiology

In the last decades, lung cancer showed a decrease in incidence and mortality. (Villalobos & Wistuba, 2017). Wherein, the incidence rate in men began to decline in the mid-1980s, while in women it was in the mid-2000s. This difference between genders is due to the fact of historical patterns of adherence and smoking cessation (Facts, 2016). However, lung cancer continues to be one of the most incident cancers, making it the second most commonly diagnosed cancer in men and women (Silva, Pastorino, & Sverzellati, 2017). According to the American Cancer Society, 222500 new cases are expected to be diagnosed this year in the United States, representing about 25% of all cancer diagnosis. In addition, it is estimated to be 155870 death in 2017 (Facts, 2017). In Portugal, lung cancer is the most lethal, accounting for 3927 deaths in 2014 (Nacional, 2015). Lung cancer is a heterogeneous disease divided into two major groups: non-small cell cancer (NSCLC) and small cell lung cancer (SCLC), that require different treatment strategies (Warth et al., 2012). NSCLC represents 85% of cases of lung cancer and is further sub-categorized as adenocarcinoma (ADC) (45%), squamous cell carcinoma (SCC) (23%) , and large cell carcinoma (LCC) (13%). Small cell lung cancer represents 14% to 15% of all lung cancer. All of these class of lung cancer have diverse profiles of morphology and genetics from the others and develop in distinct zones of the lung (Eggert, Palavanzadeh, & Blanton, 2017; Hoffman, 2017; Zheng, 2016).

Risk factors associated with lung cancer are well known and smoking is the major determinant, accounting for 75-80% of lung cancer cases in men and 50% of cases in women mondial (Gridelli et al., 2015; Sateia et al., 2017; Teh, 2014). This risk is proportional to the number of cigarettes smoked, the duration of smoking and the age of onset. Unfortunately, in China, Indonesia and several African nations, smoking has increased and a simultaneous increase in lung cancer is expected in the coming decades (Sateia et al., 2017). For smokers, the risk of developing this pathology is, on average,

ten times higher than in non-smokers throughout life. However, smokers can benefit at any age by abandoning smoking. So the probability of developing cancer decreases comparatively with those who continue to smoke. As this abstinence period increases, the risk of lung cancer decreases but never reaches the risk of a non-smoker (Alberg & Samet, 2003). Once the main cause of lung cancer has been identified, the importance of primary prevention is highlighted. This strategy involves the prevention of smoking and smoking cessation programs in order to reduce the lung cancer mortality rate. Data suggest that this really primary prevention can significantly reduce lung cancer rates. (Alberg & Samet, 2003; Hoffman, 2017; Sateia et al., 2017; Silva et al., 2017) . Passive smoking is a risk factor of increasing importance, responsible for 1.6% of lung cancers, and this risk increases to 3.6% in cases of passive smoking during childhood (Manuscript & Factors, 2009). Smoking inhalation of cigarettes, cigars, and even electronic cigarettes, can provide the origin of mutations in oncogenes and a loss heterozygosity in tumor suppressor genes, which becomes the "drivers" of carcinogenesis (Eggert et al., 2017). Occupational or environmental exposure to certain metals (nickel, cadmium, arsenic), some organic chemicals, radiation, air pollution, and reduced vegetable and fruit consumption may also be risk factors. The risk for lung cancer may also be high in people with a history of tuberculosis (Facts, 2016; Teh, 2014) . Genetic susceptibility is another contributory factor for lung cancer, especially for non-smokers who normally develop the disease at a younger age (<50 years). For smokers who are genetically predisposed to develop lung cancer, smoking increases this risk. (Eggert et al., 2017; Facts, 2016). It should be noted that all these factors can act together or independently with smoking in the development of cancer (Malhotra, Malvezzi, Negri, Vecchia, & Boffetta, 2016).

The lung cancer is a asymptomatic pathology, without characteristic symptoms that allow its early detection, reason why 75% of the cases are diagnosed late (Heuvelmans, Groen, & Oudkerk, 2017; Hoffman, 2017). This translates into a 5-year survival rate of 18.7% (Sateia et al., 2017). However, this rate can be greatly improved when lung cancer is diagnosed early. In patients diagnosed with localized lung cancer, the survival rate rises to 55.2%, with only 15% of cases being detected at the initial stage (Heuvelmans et al., 2017; Hoffman, 2017; Sateia et al., 2017). Thus, early diagnosis remains a major challenge in identifying lung cancer at an early stage,

amenable to treatment and consequent prolongation of life expectancy and reduction of mortality rate (Silva et al., 2017).

In the 70s and 80s of the 20th century, several screenings were performed, using chest X-ray (CXR) and sputum cytology, in groups at risk (Mohammad & Brennan, 2017; Nacionais, Diagn, Do, & Do, n.d.). However, these tests are not recommended as screening methods for lung cancer, alone or in combination, because they have not shown a reduction in mortality rate (Hoffman, 2017; Mohammad & Brennan, 2017). computed tomography (CT) is considerate one of the main methods of imaging and investigation of lung disease, due to the fact that it allows the lesion characterization at morphological level, nodule size measurement and monitoring of nodule growth. Moreover, due to its three-dimensional nature, it makes possible an evaluation of the chest wall, as well as obtaining information regarding the staging of the tumor. Despite these numerous advantages, CT cannot be employed as a screening method because it exposes patients to radiation that may be carcinogenic, and the development of cancer increases with the highest radiation dose. In this sense, researchers intended to create a new method using a smaller dose of radiation possible without compromising CT image quality. Hence was born the low-dose computed tomography (LDCT) using significantly less radiation than the standard-dose CT (Mohammad & Brennan, 2017). In 2002, the *National Lung Screening Trial* (NLST) began the largest lung cancer screening study in the world, which compared LDCT with CXR (Eggert et al., 2017; Heuvelmans et al., 2017). Participants were smokers aged 55 to 74 years, with a jot 30-packs-year history of tobacco smoking who were actually smoking or had quit within the past 15 years. Participants were randomly selected by NLST to undergo 3 rounds of annual examinations with LDCT or CXR (Hoffman, 2017). In August 2011, the NLST revealed the results of the study, showing that LCDT reduces lung cancer mortality by 20% compared to CXR screening (Sateia et al., 2017; Thigpen & Geraci, 2016). After publication from NLST, several institutions have authorized this approach as a screening method for lung cancer, despite the initial concern about the high false positive rate (96%) in the LCDT group, the cost of annual LCDT, and radiation exposure, although it is one-third lower than the standard-dose CT (Eggert et al., 2017; Mohammad & Brennan, 2017; Sateia et al., 2017; Thigpen & Geraci, 2016). Despite the rapid implementation of this screening method for lung cancer in the United States, European trials have not yet demonstrated the benefit of this LCDT screening, so its

implementation is still under discussion (Silva et al., 2017). In short, LDCT is currently the preferred method of screening for lung cancer. However, clinicians still require methods and technologies that may be employed in the clinic to identify early-stage lung cancer.

1.2 Histology of Lung Cancer

For many years, the treatment of lung cancer depended on anatomy-based tumor nodule metastasis staging (TNM) to determine the methods to be employed (Ma et al., 2015). However, understanding the biology associated with lung cancer, such as the identification and characterization of key events triggering carcinogenesis along with an increased insighting of the genetic background, has led to the development of specific therapies for particular subtypes of lung cancer, thus announcing the rise of the age of personalized medicine (Davidson, Gazdar, & Clarke, n.d.; Zheng, 2016). For example, targeted therapy with epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (eg, gefitinib and erlotinib) is only effective in the treatment of pulmonary ADC with mutations in the EGFR gene. The microtubule-associated protein of echinoderm 4 Gene –anaplastic lymphoma kinase (EML4- ALK) fusions appear quasi solely in ADC and are targeted by the ALK inhibitor crizotinib (Montezuma, Azevedo, & Lopes, 2013). The pemetrexed, antifolate chemotherapy drug, is only effective for the treatment of pulmonary ADC, while bevacizumab should be avoided in SCC as it can cause pulmonary haemorrhage (Langer, Besse, Gualberto, Brambilla, & Soria, 2017; Montezuma et al., 2013; Zheng, 2016). Thus, personalized medicine consists of the application of an individual treatment based on precise histological classification and biomarkers information (Zheng, 2016).

The characterization of the histological type of lung cancer plays an increasingly significant role in the multidisciplinary approach in diagnosis, as well as in the treatment selection and prognosis (Liu et al., 2017; Zheng, 2016). In addition, information derived from histological subclassification is essentially important in patients with advanced disease (NSCLC phase III and IV) and in patients with metastatic NSCLC, since most of these patients are not candidate for surgical resection of the tumor (Manuscript, 2014a). In this scenario, the lack of precise subclassification

can subject patients to unnecessary treatments with serious side effects, as well as, deny effective treatment (Liu et al., 2017).

Nevertheless, a strict classification may be difficult to obtain in small biopsy specimens or cytological specimens, in which observation of the tissue is limited, due to the scarcity of tumor cells, lack of architecture characteristic of the tumor, and the presence of high heterogeneity histology within the same lesion (Manuscript, 2014a). In the case of poorly differentiated samples, these do not present architectural characteristics of differentiation and only show a small part of the overall tumor, giving a challenge to the pathologists in their subclassification (Zheng, 2016). In these circumstances, in which the cell morphology alone is not reliable for a definitive classification, the immunohistochemical (IHC) study assists in the subclassification of the tumor (Manuscript, 2014a). IHC represents an important complementary tool for the routine diagnosis of lung cancer and for the identification of various histological types and prognostic factors. This technique focuses on the fact that the cancer is characterized by one or more tumor markers, such as: enzymes, oncogenes, tumor-specific antigens, tumor suppressor genes; whose expression are seen as signs in cancerous conditions. These markers are sometimes distinct in several types of cancer, allowing their distinction (Capelozzi, 2009). Several studies address the usefulness of IHC markers in the diagnosis and subclassification of lung cancer, with the caveat that each marker has a variable sensitivity and specificity(Carney, Kraynie, & Roggli, n.d.). However, the clinical question of how to create an IHC panel with a limited number of highly specific and sensitive markers is still under discussion. In addition, a daily challenge in clinical practice involves using a minimal amount of tumor tissue at the same time it makes an accurate and rapid diagnosis, as it may be necessary to conduct additional molecular studies to tissue (Gurda et al., 2015).

Summing up, morphological classification combined with IHC should be included in the routine for the determination of the diagnosis in order to ensure accurate treatment of patients with lung cancer (Bittermann et al., 2013).

1.2.1 Adenocarcinoma

Currently, ADC represents the most dominant histological subtype of all types of lung cancer. It is commonly identified in the female and non-smoking population and is therefore less associated with a smoking history (Davidson et al., n.d.). ADCs have been proposed to arised from club cells of bronchioles, that express secretoglobin family 1a member 1 (Scgb1a1), bronchioloalveolar stem cells (BASCs) located in the bronchioloalveolar junctions and alveolar epithelial type 2 (AT2) cells (Z. Chen, Fillmore, Hammerman, & Kim, 2014). The fact that these cells are located in the bronchiolar and alveolar epithelium is consonant with the appearance of ADCs in the most distal airways, predominantly in the periphery of the lungs. By definition, ADC is a malignant epithelial neoplasm characterized by glandular differentiation or mucin production (Zheng, 2016). The cytological characteristics of ADCs comprise clusters of tumor cells with prominent nucleoli, cytoplasm with vacuoles, predominate or overt mucin production, acinar formation and other particulars characteristic of glandular differentiation (Gurda et al., 2015).

Pulmonary ADC cells generally express pneumocytic markers, consistent with an origin in the distal airways (Micke et al., 2016; Zheng, 2016). The lung ADC markers accepted in daily practice for the identification of ADC differentiation are thyroid transcription factor 1 (TTF-1) and Napsin A, the first, exhibiting a nuclear staining, being more easily evaluated (Gurda et al., 2015; Micke et al., 2016; *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015) TTF-1 is a nuclear transcription factor that is expressed in lung epithelial cells, type II pneumocytes and club cells, and thyroid (Carney et al., n.d.; Manuscript, 2014a). In the lung, it controls the expression of genes implicated in the generation of surfactant. This marker has been the most useful IHC marker for ADC identification, with a sensitivity of 75-80% (Manuscript, 2014a). However, TTF-1 is immunoreactive in various other tumors such as: thyroid neoplasms, metastatic breast carcinoma, gastrointestinal carcinomas, neuroendocrine tumors, for example SCLC, and possibly, but with some controversy, SCC of lung (Gurda et al., 2015; Hiroshima, Iyoda, Shida, Shibuya, & Iizasa, 2006). It is also known that its expression decreases inversely with the degree of tumor differentiation, i.e., well-differentiated ADCs have a

greater possibility of expressing TTF-1 compared to poorly differentiated ADCs. Recently, Napsin A has been employed as a novel marker for pulmonary ADCs. This is an aspartic protease attached in the cytoplasm of type II pneumocytes, club cells, and alveolar macrophages in the lungs, being implicated in the post-translational modification of surfactant protein B (SP-B) (Bİr, Çelİker, Evyapan, Yaren, & Edİrne, 2016; Gurda et al., 2015; Manuscript, 2014a). Previous studies have shown that Napsin A had a sensitivity and specificity superior to TTF-1 in moderately well-differentiated ADCs. Therefore, the combined use of these two markers has been implemented in the differential diagnosis of pulmonary ADCs, with a sensitivity of 80%. In addition, both can be used as prognostic markers for pulmonary ADC (Gurda et al., 2015; *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS -*, 2015; Zheng, 2016). Cytoqueratin 7 (CK7), expressed in normal ciliated, columnar cells in the lung, has been also used as a marker of differentiation of adenocarcinoma by some groups although its use is not universally accepted, perhaps because of its frequent positivity in SCC (Carney et al., n.d.; *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS -*, 2015).

1.2.2 Squamous Cell Carcinoma

In recent decades, the incidence of SCC has been declining relative to the ADC, presumably due to changes in smoking behaviors ("p40 (D Np63), a Lung Squamous Cell Marker , Can Also Be Used to Label Breast Myoepithelial Cells," 2014; Zheng, 2016). Being strongly related to smoking history, SCC is often located in the central portion of the lung in close relation to the bronchi (Langer et al., 2017). However, the incidence in the periphery of the lung has been increasing ("p40 (D Np63), a Lung Squamous Cell Marker , Can Also Be Used to Label Breast Myoepithelial Cells," 2014). It is presumed that SCC does not arise directly from the original bronchial epithelium, but rather in areas of continuous injury and repair, progressing to squamous metaplasia, squamous dysplasia, and eventually evolving to carcinoma in situ, as a consequence of sources of irritation (Davidson et al., n.d.; Giangreco et al., 2012). It has been hypothesized that SCC originates from basal cells, given that this type of cancer is located in the proximal part of the lung, but also because they express

transcription factor transformation-related protein 63 (Trp63), SRY-box 2 (SOX2) and keratin 5 (KRT5), which are markers of the normal basal cell population (Z. Chen et al., 2014).

Histologically, tumor cells often exhibit hyperchromatic nuclei, visible to discrete nucleoli and abundant cytoplasm with delineated intercellular bridges. The tumor cells lack glandular structure or mucin production. SCCs are subdivided into keratinized and nonkeratinized (Zheng, 2016). The keratinized SCC is recognized by the presence of keratinization, pearl formation, and/ or intercellular bridges. These morphological characteristics vary with the degree of differentiation, which are prominent in well differentiated tumors, where there is a characteristic keratinization (*WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015). From the well-known recognition of these characteristics, it is possible to establish a diagnosis of SCC, even in small biopsy specimens. When the tumor is poorly differentiated and does not allow reliable morphological classification, selective squamous cell markers are used to certify squamous differentiation (Zheng, 2016).

The monoclonal p63 (4A4) antibody marker was often recommended to label SCCs because of their high sensitivity (approximately 100%). However, it has a low specificity due to its reactivity in 30% of pulmonary ADCs and other types of tumors, particularly lymphomas. This situation of strong expression of p63 could lead to a misclassification of the tumor type as well as diagnosis and consequent therapy (Bishop et al., 2011; Dvorak et al., n.d.; "p40 (D Np63), a Lung Squamous Cell Marker , Can Also Be Used to Label Breast Myoepithelial Cells," 2014). Clearly, a squamous differentiation marker that shares the sensitivity of p63, but also that is highly specific, would be extremely useful. Formerly, p63 was thought of as a single molecule, but in fact there are two variants (isoforms) differing in the structure of the N-terminal domain: a longer TAp63 isoform containing the N-terminal transactivation domain and acts as a tumor suppressor, and a truncated variant designated p40 (ΔNp63) lacking the N-terminal Domain and acts as an oncogene (figure 1) (Dvorak et al., n.d.). The anti-p63 antibody (4A4) recognizes these two isoforms. In contrast, the anti-p40 antibody recognizes only the p40 isoform (ΔNp63), resulting in a sensitivity and specificity of 100% and 98% -100%, respectively, in detection of SCCs (Davidson et al., n.d.; Manuscript, 2014a). In summary, p40 is proportional to p63 in terms of sensitivity but is

markedly higher in specificity, which eliminates the possibility of an incorrect interpretation of a p63-positive ADC or lymphoma as squamous cell carcinoma. Furthermore, of this advantage p40 is a nuclear marker, which is less prone to potential nonspecific reactivity and interpretation complexity, as can occur with cytoplasmic markers such as desmocollin-3. These evidences strongly support the routine practice of p40 rather than p63 for the diagnosis of SCC (Bishop et al., 2011). Since keratinization is a hallmark of SCC, characteristic markers related to keratinization, such as cytokeratins, are also used. The most common cytokeratins used in the identification of SCC are CK5 and CK6, intermediate sized acid cytokeratins. In normal tissues, CK5 / 6 is expressed in particular in keratinized (epithelial) and non-keratinized squamous epithelium (mucosa) (Ma et al., 2015). The less commonly used markers for squamous differentiation include desmocollin-3 and desmoglein-3, adhesion protein of the cadherin family that is expressed in the stratified squamous epithelium. Desmocollin-3 is a very specific squamous marker but is less sensitive for squamous differentiation than p40 (*WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015). P40 offers a significant advantage over p63 in that its interpretation does not require CK5 / 6 as an additional squamous marker and does not depend on TTF-1 (Bishop et al., 2011).

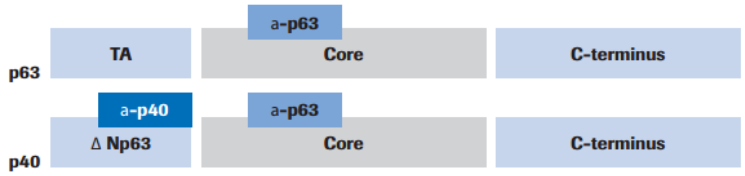


Figure 1: Scheme of p63 and p40 proteins (Dvorak et al., n.d.).

1.2.3 Large Cell Carcinoma

LCC is the third most common subtype of NSCLC, accounting for 3-9% of non-small cell lung carcinomas. LCCs are typically peripheral and are mostly found in smokers (*WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015). According to the 2004 World Health

Organization classification of lung tumors, LCC is defined as an "undifferentiated non-small cell carcinoma lacking immunohistochemical, architectural and cytological characteristics of small cell carcinoma, glandular or squamous differentiation". As LCC is a diagnosis of exclusion, it can only be performed in surgically resected tumor, since histological evaluation of the whole tumor is essential to exclude possible differentiation characteristics. Thus, a diagnosis of LCC cannot be made in small biopsies specimens or cytology. Because the lack of morphological differentiation in this type of sampling technique is usually a reflection of incomplete sampling or poor cell preservation, as opposed to a true absence of differentiating morphology throughout the tumor (Rekhtman et al., 2012). According to new recommendations by the International Association for the Study of Lung Cancer / American Thoracic Society / European Respiratory Society (IASLC / ATS / ERS), if there is an analysis in small biopsies or cytology specimens, LCCs are designated as Non- Small Cell Carcinoma-Not Otherwise Specified (NSCLC-NOS) (Zheng, 2016). The main characteristic that has been used to distinguish this type of tumor are its indeferenciated large with polygonal shape, vesicular nuclei, prominent nucleoli and with a necrotic appearance.

LCC has raised controversy over the years since this tumor may represent a truly distinct biological entity, but on the contrary, it may be an extreme in the poorly differentiated spectrum of the other main types of NSCLC: adeno- and/ or squamous. The latter hypothesis is supported by observation by electron microscopy, since LCC generally demonstrate structural features of adeno- or squamous carcinoma. In this sense, it is essential to perform auxiliary tests, such as IHC, to prevent the introduction of poorly differentiated NSCLC, such as solid adenocarcinoma and non-keratinizing SCC in this subtype of NSCLC. Based on IHC observation since LCCs express typical markers of adeno-or squamous carcinoma, clinical practice utilizes specific and sensitive differentiation markers of ADC and SCC as classifiers of morphologically undifferentiated NSCLC. These classifiers comprise TTF-1 and napsin A for ADC identification, and p40- or p63- and CK5 / 6 for the identification of SCC. The effective application of a panel of specific ADC and SCC markers in IHC to LCC is a plausible explanation for the decline of LCC in recent years, since some of the LCC cases were attributed to other categories (Manuscript, 2014b; Zheng, 2016)

1.2.4 Small Cell Lung Cancer

SCLC is a tumor with a fairly aggressive behavior that accounts for about 10 to 15% of all lung cancer. SCLC has a strong relationship with smoking, and more than 90% of patients with this type of cancer are or were smokers. (Rosell & Wannesson, 2012; Meerbeeck, Fennell, & Ruysscher, 2011) It is usually located centrally in the main airways, but can occur peripherally in the lungs in about 5% of cases (*WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015) These tumors in most cases have a very rapid growth and often spread rapidly to other parts of the body, including the brain, liver and bone (*WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015). It is in this last stage, that most patients are diagnosed, therefore, surgical resection is rarely an option. Since SLCL is considered to be a neuroendocrine malignancy, it is assumed that SCLC develops from neuroendocrine cells or neuroendocrine progenitors of the bronchial epithelium and as such may express neuroendocrine differentiation markers: chromogranin and neuron-specific enolase. In addition to these, they may also express small peptide hormones such as gastrin-releasing peptide, calcitonin and serotonin (Weynants, Humbler, Canon, & Symann, 1990).

Histologically, SCLC is characterized as a tumor with cells that have a small size, a round-to-fusiform shape, scant cytoplasm, finely granular nuclear chromatin and absent or inconspicuous nucleoli. In addition, the mitotic rate is high and there is often extensive necrosis (Travis, 2012). The 2004 WHO classification assumes the existence of SCLC subtypes: pure and combined. The combined small cell carcinoma is characterized as a pure SLCL with an additional constituent of any of the histological types of NSCLC: SCC, ADC or LCC (Davidson et al., n.d.; *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015). Because of its characteristic histological appearance, the diagnosis of SCLC can be reliably performed on small biopsy specimens and cytology. However, IHC may be necessary for confirmation of the neuroendocrine and epithelial nature of tumor cells (Davidson et al., n.d.). The panel of markers to detect neuroendocrine differentiation in SCLC are: Chromogranin A (glycoprotein ubiquitously present in neuroendocrine tissues); Synaptophysin (glycoprotein of pre-synaptic vesicles of neurons) and CD56

(neural cell adhesion molecule found in natural killer cells, natural killer-like T cells, myocytes, and seromucous glands) (Article, 2005; Hiroshima et al., 2006; Kasprzak & Zabel, 2007). CD56 is the most sensitive marker, but it is also the least specific marker. Besides these markers also a cocktail of cytokeratins is used to evaluate the epithelial nature of the tumor cells. The TTF-1 marker is also used in SCLC discrimination because it stains positively in 90-95% of cases, and napsin-A, a marker of ADC differentiation, is consistently non-reactive in SCLC (*WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015).

1.3 Genetic Mutations of Lung Cancer

The histological characteristics along with the expression of the IHC markers associated with the lung cancer subtype remain the basis of clinical diagnosis. However, recent advances in sequencing methods have allowed researchers to examine the breadth of genetic mutations in lung tumors in order to acquire new targets that aid in personalized therapy (Campbell et al., 2016). In addition, the detection of genetic alterations can predict response to treatment as well as establish a prognosis and, is imminently relevant in the early detection of lung cancer (Zheng, 2016).

One of the common mechanisms of carcinogenesis, namely in ADC, is the constitutive activation of receptor tyrosine kinases. EGFR is a tyrosine kinase receptor of the ERBB family. The EGFR gene is on the short arm of chromosome 7 at position 12.9. The binding of the extracellular ligand to the EGFR generates receptor homodimerization or heterodimerization, leading to cytoplasmic tyrosine kinase site phosphorylation and subsequent activation of various intracellular pathways: phosphatidylinositol 3-kinase (PI3K) / AKT / mammalian target of rapamycin (mTOR) and RAS / RAF / mitogen-activated protein (MAPK), which play a crucial role in cell proliferation, metastasis and prevention of apoptosis. EGFR gain of function mutations conduct to constitutive activation of downstream signaling pathways, which is critical for tumor growth (Mao, 2002; Villalobos & Wistuba, 2017). Mutations in the EGFR gene represent about 15% of adenocarcinomas, and the frequency of EGFR mutation in current or former smokers is approximately 10%, and never smokers can be up to 40-50% (*WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth*

edition - WHO - OMS -, 2015). Mutations in this receptor happen within exons 18 to 21, which encode for a portion of the EGFR kinase domain. About 90% of EGFR mutations occur in in-frame deletions in exon 19 or missense mutations in exon 21 (44% and 41% of all mutations, respectively) (Gridelli et al., 2015). Because of its frequency, this type of mutation was the first biomarker to benefit from NSCLC targeted therapy, tyrosine kinase inhibitors (TKIs). Gefitinib, erlotinib and afatinib are now recommended TKI for the first-line treatment of people with EGFR-sensitive mutations in lung cancer. Despite initial efficacy, cases of resistance to tyrosine kinase inhibitors have been reported, about 11-13 months after the onset of antitumoral therapy. In approximately 60% of cases, resistance to tyrosine kinase inhibitors (TKIs) is mediated by the presence of the secondary EGFR T790M mutation at exon 20 of the EGFR kinase domain (Popat, 2014). In the other 5-15% of cases, it usually involves amplification of the mesenchymal-epithelial transition (MET) and Human Epidermal growth factor Receptor 2 (ERBB2) gene mutations in Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform PIK3CA). In addition to these mechanisms associated with the development of resistance to TKIs, another was identified that involves the histological transformation of AD to SCLC. However, it is still controversial whether the SCLC component prior to TKI treatment is a consequence of treatment (Niederst et al., 2015; Shi, Duan, Liu, Zhou, & Liang, 2016). ALK is another tyrosine kinase receptor of the insulin receptor superfamily which is also mutated in lung cancer. The ALK gene is found on the short arm of chromosome 2 at position 23. The common form of ALK alteration, identified in 4-7% NSCLC, in particular in the ADC, is the formation of an oncogenic fusion gene with microtubule-associated echinoderm Protein-like 4 (EML4) genes. This rearrangement results in the formation of a chimeric protein which promotes malignant growth and proliferation. The ALK gene can also be rearranged with other less frequent partners such as kinesin family member 5B (KIF5B), TRK-fused gene (TFG), kinesin light chain 1 (KLC1), and huntingtin interacting protein 1 (HIP1) genes, leading to an oncogenic transformation. This type of mutation is common in non-smokers and in young patients (Mao, 2002). Identification of the ALK fusion led to the development of an ALK inhibitor: crizotinib. However, a portion of patients' patients develop resistance to this drug due to secondary mutations in the EML4-ALK kinase domain: L1196 M, C1156Y and F1174 L, among others. These results lead to the development of second generation ALK inhibitors such as alectinib, ceritinib and AP26133 and are currently under evaluation in clinical trials (Villalobos & Wistuba,

2017). Kirsten rat sarcoma oncogene (KRAS) protein stimulated pathways are downstream of the EGFR and stimulate cell growth. Mutations in the KRAS proto-oncogene are seen in 25% of the ADCs, 5% in SCC and absent in SCLCs (Z. Chen et al., 2014; Mao, 2002). Approximately 95% of the KRAS mutations are located at codon 12 and 13 of exon 2 and rarely at codon 59 and 61 of exon 3 (Campos-parra, Zuloaga, Vazquez, & Avile, 2015). In never-smokers, the most common KRAS mutations are G12D and G12V, whereas G12C is the most common mutation associated with smoking (Villalobos & Wistuba, 2017). The frequency of B-Ras proto-oncogene, serine / threonine kinase (BRAF) mutation in adenocarcinoma is 3% .BRAF is a proto-oncogene encoding a serine / threonine protein kinase which is an effector downstream protein of RAS that promotes proliferation and cell Survival BRAF mutations in NSCLC are not well characterized in the literature because of their low prevalence. The V600E mutation in BRAF, which is well characterized in melanomas, in NSCLC there are no concordant studies that associate this mutation type with smoking (Luk et al., 2015; Villalobos & Wistuba, 2017). MET amplifications are seen in 5% of NSCLC and 2% of adenocarcinomas and are commoner in male smokers (Villalobos & Wistuba, 2017). PIK3CA is a known central oncogene for the pathway phosphatidylinositide 3-kinase (PI3K) that is dysregulated in various types of cancer. PIK3CA was found altered in 1-2% of pulmonary adenocarcinomas (Mao, 2002). Fusion genes involving the ROS proto-oncogene 1, receptor tyrosine kinase (ROS1) have been found in 1-2% of NSCLC typically in no or mild smokers with adenocarcinoma (Gridelli et al., 2015). Other mutations and recurrent amplifications in many oncogenes identified in pulmonary ADCs include HER2 (also known as ERBB2) (1.7%), fibroblast growth factor receptor 1 (FGFR1) and FGFR2, neuregulin 1 (NRG1), neurotrophic tyrosine kinase receptor type 1 (NTRK1) and Ras-like protein expressed in many tissues (RIT1) (Z. Chen et al., 2014).

For lung SCC, the number of tumors for which all exome sequencing is available is lower than for ADC, which is why the ADC has a very well-characterized genetic profile (Z. Chen et al., 2014). Mutations in EGFR which present a fairly high frequency in ADC, in SCC are extremely rare. Some of the most commonly mutated genes in SCC are: TP53, ERBB4, discoidin domain-containing receptor 2 (DDR2), FGFR1, FGFR2, FGFR3 and genes in the PI3K pathway, among others. However, amplifications were also identified in SOX2, PIK3CA, BRF2, cyclin D1 protein

(CCND1), TP63, fibroblast growth factor receptor (FGFR1); and deletions in CDKN2A / B, PTEN and NF1 (Cancer & Atlas, 2012; Fukazawa, Guo, Ishida, & Yamatsuji, 2016; Weeden & Solomon, 2015).

LCC, in contrast to the other subtypes of lung cancer, does not present a specific pattern of changes in the number of copies or even mutations. However, it exhibits the same set of specific mutations and changes of other types of cancer, for example: amplification of ERBB2 and thyroid transcription factor-1 (NKX2-1) and mutations in KRAS and Serine/threonine kinase 11 (STK11) as in adenocarcinoma; amplification of L-myc-1 proto-oncogene protein (MYCL1) and retinoblastoma-associated protein (RB1) as in SCLC and amplification of CCND1, FGFR1, and SOX2 as in SCC. Thus, besides this type of cancer lacking immunohistological, architectural and cytological characteristics, it does not present a characteristic mutation panel either (Great, 2013; *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS* -, 2015)

Genetic characterization relative to small cell lung cancer is delayed compared to NSCLC cancers since this type of cancer exhibits such rapid growth and a tendency to metastasize at sites far from the lung that there is a shortage of available patient material for research purposes. The most notorious mutations found in SCLC are the loss of function of (tumor protein 53) TP53 and RB1 tumor suppressor genes. The frequency of presence of the TP53 mutation in SCLC is between 75-90%, proposing that the loss of this gene demarcates the onset of SCLC evolution (Niederst et al., 2015). A recent study has found somatic genomic rearrangements in (tumor protein 73)TP73 in a substantial fraction of cases. This genomic change consisted of a deletion of exons 2 and 3 of TP73, resulting in an oncogenic version of this gene, TP73Δ2/ 3. This finding indicates once again the p53 family's involvement in the development of SCLC, but also a new candidate with therapeutic applications (Great, 2013). The RB1 susceptibility gene encoding a protein that acts as the regulator of the transcription of genes involved in DNA synthesis and cell cycle control is the second tumor suppressor that is inactive in 90% of SCLC. The amplification of MYC familial genes, MYCL1, MYCN and MYC, has been identified in 30-20% of SCLC cases. In addition to activation of MYC signaling, activation of the oncogenic pathway of PI3K is also observed with high frequency in SCLC. It is also evidenced in this type of cancer inactivation of genes from

the tumor suppressor family NOTCH in 25% of cases of SCLC. Members of the NOTCH family are important regulators of neuroendocrine differentiation in SCLC, so their inactivation makes possible the expansion of tumor cell mass neuroendrin (Osada & Takahashi, 2002)

1.4 Exhaled Breath Condensate

Over the years, the number of minimally invasive procedures has grown exponentially in medicine, leading to greater success in the evaluation and treatment of diverse diseases (Grob, Aytekin, & Dweik, 2008). In this context, the exhaled breath condensate (EBC) is a new approach for investigating lung diseases because it contains a wide range of compounds, which may act as biomarkers (Liang, Yeligar, & Brown, 2012). Examples are markers of oxidative stress, angiogenesis and DNA, which can provide valuable information about physiologic and pathologic process in the lung. (Corradi, Gergelova, & Mutti, 2010). Many studies have analyzed the patterns of these biomarkers in exhaled breath and the composition of such biomarkers varies between healthy subjects and subjects with disease (Kubáň & Foret, 2013). Therefore, this strategy may open novel frontiers in the study of clinical diagnosis, mainly in lung cancer, in which the majority of patients are diagnosed at advanced stages (Lim & Thomas, 2013). This fact is mainly due to the lack of sensitive and specific biomarkers and tools for early screening and diagnosis but also because the 5-year survival rate remains low (Grob et al., 2008). EBC has been studied in a variety of diseases including allergic rhinitis, asthma, chronic obstructive lung disease, cystic fibrosis, and obstructive sleep apnea syndrome diseases (Grob et al., 2008).

The composition of exhaled breath is very complex, consisting of a gaseous phase and a liquid phase (Antczak & Gorski, 2002). The gaseous phase contains volatile compounds that are exhaled and absorbed by the condensing water, whereas the liquid phase, also denominated EBC, contains water vapor and aerosol particles or droplets where several nonvolatile compounds, such as DNA and proteins, have been identified (Conrad, Goyette, & Thomas, 2008; Kubáň & Foret, 2013).

The major source of the components in the EBC is believed to be airway lining fluid (ALF), mainly from the lower airway, although there is no sufficiently convincing evidence to provide clear proof of the origin of these particles (Antczak & Gorski, 2002). The ALF, which is a highly complex liquid that contain macromolecules and proteins which give it physical properties suitable for its function as a protective layer over the underlying cells. It comprises an aqueous sol phase, which lies proximal to the surface epithelium that contains cilia, and higher viscosity gel phase, which sits on top of the cilia. The principal macromolecular structural component of the gel phase is mucin glycoprotein, secreted from goblet cells of the surface epithelium and serous cells of the submucosal tissue. The cilia-driven upward movement in the low-viscosity sol phase produces a metachronal wave motion that moves the viscoelastic gel toward the exit of the respiratory tract, in direction to the pharynx (Chung & Widdicombe, 2008; *Smith and Williams' Introduction to the Principles of Drug Design and Action, Fourth Edition*, 2017). This mechanism, also named mucociliary escalator, constitutes the mucociliary clearance mechanism of the lung, whose function is to clear deposited matter, such as pathogens and irritants, trapped within the mucus barrier (*Toxicology of the Lung, Fourth Edition*, 2017).

In addition to the lack of knowledge about the precise anatomic and molecular source of the compounds present in EBC, the exact mechanism of EBC formation are also still a matter of dispute (Dodig & Cepelak, 2013).

Unlike traditional methods for sampling ALF of the lower respiratory tract, such as bronchoalveolar lavage (BAL) and spuctum induction (SI), EBC is a noninvasive method that can be repeated within short time intervals without causing airway inflammation and discomfort for the patients (Kubáň & Foret, 2013). EBC differs from these methods in that it does not contain respiratory cells but EBC biomarkers, which are believed to be directly influenced by these cells and are thus capable of reflecting changes in cellular composition and activity (Ahmadzai et al., 2013). EBC also represents an inexpensive way to rapidly screen for certain diseases, instead of CXR and CT, which also involve radiation exposure as a major drawback (Lim et al., 2013). Additionally, the content of EBC includes various biomarkers of respiratory function that are not present in serum or urine (M. Zhou, Liu, & Duan, 2012). Another advantage of this method relies on the fact that its collection devices are portable, and children or patients with respiratory impairments are able to successfully participate in

this technique (Grob et al., 2011). It is not yet possible to accurately identify individuals with lung cancer in risk populations even with this procedure, although exhaled breath is thought to contain material from the lungs and the lower part of the respiratory tract, reproducing a good surrogate for the molecular analysis of his disease (Kubáň & Foret, 2013).

1.4.1 Lung Cancer Driver Gene Mutations in EBC

Cancer occurs when genetic mutations build up in critical genes, allowing cells to grow and divide in uncontrolled ways (Osada & Takahashi, 2002). Molecular biological studies have demonstrated that lung cancer develops through sequential morphological steps with the accumulation of multiple genetic and epigenetic alterations leading to the inactivation of tumor suppressor genes and activation of dominant oncogenes (Osada & Takahashi, 2002). Many of the tumor suppressor genes and oncogenes altered in lung cancer are known to play a role in the regulation of cell cycle progression in either a direct or an indirect manner, and a considerable proportion of the lung cancer-related genes are a component of checkpoint mechanisms. (Baldi et al., 2011) Alteration in these genes confers special abilities to lung cells, including an escape from growth inhibitory signals, over shortening of telomeres, resistance to apoptosis, sustained stimuli for proliferation and angiogenesis, and metastatic characteristics (Massion & Carbone, 2003). These cellular alterations have been recognized in the carcinogenic process of lung cancer (Massion & Carbone, 2003). It is widely reported that cigarette smoking is a dominant risk factor for this type of cancer, and that carcinogens in the smoke are its main initiators by inducing multiple genetic alterations, mainly through the formation of DNA adducts (Dhar, Squire, Hande, Wellinger, & Pandita, 2000). Fingerprints of such genetic insults may be seen, for example, in the mutational spectra of p53 and KRAS (Osada & Takahashi, 2002). A summary of the genetic alterations detected in EBC from Lung Cancer patients is provided in Table 1.

1.4.1.1 EGFR Mutations in EBC

EGFR is a cell signaling trans-membrane protein intimately involved in proliferation and is found at abnormally high levels on many types of cancer cells such as NSCLCs (Normanno et al., 2006). Mutations associated with this receptor occur in the tyrosine kinase domain and lead to unregulated phosphorylation and activation of cell survival / proliferation pathways. There are several mutations associated with this domain: in-frame deletions in exon 19, L858R in exon 21, nucleotide substitutions in exon 18 (for example, G719C or G719S), in-frame insertions in exon 20. The mutation T790M in exon 20 is associated with acquired resistance to TKI therapy, gefitinib and erlotinib (S V Sharma, Bell, Settleman, & Haber, 2007; Sordella, Bell, Haber, & Settleman, 2004). In addition to these two inhibitors, monoclonal antibodies have also been used against extracellular EGFR, such as cetuximab and panitumumab (Greenberg & Lee, 2007). More recent studies indicate there may be other important parameters involved in cancer genesis, such as EGFR copy numbers and mRNA expression (Pao et al., 2004). Both EGFR mutations and EGFR copy numbers can be detected in biopsy specimens, and may aid the clinician to choose between therapeutic agents (Greenberg & Lee, 2007).

Paradiso et al. (2008) investigated the possibility of detecting EGFR mutations in EBC from patients with histological evidence of NSCLC, in order to contribute for noninvasive diagnosis of lung cancer (Paradiso, Tommasi, Pinto, Carpagnano, & Foschino-Barbaro, 2008). But the authors were unable to evidence the pathological deletion in exon 19 and polymorphisms in introns 19 and 20, and exon 20 of this receptor gene, using this technique, which is clearly evident in the tumor tissue. In addition, alterations in microsatellite instability and loss of heterozygosity in the five loci from chromosome 3p is confirmed in both EBC and tumor tissue. This suggests that the material present in EBC is not representative of the tumor and therefore, EBC is not suitable for investigating EGFR mutations (Paradiso et al., 2008).

Zhang et al (2011) on the other hand, were able to detect the EGFR mutation in exon 19 in EBC from a heavy smoker with SCC of the lung (Zhang et al., 2011). In addition, DNA from a trans-bronchial biopsy confirmed the presence of this mutation. This patient was treated with gefitinib and after 8 months, tumor regression is observed and no tumor progression symptom is detected, thus indicating that mapping of EGFR mutations in EBC can be used to follow the status of lung cancer (Zhang et al., 2011).

1.4.1.2 p53 Mutations in EBC

p53 tumor suppressor gene is regarded as a guardian of the genome, given its crucial role in maintaining genetic stability and prevention of cancer formation. p53 protein functions as a nuclear transcription factor that regulates expression of various genes encoding proteins involved in cell-cycle checkpoints, apoptosis, and DNA repair (Zilfou & Lowe, 2009). Alterations in p53 occur at an early stage of lung cancer. Generally, p53 mutations arise mainly from exons 5–8, being that p53 mutations in exon 8 correlated most strongly with survival, whereas mutations in exon 5 are associated with poor prognosis (Skaug et al., 2000).

Gessner et al. (2004) observe that p53 mutations are present in 4 out of the 11 EBC samples from NSCLC patients, while none of the healthy subjects showed mutations in p53 from EBC samples. One interesting finding is that the p53 mutations that were found in EBC from NSCLC patients were non matching with the DNA of the corresponding tumor tissue, suggesting that breath condensate represents a summed signal from the airways and lung parenchyma, while biopsies are only representative of the site where it is performed (Gessner et al., 2004).

The P53 gene has hotspot codons that have been associated to the presence of carcinogenic chemicals in tobacco smoke, such as benzo(α)pyrene. Because these codons contain CpG islands and the presence of 5-methyl cytosine greatly enhances binding benzo[a]-pyrene diol epoxide to N2 of guanine, this leads to the formation of trans adducts. In relation to the study by Gessner et al., (2004), people with cancer who had mutations were distinguishable from healthy volunteers, so p53 mutations in EBC samples might be used as a marker of direct tobacco-related DNA damage (Gessner et al., 2004).

1.4.1.3 KRAS Mutations in EBC

KRAS mutation remains the most common form of mutation in patients with NSCLC and confers a poor prognosis (Bhattacharya, Socinski, & Burns, 2015). The prognostic value of these mutations is still unclear, but they have been correlated with survival more frequently in resected stage I NSCLC patients. KRAS is one of the 3

proto-oncogenic members of the RAS family with intrinsic GTPase activity that regulate cell growth, differentiation, and apoptosis (D'Arcangelo & Cappuzzo, 2012). A single amino acid substitution at codon 12, 13, or less common in 61, affects the GTPase activity, resulting in the accumulation of an active GTP-bound form and the constitutive activation of downstream targets, including the RAF-MEK-ERK (MAPK) signaling pathway and AKT-PI3K-MTOR (Bhattacharya et al., 2015). This oncogenic Ras activation may result in a wide variety of biological responses, ranging from the activation of a senescence program to increased cell proliferation and inhibition of apoptosis, depending on the cellular and molecular context (Porta et al., 2009). Oncogenic RAS proteins can interfere with metabolism of tumour cells, microenvironment remodeling, evasion of immune response, and can contribute to the metastatic process (Yu et al., 2015).

Kordiak et al., 2012 shows a significant decrease in KRAS mutation in exhaled breath DNA of NSCLC patients that were previously analyzed for KRAS mutation at codon 12, after surgery (Kordiak, Szemraj, Hamara, Bialasiewicz, & Nowak, 2012). Some of the patients even show a complete depletion of KRAS mutation 30 days after surgery. This suggests that the resected lesion was the source of extracellular DNA bearing the KRAS mutation in EBC. However, the presence of KRAS in EBC after complete tumor resection could be the result of relatively long half-time of extra-cellular DNA bearing KRAS mutations in ALF, since the DNA released from pulmonary lesions can adhere to airways epithelium and is gradually released from its surface to ALF after tumor resection (Kordiak et al., 2012). In addition, this study showed that smoking cigarettes is not a necessary condition to induce mutations in the KRAS respiratory tract gene, since half of the individuals control, without any lung pathology, had KRAS mutations in EBC-DNA, because they were current smokers. However, 15% of the lung ADC of patients who never smoked had KRAS mutations, despite this pathology most frequently hosts KRAS mutations in patients who smoke. In this sense, further studies are needed to evaluate the predictability of KRAS mutation mapping from EBC-DNA. Moreover, more than 50% of screened patients with lung tumor are negative for KRAS mutation. This indicates rather low specificity and sensitivity of EBC KRAS as a single genetic marker of lung cancer. But it cannot be excluded that determination of EBC KRAS, in combination with other genetic markers, could be useful as a potential marker of the recurrence of pulmonary malignancy in follow-up after surgical treatment.

1.4.1.4 P16 Mutations in EBC

Chen et al., (2015) detected a mutation rate in exon 1 and 2 of the p16 gene in NSCLC patients, whereas no mutation was found in healthy individuals, indicating a high specificity of the detection of these mutation for diagnosis of patients with this type of lung cancer. Moreover, the mutation rate of p16 showed a direct association with tumor stage, revealing an increase with cancer progression and thus, the potential to be used to evaluate patient's condition (Jin‐Liang Chen et al., 2015).

1.4.2 Gene Promoter Methylation in EBC

DNA methylation is related with histone modifications and the interplay of these epigenetic modifications is crucial to regulate the functioning of the genome by changing chromatin architecture (Shikhar Sharma, Kelly, & Jones, 2009). DNA methylation occurs by covalent addition of a methyl group at the 5' carbon of the cytosine ring, resulting in 5-methylcytosine, process catalyzed by DNA methyltransferases (Balgkouranidou, Liloglou, & Lianidou, 2013). This covalent bond only occurs in cytokines that are directly attached to a guanine by a phosphodiester bond, forming a CpG dinucleotide pair. Regions with a high frequency of CpG sites form the CpG islands, which are typically found in or near promoter regions of genes, where transcription is initiated (Balgkouranidou et al., 2013). Methylation causes a change in DNA structure by altering the ability of transcription factors to interact with the promoter leading to inhibition of transcription (Kulis & Esteller, 2010). This process is commonly dysregulated in tumor cells and it is believed that epigenetic silencing is at least as common, if not most occurring, than mutational events in the development of cancer (Baylin, 2005; B. Jin, Li, & Robertson, 2011).

A wide variety of epigenetic biomarkers have been evaluated so far in lung cancer diagnostics, either as individual genes or as gene combinations (Hatzimichael, Lagos, Sim, & Briasoulis, 2014). p16 was the first tumor suppressor gene found inactivated in lung cancer predominantly through aberrant hypermethylation and its inactivation is likely due to the critical function of this gene in inhibiting cyclin dependent kinases that bind cyclin D1 and phosphorylate the retinoblastoma gene

(Liggett & Sidransky, 1998). Several researchers have tried to improve the diagnostic utility of epigenetic biomarkers in lung cancer by analyzing the methylation status of multiple genes and defining gene promoter methylation signatures as diagnostic tools, such as O-6-methylguanine-DNA methytransferase (MGMT), Ras association domain-containing protein 1 (RASSF1A), telomerase reverse transcriptase (TERT), Wilms tumor protein (WT1), death associated protein kinase (DAPK) and deleted in colorectal cancer (DCC) (Hatzimichael et al., 2014). Han et. al., 2009 demonstrates that the methylation density of *RASSF1A*, associated with the signaling pathway of Ras-dependent apoptosis, is statistically higher in former smokers relatively to never and current smokers (Han, Wang, Reilly, Keller, & Spivack, 2009). Additionally, this study reveals that methylation of CpG at -63 in DAPK promoter, a gene associated with apoptosis, and CpG at +52 PAX5β promoter, a gene related to the regulation of the cell cycle, is associated with long cancer status. Finally, the promoter DAPK presents a regional methylation pattern with two blocks, being in block two that differences in methylation status between lung cancer patients and healthy individuals were found (Han et al., 2009). In another study, Xiao et al. (2014) shows the presence of an aberrant promotor methylation of p16 with a positivity rate of 86.66% in tumor tissue, 50% in blood plasma and 40% in EBC, in opposition to healthy individuals, were no aberrant methylation was found (Xiao et al., 2014).

1.4.3 Microsatellite Alteration in EBC

Microsatellite alteration (MA), namely loss of heterozygosity (LOH) and microsatellite instability (MSI), has been reported as an early detection marker for lung cancer (Fleisher et al., 1999). Microsatellites are nucleotide sequence repeats of various lengths, which are scattered throughout the genome, between and within genes, exhibiting variations among individuals (Yong et al., 2000). Any change in length of microsatellite sequences, as a result of base deletion or insertion, is termed microsatellite instability. MSI is an indicator of deficient mismatch repair (MMR) system, which is a multi-protein complex responsible for correction of errors arising during DNA replication and cell division (Field et al., 1999). Loss of heterozygosity involves the mutation of one allele followed by the deletion of the second allele, reflecting an allelic imbalance arisen from chromosomal instabilities (Boland & Goel, 2010; Field et al., 1999). The comparative analysis of microsatellite loci in tumor and

normal counterpart tissue has become the most widely used method to determine such genetic alterations. Comparing polymorphic loci in normal and tumor DNA from the same patient, allows mapping of sequence deletions through the identification of sites of LOH (Merlo, Gabrielson, Askin, & Sidransky, 1994; Miozzo et al., 1996; Neville et al., 1996).

Previous studies show that the presence of MA in patients with NSCLC depends on the number of loci studied, their location and of the clinical and pathological characteristics of these patients (Yong et al., 2000). The high incidence of LOH and MSI on chromosomes has been described in lung cancer (J. Zhou et al., 2001), namely in the short arm of chromosome 3p, which contains several tumor suppressor genes (Ingvarsson, 2005). In this context, Carpagnano et al. (2005) reports that DNA found in EBC of 89% of NSCLC patients shows MSI or LOH while only 35 % of healthy individuals exhibit these alterations. Moreover, they also demonstrate that there is a proportional relation between the frequency of MA and the progression of the cancer, and also a direct relation between the number of MA and the consumption of tobacco in EBC of NSCLC patients (Carpagnano et al., 2005). Carpagnano et al. (2008) demonstrate that microsatellite alterations found in DNA from EBC and tumor tissue of each NSCLC patient have similar profiles of MI e LOH. This study also shows that the frequency of MA was significantly higher in heavy smokers and current smokers when compared with ex-smokers and non-smoker, what suggests that this marker might be related with carcinogen exposure. Therefore, the evaluation of microsatellite 3p can be relevant for the follow-up in NSCLC patients, screening of high-risk populations and the early diagnostic of lung cancer (Carpagnano et al., 2008).

Excision Repair Cross-Complementation group 1 (ERCC-1) and ERCC-2 are located at the long arm of chromosome 19 and MA in this chromosome has been recently associated with NSCLC pathogenesis and reduced survival. These genes encode enzymes critical for the mechanism of nucleotide excision and repair (Carpagnano, Palladino, Gramiccioni, Foschino Barbaro, & Martinelli, 2010; Simon, Sharma, Cantor, Smith, & Bepler, 2005). A reduced expression of both ERCC-1 and ERCC-2 has been detected in lung cancer and is associated with a deletion in chromosome 19 that gives rise to LOH (Bicher et al., 1997; Carpagnano et al., 2010). Cargagnano et al, (2010) demonstrates that 16% of exhaled MAs was found in healthy smokers and 25% of exhaled MAs in NSCLC patients, while healthy non-smokers do not show any MA. This results confirm the specificity of these genetic alterations for

lung cancer and a direct relation with tobacco consumption. MA at ERCC-1 showed to be predictive of survival in lung cancer patients, since patient survival decreased significantly in cases of microsatellite alteration present in this gene (Carpagnano et al., 2010).

1.4.4 MicroRNAs in EBC

MicroRNAs (miRNAs) are a class of small noncoding RNAs of 22 to 25 nucleotides in length which modulate the expression of proteins by binding to a short recognition sequence in the 3´ untranslated region (UTR) of its target mRNA (Cherni & Weiss, 2011). This binding can be complementary or partially complementary and thereby target the mRNA for degradation or translational inhibition (Felekkis, Touvana, Stefanou, & Deltas, 2010). This class of biomolecules has received particular interest in the field of cancer biology because of their involvement in multiple cellular processes, such as cell differentiation, proliferation, metabolism and apoptosis, some of which are commonly deregulated in malignancy. Several studies suggest that miRNA profiles allow distinction of tumor from normal lung tissue and that patterns of altered miRNA expression are consistently associated to clinical outcome, thus representing a useful diagnostic and prognostic tool (Cherni & Weiss, 2011). The precise molecular mechanisms associated with the altered expression of miRNAs in lung cancers are unclear (Yanaihara et al., 2006). However, abnormal expression of miRNAs could be caused by epigenetic change, such as DNA methylation and alterations of chromatin structure, which are important processes of transcriptional silencing and alteration in many genes (Shikhar Sharma et al., 2009). It has been shown that a number of miRNA genes are located near sites of translocation breakpoints or deletions in various cancers what can lead to overexpression or downregulation of miRNAs (Kumarswamy, Volkmann, & Thum, 2011). Thus, miRNAs can act as tumor suppressors or oncogenes leading to the development and progression of cancer (Lages et al., 2012).

A particular interest has arisen in order to detect the presence of miRNAs in EBC as lung cancer biomarkers, such as miRNA-21 and miRNA-486 (Shen et al., 2011). miRNA-21 expression has been found to be deregulated in lung cancer and a number of other solid tumors, and may be associated with apoptosis inhibition and acquisition of invasive properties, likely mediated by its downregulating effects on the

expression of target tumor suppressors phosphatase and tensin homolog (PTEN), Tropomyosin-1 (TPM1) and programmed cell death 4 (PDCD4) (Lages et al., 2012; Zhu, Si, Wu, & Mo, 2007). That way, miR-21 itself displays oncogenic activity and can be classified as an oncomir (Pezzolesi, Platzer, Waite, & Eng, 2008). miRNA-486 was found to be underexpressed in various forms of cancer, including lung cancer(Navon et al., 2009). Moreover, this miRNA is down-regulated in oral squamous cell carcinomas (OSCC) in comparison with normal tongue tissue and, in gastric cancer (Mozzoni et al., 2013). This suggests that miRNA-486 might be a potential tumor suppressor and that its inactivation is required for the expression of prooncogenic traits, which may involve the direct targeting and inhibition of the anti-apoptotic factor OLFM4 (Oh et al., 2011).

Two studies showed that these two biomolecules are present in EBC in non-cancerous pulmonary diseases (Pinkerton et al., 2013; Sinha et al., 2013). Following this study, Mozzoni et al., 2013 analyzed the levels of miRNA-21 and miRNA-486 in EBC and tumor tissue of NSCLC patient and found that miRNA-21 was expressed in higher levels in both tumor tissue and EBC, whereas miRNA-486 was significantly lower in these samples in comparison with the controls. This study also showed a correlation between the expression of EBC miRNA-21 and the expression of miRNA-21 in pathological tissue, suggesting that the genetic material in EBC is of cancerous origin. The difference found in miRNA-21 and miRNA-486 expression in NSCLC patients reveals a high diagnostic value and it could be used clinically as a screening test (Mozzoni et al., 2013).

Chen et al., 2016 showed that the mi-RNA-21 levels in both serum and EBC of NSCLC patients were higher than in healthy individuals, indicating a viability in using EBC to attest for altered miRNA-21 levels in lung cancer patients. Furthermore, EBC and serum miRNA-21 levels in stage III of the NSCLC group were higher than those in stages I and II, indicating that the progress of the cancer can also be monitored through miRNA-21 levels in EBC of NSCLC patients. Thus, tracking patterns of miRNA-21 expression in EBC has a high prognostic and diagnostic value, which as an additional significance as a non-invasive method (Jin-liang Chen et al., 2016).

1.4.5 Mitochondrial DNA Alterations in EBC

The human mitochondrial genome consists of 16,569 base pairs and is a supercoiled, double-stranded molecule that contains 37 genes coding for 13

polypeptides of the mitochondrial electron respiratory chain, 22 transfer ribonucleic acid (tRNAs) and 2 ribosomal ribonucleic acid (rRNAs) necessary for the synthesis of polypeptides (Penta, Johnson, Wachsman, & Copeland, 2001). Compared to nuclear DNA, mitochondrial DNA (mtDNA) is more susceptible to oxidative damage due to its close proximity to the electron transport system, where reactive oxygen species (ROS) are continuously generated, but also due to inefficient repair mechanisms and absence of protective histones (Ott, Gogvadze, Orrenius, & Zhivotovsky, 2007). In addition, these factors also contribute to the advent of mtDNA mutations (H. Sharma, Singh, Sharma, Jain, & Singh, 2005), which have been postulated to play a role in cancer development through several pathways such as mitochondrial respiratory chain dysfunction with increases levels of ROS, abrogation of apoptosis or mitochondrial respiratory chain dysfunction (Yang Ai et al., 2013). The D-loop, a non-coding region, acts as a promoter for both the heavy and light strands of the mtDNA, and contains essential transcription and replication elements (H. Sharma et al., 2005; Taanman, 1999). Several studies in lung cancer have demonstrated prevalent mutations in this mtDNA region in lung tissue, sputum and BAL, and it has been alleged that their number may correlate with prognostic factors such as tumor stage and nodal metastasis (Jakupciak et al., 2005; X. Jin et al., 2007).

Yang et al. 2013 was the first group to demonstrate the feasibility of detecting mtDNA changes in EBC from patients with lung cancer and healthy controls (including smokers and non-smokers) (Yang Ai et al., 2013). The authors reported a higher prevalence of mitochondrial D-loop mutations in EBC of lung cancer patients compared to the control group (7 vs 3.5 for smokers/ex-smokers, and 7 vs. 4 for non-smokers). Albeit the number of mutations in heavy smokers was higher than in light smokers, there was no significant difference with respect to the mtDNA mutation frequency for number of cigarettes smoked per day, pack-years and age at smoking initiation (Yang Ai et al., 2013). Furthermore, the group found a mutation (T16217C), in D-loop of DNA mitochondrial, that was specific for the lung cancer group. The D-loop region of mitochondrial DNA contains the major control elements responsible for the transcription and replication of mtDNA, and the abundance of specific mutations may be associated with the clonal expansion of neoplastic cells. Thus, this mutation is worthy of future studies as it may be an excellent biomarker of lung cancer (Yang Ai et al., 2013).

Table1: Genetic Alterations in EBC Collected from Lung Cancer Patients

Genetic Markers in EBC	Reference	Number of individuals		Collection device	Duration of Collection	Detection method
		Cancer patients	Healthy controls			
Driver Gene Mutation						
EGFR	Paradiso et al.2008	23 NSCLC	NR	NR	NR	NR
	Zhang et al. 2011	1 squamous cell carcinoma	-	Rtube	15 Min.	PCR
P53	Gessner et al. 2004	18 NSCLC	18 Healthy non-smokers	EcoScreen	20 Min.	Nested PCR, sequencing
KRAS mutations	Kordiak et al.2012	57 (46 NSCLC and 11 Benign lesions)	52 Healthy controls	EcoScreen	25 Min.	Mutant-enriched PCR
p16 mutations	Chen et al.2015	58 NSCLC	30 Healthy controls	EcoScreen	20 Min.	PCR, sequencing
Gene Promoter Methylation						
DAPK, RASSF1A, PAX5β	Han et al.2009	17 (NSCLC and SCLC)	37 (Smokers and Non-smokers)	RTube™ (Charlottesville,VA)	10-15 Min.	Multiplex PCR, (tBGS), (qMSP)

p16 methylation	Xiao et al. 2014	30 NSCLC	30 Healthy controls	F-qMSP	15-20 Min.	F-qMSP
Microsatellite Alteration						
3p MAs	Carpagnano et al.2005	30 NSCLC	20 Healthy controls	EcoScreen	20 Min.	PCR
3p microsatellite signature	Carpagnano et al.2008	41 NSCLC	14 Controls	EcoScreen	20 Min.	PCR
MA at chromosome 19	Carpagnano et al. 2010	34 NSCLC	33 (14 Smokers and 19 non-smokers)	EcoScreen	20 Min.	Fluorescent PCR
MicroRNAs						
miRNA-21 and miRNA-486	Mozzoni et al,2013	54 NSCLC 30 NSCLC	46 Non-cancerous patients as controls	TURBO-DECCS system	15 Min.	qPCR
miRNA-21	Chen et al. 2016		30 Healthy controls	EcoScreen	20 Min.	qPCR
Mitocondrial DNA						
mtDNA mutations	Yang Ai et al.2013	9 NSCLC	16 (10 Smokers and 6 Non-smokers)	Custom-made glass condenser	20 Min.	PCR and capillary sequencing

NSCLC, non-small cell lung carcinoma; SCLC, small cell lung carcinoma; NR, not reported; MAs, microsatellite alterations; F-qMSP, fluorescent quantitative methylation-specific polymerase chain reaction; tBGS, tag-adapted bisulfite genomic DNA sequencing; Min, minutes; mtDNA, Mitochondrial DNA; miRNA, micro RNA;

1.5 EBC Collection Devices

During EBC collection, the warm air exhaled from the lungs reaches a temperature below its dew point by transfer of heat to a chilled condenser surface. This temperature difference causes the condensation of the aerosolized particles to form droplets on the condenser wall (Carter, Davis, & Kovacs, 2012). Depending on the cooling temperature used, the condensate will be liquid, solid or a mixture of both. Importantly, temperature at which EBC is accomplished has been shown to impact collection volumes as well as the relative concentration of biomarkers, depending on their characteristics. Moreover, the volume collected also depends on total expired air volume, condenser material and turbulence characteristics. EBC is collected during tidal breathing, as a product of cooling and condensation of the exhaled aerosol (Konstantinidi, Lappas, Tzortzi, & Behrakis, 2015).

Nowadays, several research groups use different condenser equipment, all based on the same principle of cooling exhaled air (Horvath & Jongste, 2010). Various noncommercial designs have been described, based on two simple layouts: a double wall borosilicate glass condenser system cooled by circulating icy water using a pump: in this device, the EBC collected on the walls of the cooled glass tube falls into the container below by gravity (a), and also by immersion of a Teflon-lined tubing in an ice-filled bucket (b) (figure 2). In both devices, the exhaled breath is directed through the cooling tube and exits freely to the environment, yielding nearly 50% of the collection efficiency (Konstantinidi et al., 2015; Kubáň & Foret, 2013). The condensers currently marketed for the EBC collecting are the EcoScreen (Erich Jaeger GmbH, Hoechberg, Germany), the RTube breath condensate collector (Respiratory Research Inc., Charlottesville, VA, USA), the ANACON condensor (Biostec, Valencia, Spain) and the TURBO-DECCS (Transportable Unit for Research on Biomarkers Obtained from Disposable Exhaled Condensate Collection Systems: ItalChill, Pharma Italy) (Carter et al., 2012; Hoffmeyer, Raulf-Heimsoth, Harth, Bünger, & Brüning, 2009). Although all devices have the same basic principle, they differ widely in their designs, cooling system and coolant materials. (Youssef et al., 2016) Furthermore, the design of each device and its physical characteristics can significantly affect the level of the biomarker of interest and explain the disparity in the results obtained with different devices (Hoffmeyer et al., 2009). pH values of the collected have been shown to differ from different condenser devices and the surface of the devices also influence the

adhesion capacity of the biomarkers (Carter et al., 2012; Czebe et al., 2008). So an appropriate choice of the collecting device relies on the kind of the biomarker to be examined (Youssef et al., 2016).

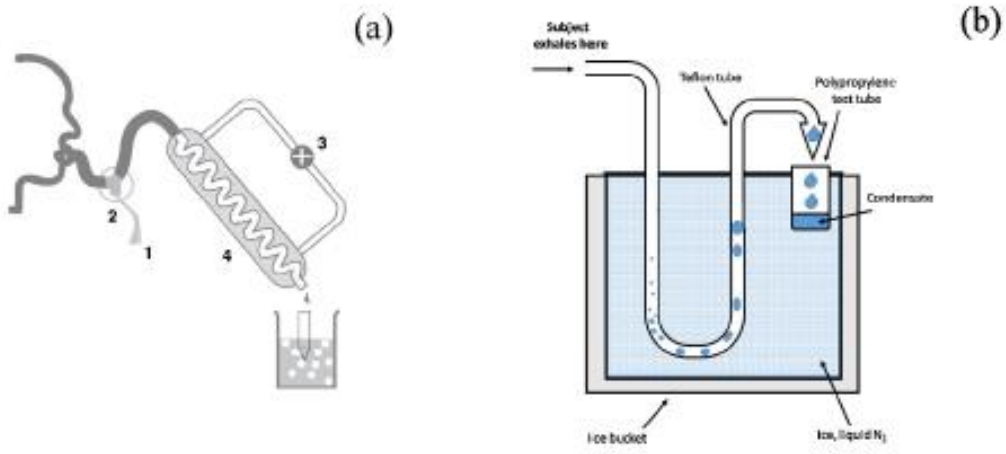


Figure 2: Schematic representation of non-marketed EBC collection devices. (a) The subject breaths tidally and inhales room air (1) through a two-way non-rebreathing valve (2). Exhaled air is guided by the tubing through a double-wall glass condenser (4), which is cooled by a counter-current circulating ice-water pump (3). The two-way non-rebreathing valve (2) and tubing also serve as a saliva trap. (b) Device that uses a Teflon tube inserted in a bucket filled with ice (Kubáň & Foret, 2013).

The EcoScreen electrically cools the exhaled air conducted through a polytetrafluoroethylene (PTFE) coated aluminum double lumen lamellar system. The cooled air is collected in a disposable polypropylene collecting cup inserted into the cooling cuff (electric refrigerator) (Dodig & Cepelak, 2013; Konstantinidi et al., 2015). The condenser maintains a temperature of -10ºC during a 10-15 min collection period and the refrigeration through the electric system is turned on at least 40 minutes before sampling (Czebe et al., 2008; Davidsson & Schmekel, 2010). This device uses a saliva trap, near the mouthpiece, that keeps the collected fluid amylase-free (figure 3) (Janicka, Kot-Wasik, Kot, & Namieśnik, 2010). The Ecoscreen I is not currently marketed owing to several technical malfunctions, including absence of manual control of condensing temperature and cleaning requirements of the device between consequent assays (Konstantinidi et al., 2015). Moreover, it was reported that nitric oxide (NO) measurements might be confounded by the device and partly reflect contamination with NO coming from the device itself (Janika et al., 2010). Thus, the EcoScreen I has been substituted by Ecscreen II, a temperature-controlled device (Hoffmeyer et al., 2009).

The EcoScreen II allows the fractionated collection of EBC into two disposable polyethylene bags, inserted into a refrigerating system, through a valve system (Konstantinidi et al., 2015). The device also includes a built in spirometer, allowing the measurement of exhaled volumes and time to be controlled in relation with a standardized collection (figure 3) (Carter et al., 2012). According to its construction, saliva contamination is highly unlikely and was excluded in testing of amylase activity. Opposed to its precedent, EcoScreen II is a single-use disposable condensing and collection system (Hoffmeyer et al., 2009). Due to its characteristics, this device might be useful for studying the origin of biomarkers in the respiratory system and detect differences between healthy and diseased subjects, but no published studies are available for this condenser (*Studies on Experimental Models | Samar Basu | Springer*, 2017). Both Ecoscreen devices are non-portable and weigh approximately 20 kg (Konstantinidi et al., 2015). A comparative study demonstrated that EcoScreen II collected larger sample volumes and greater concentrations of biomarker for analysis compared to EcoScreen I (Carter et al., 2012). Another study has showed that the ECoScreen II plastic collecting surface is more suitable for collecting lipid-derived compounds than the Teflon coated metal surface of EcoScreen I (Hoffmeyer et al., 2009). Both devices can be adapted for mechanically ventilated patients, although for the EcoScreen II this characteristic is less well defined in the literature (Carter et al., 2012).

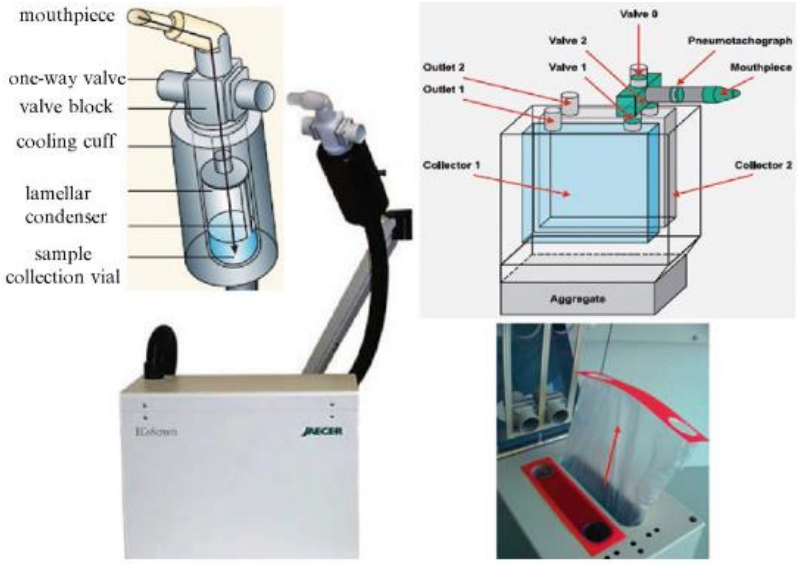


Figure 3: Schematic representation of the EcoScreen I and EsoScreen II. (Elsevier, 2017)

The RTube™ breath condensate collection device, also called Portable Collection Device, was designed for easy patient use without supervision at home, workplace, laboratory, hospital or clinic (RTubeTM -Respiratory Research). RTube disposable collection system employs a large Tee section, made of polypropylene, which separates saliva from the exhaled breath through gravity, thus preventing salivary contaminations. Furthermore, this device uses an exhalation valve, made of silicone rubber, which also serves as a syringe-style plunger to pull fluid off the condenser walls, allowing a reduction in collection time (Konstantinidi et al., 2015; RTube™ - Respiratory Research). Cooling is reached by placing an aluminum cooling sleeve over the disposable polypropylene condensation chamber, which can also work as the sample collector container (figure 4). The temperature of the cooling sleeve can be defined by the investigator (Grob et al., 2008). One of the possible disadvantages is that the condensation temperature may alter when collecting large volumes of EBC over a long time (Kubáň & Foret, 2013).

Figure 4: RTube device for exhaled breath condensate (EBC) collection. RTube disposable collection system consists of a unidirectional valve crafted from silicone rubber, a large polypropylene Tee section separating the saliva from the exhaled air; and a polypropylene collecting tube that is cooled by an aluminum sleeve placed around it. (RTube™ -Respiratory Research)

TURBO-DECCS (Transportable Unit for Research on Biomarkers Obtained from Disposable Exhaled Condensate Collection Systems) is a portable condenser consisting of a refrigerating system (TURBO) that thermostatically controls the temperature, and a disposable polyethylene DECCS collection system (Caglieri et al., 2006; Goldoni et al., 2005). It is composed of a mouthpiece equipped with a one-way

aspiration valve and saliva trap, connected to a collecting vial inserted in a Peltier-type electrical cooling system. This connection is made through a tube with a special stopper (figure 5) (Caglieri et al., 2006). The DECCS ensures the absence of chemical or bacterial contamination of the collection. Similarly to the EcoScreen devices, it provides an adjustable condensation temperature through electronic control but it differs from the latter by giving rise to the Peltier effect (Kubáň & Foret, 2013). The collecting temperature is regulable from -10ºC to room temperature (Horvath & Jongste, 2010).

Figure 5: TurboDECCS device for exhaled breath condensate (EBC) collection is comprised of a portable Turbo Unit and a disposable DECCS collection system. DECCS is composed by a mouthpiece, a one-way aspiration valve, a tube, and a collection cell inserted in a Peltier-type electrical cooling system (Leese, Morton, Gardiner, & Carolan, 2016).

The condenser ANACON has been used by several research groups, mainly in mechanically ventilated patients (Konstantinidi et al., 2015). This device is attached to the expiratory branch of the ventilator circuit via adaptors that are part of the device. It is also composed of a thermoelectric pump to cool a glass-surface condenser, which can be adjusted by the investigator (figure 6) (Carter et al., 2012; Konstantinidi et al., 2015). The condensation temperature can be constantly controlled within the limits of −15°C to −5°C (Czebe et al., 2008; Romero et al., 2006).

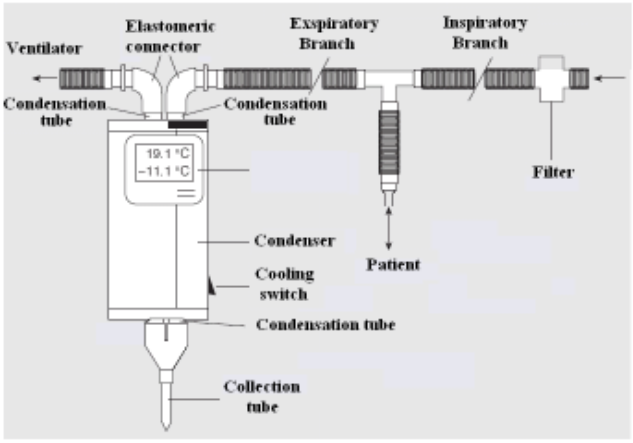


Figure 6: "ANACON condenser integrated in the mechanical ventilation circuit. The condenser is inserted in the expiratory branch of the ventilation circuit via 2 adaptors and 2 elastomeric connectors. The exhaled air passes towards the condensation tubes that pass through the body of the condenser. A Y piece closes the circuit with the collection tube for the exhaled breath condensate. A thermometer allows the condensation temperature to be monitored. The apparatus also contains a cooling switch" (Romero et al., 2006).

1.6 Oncosopro

The startup "Infogene" has developed a portable device, which was named Oncosopro, which allows the collection, storage, transport and subsequent analytical determination of biological material carried by exhaled breath, including biomarkers of lung cancer. This device comprises a collector recipient with a receptacle inside, which together are operatively connected to a cover, through a closure-aperture system, integrating a screw system. The receptacle has a conical shape with an orifice at its lower end, facing the cover, resulting in a more efficient recovery of the exhaled biological material (figure 7). The cover is filled with a filter composed of a solid matrix which is purchased to Flinders Technology Associates (FTA^{TM}),that captures, stores and preserves the exhaled nucleic acids (M. Da Costa & Pt, 2013; M. D. A. Costa & Pedro, 2014). The filter is an absorbent cellulose-based paper with an embedded chemical formula containing: (Carpi, Pietro, Vincenzetti, & Mignini, 2011),.(Application & Data, 1996)

(1) A monovalent weak base: Creates an alkaline environment with pH's between 8.0-9.5 to ensure suitable functionality of the chelating agent, a

constituent of the paper, and avoid the action of acid nucleases, which activity are not totally dependent on divalent metals. The monovalent weak base is mainly constituted of organic material, such as TRIS (tris-hyroxymethyl methane), but it can also be made of inorganic material, such as sodium bicarbonate.

(2) **A chelating agent:** Binds to divalent metal ions, such as calcium or magnesium, which act as co-factors of enzymes that promote DNA degradation. In addition, it may also bind to transition metal ions, such as iron, which easily undergoes oxidation-reduction reactions producing free radicals that damage DNA. Within a range of strong chelating agents available Ethylene Diamine Tetracetic Acid (EDTA) is the most used in this card.

(3) **An anionic detergent:** The most commonly used anionic detergent is Sodium Dodecyl Sulphate (SDS) and acts as a denaturation agent. In the presence of pathogens, the anionic detergent promotes unspecific destruction of the secondary structure of its coating proteins and internal proteins. However, the anionic detergent is ineffective against more resistant bacterial spores.

(4) **Facultative, uric acid or a urate salt:** It's vital for the long-term storage because it helps maintain an alkaline pH, since it acts as a component of a buffer system. In addition, it acts as a free-radical trap that prevents the damage of the DNA guanine base.

Summing up, the presence of all these components culminates in the protein denaturation, including nucleases; in the lysis of the cell membranes, leading to the release of the nucleic acids that are physically entrapped in the cellulose fibers of the matrix; immobilization of DNA and preservation for long-term storage at room temperature in the filter (Jignal, Mg, & Darshan, 2014; Petras, Lefferts, Ward, Suriawinata, & Tsongalis, 2011). Other forms of protection and stability offered by this filter are inhibition of fungi and other microorganism's growth, inactivation of oxidation processes, as well as, possible UV interference. (Rajendram, 2007).

Figure 7: Schematic representation of the OncoSopro device.

During collection of exhaled breath, the user deeply inhales and then exhales as much air as he can into the device. After expiration, the user withdraws the device from the mouth and recovers the breath naturally. This procedure is repeated for 10 min and results in air condensation on the walls of the collector recipient. During the exhalation, the exhaled air leaves the device through two perforations that are in the upper portion of the collector recipient, near the place where the subject inserts the mouth. If the user notices the accumulation of saliva in the mouth, swallows it and prevents it from going into the device.

CHAPTER II

2. OBJECTIVES

Exhaled breath condensate is an innovative approach to creating screening methods for the diagnosis of lung cancer because contains a wide range of compounds that can function as biomarkers. These biomarkers that can go from: miRNA, DNA, mtDNA, etc and provide precious knowledge about the physiological and pathological conditions of the lung, since studies have shown that the pattern of these biomarkers differs between healthy individuals and sick individuals (Corradi et al., 2010; Kubáň & Foret, 2013). In addition, screening methods based on this type of approach are convenient, totally non-invasive and accessible (Lim & Thomas, 2013). The Oncosopro device is based on this methodology and permits the collection and storage of biological material containing genetic biomarkers that help us in the early detection of lung cancer. Moreover, it also may include biomarkers that can predict response to biological therapies by assisting in the selection of the drug of choice for the first line of treatment for lung cancer treatment. The purpose of this thesis is optimize the laboratory protocol of recovery of biological material of healthy volunteers from the device's filters ensuring the highest recovery efficiency and the largest possible amount of biological sample available for subsequent laboratory analytical methods. The efficiency of the several protocols was evaluated based on the quality of the DNA obtained from the analysis of 2 parameters:

(I) Amount and purity of DNA obtained by absorbance analysis at 260, 280 and 230 nm by Nanodrop 2000 Sprectrophotometer and fluorometric method: Qubit

(II) Integrity and quality of DNA through PCR and agarose gel electrophoresis

In order to ascertain the efficacy and reproducibility of the optimized extraction protocol, a population control-based study was carried out by healthy individuals, smokers and ex-smokers, over 50 years old.

CHAPTER III

3.METHODS

3.1 Extraction of Genomic DNA from Filters Using Heat and pH Treatment

The genomic DNA, which is strongly adsorbed to the cellulose fibers of the filter, was eluted through a heat pH treatment that involves two steps. The first step consists of a purification of the DNA captured in the filter, which requires the use of two reagents: filter Purification Reagent and TE Buffer(10mM Tris-HCL, 0.1mM EDTA, pH 8.0). Filter Purification Reagent exhibits chemical properties that removes PCR inhibitors and potential contaminants, ensuring superior DNA quality for subsequent analytical methods. The TE Buffer removes excess Reagent Purification on the filter. The second step consists of an elution through a pH treatment, which employs two reagents: an alkaline elution buffer (0.1M NaOH, 0.3mM EDTA, pH 13.0), that sets up a highly alkaline environment, desnaturating and releasing DNA from the solid matrix into the elution buffer; and a neutralization buffer (0.1 M Tris-HCl, pH 7.0) to re-stabilize the DNA for later storage in solution. This DNA extraction protocol can be performed in a single microtube, greatly facilitating the process and reducing the risk of possible cross-contaminations.

In pratice, filters (10 mm diameter) of dry exhaled breath samples were cut in half with scissors and tweezers and were then transferred to individual 1.5 mL microcentrifuge tubes. The scissors and tweezers were cleaned between each cutting process by rinsing the tips with 70% ethanol to mimimize cross contamination. To wash the filter, 1000µL Purification Reagent was added, followed by a flash vortex 10 times and an incubation for 5 min at room temperature. The added Purification Reagent was then removed using a pipette and discarded. This process was repeated for a total of three washes with this reagent, and in the two following incubations at room temperature the microtube was slowly vortexing during 5 min. Subsequently, 1000µL TE buffer was added followed by a slow voxtex incubation for 5 min at room temperature. The TE buffer added was discarded and the procedure was repeated three times, with the remaining incubations resting during the 5 min. To elute the DNA samples, 70µL of the alkaline elution buffer was added to each tube, followed by

maceration of the filter and incubation at 65 ° C on a heating block for 15 minutes. Posteriorly, 130 µL of neutralization buffer was added, followed by vortex flash 5 times to mix the solution, and a incubation during 10 minutes at room tempertaure. The microtube was again flash vortexed 10 times. The filter was then removed and squeezed to recover the maximum volume of dna-containing eluate which, subsquently, was subjected to a centrifugation at 10000 rpm for 2 min to deposit possibly present cellulose fibers. Lastly, the supernatant, TE buffer, cointaning DNA is transferred to a 1mL microtube and stored at -20 ° C or used directly for subsequent analysis.

3.2 Measurement of DNA concentration

The concentration of the DNA extracted by the different protocols was assessed using NanoDrop 2000 Sprectrophotometer (Thermo Fisher Scientific) and Qubit 2.0 Fluorometer.

3.2.1 NanoDrop 2000 Sprectrophotometer

Nanodrop 2000 Sprectrophotometer calculates the DNA concentration by measuring the absorbance of the DNA sample at 260 nm, zone where the heterocyclic rings of the nucleotides absorb ultraviolet (UV) light. The purity of the DNA is evaluated by the ratios A260/280 and A260/230. The 260/280 ratio having a value of approximately 1.8 is generally accepted as "pure" for the DNA, while a ratio of about 2.0 is defined as “ pure” RNA. If the proportion is appreciably lower, it may indicate the presence of proteins, phenol or other contaminants that intensely absorb at or near 280nm. The A260 / 230 ratio is a secondary measure of nucleic acid purity. Values relative to this ratio for pure samples are often greater than the respective A260 / 280 ratio values, usually in the range of 2.0-2.2. If the A260 / 230 ratio is below, this range it is indicative of the presence of contaminants such as carbohydrates, EDTA or other impurities that absorb at 230nm.

Prior to measurement, the NanoDrop2000 was calibrated using 2 µL “blank” sample: TE buffer. 2µl of each DNA sample were pipetted onto the pedestal of

platform, the absorption was measured at 260nm, 280nm, respectively, and the concentration was showed in ng/µl. The ratio 260/ 280 nm indicated the purity of the tested sample

3.2.2 Qubit® 2.0 Fluorometer

The Qubit® 2.0 Fluorometer is a benchtop fluorometer that allows the quantification of DNA, RNA and protein, based on fluorescence. This system employs a specific fluorescent dye which has low fluorescence until it binds to the target DNA. After binding with the target DNA, by intercalation between the bases, the dye has a more rigid form and emits a greater fluorescence.

Qubit Fluorometer (invitrogen) calculates the concentration of DNA from the fluorescence signal emitted by the dye, which is in a certain quatity. The fluorometer picks up this fluorescence signal and converts it to a measure of DNA concentration using standards of known concentration. These patterns allow to establish a relationship between DNA concentration and fluorescence This type of quantification, based on the detection of specific fluorescence of the target, is more sensitive and precise than the quantification based on the UV absorbance

DNA was quantified by Qubit 2.0 Fluorometer using a dsDNA HS Assay Kit. The Qubit working solution was prepared including HS reagent diluted 1:200 in HS buffer. The two standards and the samples were mixed 1:20 and 1:100 to a final volume of 200µl in working solution, respectively, and subsequently incubated for 2 min at room temperature. The fuorescence of the dye was subsequently measured and the DNA concentration was calculated from the following formula:

Concentration of the sample in ng/ml = QF value x $\frac{200}{x}$

Where QF value = the value given by the Qubit® 2.0 Fluorometer (ng/mL)

x value = the number of µL of sample added to the assay tube

3.3 Polymerase Chain Reaction

The polymerase chain reaction (PCR) is a powerful molecular biology technique that permits the exponential enzymatic amplification of specific fragment of double-stranded DNA. The enzymatic amplification involved in the PCR reaction is performed by a heat- stable DNA polymerase. This polymerase assembles a new DNA strand using four deoxyribonucleotide triphosphates (dNTPs), wherein the primers serve as starting points. The term 'chain reaction' which makes use of the PCR technique refers to various repeated heating and cooling cycles which implicated three principal steps as for one PCR cycle: desnaturation, annealing, extension. Through the successive cycles of heating and cooling, double-stranded DNA is separated and replicated, producing copies of a specific DNA fragment.

The quality / integrity of genomic DNA eluted was evaluated through PCR amplification with a set of primers for the human Beta-Globin gene. PCR was carried out in 25 µl total reaction volumes, each containing 10/20 µl template DNA, 0.4 µl of each primer X3, 1x reaction buffer (nzytech), 0.2 mM dNTPs, 75 mM MgCl2 (nzytech), and 1.5U DreamTaq DNA Polymerase. Amplification procedure was carried out in PCR termocycler (Bio-Rad) about 2: 30 hours. The reaction mixture was heated to 95ºC for 5 min, followed by 15 cycles, each consisting of desnaturation at 95ºC for 30 seconds, annealing at 53ºC for 30 seconds, extension at 72ºC for 1 min, followed by a last cycle consisting of a desnaturation at 95ºC for 30 seconds, annealing at 48ºC for 30 seconds and extension at 72ºC for 1 min

3.4 Electrophoresis

The integrity of DNA extracted by each protocol was assessed by agarose gel electrophoresis.

Agarose gel electrophoresis is an analytical method that allows the separation of nucleic acids based on their size, electrical charge and structure, as well as, the subsequent identification and purification of the separated DNA fragments. The method consists on the migration of electrically charged DNA molecules under the influence of

an electric field in an inert agarose matrix. The electrophoretic mobility of DNA fragments principally depends of the size DNA fragment,the concentration of agarose used as well as applied voltage and electrophoresis buffer utilized .

The PCR products were separated on a 2% agarose eletrophoretic gel, in which 3g of agarose were dissolved in 150 ml of TAE by heating. Thereafter was added 3 µL of Green Safe Premium (nzytech). After solidification at room temperature, the gel was immersed with 1x TAE-buffer. Samples were prepared for electrophoresis by mixing 9µL of the PCR product with 3µL loading buffer. FastRuler Low Range DNA ladder (Thermo Scientific) was used as a molecular weight standard, wherein 9µL of this Ladder was mixed with 3µL Loading buffer. The gel was subjected to 140V current and the separation of the DNA fragments was visualized by a UV transilluminator for during periods of 20 minutes to detect the DNA bands. The photographs of the gel were taken for documentation.

CHAPTER IV

4. RESULTS

4.1 Optimization of the biological material collection protocol

The exhaled breath condensate contains a myriad of compounds that can function as biomarkers (Corradi et al., 2010). Several studies have evaluated this approach in several diseases, such as allergic rhinitis, asthma, chronic obstructive pulmonary disease, and have found that its composition varies between healthy individuals and individuals with disease (Kubáň & Foret, 2013). The results obtained from diverse studies in the field of exhaled breath condensate have triggered the production of a blow device whose main objective is to detect genetic markers that aid in the early detection of lung cancer. This thesis project focuses on optimizing a protocol for extracting biological material captured by a filter, so that in the future, genetic biomarkers associated with lung cancer can be detected.

The extraction protocol involves a washing process followed by a pH treatment. The optimization of the extraction protocol had as a starting point the conventional protocol from which several modifications were made changing the number of washes with Purification Reagent and TE buffer, the non-maceration of the filter at the time of the pH treatment and finally, the elution volume of the DNA (table 2).

Table 2: Protocols tested with the respective procedures

Protocols	Procedure
Conventional Protocol	3x washes with Purification Reagent 3x washes with TE Buffer Elution with Solution 1 and solution 2 –with maceration
Protocol 1	**1x** washes with Purification Reagent 3x washes with TE Buffer Elution with Solution 1 and solution 2 –with maceration
Protocol 2	1x wash with Purification Reagent **1x** wash with TE Buffer Elution with Solution 1 and Solution 2– with maceration
Protocol 3	1x wash with Purification Reagent 1x wash with TE Buffer Elution with Solution 1 and Solution 2– **without** maceration
Protocol 4	Elution with Solution 1 (70µL) and Solution 2 (130µL) : without maceration
Optimized Protocol	**2x** Elution with Solution 1 (140µL) and Solution 2 (260µL): without maceration

The concentration and purity measured by Nanodrop 2000 Sprectrophotometer of the DNA obtained in each of the protocols being tested was one of the parameters analyzed.

Generally, the reduction in the number of washes with Purification Reagent and TE buffer (protocol 2, protocol 3), as well as the non-washing of the filter (protocol 4, optimized protocol), led to an increase in the DNA concentration obtained compared to the Conventional protocol. The reduction to only one wash with the Purification Reagent in the extraction protocol (protocol 1) had no difference in the concentration of DNA obtained compared to the conventional protocol. However, the condition of

reducing 3 washes with TE buffer to only one wash with this reagent (protocol 2), by itself, led to a significant increase in DNA concentration as compared to the conventional protocol. The maceration (conventional protocol, protocol 1, protocol 2) or no maceration of the filter (protocol 3) had no significant change in DNA concentration. The lack of washings in the extraction protocol (protocol 4 and optimized protocol) generates a large amount of DNA compared to the conventional protocol. However, protocol 4 presents a large standard deviation, so it may not be a very conclusive result, requiring a greater number of extractions. Double the elution volume in the pH treatment process, without any previous filter washing (optimized protocol) led to a very significant increase in DNA concentration compared to the remaining protocols being tested. The optimized protocol is the one with the highest DNA concentration obtained (figure 8).

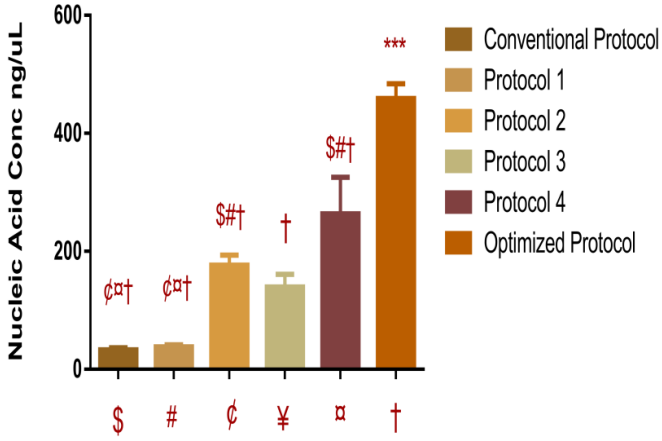


Figure 8: Nucleic Acid Concentration (ng/mL) of exhaled breath samples extracted by various protocols. Statistical analysis was performed by one-way ANOVA followed by Bonferroni test ($¢ #¢ ¤† p <0.01, $ † $¤ #¤ #† † ¢ *** p <0.001).

Relative to the absorbance ratio 260/280, one of the measures that evaluate the purity of the DNA, a significant decrease of this ratio is observed when the washing is stopped. The decrease of the ratio 260/280 ratio is notorious when only one washing with TE buffer is carried out (protocol 2), having as comparison protocol 1 and conventional protocol. The act of maceration decreases the ratio 260/280. In addition, there is not a very significant difference in the level of this ratio between the conventional protocol and the optimized protocol (figure 9).

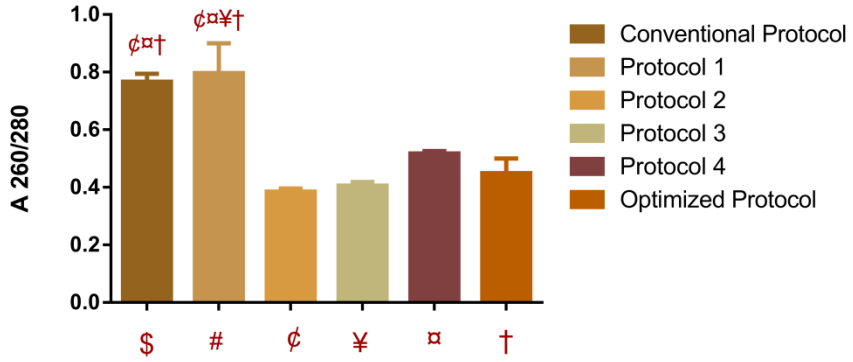


Figure 9: Ratio A260/280 of exhaled breath samples extracted by various protocols. Statistical analysis was performed by one-way ANOVA followed by Bonferroni test (#† #¤ $† p <0,05 $¥ ¥# p <0.01, $¢ #¢ p <0.001).

According to the remaining data provided by Nanodrop, a progressive increase of A260 and A280 was observed with a reduction in the number of washes (protocol 1, protocol 2, protocol 3) and also in the absence of washes (protocol 4 and optimized protocol) compared to the conventional protocol. The increase of these two parameters (A260 and A280) is notorious when comparing the conventional protocol with optimized protocol. It should be noted that the standard deviations of the A280 for the protocols 2, 3, 4 and optimized protocol and the standard deviation of the A260 for the protocol 4 are large compared to the other protocols. So a larger amount of samples for these protocols should result on smaller standard deviation values. In relation to the secondary measurement of DNA purity, A260/230, this does not present a great difference between the several protocols under study, being that the optimized protocol presents a slightly increased value (table 3).

Table 3: Nucleic Acid Concentration, absorbance at 260 and 280 and absorbance ratios: A260/280 and A260/230 of DNA extracted by the protocols tested, using NanoDrop spectrophotometer

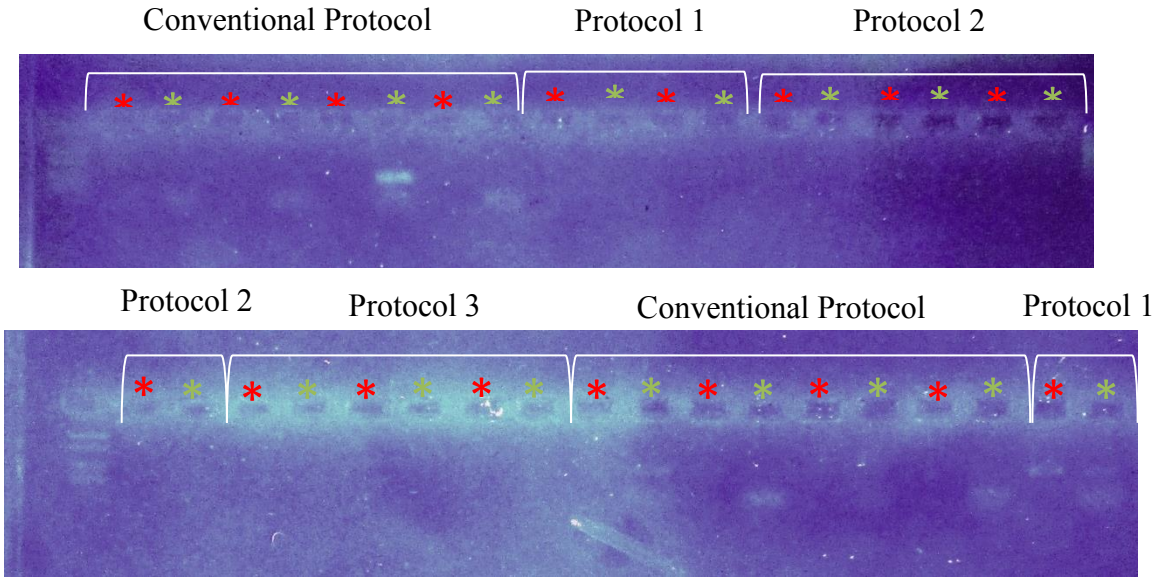
Extraction Protocols	N	Nucleic Acid Conc ng/µl	A 260	A280	A260/280	A260/230
Conventional Protocol	8	33.25 ± 9.86	0.67 ± 0.02	0.86 ± 0.21	0.77 ± 0.07	0.32 ± 0.07
Protocol 1	7	38.01 ± 9.48	0.76 ± 0.19	1.00 ± 0.22	0.80 ± 0.25	031 ± 0.03
Protocol 2	7	176.64 ± 41.68	3.53 ± 0.83	8.23 ± 2.51	0.39 ± 0.02	0.28 ± 0.03
Protocol 3	3	139.87 ± 30.19	2.80 ± 0.60	6.86 ± 1.23	0.41 ± 0.02	0.25 ± 0.01
Protocol 4	5	263.86 ± 123.30	5.28 ± 2.47	10.38 ± 5.18	0.52 ± 0.02	0.33 ± 0.11
Optimized Protocol	3	459.37 ± 35.05	9.19 ± 0.70	21.19 ± 5.45	0.45 ± 0.07	0.42 ± 0.08

Qubit quantification was performed for the samples obtained from protocol 4 and optimized protocol (table 4). According to the table, the mean QF value of the optimized protocol, 0.99, is slightly higher compared to protocol 4, 0.89. The mean concentration obtained for both protocols are quite similar, but the mean total amount obtained is higher in the optimized protocol. Given this, the optimized protocol seems to be more efficient.

Table 4: Qubit quantification of the exhaled breath samples extracted with protocol 4 and optimized protocol.

Protocols	QF	Conc. (ng/µL)	Mass (ng)
Protocol 4	0.9	0.09	36
	1.36	0.14	54
	0.99	0.10	40
	0.5	0.05	20
	0.72	0.07	29
MEAN	0.89 ± 0.29	0.09 ± 0.03	35.76 ± 11.48
Optimized Protocol	0.5	0.05	20
	1.15	0.12	46
	1.33	0.13	53
MEAN	0.99 ±0.36	0.10 ± 0.04	39.73 ± 14.26

The second parameter analyzed to evaluate the efficiency of the various protocols under test is the integrity of the DNA by agarose gel electrophoresis. According the images below we can observe that some samples with 10/20 µl DNA of conventional protocol and a protocol 1 has bands. However, the samples with 10/20 µL DNA of Protocol 4 and the optimized protocol present more defined bands at 350 bp (figure 10).



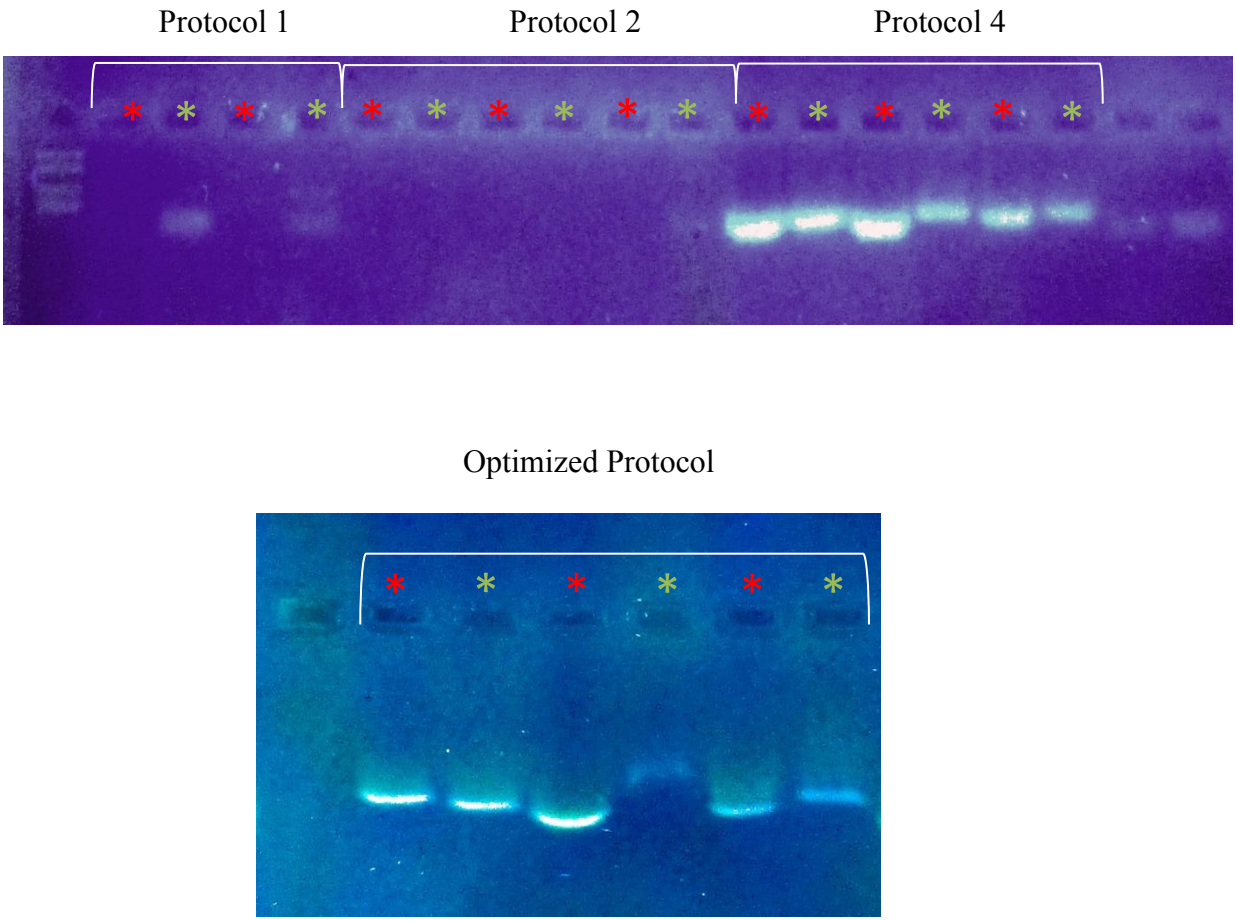


Figure 10: Agarose gel eletrophoresis (2% agarose) of PCR amplified products with 10µL (*) or 20 µL (*) DNA of the various protocols in tests, using a set of primers for the human beta-globin gene.

4.0 Analysis of breath samples in Population Controls

The efficacy and reproducibility of the optimized extraction protocol was verified by the study of a control group composed of 24 individuals over 50 years of age from the general population. In the table below the QF values of the Qubit quantifications obtained in each of the exhaled breath samples of the study population are presented (table 5). From the value of QF given by the Qubit concentration and total mass was calculated. A mean QF value of 3.05 ng/mL was obtained, the minimum of this value was 0.50 ng/mL and the maximum was 10.10 ng/mL. Regarding the

concentration, a mean value of 0.31 ng/uL was obtained. Regarding the total amount, a mean mass value of 122 ng was measured, with the maximum mass of 404 ng and the minimum of 20 ng.

Table 5: Quantification Qubit Fluorometer with 2 µl of DNA. The mass calculation was from an elution volume of 400µL.

N	QF	Conc. (ng/µL)	Mass (ng)
A	1.15	0.12	46
B	7.9	0.79	316
C	2.05	0.21	82
D	3.93	0.39	157
E	10.10	1.01	404
F	10.10	1.01	404
G	1.55	0.16	62
H	9.41	0.94	376
I	4.88	0.49	195
J	0.50	0.05	20
K	0.50	0.05	20
L	6.27	0.63	251
M	1.14	0.11	45.6
N	0.50	0.05	20
O	1.17	0.12	47
P	7.58	0.76	303
Q	0.67	0.07	27
R	0.52	0.05	21
S	0.50	0.05	20
T	0.54	0.05	22
U	0.50	0.05	20
V	0.50	0.05	20
W	0.50	0.05	20
X	0.84	0.08	34
MEAN	3.05 ± 3.53	0.31 ± 0.34	122 ± 136.92

The integrity of the DNA was evaluated by electrophoresis, in which 350bp bands corresponding to the human beta-globin fragment, indicative of non-degradation of DNA (figure 11).

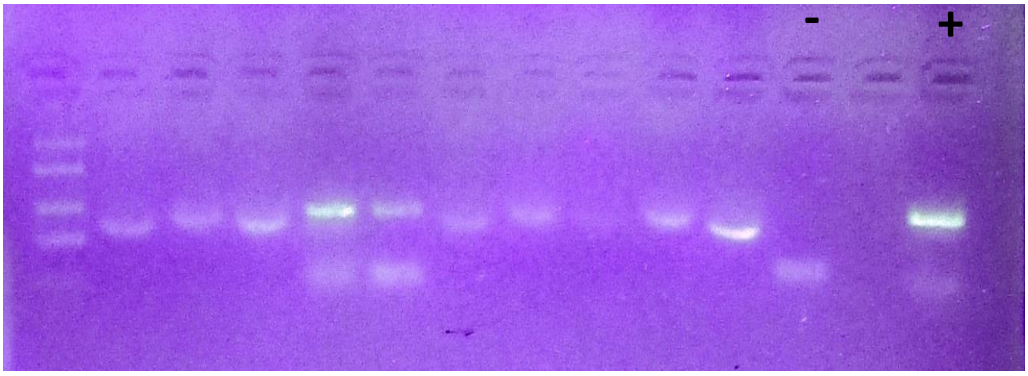


Figure 11: Agarose gel electrophoresis (2% agarose) of PCR amplified products using a set of primers for the human beta-globin gene.

CHAPTER V

5. DISCUSSION AND CONCLUSION

Given the current dismal survival rate along with the absence of noticeable symptoms in the early stage of lung cancer, early detection tools in high-risk populations of asymptomatic individuals is a challenge that needs to be addressed in the next years or decades (Dent, Sutedja, & Zimmerman, 1999). Nowadays, the screening methods employed in the clinical practice either involve radiation exposure or are invasive and have a high rate of false positives (Greenberg & Lee, 2007). The need to discover biomarkers that increase the possibility of detecting lung cancer in the early stages has been leading to several published studies reporting intriguing results regarding the detection of biomarkers in tumor samples, but also in readily accessible samples such as EBC. EBC is a source of proteins, lipids, markers of oxidative stress, and human DNA from the fluid lining of airway (Chorostowska-wynimko & Szpechcinski, 2007). All of these compounds may be useful indicators in understanding processes of diseases in the lung, namely, cancer (Dent et al., 1999). Several studies have already announced the potential of this approach in the detection of genetic biomarkers associated with lung cancer, such as mutations in genes: p53, EGFR, p16, KRAS; alterations of microsatellite DNA of chromosomal region 19q and 3p; microRNAs: miRNA-21 and miRNA-486; gene promoter methylation: DAPK, RASSF1A, PAX5β and p16 and, finally, mtDNA mutations. In the studies in which 3p microsatellite alterations were detected, the number of these alterations correlated with the survival of patients with NSCLC, as well as with smoking history. Furthermore, authors have demonstrated that the DNA detected in EBC exhibited a similar spectrum of microsatellite alterations compared to tumor tissues and that this spectrum is significantly different between patients with and control individuals (Carpagnano et al., 2005, 2010). Regarding mutations detected in the EGFR gene, only one study has so far been able to detect this type of mutation and in only one patient. Given this inconsistency, further studies will be needed to prove the reproducibility of its detection, since this gene is quite prevalent in lung cancer adenocarcinomas. In addition, changes in EGFR have clinical value in predicting the response to therapy with tyrosine kinase inhibitors (Paradiso et al., 2008). The value of the KRAS mutation as a potential marker in determining the recurrence of pulmonary malignancy in a follow-up setting after surgical treatment has also been demonstrated (Kordiak et al., 2012). Therefore,

detection of biomarkers associated with lung cancer in exhaled breath condensate have several potential clinical applications. They can be used in risk stratification, early detection of lung cancer, treatment selection, prognosis and monitoring for recurrence (Greenberg & Lee, 2007). These genetic-based cancer biomarkers are detected by readily accessible and sensitive molecular biology techniques such as polymerase chain reaction (PCR), reverse transcriptase polymerase chain reaction (RT-PCR) and DNA sequencing, which can be used to detect the genetic alterations occurring in the cancerous stage (Sung & Cho, n.d.).

All these smiling results have prompted the development of portable, self-collection devices that allows the recovery and storage of biological material from the exhaled breath, in order to detect genetic biomarkers that could ultimately assist in the early detection of lung cancer. The capture of biological material is performed by using a filter. The optimization of a protocol for collecting biological material from the exhaled breath captured by this filter, in order to obtain the highest possible amount of nucleic acids for further detection of genetic biomarkers, was the major objective of the present thesis. The efficiency of each of the protocols was evaluated based on (i) the concentration and purity of DNA obtained by spectrophotometric method (Nanodrop 2000 Sprectrophotometer) and fluorometric method (Qubit); (ii) the integrity of the DNA amplified by Human Beta-Globin primers when analyzed by PCR and agarose gel.

The conventional extraction protocol involves a washing step, consisting of 3 washes with purification reagent and 3 washes with TE buffer; followed by a heat pH treatment by the addition of two solutions, as well as maceration of the filter. The optimization of the extraction protocol was based on the conventional protocol, in which changes were made in the number of washes, non-maceration of paper and modification of the volume of DNA elution (table 2). According to the data obtained, it was observed that the reduction in the number of washes (protocol 1, protocol2, protocol 3), as well as its non-procedure (protocol 4, optimized protocol) led to an increase in the concentration of DNA obtained (figure 8), however with lower purity (figure 9), based on the comparison with the conventional protocol. This result demonstrates that the washing steps remove DNA which is attached to the cellulose fibers. On the other hand, the fact that not washing the filter does not remove contaminants, which is why the optimized protocol has a ratio A260/280 = 0.45 and the

conventional protocol a ratio A260/280 = 0.77. Protocol 1 shows a DNA concentration and A260/280 slightly similar to that of the conventional protocol. This means that only a wash with purification reagent had no effect on the optimization of the extraction protocol, leading to the belief that washing with TE buffer could lead to the removal of the DNA attached to the filter. However, when one washing with purification reagent and TE buffer (protocol 2), there is a sudden increase in the DNA concentration obtained (figure8), as well as a sudden decrease of A260/280 (figure 9), compared to the conventional protocol and protocol 1. It is inferred that this protocol does not exhibit efficient washing, ie, it does not remove many contaminants from the filter, but it also does not remove much DNA. Therefore, this protocol presents pros and cons. When the filter maceration is stopped at the time of the heat pH treatment (protocol 3), there is not much change in the concentration of DNA obtained and in the ratio A260/280, compared to protocol 2. Since, there is no change in the DNA concentrations obtained, the maceration was not performed, for reasons of filter processing time during extraction, especially for studies with large amount of samples. The absence of any filter washing (protocol 4) led to a significant increase in DNA concentration. However, the A260/280 ratio remains lower than the conventional protocol and protocol 1, but it remains in the same range as protocol 2 and 3. This result demonstrates that the wash step removes immense quantity of DNA from the filter, a result consistent with the above. By placing twice of the volume of solutions 1 and 2 stipulated by the conventional protocol along with the absence of washing procedure (optimized protocol), an increase in DNA concentration was observed, which was a significant increase in comparison with all the remaining protocols. Regarding the A260/280 ratio, it remains well below the conventional protocol and protocol 1. According to the Qubit quantifications in filters processed by protocol 4 and optimized protocol, it was observed that despite the similar concentration, the optimized protocol has a higher total amount of DNA (table 4). The combination of the results obtained by the NanoDrop and Qubit quantification leads us to believe that the optimized protocol is the best option compared to the others, even though it presents an A260/280 lower than the initial protocols.

According to the PCR and electrophoresis data, which allows the evaluation of DNA integrity, it is possible to observe that the DNA from extraction with protocol 4 and optimized protocol are the ones that present clearer bands at 350bp, demonstrating

the integrity of DNA for amplification (figure 10). On the other hand, the remaining protocols do not present almost any band, and those that are observed are very subtle, which leads to the conclusion that the remaining protocols may lead to a degradation of the DNA, or may also be due to the fact that there is little concentration of DNA or presence of impurities. Again, the optimized protocol seems to be the best option for extracting DNA from exhaled breath samples captured by the filter, although it has a low A260/280 purity ratio. The results demonstrate that the optimization of extraction protocols led to an increase in the amount of DNA extracted from the filters, especially protocol 4 and the optimization protocol (table 3). However, it is expected that the increase in the extraction of proteins and other biomolecules present in the filter has also occurred, and that they may contribute to the absorbance at 280 nm verified by the NanoDrop. This may explain the low A260/280 ratios observed for the samples but whose PCR amplification performance was not affected by the degree of purity of the DNA obtained. As these samples are from exhaled breath it is well possible that the amount of DNA from this type of sample is much lower than that of DNA samples from other sources. In this sense, it was chosen to primarily take into account the concentration/quantity of DNA obtained.

There are no results in the literature that support our results obtained with the optimized protocol. However, there are studies that have performed a protocol similar to the conventional protocol and protocol 1, but do not always demonstrate or perform NanoDrop quantifications. Hansen et al., 2007 evaluated the yield and purity of the genomic DNA by NanoDrop and agarose gel electrophoresis, using several collection methods, including buccal cells in 3mm of filter. The extraction of buccal cells in the filter consisted of 2 washes with 200 µL of purification reagent, a wash with 200 µL TE buffer followed by an elution at room temperature with 35 µL Solution 1 and 65 µL Solution 2. This assay resulted in a mean concentration of 3.6 ng/µL and a mean values of A 260/280 and A260/230 raios of 0.91 and 0.21, respectively. It is noteworthy that they use a 3mm punch and we use a 10mm punch. However, if the proportion of the DNA concentration obtained with the conventional protocol for a 3 mm punch, we obtain a concentration of 5 ng/µL, slightly higher than the results obtained by Hansen et al.2007. If we do the same reasoning for the optimized protocol, we will obtain a concentration of 69 ng/µL, much higher than the study in comparison. Additionally, they obtained a A260/280 ratio higher than our study, but we obtained an upper

A260/230. However, the value of A260 / 280 and A260 / 230 are below the desired values, 1.8 and 2.0, respectively. It should be noted that no band in the study of Hansen et al., 2007 was observed after PCR and electrophoresis with 5μL of DNA (Hansen, Simonsen, Nielsen, & Hundrup, 2007). Petras et al., 2011 carried out a study that aimed to detect KRAS mutations in colonic tumors by DNA extraction from a similar filter. For that they used two 2.0 mm punch of filter. DNA extraction consisted of 3 washes with 200 μL purification reagent, 2 washes with 200 μL TE buffer, followed by elution at room temperature with 35 μL Solution 1 and 65 μL Solution 2. The authors had a mean concentration of 23.5 ng/μL and A260/280 = 1.1. These results, at the DNA concentration level, are superior to the conventional protocol, when calculating the proportion of DNA obtained from our study for a two 2 mm punch, in which we obtain a concentration of 6.65ng/μL. At the same time the results of Petras *et al.*, 2011 show lower amounts relative to our results with the optimized protocol, in with which we obtain a concentration of 91.87ng/μL. However, Petras *et al.*, 2011 reported a ratio 260/280 far superior to ours, but still below to the desirable (Petras et al., 2011). Lema *et al.*, 2006 optimized a pH method for DNA elution from buccal cells collected in 6mm of similar filters. In this study they extracted the DNA from 2 washes of 200 μL of Purification Reagent in ice and a wash with 150 μL TE buffer, followed by an elution with 200 μL of 0.1 NaOH, 0.1 M Tris-HCl, 0.3 mM EDTA, pH = 8.5. With this extraction method that conjugates only one solution to elute DNA, these authors obtained a mean DNA yield of 3.72 μg (SD = 0.45) and mean A260/280 ratio of 0.595 (SD = 0.065). Making a proportion, the results we obtained with the conventional protocol for 6mm would be about 2μg and with the optimized protocol, about 55 μg. However, in terms of A260/280 the conventional protocol has an index greater than the protocol under study, with 0.77, and the optimized protocol with 0.45. Thus, with the use of our optimized protocol the concentration of DNA obtained is much higher than the protocol of Lemas *et al.*, 2006. Again, the study does not reach a desirable A260/280 ratio (Lema, Kohl-white, Lewis, & Dao, 2006).

The population control based-study of healthy individuals, some of them smokers and some former smokers, whose objective was to verify the reproducibility and efficiency of the optimized protocol, had a mean DNA concentration of 0.31 ng/μL, as measured by the Qubit method (table5). According to the PCR and electrophoresis,

all exhaled breath samples from presented bands at 350 bp, indicating the reproducibility of our optimized protocol (figure 11).

In conclusion, the present work has successfully achieved the optimization of a protocol for recovery of biological material from exhaled breath for subsequent detection of genetic biomarkers. This is the first step towards the use of a breath test in early diagnosis of lung cancer.

CHAPTER VI

6. FUTURE PERSPECTIVES

The results that have been achieved with this work, along with the results of the studies carried out in the field of exhaled breath condensate prompt for the testing of the device Oncosopro in a clinical context. Clinical evaluation of the Oncosopro device would be done through a case-control study, comprising groups of healthy individuals, both smokers and non-smokers, patients with benign lung diseases patients with lung cancer and patients with other types of cancer. In the lung cancer patient group a sample of exhaled breath would be collected prior to any therapeutic approach as well as during follow up. This study would have as its initial task the identification of a panel of sensitive and specific markers, namely mutations in oncogenes and methylation in tumor suppressor genes. Within these markers, it would be interesting to include a spectrum of molecular alterations that have relevance in prediction of the therapeutic response. The determination of this genetic and epigenetic panel would be performed by the Ion Torrent sequencing platform, and each of the detected mutations would be validated by specific Real-Time PCR assays. The profile of mutations and methylations detected by both techniques would be compared to the tumor DNA to verify if the profile of changes detected in the exhaled breath reflects the changes present in the tumor. Clinical validation of this device involves determining the specificity, sensitivity, negative predictive value and positive predictive value of the markers panel in the early detection of lung cancer. In addition, a comparison in terms of sensitivity and specificity, this method with low-dose computed tomography (LCDT), the only screening method for lung cancer implemented in the United States, will be required. If the clinical utility of this blow device is approved, this method can be integrated into asymptomatic population screening programs and in a more hospital setting, it can be used to monitor disease regression or progression in post-surgery patients.

7. BIBLIOGRAPHY

Ahmadzai, H., Huang, S., Hettiarachchi, R., Lin, J. L., Thomas, P. S., & Zhang, Q. (2013). Exhaled breath condensate: A comprehensive update. *Clinical Chemistry and Laboratory Medicine*, *51*(7), 1343–1361. http://doi.org/10.1515/cclm-2012-0593

Alberg, A. J., & Samet, J. M. (2003). Epidemiology of Lung Cancer *. *CHEST*, *123*(1), 21S–49S. http://doi.org/10.1378/chest.123.1

Antczak, A., & Gorski, P. (2002). Markers of pulmonary diseases in exhaled breath condensate. *Int J Occup Med Environ Health*, *15*(4), 317–323. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12608619

Application, F., & Data, P. (1996). |lllllllllllllllIlllllllllllllllllllllllllllllllllllll|||lllllllllllllllllllllll.

Article, O. (2005). carcinomas on biopsies with extensive crush artefact, 978–981. http://doi.org/10.1136/jcp.2004.023044

Baldi, A., De Luca, A., Esposito, V., Campioni, M., Spugnini, E. P., & Citro, G. (2011). Tumor suppressors and cell-cycle proteins in lung cancer. *Pathology Research International*, *2011*, 605042. http://doi.org/10.4061/2011/605042

Balgkouranidou, I., Liloglou, T., & Lianidou, E. S. (2013). Lung cancer epigenetics: emerging biomarkers. *Biomarkers in Medicine*, *7*(1), 49–58. http://doi.org/10.2217/bmm.12.111

Baylin, S. B. (2005). DNA methylation and gene silencing in cancer. *Nature Clinical Practice. Oncology Oncology*, *2 Suppl 1*(December), S4–11. http://doi.org/10.1038/ncponc0354

Bhattacharya, S., Socinski, M. A., & Burns, T. F. (2015). KRAS mutant lung cancer: progress thus far on an elusive therapeutic target. *Clinical and Translational Medicine*, *4*(1), 35. http://doi.org/10.1186/s40169-015-0075-0

Bicher, a, Ault, K., Kimmelman, a, Gershenson, D., Reed, E., & Liang, B. (1997). Loss of heterozygosity in human ovarian cancer on chromosome 19q. *Gynecologic Oncology*, *66*(1), 36–40. http://doi.org/10.1006/gyno.1997.4709

Bİr, F., Çelİker, D., Evyapan, B. F., Yaren, A., & Edİrne, T. (2016). New immunohistochemical markers in the differential diagnosis of nonsmall cell lung carcinoma, 1854–1861. http://doi.org/10.3906/sag-1501-68

Bishop, J. A., Teruya-feldstein, J., Westra, W. H., Pelosi, G., Travis, W. D., & Rekhtman, N. (2011). p40 (D Np63) is superior to p63 for the diagnosis of pulmonary squamous cell carcinoma. *Modern Pathology*, *25*(3), 405–415.

http://doi.org/10.1038/modpathol.2011.173

Bittermann, N., Rawluk, J., Kayser, G., Csanadi, A., Otto, C., Plo, T., … Werner, M. (2013). Simultaneous Multi-Antibody Staining in Non-Small Cell Lung Cancer Strengthens Diagnostic Accuracy Especially in Small Tissue Samples, *8*(2), 1–10. http://doi.org/10.1371/journal.pone.0056333

Boland, R. C., & Goel, A. (2010). Microsatellite Instability in Colorectal Cancer. *Gastroenterology*, *138*(6), 2073–2087. http://doi.org/10.1053/j.gastro.2009.12.064.Boland

Caglieri, A., Goldoni, M., Acampa, O., Andreoli, R., Vettori, M. V., Corradi, M., … Mutti, A. (2006). The effect of inhaled chromium on different exhaled breath condensate biomarkers among chrome-plating workers. *Environmental Health Perspectives*, *114*(4), 542–546. http://doi.org/10.1289/ehp.8506

Campbell, J. D., Alexandrov, A., Kim, J., Wala, J., Berger, A. H., Pedamallu, C. S., … Govindan, R. (2016). Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas, (April). http://doi.org/10.1038/ng.3564

Campos-parra, A. D., Zuloaga, C., Vazquez, E., & Avile, A. (2015). O RIGINAL A RTICLE KRAS Mutation as the Biomarker of Response to Chemotherapy and EGFR-TKIs in Patients With Advanced Non – Small Cell Lung Cancer Clues For Its Potential Use in Second-Line Therapy Decision Making, *38*(1), 33–40. http://doi.org/10.1097/COC.0b013e318287bb23

Cancer, T., & Atlas, G. (2012). Comprehensive genomic characterization of squamous cell lung cancers. http://doi.org/10.1038/nature11404

Capelozzi, V. L. (2009). Review Article, *35*(December 2008), 375–382.

Carney, J. M., Kraynie, A. M., & Roggli, V. L. (n.d.). Immunostaining in Lung Cancer for the Clinician, (5), 429–435. http://doi.org/10.1513/AnnalsATS.201501-004FR

Carpagnano, G. E., Foschino-Barbaro, M. P., Mulé, G., Resta, O., Tommasi, S., Mangia, A., … Paradiso, A. (2005). 3P Microsatellite Alterations in Exhaled Breath Condensate From Patients With Non-Small Cell Lung Cancer. *American Journal of Respiratory and Critical Care Medicine*, *172*(6), 738–744. http://doi.org/10.1164/rccm.200503-439OC

Carpagnano, G. E., Foschino-Barbaro, M. P., Spanevello, A., Resta, O., Carpagnano, F., Mulé, G., … Paradiso, A. (2008). 3P Microsatellite Signature in Exhaled Breath Condensate and Tumor Tissue of Patients With Lung Cancer. *American Journal of Respiratory and*

Critical Care Medicine, *177*(3), 337–41. http://doi.org/10.1164/rccm.200707-1136OC

Carpagnano, G. E., Palladino, G. P., Gramiccioni, C., Foschino Barbaro, M. P., & Martinelli, D. (2010). Exhaled ERCC-1 and ERCC-2 microsatellite alterations in NSCLC patients. *Lung Cancer*, *68*(2), 305–307. http://doi.org/10.1016/j.lungcan.2010.01.020

Carpi, F. M., Pietro, F. Di, Vincenzetti, S., & Mignini, F. (2011). Human DNA Extraction Methods : Patents and Applications, 1–7.

Carter, S. R., Davis, C. S., & Kovacs, E. J. (2012). Exhaled breath condensate collection in the mechanically ventilated patient. *Respiratory Medicine*. http://doi.org/10.1016/j.rmed.2012.02.003

Chen, J., Chen, J., Han, H., Zhou, F., Lv, X., & Ma, H. (2016). Clinical significance of miRNA21 in exhaled breath condensate of non-small-cell lung cancer. *International Journal of Clinical and Experimental Medicine*, *9*(9), 17232–17238.

Chen, J., Chen, J., Huang, F., Tao, G., Zhou, F., & Tao, Y. (2015). Analysis of p16 gene mutations and their expression using exhaled breath condensate in non‐small‐cell lung cancer. *Oncology Letters*, 1477–1480. http://doi.org/10.3892/ol.2015.3426

Chen, Z., Fillmore, C. M., Hammerman, P. S., & Kim, C. F. (2014). Non-small-cell lung cancers : a heterogeneous set of diseases. *Nature Publishing Group*, *14*(8), 535–546. http://doi.org/10.1038/nrc3775

Cherni, I., & Weiss, G. J. (2011). miRNAs in lung cancer: large roles for small players. *Future Oncology (London, England)*, *7*(9), 1045–55. http://doi.org/10.2217/fon.11.74

Chorostowska-wynimko, J., & Szpechcinski, A. (2007). The Impact of Genetic Markers on the Diagnosis of Lung Cancer : A Current Perspective. *Journal of Thoracic Oncology*, *2*(11), 1044–1051. http://doi.org/10.1097/JTO.0b013e318158eed4

Chung, K. F., & Widdicombe, J. G. (2008). *Cough: Causes, Mechanisms and Therapy*. *Cough: Causes, Mechanisms and Therapy*. http://doi.org/10.1002/9780470755846

Conrad, D. H., Goyette, J., & Thomas, P. S. (2008). Proteomics as a method for early detection of cancer: A review of proteomics, exhaled breath condensate, and lung cancer screening. *Journal of General Internal Medicine*, *23*(1 SUPPL.), 78–84. http://doi.org/10.1007/s11606-007-0411-1

Corradi, M., Gergelova, P., & Mutti, A. (2010). Use of exhaled breath condensate to investigate occupational lung diseases. *Current Opinion in Allergy and Clinical Immunology*, *10*(2),

93–98. http://doi.org/10.1097/ACI.0b013e3283357fb7

Costa, M. Da, & Pt, M. (2013). (19) United States (12), *1*(19).

Costa, M. D. A., & Pedro, R. (2014). Tepzz 68 _88a_t (11), *1*(19), 1–16.

Czebe, K., Barta, I., Antus, B., Valyon, M., Horváth, I., & Kullmann, T. (2008). Influence of condensing equipment and temperature on exhaled breath condensate pH, total protein and leukotriene concentrations. *Respiratory Medicine*, *102*(5), 720–725. http://doi.org/10.1016/j.rmed.2007.12.013

D'Arcangelo, M., & Cappuzzo, F. (2012). K-Ras Mutations in Non-Small-Cell Lung Cancer: Prognostic and Predictive Value. *ISRN Molecular Biology*, *2012*, 1–8. http://doi.org/10.5402/2012/837306

Davidson, M. R., Gazdar, A. F., & Clarke, B. E. (n.d.). Review Article The pivotal role of pathology in the management of lung cancer. http://doi.org/10.3978/j.issn.2072-1439.2013.08.43

Davidsson, A., & Schmekel, B. (2010). Efficacy of two breath condensers. *Journal of Clinical Laboratory Analysis*, *24*(4), 219–223. http://doi.org/10.1002/jcla.20389

Dent, A. G., Sutedja, T. G., & Zimmerman, P. V. (1999). Review Article Exhaled breath analysis for lung cancer, *1997*(6). http://doi.org/10.3978/j.issn.2072-1439.2013.08.44

Dhar, S., Squire, J. A., Hande, M. P., Wellinger, R. J., & Pandita, T. K. (2000). Inactivation of 14-3-3sigma influences telomere behavior and ionizing radiation-induced chromosomal instability. *Molecular and Cellular Biology*, *20*(20), 7764–72. http://doi.org/10.1128/MCB.20.20.7764-7772.2000

Dodig, S., & Cepelak, I. (2013). Exhaled breath condensate--from an analytical point of view. *Biochemia Medica*, *23*(3), 281–95. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3900073&tool=pmcentrez&rendertype=abstract

Dvorak, K., Yambert, J., Palting, J., Reinhardt, K., Roessler, C., Mcginn, S., & Marner, E. (n.d.). Lung cancer classification using an immunohistochemical assay with the anti-p40 (BC28) Mouse Monoclonal Primary Antibody, *40*.

Eggert, J., Palavanzadeh, M., & Blanton, A. J. (2017). S CREENING AND E ARLY. *Seminars in Oncology Nursing*, 1–12. http://doi.org/10.1016/j.soncn.2017.03.001

Elsevier. (2017). *Volatile Biomarkers - 1st Edition*. Retrieved from https://www.elsevier.com/books/volatile-biomarkers/amann/978-0-444-62613-4

Facts, C. (2016). Cancer Facts & Figures.

Facts, C. (2017). Cancer Facts & Figures 2017.

Felekkis, K., Touvana, E., Stefanou, C., & Deltas, C. (2010). MicroRNAs: A newly described class of encoded molecules that play a role in health and disease. *Hippokratia*, *14*(4), 236–240.

Field, J. K., Liloglou, T., Xinarianos, G., Prime, W., Fielding, P., Walshaw, M. J., & Turnbull, L. (1999). Genetic alterations in bronchial lavage as a potential marker for individuals with a high risk of developing lung cancer. *Cancer Res*, *59*(11), 2690–2695. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10363993

Fleisher, A. S., Esteller, M., Wang, S., Tamura, G., Suzuki, H., Yin, J., … Meltzer, S. J. (1999). Hypermethylation of the hMLH1 gene promoter in human gastric cancers with microsatellite instability. *Cancer Res*, *59*(5), 1090–1095.

Fukazawa, T., Guo, M., Ishida, N., & Yamatsuji, T. (2016). SOX2 suppresses CDKN1A to sustain growth of lung squamous cell carcinoma. *Nature Publishing Group*, (February), 1–12. http://doi.org/10.1038/srep20113

Gessner, C., Kuhn, H., Toepfer, K., Hammerschmidt, S., Schauer, J., & Wirtz, H. (2004). Detection of p53 gene mutations in exhaled breath condensate of non-small cell lung cancer patients. *Lung Cancer*, *43*(2), 215–222. http://doi.org/10.1016/j.lungcan.2003.08.034

Giangreco, A., Lu, L., Vickers, C., Teixeira, V. H., Groot, K. R., Ekaterina, C. R. B., … Janes, S. M. (2012). β -Catenin determines upper airway progenitor cell fate and preinvasive squamous lung cancer progression by modulating epithelial – mesenchymal transition, 575–587.

Goldoni, M., Caglieri, A., Andreoli, R., Poli, D., Manini, P., Vettori, M. V., … Mutti, A. (2005). Influence of condensation temperature on selected exhaled breath parameters. *BMC Pulmonary Medicine*, *5*, 10. http://doi.org/10.1186/1471-2466-5-10

Great, L. T. (2013). A Genomics-Based Classification of Human Lung Tumors, *153*. http://doi.org/10.1126/scitranslmed.3006802

Greenberg, A. K., & Lee, M. S. (2007). Biomarkers for lung cancer: clinical uses.

Curr.Opin.Pulm.Med, *13*(4), 249–255. http://doi.org/10.1097/MCP.0b013e32819f8f06

Gridelli, C., Rossi, A., Carbone, D. P., Guarize, J., Karachaliou, N., Mok, T., … Rosell, R. (2015). Non-small-cell lung cancer, *1*, 1–16. http://doi.org/10.1038/nrdp.2015.9

Grob, N. M., Aytekin, M., & Dweik, R. A. (2008). Biomarkers in exhaled breath condensate: a review of collection, processing and analysis. *J Breath Res*, *2*(3), 37004. http://doi.org/10.1088/1752-7155/2/3/037004

Gurda, G. T., Zhang, L., Wang, Y., Chen, L., Geddes, S., Cho, W. C., … Li, Q. K. (2015). Utility of five commonly used immunohistochemical markers TTF-1 , Napsin A , CK7 , CK5 / 6 and P63 in primary and metastatic adenocarcinoma and squamous cell carcinoma of the lung : a retrospective study of 246 fine needle aspiration cases. ??? http://doi.org/10.1186/s40169-015-0057-2

Han, W., Wang, T., Reilly, A. a, Keller, S. M., & Spivack, S. D. (2009). Gene promoter methylation assayed in exhaled breath, with differences in smokers and lung cancer patients. *Respiratory Research*, *10*(7), 86. http://doi.org/10.1186/1465-9921-10-86

Hansen, T. O., Simonsen, M. K., Nielsen, F. C., & Hundrup, Y. A. (2007). Collection of Blood , Saliva , and Buccal Cell Samples in a Pilot Study on the Danish Nurse Cohort : Comparison of the Response Rate and Quality of Genomic DNA, *16*(October), 2072–2077. http://doi.org/10.1158/1055-9965.EPI-07-0611

Hatzimichael, E., Lagos, K., Sim, V. R., & Briasoulis, E. (2014). Review article : EPIGENETICS IN DIAGNOSIS , PROGNOSTIC ASSESSMENT AND TREATMENT OF CANCER : AN UPDATE. *Excli Journal*, *13*, 954–976.

Heuvelmans, M. A., Groen, H. J. M., & Oudkerk, M. (2017). Early lung cancer detection by low-dose CT screening : therapeutic implications. *Expert Review of Respiratory Medicine*, *0*(0). http://doi.org/10.1080/17476348.2017.1276445

Hiroshima, K., Iyoda, A., Shida, T., Shibuya, K., & Iizasa, T. (2006). Distinction of pulmonary large cell neuroendocrine carcinoma from small cell lung carcinoma : a morphological , immunohistochemical , and molecular analysis, 1358–1368. http://doi.org/10.1038/modpathol.3800659

Hoffman, R. M. (2017). L u n g C a n c e r S c re e n i n g. *Medical Clinics of NA*, *101*(4), 769–785. http://doi.org/10.1016/j.mcna.2017.03.008

Hoffmeyer, F., Raulf-Heimsoth, M., Harth, V., Bünger, J., & Brüning, T. (2009). Comparative

analysis of selected exhaled breath biomarkers obtained with two different temperature-controlled devices. *BMC Pulmonary Medicine*, *9*(1), 48. http://doi.org/10.1186/1471-2466-9-48

Horvath, I., & Jongste, J. C. de. (2010). *Exhaled Biomarkers*. European Respiratory Society. http://doi.org/10.1183/1025448x.erm4910

Ingvarsson, S. (2005). Tumor Suppressor Genes on Human Chromosome 3 and Cancer Pathogenesis. *Cancer Genomics Proteomics*, *2*(4), 247–254.

Jakupciak, J. P., Wang, W., Markowitz, M. E., Ally, D., Coble, M., Srivastava, S., … O'Connell, C. D. (2005). Mitochondrial DNA as a cancer biomarker. *J Mol Diagn*, *7*(2), 258–267. http://doi.org/10.1016/S1525-1578(10)60553-3

Janicka, M., Kot-Wasik, A., Kot, J., & Namieśnik, J. (2010). Isoprostanes-biomarkers of lipid peroxidation: Their utility in evaluating oxidative stress and analysis. *International Journal of Molecular Sciences*, *11*(11), 4631–4659. http://doi.org/10.3390/ijms11114631

Jignal, P., Mg, S., & Darshan, M. (2014). Forensic Conception : DNA typing of FTA Spotted Samples, *2*(04), 21–29. http://doi.org/10.7324/JABB.2014.2404

Jin, B., Li, Y., & Robertson, K. D. (2011). DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes & Cancer*, *2*(6), 607–17. http://doi.org/10.1177/1947601910393957

Jin, X., Zhang, J., Gao, Y., Ding, K., Wang, N., Zhou, D., … Cheng, S. (2007). Relationship between mitochondrial DNA mutations and clinical characteristics in human lung cancer. *Mitochondrion*, *7*(5), 347–353. http://doi.org/10.1016/j.mito.2007.06.003

Kasprzak, A., & Zabel, M. (2007). Selected Markers (Chromogranin A , Neuron-Specific Enolase , Synaptophysin , Protein Gene Product 9 . 5) in Diagnosis and Prognosis of Neuroendocrine Pulmonary Tumours, 23–33.

Konstantinidi, E. M., Lappas, A. S., Tzortzi, A. S., & Behrakis, P. K. (2015). Exhaled Breath Condensate: Technical and Diagnostic Aspects. *The Scientific World Journal*, *2015*, 1–25. http://doi.org/10.1155/2015/435160

Kordiak, J., Szemraj, J., Hamara, K., Bialasiewicz, P., & Nowak, D. (2012). Complete surgical resection of lung tumor decreases exhalation of mutated KRAS oncogene. *Respiratory Medicine*, *106*(9), 1293–1300. http://doi.org/10.1016/j.rmed.2012.06.019

Kubáň, P., & Foret, F. (2013). Exhaled breath condensate: Determination of non-volatile

compounds and their potential for clinical diagnosis and monitoring. A review. *Analytica Chimica Acta*, *805*, 1–18. http://doi.org/10.1016/j.aca.2013.07.049

Kulis, M., & Esteller, M. (2010). 2 – DNA Methylation and Cancer. *Advances in Genetics*, *70*(10), 27–56. http://doi.org/10.1016/B978-0-12-380866-0.60002-2

Kumarswamy, R., Volkmann, I., & Thum, T. (2011). Regulation and function of miRNA-21 in health and disease. *RNA Biology*, *8*(5), 706–713. http://doi.org/10.4161/rna.8.5.16154

Lages, E., Guttin, A., Nesr, H., Lages, E., Guttin, A., & Nesr, H. (2012). MicroRNAs : molecular features and role in cancer . To cite this version :

Langer, C. J., Besse, B., Gualberto, A., Brambilla, E., & Soria, J. (2017). J OURNAL OF C LINICAL O NCOLOGY The Evolving Role of Histology in the Management of Advanced Non – Small-Cell Lung Cancer, *28*(36), 5311–5320. http://doi.org/10.1200/JCO.2010.28.8126

Leese, E., Morton, J., Gardiner, P. H. E., & Carolan, V. A. (2016). Development of a method for the simultaneous detection of Cr(iii) and Cr(vi) in exhaled breath condensate samples using μLC-ICP-MS. *J. Anal. At. Spectrom.*, *31*(4), 924–933. http://doi.org/10.1039/C5JA00436E

Lema, C., Kohl-white, K., Lewis, L. R., & Dao, D. A. T. D. (2006). Optimized pH Method for DNA Elution from Buccal Cells Collected in Whatman FTA ® Cards, *10*(2), 126–130.

Liang, Y., Yeligar, S. M., & Brown, L. A. (2012). Exhaled breath condensate: a promising source for biomarkers of lung disease. *ScientificWorldJournal*, *2012*, 217518. http://doi.org/10.1100/2012/217518

Liggett, W. H., & Sidransky, D. (1998). Role of the p16 tumor suppressor gene in cancer. *Journal of Clinical Oncology*, *16*(3), 1197–1206. http://doi.org/10.1200/JCO.1998.16.3.1197

Lim, M., & Thomas, P. (2013). Biomarkers in Exhaled Breath Condensate and Serum of Chronic Obstructive Pulmonary Disease and Non-Small-Cell Lung Cancer. *International Journal of Chronic Diseases*, *2013*.

Liu, L., Teng, J., Zhang, L., Cong, P., Yao, Y., Sun, G., … Liu, M. (2017). The Combination of the Tumor Markers Suggests the Histological Diagnosis of Lung Cancer, *2017*. http://doi.org/10.1155/2017/2013989

Luk, P. P., Yu, B., Ng, C. C., Mercorella, B., Selinger, C., Lum, T., … Cooper, W. A. (2015).

BRAF mutations in non-small cell lung cancer, *4*(7), 142–148. http://doi.org/10.3978/j.issn.2218-6751.2014.08.08

Ma, Y., Fan, M., Dai, L., Kang, X., Liu, Y., Sun, Y., & Xiong, H. (2015). Expression of p63 and CK5 / 6 in early-stage lung squamous cell carcinoma is not only an early diagnostic indicator but also correlates with a good prognosis, *6*, 288–295. http://doi.org/10.1111/1759-7714.12181

Malhotra, J., Malvezzi, M., Negri, E., Vecchia, C. La, & Boffetta, P. (2016). Risk factors for lung cancer worldwide. http://doi.org/10.1183/13993003.00359-2016

Manuscript, A. (2014a). NIH Public Access, *45*(5), 926–934. http://doi.org/10.1016/j.humpath.2014.01.005.The

Manuscript, A. (2014b). NIH Public Access, *5*(209), 1–28. http://doi.org/10.1126/scitranslmed.3006802.A

Manuscript, A., & Factors, R. (2009). NIH Public Access, *83*(5), 584–594.

Mao, L. (2002). Recent advances in the molecular diagnosis of lung cancer, 6960–6969. http://doi.org/10.1038/sj.onc.1205564

Massion, P. P., & Carbone, D. P. (2003). The molecular basis of lung cancer : molecular abnormalities and therapeutic implications, *15*, 1–15.

Merlo, A., Gabrielson, E., Askin, F., & Sidransky, D. (1994). Frequent loss of chromosome 9 in human primary non-small cell lung cancer. *Cancer Res*, *54*(3), 640–642. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8306323

Micke, P., Sofia, J., Mattsson, M., Djureinovic, D., Nodin, B., Jirström, K., … Brunnström, H. (2016). Department of Pathology. *Journal of Thoracic Oncology*. http://doi.org/10.1016/j.jtho.2016.01.020

Miozzo, M., Sozzi, G., Musso, K., Pilotti, S., Incarbone, M., Pastorino, U., & Pierotti, M. A. (1996). Microsatellite alterations in bronchial and sputum specimens of lung cancer patients. *Cancer Research*, *56*(10), 2285–2288.

Mohammad, B. Al, & Brennan, P. C. (2017). A review of lung cancer screening and the role of computer-aided detection. *Clinical Radiology*. http://doi.org/10.1016/j.crad.2017.01.002

Montezuma, D., Azevedo, R., & Lopes, P. (2013). A panel of four immunohistochemical

markers (CK7 , CK20 , TTF-1 , and p63) allows accurate diagnosis of primary and metastatic lung carcinoma on biopsy specimens, 749–754. http://doi.org/10.1007/s00428-013-1488-z

Mozzoni, P., Banda, I., Goldoni, M., Corradi, M., Tiseo, M., Acampa, O., … Mutti, A. (2013). Plasma and EBC microRNAs as early biomarkers of non-small-cell lung cancer. *Biomarkers*, *18*(8), 679–686. http://doi.org/10.3109/1354750X.2013.845610

Nacionais, E. S., Diagn, P., Do, T., & Do, C. (n.d.). Recomendações nacionais para diagnóstico e tratamento do cancro do pulmão 09.

Nacional, P. (2015). Doenças Oncológicas em Números – 2015.

Navon, R., Wang, H., Steinfeld, I., Tsalenko, A., Ben-Dor, A., & Yakhini, Z. (2009). Novel rank-based statistical methods reveal microRNAs with differential expression in multiple cancer types. *PLoS ONE*, *4*(11). http://doi.org/10.1371/journal.pone.0008003

Neville, E. M., Stewart, M. P., Swift, A., Liloglou, T., Ross, H., & Gosney, J. R. (1996). Allelotype of non-small cell lung cancer, 533–539.

Niederst, M. J., Sequist, L. V, Poirier, J. T., Mermel, C. H., Lockerman, E. L., Garcia, A. R., … Engelman, J. A. (2015). RB loss in resistant EGFR mutant lung adenocarcinomas that transform to small-cell lung cancer, 199–203. http://doi.org/10.1038/ncomms7377

Normanno, N., De Luca, A., Bianco, C., Strizzi, L., Mancino, M., Maiello, M. R., … Salomon, D. S. (2006). Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene*, *366*(1), 2–16. http://doi.org/10.1016/j.gene.2005.10.018

Oh, H. K., Tan, A. L. K., Das, K., Ooi, C. H., Deng, N. T., Tan, I. B. H., … Tan, P. (2011). Genomic loss of miR-486 regulates tumor progression and the OLFM4 antiapoptotic factor in gastric cancer. *Clinical Cancer Research*, *17*(9), 2657–2667. http://doi.org/10.1158/1078-0432.CCR-10-3152

Osada, H., & Takahashi, T. (2002). Genetic alterations of multiple tumor suppressors and oncogenes in the carcinogenesis and progression of lung cancer. *Oncogene*, *21*(48), 7421–7434. http://doi.org/10.1038/sj.onc.1205802

Ott, M., Gogvadze, V., Orrenius, S., & Zhivotovsky, B. (2007). Mitochondria, oxidative stress and cell death. *Apoptosis*, *12*(5), 913–922. http://doi.org/10.1007/s10495-007-0756-2

p40 (D Np63), a Lung Squamous Cell Marker , Can Also Be Used to Label Breast Myoepithelial Cells. (2014), *138*(May), 5858. http://doi.org/10.5858/arpa.2013-0414-LE

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., … Varmus, H. (2004). EGF 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*, *101*, 13306–13311. http://doi.org/10.1073/pnas.0405220101

Paradiso, A., Tommasi, S., Pinto, R., Carpagnano, G. E., & Foschino-Barbaro, M. P. (2008). Exhaled breath condensate is not suitable to detect EGFR somatic mutations. http://doi.org/10.1183/09031936.00083008

Penta, J., Johnson, F., Wachsman, J., & Copeland, W. (2001). Mitochondrial DNA in human malignancy. *Mutat Res.*, *488*(2), 119–33. http://doi.org/10.1016/S1383-5742(01)00053-9

Petras, M. L., Lefferts, J. A., Ward, B. P., Suriawinata, A. A., & Tsongalis, G. J. (2011). O RIGINAL A RTICLE KRAS Detection in Colonic Tumors by DNA Extraction From FTA Paper : The Molecular Touch-Prep, *20*(4), 189–193.

Pezzolesi, M. G., Platzer, P., Waite, K. A., & Eng, C. (2008). Differential Expression of PTEN-Targeting MicroRNAs miR-19a and miR-21 in Cowden Syndrome. *American Journal of Human Genetics*, *82*(5), 1141–1149. http://doi.org/10.1016/j.ajhg.2008.04.005

Pinkerton, M., Chinchilli, V., Banta, E., Craig, T., August, A., Bascom, R., … Ishmael, F. T. (2013). Differential expression of microRNAs in exhaled breath condensates of patients with asthma, patients with chronic obstructive pulmonary disease, and healthy adults. *Journal of Allergy and Clinical Immunology*, *132*(1), 217–219.e2. http://doi.org/10.1016/j.jaci.2013.03.006

Popat, S. (2014). Toward precision medicine with next-generation EGFR inhibitors in non-small-cell lung cancer, 285–295.

Porta, M., Crous-Bou, M., Wark, P. A., Vineis, P., Real, F. X., Malats, N., & Kampman, E. (2009). Cigarette smoking and K-ras mutations in pancreas, lung and colorectal adenocarcinomas: Etiopathogenic similarities, differences and paradoxes. *Mutation Research - Reviews in Mutation Research*. http://doi.org/10.1016/j.mrrev.2009.07.003

Rajendram, D. (2007). Long-term storage and safe retrieval of DNA from microorganisms for molecular analysis using FTA matrix cards Long-term storage and safe retrieval of DNA from microorganisms for molecular analysis using FTA matrix cards, (January). http://doi.org/10.1016/j.mimet.2006.05.010

Rekhtman, N., Tafe, L. J., Chaft, J. E., Wang, L., Arcila, M. E., Colanta, A., … Ladanyi, M.

(2012). Distinct profile of driver mutations and clinical features in immunomarker-defined subsets of pulmonary large-cell carcinoma, *26*(4), 511–522. http://doi.org/10.1038/modpathol.2012.195

Romero, P. V, Rodríguez, B., Martínez, S., Cañizares, R., Sepúlveda, D., & Manresa, F. (2006). Analysis of oxidative stress in exhaled breath condensate from patients with severe pulmonary infections. *Archivos de Bronconeumologia*, *42*(3), 113–119. http://doi.org/10.1016/S1579-2129(06)60128-6

Sateia, H. F., Choi, Y., Rosalyn, W., Peairs, K. S., Peairs, K. S., Cancer, L., & Oncology, S. (2017). Author ' s Accepted Manuscript Screening for Lung Cancer To appear in : Seminars in Oncology. *Seminars in Oncology*. http://doi.org/10.1053/j.seminoncol.2017.02.003

Sharma, S. V, Bell, D. W., Settleman, J. E., & Haber, D. A. (2007). Epidermal growth factor receptor mutations in lung cancer. *Nat. Rev. Cancer*, *7*(3), 169–181. http://doi.org/10.1038/nrc2088

Sharma, H., Singh, A., Sharma, C., Jain, S. K., & Singh, N. (2005). Mutations in the mitochondrial DNA D-loop region are frequent in cervical cancer. *Cancer Cell International*, *5*, 34. http://doi.org/10.1186/1475-2867-5-34

Sharma, S., Kelly, T. K., & Jones, P. A. (2009). Epigenetics in cancer. *Carcinogenesis*, *31*(1), 27–36. http://doi.org/10.1093/carcin/bgp220

Shen, J., Todd, N. W., Zhang, H., Yu, L., Lingxiao, X., Mei, Y., … Jiang, F. (2011). Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, *91*(4), 579–87. http://doi.org/10.1038/labinvest.2010.194

Shi, X., Duan, H., Liu, X., Zhou, L., & Liang, Z. (2016). Genetic alterations and protein expression in combined small cell lung cancers and small cell lung cancers arising from lung adenocarcinomas after therapy with tyrosine kinase inhibitors, *7*(23).

Silva, M., Pastorino, U., & Sverzellati, N. (2017). Lung cancer screening with low-dose CT in Europe : strength and weakness of diverse independent screening trials. *Clinical Radiology*. http://doi.org/10.1016/j.crad.2016.12.021

Simon, G. R., Sharma, S., Cantor, A., Smith, P., & Bepler, G. (2005). ERCC1 expression is a predictor of survival in resected patients with non-small cell lung cancer. *Chest*, *127*(3), 978–983. http://doi.org/10.1378/chest.127.3.978

Sinha, A., Yadav, A. K., Chakraborty, S., Kabra, S. K., Lodha, R., Kumar, M., … Agrawal, A. (2013). Exosome-enclosed microRNAs in exhaled breath hold potential for biomarker discovery in patients with pulmonary diseases. *Journal of Allergy and Clinical Immunology*, *132*(1). http://doi.org/10.1016/j.jaci.2013.03.035

Skaug, V., Ryberg, D., Arab, E. H. K. M. O., Stangeland, L., Myking, A. O., & Haugen, A. (2000). p53 Mutations in Defined Structural and Functional Domains Are Related to Poor Clinical Outcome in Non-Small Cell Lung Cancer Patients. *Clinical Cancer Research*, *6*(3), 1031–1037. Retrieved from http://clincancerres.aacrjournals.org/content/6/3/1031.abstract

Smith and Williams' Introduction to the Principles of Drug Design and Action, Fourth Edition. (2017). Retrieved from https://books.google.pt/books/about/Smith_and_Williams_Introduction_to_the_P.html?hl=pt-PT&id=P2W6B9FQRKsC

Sordella, R., Bell, D. W., Haber, D. A., & Settleman, J. (2004). Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science*, *305*(5687), 1163–1167. http://doi.org/10.1126/science.1101637

Studies on Experimental Models | Samar Basu | Springer. (2017). Retrieved from http://www.springer.com/us/book/9781607619550

Sung, H., & Cho, J. (n.d.). Biomarkers for the lung cancer diagnosis and their advances in proteomics, 615–625.

Taanman, J.-W. (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1410*, 103–123. http://doi.org/10.1016/S0005-2728(98)00161-3

Teh, E. (2014). Lung cancer : diagnosis , staging and treatment. *Surgery*, *32*(5), 242–248. http://doi.org/10.1016/j.mpsur.2014.02.015

Thigpen, S. C., & Geraci, S. A. (2016). Author ' s Accepted Manuscript. *The American Journal of the Medical Sciences*. http://doi.org/10.1016/j.amjms.2016.06.001

Toxicology of the Lung, Fourth Edition. (2017). Retrieved from https://www.crcpress.com/Toxicology-of-the-Lung-Fourth-Edition/Gardner/p/book/9780849328350

Travis, W. D. (2012). Update on small cell carcinoma and its differentiation from squamous cell

carcinoma and other non-small cell carcinomas. *Modern Pathology*, *25*(S1), S18–S30. http://doi.org/10.1038/modpathol.2011.150

Villalobos, P., & Wistuba, I. I. (2017). Lung Cancer Biomarkers Lung cancer Genotyping Biomarkers Molecular targets. *Hematology/Oncology Clinics of NA*, *31*(1), 13–29. http://doi.org/10.1016/j.hoc.2016.08.006

Warth, A., Muley, T., Herpel, E., Meister, M., Herth, F. J. F., Schirmacher, P., … Schnabel, P. A. (2012). Large-scale comparative analyses of immunomarkers for diagnostic subtyping of non-small-cell lung cancer biopsies, 1–9. http://doi.org/10.1111/j.1365-2559.2012.04308.x

Weeden, C. E., & Solomon, B. (2015). FGFR1 inhibition in lung squamous cell carcinoma : questions and controversies. *Nature Publishing Group*, (September), 1–5. http://doi.org/10.1038/cddiscovery.2015.49

Weynants, P., Humbler, Y., Canon, J. L., & Symann, M. (1990). Biology of small cell lung cancer : an overview.

WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Fourth edition - WHO - OMS -. (2015). Retrieved from http://apps.who.int/bookorders/anglais/detart1.jsp?codlan=1&codcol=70&codcch=4007

Xiao, P., Chen, J., Zhou, F., Lu, C., Yang, Q., Tao, G., … Gene, P. (2014). Lung Cancer Methylation of P16 in exhaled breath condensate for diagnosis of non-small cell lung cancer. *Lung Cancer*, *83*(1), 56–60. http://doi.org/10.1016/j.lungcan.2013.09.008

Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., … Harris, C. C. (2006). Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell*, *9*(3), 189–198. http://doi.org/10.1016/j.ccr.2006.01.025

Yang Ai, S. S., Hsu, K., Herbert, C., Cheng, Z., Hunt, J., Lewis, C. R., & Thomas, P. S. (2013). Mitochondrial DNA mutations in exhaled breath condensate of patients with lung cancer. *Respiratory Medicine*, *107*(6), 911–918. http://doi.org/10.1016/j.rmed.2013.02.007

Yong, J., Jeon, H., Ha, S., In, T., Son, J., Ho, C., … San, I. (2000). Microsatellite alteration in histologically normal lung tissue of patients with non-small cell lung cancer, *30*, 83–89.

Youssef, O., Sarhadi, V. K., Armengol, G., Piirilä, P., Knuuttila, A., & Knuutila, S. (2016). Exhaled Breath Condensate as a Source of Biomarkers for Lung Carcinomas. A Focus on Genetic and Epigenetic Markers - A Mini-Review. *Genes, Chromosomes & Cancer*.

http://doi.org/10.1002/gcc.22399

Yu, B., Swatkoski, S., Holly, A., Lee, L. C., Giroux, V., Lee, C.-S., … Luo, J. (2015). Oncogenesis driven by the Ras/Raf pathway requires the SUMO E2 ligase Ubc9. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(14), E1724–33. http://doi.org/10.1073/pnas.1415569112

Zhang, D., Takigawa, N., Ochi, N., Tanimoto, Y., Noujima, D., Chen, Y. Y., … Kiura, K. (2011). Detection of the EGFR mutation in exhaled breath condensate from a heavy smoker with squamous cell carcinoma of the lung. *Lung Cancer (Amsterdam, Netherlands)*, *73*(3), 379–80. http://doi.org/10.1016/j.lungcan.2011.05.018

Zheng, M. (2016). Classification and Pathology of Lung Cancer. *Surgical Oncology Clinics of NA*, *25*(3), 447–468. http://doi.org/10.1016/j.soc.2016.02.003

Zhou, J., Nong, L., Wloch, M., Cantor, A., Mulshine, J. L., & Tockman, M. S. (2001). Expression of early lung cancer detection marker: hnRNP-A2/B1 and its relation to microsatellite alteration in non-small cell lung cancer. *Lung Cancer (Amsterdam, Netherlands)*, *34*(3), 341–350. http://doi.org/10.1016/S0169-5002(01)00254-9

Zhou, M., Liu, Y., & Duan, Y. (2012). Breath biomarkers in diagnosis of pulmonary diseases. *Clinica Chimica Acta*, *413*(21-23), 1770–1780. http://doi.org/10.1016/j.cca.2012.07.006

Zhu, S., Si, M. L., Wu, H., & Mo, Y. Y. (2007). MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). *Journal of Biological Chemistry*, *282*(19), 14328–14336. http://doi.org/10.1074/jbc.M611393200

Zilfou, J. T., & Lowe, S. W. (2009). Tumor suppressive functions of p53. *Cold Spring Harbor Perspectives in Biology*, *1*(5), 1–12. http://doi.org/10.1101/cshperspect.a001883