Ana Barbosa de Sousa Nogal

Proteomic characterization of lung cancer and chronic obstructive pulmonary disease: a bronchoalveolar lavage fluid analysis

Tese de Candidatura ao grau de Doutor em Ciências Biomédicas submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – Doutor Luis Paz-Ares

Categoria – Director do Serviço de Oncologia Médica e Professor Titular de Medicina.

Afiliação – Serviço de Oncologia Médica (Hospital Universitário Virgen del Rocío - HUVR, Sevilha). Instituto de Biomedicina de Sevilha (HUVR, Conselho Superior de Investigação Científica - CSIC, Universidade de Sevilha)

Co-orientador – Doutora Maria Dolores Pastor

Categoria – Investigadora pós-doutoral

Afiliação – Instituto de Biomedicina de Sevilha (HUVR, Conselho Superior de Investigação Científica - CSIC, Universidade de Sevilha)

Co-orientador – Doutor Rui Manuel Medeiros Silva

Categoria – Professor associado com agregação

Afiliação – Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto. Instituto Português de Oncologia do Porto FG, EPE.

Depois da tempestade…

TABLE OF CONTENTS

ACKNOWLEDGMENTS

Ao Dr. Luis Paz-Ares, meu orientador, por me ter permitido cumprir um sonho de juventude, pela oportunidade de fazer parte de uma nova realidade, de um novo grupo de investigação, pelos braços abertos com que me recebeu, pela sua capacidade de liderança que inspira todos a alcançar sempre algo melhor. Obrigada por toda a sua ajuda, simpatia e disponibilidade.

À Dra. Maria Dolores Pastor, minha co-orientadora, não existem linhas suficientes para agradecer tudo o que fez por mim: muito mais do que imagina ter feito. Muito obrigada pela compreensão, pela paciência com o meu mau espanhol (de início), pela amizade e simpatia extremas, pelo carinho nos momentos mais complicados e pela rapidez com que os resolvia, pela sua determinação e diligência e sobretudo obrigada pelo exemplo, pela inspiração e pela vontade de ser algum dia algo parecido com ela.

Ao Dr. Rui Medeiros, meu co-orientador, que me permitiu entrar no campo da investigação quando era apenas uma "jovem" em 2005, muito obrigada pelo apoio e amizade demonstrados ao longo destes anos, pelas oportunidades e portas que me abriu, pela disponibilidade que demonstra sempre que lhe entro no gabinete por qualquer motivo.

À Fundação para a Ciência e Tecnologia, que me concedeu a bolsa de doutoramento e me permitiu fazer, durante 4 anos, aquilo que mais gosto.

Ao fabuloso grupo de "Oncología molecular y nuevas terapias", o meu gigante agradecimento. Obrigada Sonia, Rocío, Ana, Ricardo e demais colegas do HUVR e IBIS pelo companheirismo, pela força que me deram, pela compreensão e paciência quando me explicavam as coisas muito lentamente, pela momentos divertidos, pelas "sevillanas maneras" que me foram ensinando, pelo trabalho que desenvolvem diariamente e do qual sinto muito orgulho e saudade. Obrigada sobretudo por me fazerem sentir em casa e por me fazerem esquecer durante a semana que estava longe...

À Ana Coelho, Dr. Araújo e Raquel Catarino, pela orientação que me dão nestas andanças, mesmo sem achar que o estão a fazer. Obrigada pelo apoio, amizade, por fazerem crescer como pessoa e como investigadora.

Às minhas companheiras de luta, Ana Luísa, Mónica e Carina, pelo apoio que me deram nos momentos mais complicados desta jornada, pela motivação, pela amizade, pelo carinho, pelos momentos de galhofa que nos permitiram manter uma réstia de sanidade mental.

Aos meus amigos, pela força que me transmitiram, pelos pequenos grandes gestos de apoio demonstrados numa simples mensagem, email ou telefonema com notícias ou conversas simples que me iam mantendo motivada, concentrada e perto de casa mesmo a 700 km de distância... O meu sincero obrigada por me manterem de pé.

Aos meus pais, irmã, cunhado, sobrinhos, avós, os meus pilares, os meus exemplos a seguir, pelo apoio incondicional, pela paciência, pelos sorrisos e gargalhadas, por estarem sempre comigo e por me fazerem querer sempre ser melhor... Não teria chegado onde cheguei sem o vosso toque de magia...

A ti, não sei como agradecer... Tento todos os dias fazê-lo e retribuir tudo o que me dás... Agradeço por te ter todos os dias ao meu lado, a apoiar-me nas minhas pequenas loucuras, a percorrer os caminhos pelos quais às vezes nos meto, a usufruir dos momentos espetaculares pelos quais passamos e ainda vamos passar, a fazer-me crescer e a tornar-me uma pessoa mil vezes melhor... Obrigada pelo amor, carinho, amizade, paciência, por seres o meu porto de abrigo... Obrigada por tudo!

LIST OF ABBREVIATIONS

2D-DIGE	Two-dimensional difference gel electrophoresis
2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
ADC	Adenocarcinoma
AIF	Apoptosis-inducing factor
AJCC	American Joint Committee on Cancer
AKR1B10	Aldo-keto reductase family 1 member B10
AKR1C3	Aldo-keto reductase family 1 member C3
Akt	Protein Kinase B
ALDH3A1	Aldehyde dehydrogenase
ALDOA	Aldolase A
ALK	Anaplastic lymphoma kinase
AMY	Alpha amylase
ANXA	Annexin
AP-1	Activator protein 1
AR	Androgen receptor
ARHGDIB	Rho GDP dissociation inhibitor 2
AUC	Area under the curve
B-NGF	Beta nerve growth factor
BALF	Bronchoalveolar lavage fluid
BaP	Benzo[a]pyrene
BCA	Bicinchoninic acid assay
bFGF	Basic fibroblast growth factor
BLC	B-cell lymphoma
BMP	Bone morphogenetic protein
BNDF	Brain-derived neurotrophic factor
BSA	Bovine serum albumin
C3	Complement C3
CA1	Carbonic anhydrase 1
CAPS2	Calsyphosin 2
CAT	Catalase
CCL1	Chemokine (C-C motif) ligand 1
CCR8	Chemokine (C-C motif) receptor 8
CEA	Carcinoembryonic antigen

CFL1	Cofilin-1
CI	Confidence interval
COPD	Chronic obstructive pulmonary disease
CRP	C reactive protein
CT	Computerized tomography
CTSD	Cathepsin D pro-protein
CYFRA21-1	Cytokeratin 19 fragment antigen 21-1
DTT	Dithiothreitol
EG-VEGF	Endocrine gland-derived vascular endothelial growth factor
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ELISA	Enzyme linked immunosorbent assay
EML4	Echinoderm microtubule-associated protein like 4
EMT	Epithelial to mesenchymal transition
ENO1	Alpha enolase
ERK1/2	Extracellular-signal-regulated kinases 1/2
ESI	Electrospray ionization
EZR	Ezrin
FBP1	Fructose-1,6-biphosphatase
FEV1	Forced expiratory volume in 1 second
FGF	Fibroblast growth factor
FVC	Forced vital capacity
G-CSF	Granulocyte colony-stimulating factor
GDF15	Growth differentiation factor 15
GDNF	Glial cell-derived neurotrophic factor
GH	Growth hormone
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GOLD	Global initiative for chronic obstructive pulmonary disease
GSR	Glutathione reductase
GSTA	Glutathione S-transferase alpha
GSTP1	Glutathione S-transferase pi 1
HB-EGF	Heparin-binding EGF-like growth factor
HER2	Human epidermal growth factor receptor 2
HGF	Hepatocyte growth factor
HIF	Hypoxia-inducible factor
HSP70	Heat shock protein 70

ICAM-1	Inter-Cellular Adhesion Molecule 1
ICAT	Isotope-coded affinity tags
IDH1	Isocitrate dehydrogenase 1
IEF	Isoelectric focusing
IFN γ	Interferon-gamma
IGF-1	Insulin-like growth factor 1
IGF-1R	Insulin-like factor-1 receptor
IGFBP	Insulin-like growth factor-binding protein
IHC	Immunohistochemistry
IL	Interleukin
IP	Isoelectric point
IPA	Ingenuity pathway analysis
iTRAQ	Isobaric tags for relative and absolute quantitation
JAK	Janus kinase
JNK	c-Jun N-terminal kinases
LC	Lung cancer
LCC	Large cell carcinoma
LCN2	Lipocalin 2
LPS	Lipopolysaccharides
m/z	Mass to charge ratio
MALDI	Matrix-assisted laser desorption ionization
MAPK	Mitogen-activated protein kinase
MCFR	Macrophage chemotactic factor receptor
MCP-1/CCL2	Monocyte chemotactic protein-1/Chemokine (C-C motif) ligand 2
MCSF	Mouse stem cell factor
MET	Hepatocyte growth factor receptor
MIG/CXCL9	Monokine induced by IFN-Gamma/Chemokine (C-X-C motif) ligand 9
MIP1α/CCL3	Macrophage inflammatory protein 1 α /Chemokine (C-C motif) ligand 3
MIP1β/CCL4	Macrophage inflammatory protein 1 β /Chemokine (C-C motif) ligand 4
MIP-1δ	Macrophage inflammatory protein 1 delta
MMP	Metalloproteinases
MS	Mass spectrometry
mTOR	Mammalian target of rapamycin
MW	Molecular weight
NDMA	N-nitrosodimethylamine
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells

NGFR	Nerve growth factor receptor
NNK	Nicotine-derived nitrosamine ketone
NNN	N'-nitrosonornicotine
NPV	Negative predictive value
NSCLC	Non-small cell lung cancer
NSE	Neuron specific enolase
NT	Neurotrophin
OPG	Osteoprotegerin
OR	Odds ratio
PBS	Phosphate buffered saline
PDGF	Platelet-derived growth factor
PEBP4	Phosphatidylethanolamine-binding protein 4
PI3K	Phosphatidylinositide 3-kinase
PIGF	Phosphatidylinositol-glycan biosynthesis class F
PKM2	Pyruvate kinase 2
PPIA	Peptidylprolyl isomerase A
PPV	Positive predictive value
PRDX	Peroxiredoxin
PVDF	Polyvinylidene fluoride
PYGM	Glycogen phosphorylase
Rantes/CCL5	Regulated upon Activation, Normal T-cell Expressed, and Secreted/ Chemokine (C-C motif) ligand 5
ROC	Receiver operating characteristic
ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase polymerase chain reaction
SCC	Squamous cell carcinoma
SCCA	Squamous cell carcinoma antigen
SCF	Stem cell factor
SCFR	Stem cell factor receptor
SCLC	Small cell lung cancer
SD	Standard deviation
SDS	Sodium docedyl sulfate
SDS-PAGE	Sodium docedyl sulfate polyacrylamide gel electrophoresis
SELDI	Surface-enhanced laser desorption/ionization
SELENBP1	Selenium-binding protein 1
SERPINB1	Serpin peptidase inhibitor, clade B, member 1

SILAC	Stable isotope labelling with amino acids in cell culture
STAT	Signal transducer and activator of transcription
TBS	Tris buffered saline
TBST	Tris buffered saline tween
TGF	Transforming growth factor
TIMP	Tissue inhibitor of metalloproteinases
TKT	Transketolase
TNF	Tumour necrosis factor
TNFR	Tumour necrosis factor receptor
TOF	Time of flight
TPPP3	Tubulin polymerization-promoting protein family member 3
TXN	Thioredoxin
UCHL1	Ubiquitin carboxyl-terminal hydrolase isozyme L1
UICC	Union for International Cancer Control
UPGMA	Unweighted pair group method with arithmetic mean
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
WHO	World Health Organization

LIST OF PUBLICATIONS

According to article 31 of the Decree-Law nº230/2009, the present Thesis has already produced the following publications in peer-reviewed journals in addition to a book chapter:

Nogal A, Pastor MD, Molina-Pinelo S, Carnero A, Paz-Ares L. Proteomic biomarkers in lung cancer. Clinical & translational oncology: official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2013;15(9):671-82 (M. D. Pastor and A. Nogal contributed equally to the present work).

Nogal A, Pastor MD, Molina-Pinelo S, Melendez R, Romero-Romero B, Mediano MD, et al. Identification of oxidative stress related proteins as biomarkers for lung cancer and chronic obstructive pulmonary disease in bronchoalveolar lavage. International journal of molecular sciences. 2013;14(2):3440-55 (The first two authors contributed equally to this work).

Nogal A, Pastor MD, Molina-Pinelo S, Melendez R, Salinas A, Gonzalez De la Pena M, et al. Identification of proteomic signatures associated with lung cancer and COPD. Journal of proteomics. 2013;89:227-37 (The first two authors contributed equally to this work).

Ma Dolores Pastor, **Ana Nogal**, Sonia Molina-Pinelo, Luis Paz-Ares and Amancio Carnero (2013). Oncoproteomic Approaches in Lung Cancer Research, Oncogenomics and Cancer Proteomics - Novel Approaches in Biomarkers Discovery and Therapeutic Targets in Cancer, Dr. Cesar Lopez (Ed.), ISBN: 978-953-51-1041-5, InTech, DOI: 10.5772/53873. Available from: http://www.intechopen.com/books/oncogenomics-and-cancer-proteomics-novel-approaches-in-biomarkers-discovery-and-therapeutic-targets-in-cancer/oncoproteomic-approaches-in-lung-cancer-research

I hereby declare that I have actively participated in the gathering and study of the material included in each of the publications presented and have written the manuscripts in collaboration with the other authors.

ABSTRACT

ABSTRACT

Cancer is a major worldwide public health issue and continues to be a leading cause of death in developed countries. On the top of incidence and mortality rates lies lung cancer (LC). This disease, caused mainly by tobacco smoking, has one of the lowest 5-year survival rates among all types of cancers: only 15%. Despite the recent advances in systemic therapies, such as chemotherapy and targeted therapies, prognosis remains strongly linked to early diagnosis. Unfortunately, most patients are diagnosed in advanced stages, when curative options are limited. Consequently, there is a pressing need for detection methods that may allow the diagnosis of LC early in its natural history, especially among high-risk individuals such as smokers, particularly if chronic obstructive pulmonary disease (COPD) is also present. The primary risk factor for COPD is, as for LC, tobacco smoking. However, this inflammatory disease is an independent risk factor for the development of LC and some pathogenic processes are thought to be important in the development of both diseases. A better understanding of the molecular and cellular mechanisms by which chronic inflammatory diseases, such as COPD, lead to the initiation of LC, may identify early diagnostic biomarkers and screening tests, valuable targets, and more effective therapeutic strategies.

The search for early detection markers has gained an ally in the last several years. Proteomic technologies have begun to uncover the molecular complexity of lung tumours by allowing a rapid and complete analysis of the proteins that are expressed in the context of this disease. The analysis of lung samples from LC patients and high-risk patients by proteomic methodologies might allow us to gain more knowledge on the mechanisms of both diseases and possibly identify LC diagnostic biomarkers that could help in the clinical practice. The major aim of the present work was to study the protein profile of bronchoalveolar lavage fluid (BALF) in patients with LC and/or COPD using proteomic methodologies in an attempt to identify new LC biomarkers among high-risk patients.

Patients who had required flexible bronchoscopy for diagnostic purposes at the Hospital Universitario Virgen del Rocío (Seville, Spain) were recruited for this study. BALF samples were treated and submitted to different proteomic technologies such as two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), mass spectrometry (MS), western blot, antibody arrays, and enzyme-linked immunosorbent assay (ELISA).

The use of a 2D-PAGE MALDI-TOF/TOF MS methodology allowed the detection of 40 differentially expressed proteins between the control group (without LC or COPD) and the LC, COPD, and LC&COPD groups. Ingenuity pathway analysis (IPA) of these proteins revealed three top molecular mechanisms occurring in both LC and COPD: glycolysis and gluconeogenesis, free radical scavenging and oxidative stress response,

and inflammation. The major connector of all three pathways was the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), a key transcription factor in the inflammatory process and carcinogenesis. The importance of inflammation in LC and COPD was further evaluated by application of BALF samples to cytokine and growth factor antibody arrays. This analysis revealed IL-11 and CCL1 as markers for adenocarcinoma of the lung with areas under the curve (AUC) of 93 and 83%, respectively. Sensitivity and specificity values were 90 and 88% when evaluating IL-11 and 80 and 74% when considering CCL1.

In conclusion, the present work has revealed the proteomic differences and similarities between LC and COPD. In addition, the proteomic methodology used has allowed the identification of two possible adenocarcinoma biomarkers that could be used to detect this type of LC. The application of a simple ELISA assay could be used to improve the diagnosis of lung adenocarcinoma in smokers, despite the presence or absence of COPD. The use of this assay in routine clinical practice, if validated in further studies, could in the future improve the early detection of lung adenocarcinoma, currently the most common form of NSCLC, and hopefully contribute to an increase of the 5-year survival rate associated with this disease.

RESUMO

RESUMO

O cancro é um problema de saúde pública em todo o mundo e continua a ser uma das principais causas de morte nos países desenvolvidos. No topo das taxas de incidência e mortalidade encontra-se o cancro do pulmão. Esta neoplasia, causada principalmente pelo fumo do tabaco, tem uma das menores taxas de sobrevivência a 5 anos entre todos os tipos de cancro: apenas 15%. Apesar dos recentes avanços nos tratamentos sistémicos como a quimioterapia e medicamentos direcionados, o prognóstico permanece fortemente ligado ao diagnóstico precoce. Infelizmente, a maioria dos pacientes é diagnosticada em estadios avançados da doença, onde as opções terapêuticas são limitadas. Consequentemente, há uma necessidade premente de encontrar métodos de detecção que permitiam o diagnóstico do cancro do pulmão no início de sua história natural, especialmente entre os indivíduos de alto risco, como fumadores e pessoas com doença pulmonar obstrutiva crónica. O principal fator de risco para a doença pulmonar obstrutiva crónica é, tal como para o cancro do pulmão, o fumo do tabaco. No entanto, esta doença inflamatória é um factor de risco independente para o desenvolvimento de cancro do pulmão e pensa-se que alguns processos patogénicos possam ser importantes para o desenvolvimento de ambas as doenças. Uma melhor compreensão dos mecanismos moleculares e celulares pelos quais doenças inflamatórias crónicas, como a doença pulmonar obstrutiva crónica, levam ao desenvolvimento de cancro do pulmão, pode identificar biomarcadores de diagnóstico precoce, alvos terapêuticos, e estratégias de combate à doença mais eficazes.

A procura de marcadores de detecção precoce ganhou um forte aliado nos últimos anos. O uso de tecnologias proteómicas permitiu começar a descobrir a complexidade molecular dos tumores pulmonares, já que permitem uma análise rápida e completa das proteínas que são expressas no contexto desta doença. A análise de amostras de pacientes com cancro de pulmão e pacientes de alto risco através de técnicas proteómicas pode permitir um maior conhecimento sobre os mecanismos de ambas as doenças e possivelmente identificar biomarcadores de diagnóstico de cancro do pulmão que possam auxiliar na prática clínica.

O principal objetivo do presente trabalho foi estudar o perfil proteómico de lavado bronco-alveolar em pacientes com cancro do pulmão e/ou doença pulmonar obstrutiva crónica através de métodos proteómicos, na tentativa de identificar novos biomarcadores diagnósticos de cancro do pulmão em pacientes de alto risco.

Para o estudo foram recrutados pacientes aos quais tinham sido efectuadas broncoscopias para fins de diagnóstico no Hospital Universitário Virgen del Rocío (Sevilha, Espanha). Amostras de lavado bronco-alveolar foram tratadas e submetidas a

diferentes técnicas proteómicas, como a electroforese bidimensional em gel de poliacrilamida (2D-PAGE), espectrometria de massa (MS), western blot, arrays de anticorpos e ELISA.

A utilização de uma metodologia 2D-PAGE MALDI-TOF/TOF MS permitiu a detecção de 40 proteínas com diferenças de expressão entre o grupo controlo (sem cancro do pulmão ou doença pulmonar obstrutiva crónica) e os grupos de pacientes com cancro de pulmão, doença pulmonar obstrutiva crónica e ambas as doenças. A análise através do programa IPA destas proteínas revelou três mecanismos moleculares importantes a ocorrer em ambas as doenças: glicólise e gluconeogénese, resposta a stress oxidativo e radicais livres e inflamação. A principal ligação entre as três vias foi o NF-kB, um fator de transcrição chave no processo inflamatório e na carcinogénese. A importância da inflamação no desenvolvimento de cancro do pulmão e doença pulmonar obstrutiva crónica foi avaliada mais em profundidade pela aplicação de amostras de lavado bronco-alveolar a arrays de citoquinas e factores de crescimento. Esta análise revelou as proteínas IL-11 e CCL1 como marcadores para adenocarcinoma de pulmão com AUC de 93 e 83%, respectivamente. Os valores de sensibilidade e especificidade foram de 90 e 88% quando se avaliaram os níveis de IL-11 e 80 e 74% quando se consideraram os níveis de CCL1 .

O presente estudo proteómico revelou as diferenças e semelhanças entre o cancro do pulmão e a doença pulmonar obstrutiva crónica. Além disso, a metodologia utilizada permitiu a identificação de dois possíveis biomarcadores de adenocarcinoma que poderiam ser utilizados para detectar este tipo de cancro. O uso de um ensaio simples como o de ELISA pode melhorar o diagnóstico de adenocarcinoma do pulmão em fumadores, tendo estes doença pulmonar obstrutiva crónica ou não. Em conclusão, a futura utilização deste ensaio na rotina clínica poderia melhorar a detecção precoce de adenocarcinoma de pulmão, atualmente a forma mais comum cancro do pulmão e eventualmente contribuir para um aumento na taxa de sobrevivência a 5 anos associada a esta doença.

INTRODUCTION

1. INTRODUCTION

1.1. Lung cancer

1.1.1. Epidemiology and causes

The World Health Organization (WHO) declared cancer the leading cause of death in developed countries (1). According to the latest estimates from this organization, approximately 12.7 million people were diagnosed with cancer in 2008 (Figure 1) (2). Global cancer deaths are projected to increase from 7.1 million in 2002 to 11.5 million in 2030 (3)

Globally, lung cancer (LC) is on the top of the incidence and mortality rates around the world. Approximately 1.6 million people were diagnosed with lung cancer in 2008 and 55% of new cases occurred in developing countries (4). In the same year, lung cancer was the leading cause of cancer related death in men and the second in women, with approximately 1.4 million reported deaths. The majority of cases and deaths occurred in men, although female rates have been rising along the years (4). Projections for 2030 state that 3.1% of all deaths will be caused by LC, making it the 6th most common cause of death worldwide (3).

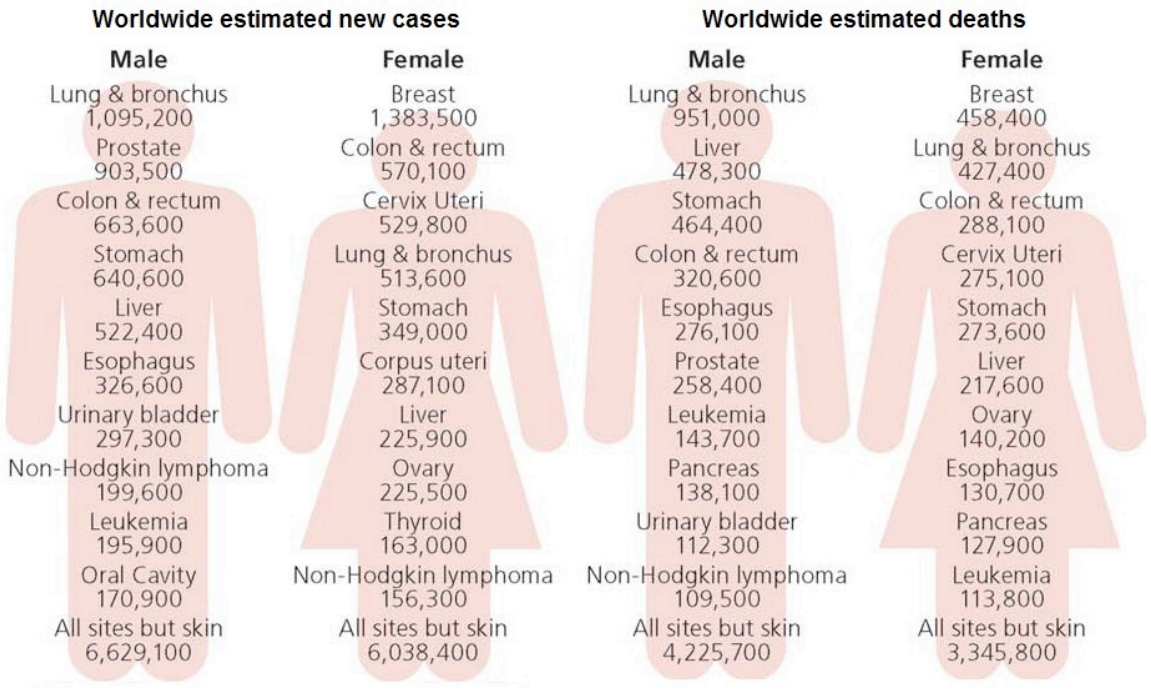


Figure 1: Global estimates for new cancer cases and deaths in males and female, from Jemal et al. (2011).

Lung cancer was a rare disease in the beginning of the 20th century but the introduction of manufactured cigarettes with addictive properties caused a steep rise in cases beyond that period of time, revealing that the most important risk factor for the development of LC was tobacco smoking (5). This relationship was clearly established by the works of Doll and Mills in 1950 and later confirmed by the US Surgeon General in 1964 (5-7). Epidemiological evidence demonstrates that two to three decades after a peak in smoking prevalence in a given population, there is a peak in lung cancer related deaths (8). Tobacco smoking causes approximately 85-90% of all lung cancer cases worldwide and smokers present a 15 to 30 fold increased risk to develop lung cancer compared to non-smokers (9-11).

Tobacco smoke targets the central and peripheral compartments of the lung and chronically presents over 4800 different compounds, of which 66 are known carcinogens, to the airway epithelial cells. Substances like polycyclic aromatic hydrocarbons, tobacco specific nitrosamines, and nitric oxide, among many others, can initiate, promote and amplify oxidative DNA damage, and create DNA adducts (12). Cells are capable of removing and correcting the injuries caused by tobacco smoke components such as benzo[a]pyrene (BaP), NNK (Nicotine-derived nitrosamine ketone), NDMA (N-nitrosodimethylamine), NNN (N'-nitrosonornicotine), ethylene oxide, and 4-aminobiphenyl, but the sustained exposure to these carcinogens can lead to the development of permanent molecular alterations and ultimately to LC and other airway diseases. The duration of smoking is the strongest determinant of risk to develop LC and increases proportionally with the number of cigarettes smoked (12). Although it is the number one cause of LC worldwide, only approximately 15% of smokers will eventually develop LC (13).

Cancer is a complex multistep disease that develops over time, progressively accumulating genetic alterations that transform normal cells into malignant cells. The ability to escape apoptosis, the insensitivity to growth-inhibitory signals, limitless replicative potential, self-sufficiency in growth signals, sustained angiogenesis, and the ability to invade surrounding tissues and generate metastasis, are considered hallmarks of cancer development and are acquired by cells through a variety of mechanisms (Figure 2). Evidence that cancer cells have deregulated cellular energy pathways and mechanisms to avoid their destruction by the immune system has been accumulating over the last years and are now also considered emerging hallmarks. Genomic instability and mutations and also inflammation are now considered to be enabling characteristics that drive the carcinogenic process (14). The acquisition of these abilities and the interactions between the altered cells and the surrounding microenvironment are relevant to all stages of cancer development, such as cancer initiation, promotion, and progression.

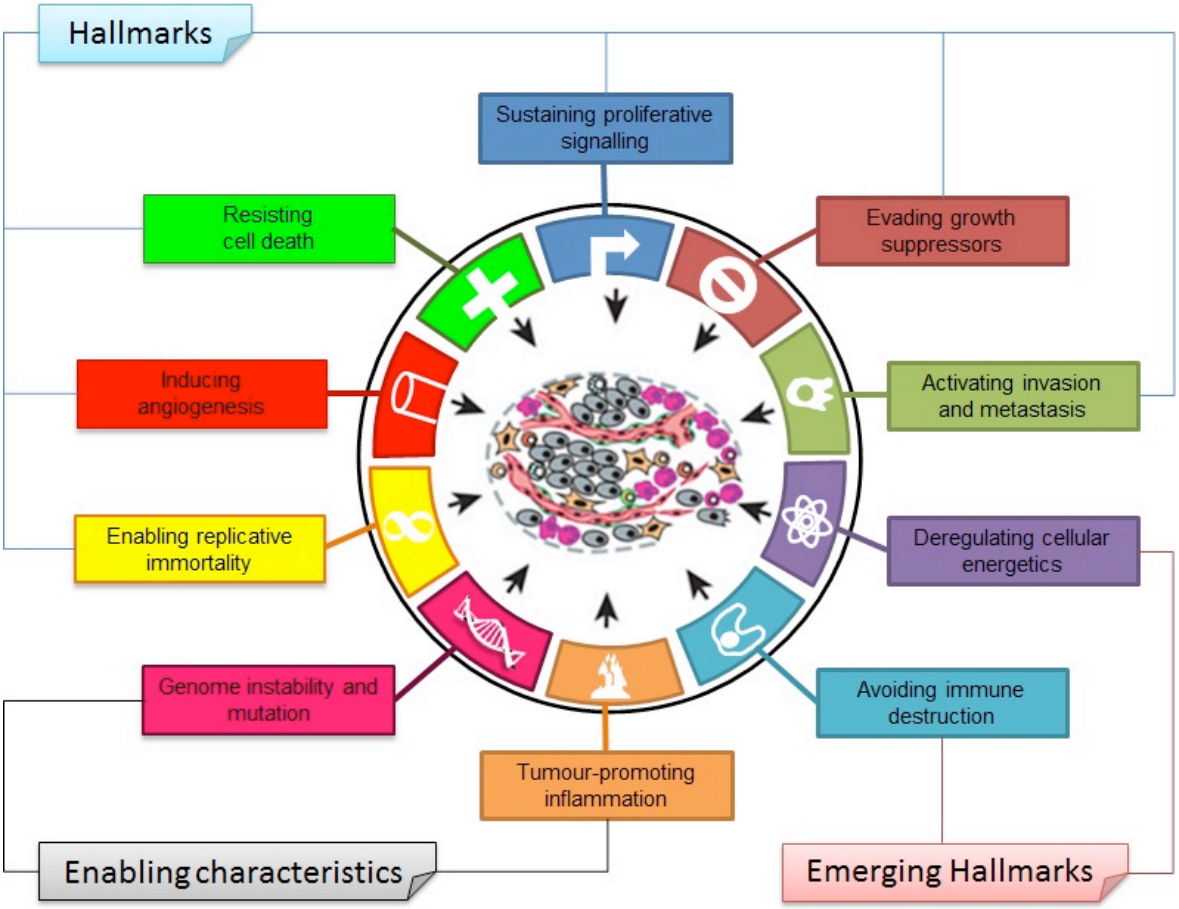


Figure 2: Hallmark capabilities and enabling characteristics of cancer, adapted from Hanahan and Weinberg (2010).

1.1.2. Classification and staging

The classification of lung cancer is based on histological characteristics identified by light microscopy. The most important and most relevant division is made between small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Between 15-25% of all lung cancer cases are classified as SCLC (15, 16). This histological subtype is characterized by small, oval shaped cells with limited cytoplasm and is considered to be a very aggressive carcinoma, clinically and biologically different from NSCLC (Figure 3A) (17, 18). It has a strong connection to smoking given that 98% of all SCLC cases are caused by tobacco smoke (19).

The NSCLC category is a heterogeneous group that represents 75-85% of all diagnosed cases of lung cancer. This group is divided into three major histological types: adenocarcinoma, squamous cell carcinoma (SCC), and large cell carcinoma (LCC). In the last 20 years, the adenocarcinoma incidence has been rising making it the most predominant histological type in many countries, representing up to 50% of all NSCLC

cases. Adenocarcinomas are currently more often diagnosed in women, young males (<50 years old), non-smokers and Asians (20). This histological type is defined as a malignant epithelial tumour with glandular differentiation or mucin production and tends to arise in the periphery of the lungs and the alveoli (Figure 3B). There are several subtypes of adenocarcinoma, such as acinar, papillary, bronchioloalveolar, solid with mucin production, and mixed. However, most adenocarcinomas are histologically heterogeneous and consequently classified as mixed (20-22).

Malignant epithelial tumours that show keratinisation and/or intercellular bridges are defined as SCC (23). Also known as epidermoid carcinomas, this group of lung tumours was the most frequent but now represents approximately 30% of all diagnosis (16). They are more common in men and smokers and the majority occur as slow growing tumours in the bronchial epithelium of the central airways (Figure 3C) (21). All NSCLCs that lack the cytological and architectural characteristics of the adenocarcinomas or SCCs are classified as LCCs (24). They represent approximately 10% of all lung cancer cases and their incidence is decreasing most likely due to improved diagnostic techniques (Figure 3D) (25). Although the accurate classification of a lung tumour is extremely important step in the evaluation of a patient, the correct determination of extent of the patients' disease is also necessary to determine the course of treatment and prognosis.

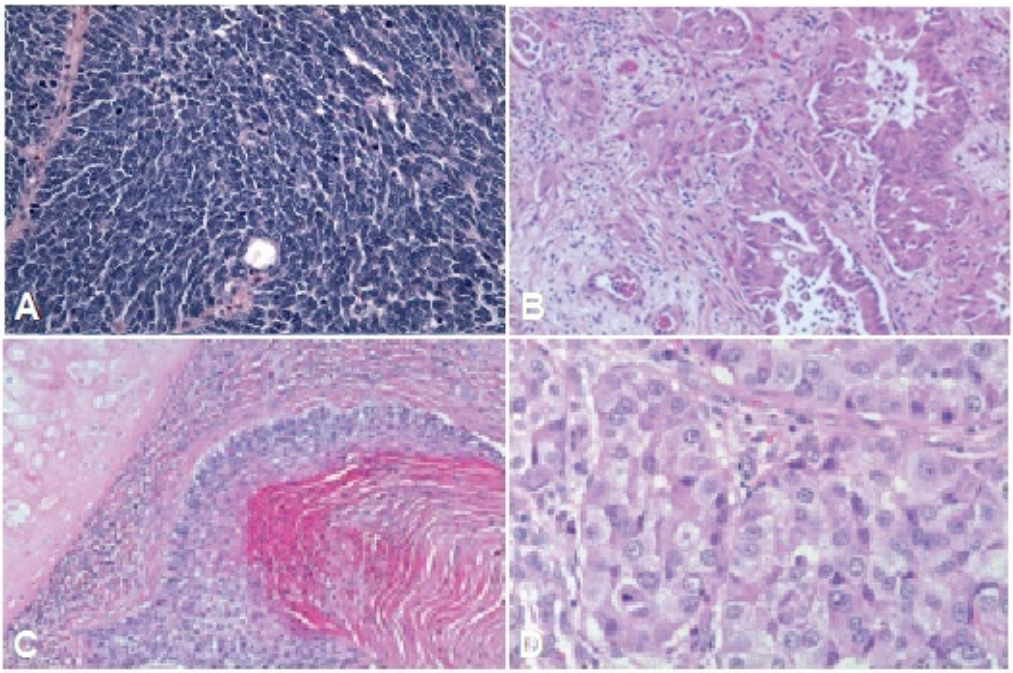


Figure 3: Histological patterns of LC. (A) Small cell lung cancer, from Jackman et al. (2005); (B) adenocarcinoma, from Colby et al. (2004); (C) squamous cell carcinoma, from Hammar et al. (2004); (D) large cell carcinoma, from Hammar et al. (2004).

Lung cancer staging is based on the TNM classification for malignant tumours (TNM). This classification was originally proposed by Mountain and adopted by the American Joint Committee on Cancer (AJCC) in 1973 and one year later by the Union for International Cancer Control (UICC) (26). The TNM system is used for the majority of malignant tumours and is based on the extent of the disease. Its goal is to assist treatment planning, evaluation of treatment outcomes, and prognosis determination (27). The letter T stands for primary tumour and determines the size and degree of locoregional invasion. The N indicates the extent of regional lymph node involvement, and the M reveals the presence or absence of intrathoracic or distant metastasis (27, 28). The latest edition of the TNM classification of lung cancer can be found on Figure 4.

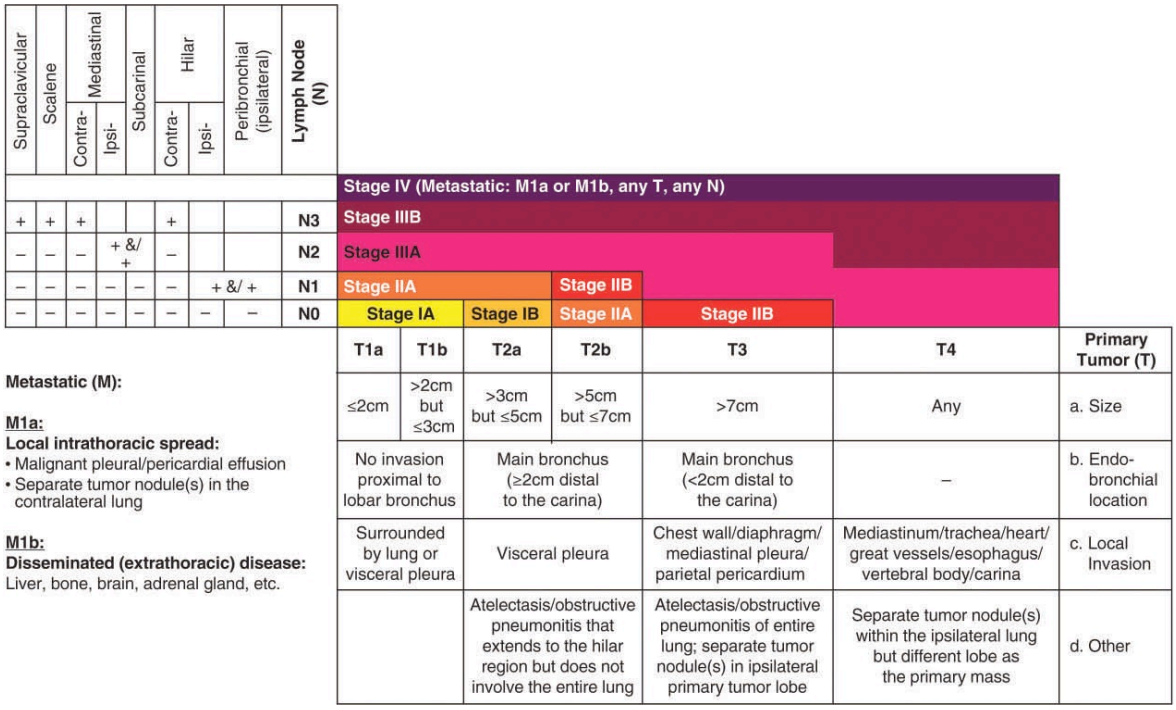


Figure 4: Chart of the latest TNM classification of lung cancer, from Uybico et al. (2010).

1.1.3. Diagnosis and treatment

Lung cancer diagnosis includes evaluation of clinical history, several routine blood exams such as anaemia and liver enzymes, bronchoscopy to allow the evaluation of the extent of the disease in the tracheobronchial tree, and the tumour identification after biopsy (29). Information on tumour size and local invasion are obtained by chest radiography (30). Computerized tomographies (CT) have been of great help in LC staging, given that they offer 3D information of the disease extent such as mediastinal invasion, chest wall invasion, or the presence of other pulmonary nodules (29). Mediastinum

staging is further amended by endobronchial ultrasound, endoscopic ultrasound, mediastinoscopy, and positron emission tomography. The latter is also a sensitive method to detect metastatic spread of the tumour. Following the classification and staging of a patients' tumour, course of treatment is decided, selecting between the available therapeutic options: surgery, radio and chemotherapy, used separately or in different combinations.

Lung cancer is usually asymptomatic during its development and this causes the majority of patients to be diagnosed at advanced stages of the disease. The accurate diagnosis of LC, including classification and staging, is essential to the correct treatment and prognosis of a given patient.

Tumour resection by surgery is the most successful curative option available for patients with early stage disease: stage I, II, and selected patients with stage IIIA. Nonetheless, the survival rate of these patients varies greatly depending on their disease stage (28). Stages IIIB and IV, the grand majority of diagnoses, are considered inoperable and other treatments, such as radiotherapy and chemotherapy, are applied.

Radiotherapy is a valuable therapeutic option for patients who, despite the presence of early disease, are medically unfit or refuse to undergo surgical resection. It is also of use as adjuvant therapy for patients with incomplete resection or node-positive disease and as palliative therapy, providing symptomatic improvement (31).

Chemotherapy was initially used in patients with advanced metastatic disease as a palliative measure. Since then, the application of drugs (cisplatin, taxanes, antimetabolites, topoisomerase inhibitors, among others) has been used with curative intent combined with surgery, neoadjuvant or adjuvant, and/or radiotherapy, in an attempt to improve the survival of NSCLC patients (32). In addition to these traditional drugs, targeted therapy has been improving in the last years, gathering information from molecular studies that identified specific alterations in groups of LC patients. The greatest achievements of this targeted therapy have been the development of monoclonal antibodies and tyrosine kinase inhibitors against the epidermal growth factor receptor (EGFR), such as cetuximab and erlotinib, and the vascular endothelial growth factor (VEGF), such as bevacizumab and sorafenib, among others (Figure 5) (33). In addition, novel agents that target the echinoderm microtubule-associated protein-like 4 (EML4) and the anaplastic lymphoma kinase (ALK) gene fusion, for example crizotinib, the insulin-like growth factor-1 receptor (IGF-1R), such as cixutumumab or linsitimib, the mammalian target of rapamycin (mTOR), for example deforolimus, and the hepatocyte growth factor receptor (MET), such as tivantinib, are currently being investigated in NSCLC clinical trials (Figure 5) (33).

Although a lot of progress has been made, the 5-year survival rate for lung cancer is still approximately 15%, reinforcing the need for diagnostic markers that can aid in the early detection of this tumour.

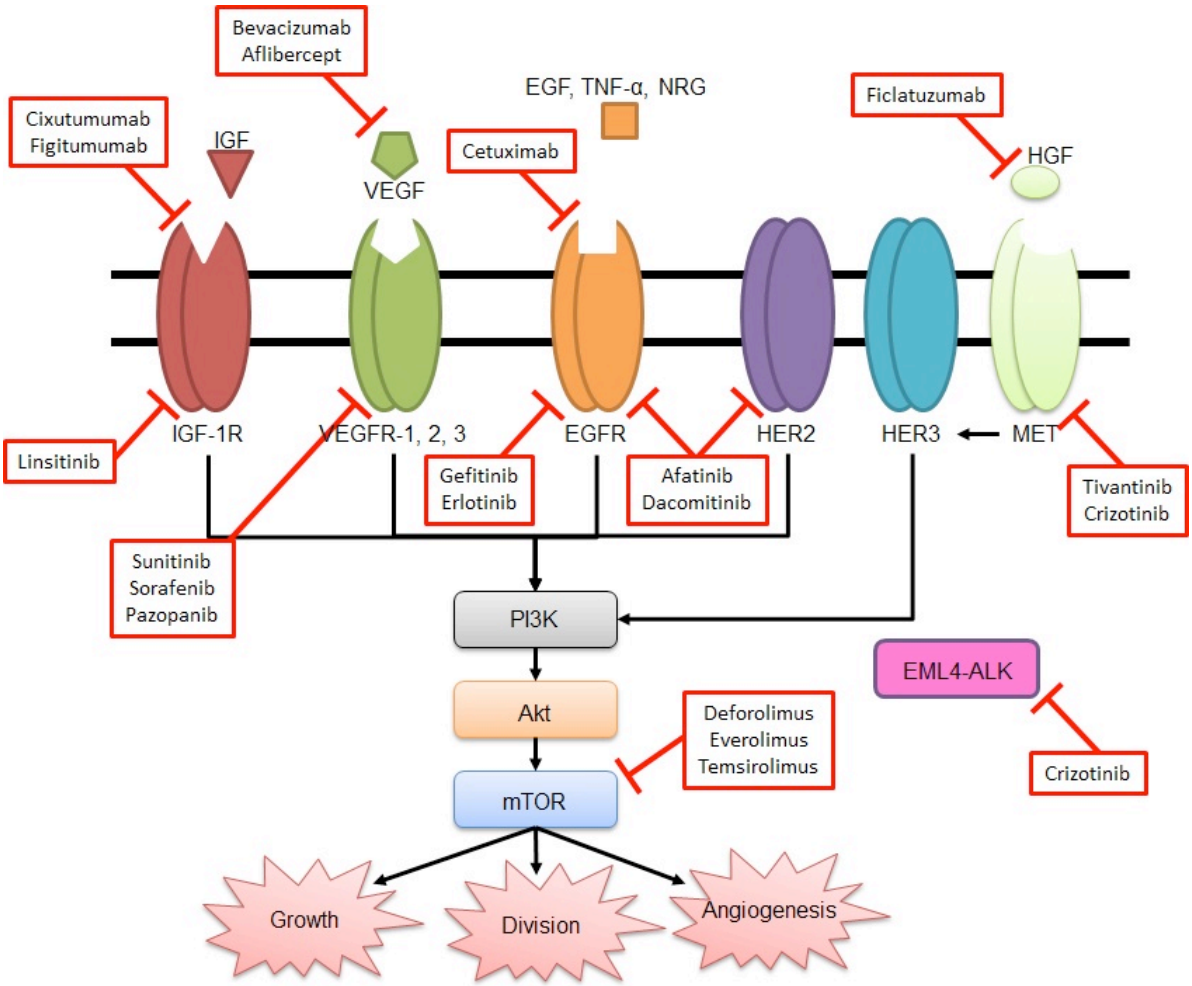


Figure 5: Most relevant pathways in NSCLC and developed targeted therapies, adapted from Pal et al. (2010).

1.2. Chronic obstructive pulmonary disease

1.2.1. Epidemiology and causes

Chronic obstructive pulmonary disease (COPD) is, according to the Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD), a "preventable and treatable disease with significant extrapulmonary effects. The pulmonary component of COPD is characterized by a progressive airflow limitation associated with an atypical inflammatory response of the lungs to noxious particles and gases" (34). The airflow limitation present in COPD patients is the result of damages to the small and large airways (chronic bronchiolitis and bronchitis, Figure 6A) and to the pulmonary

parenchyma (emphysema, Figure 6B) (35). The majority of patients present signs of both bronchitis and emphysema instead of one component exclusively (36).

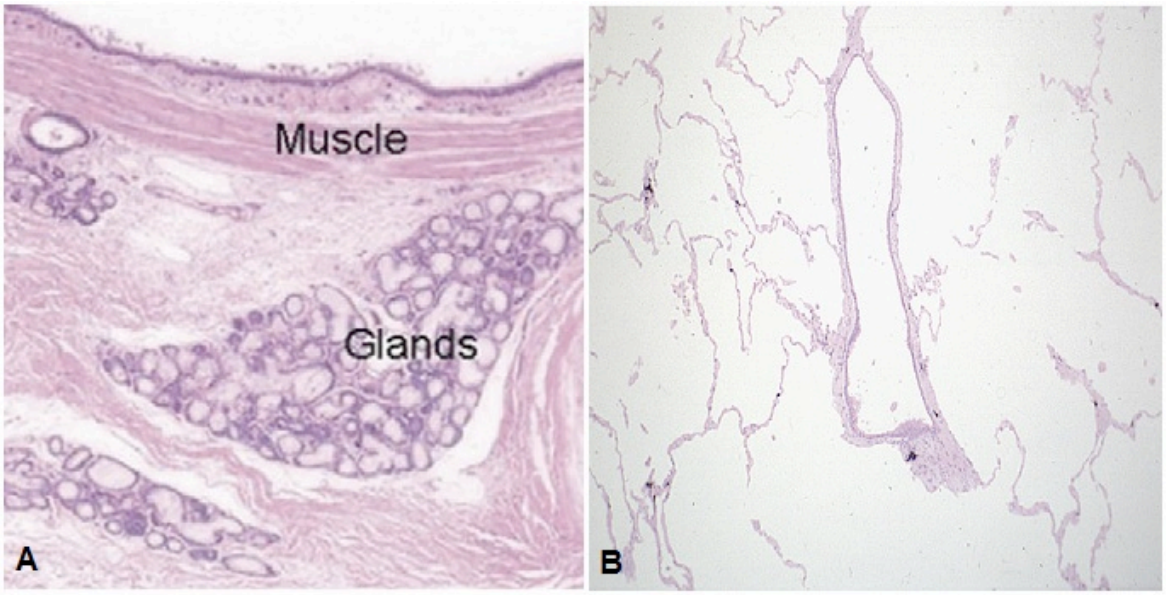


Figure 6: Histological patterns of bronchitis with increased number of glands and thicker smooth muscle bundles (A) and emphysema with enlarged airway spaces (B), from Macnee et al. (2007).

COPD is a leading cause of worldwide morbidity and mortality, apart from being a major economic and social burden. The WHO estimated that COPD was the fifth leading cause of death in 2006 and is expected to rise to fourth place by 2020 (3). It is clearly more frequent in men and the elderly, but incidence is increasing in women reflecting changes in their exposure to COPD risk factors since the middle of the 21st century

The primary risk factor for COPD is, as for LC, tobacco smoking. The WHO assessed that 40 to 70% of all COPD mortality is related to smoking (37). Age of smoking initiation, pack-years smoked, and current smoking status, influence COPD development, prognosis, and mortality (38). Even so, reports state that 15-50% of smokers will develop some form of COPD over time (36, 37). There is a natural decline in lung function with age, but active smokers have an accelerated rate of deterioration that can be reversed to normal levels of decline with sustained smoking cessation.

1.2.2. Diagnosis, classification, and treatment

COPD is a multicomponent and heterogeneous disease that varies within a group of patients with respect to lung pathology, disease progression, and comorbidities. Given that the clinical presentation is far from uniform, healthcare practitioners should consider the diagnosis of COPD if any patient is over 40 years of age, has been exposed to the

aforementioned risk factor, and presents symptoms such as breathlessness during exercise, increased effort to breath, cough, or excessive mucus production (39, 40). If these symptoms are present, physical examination and spirometric tests are required to make the accurate diagnosis and disease staging.

Spirometry is the gold standard for the detection of COPD, given its high reproducibility and availability (38). It is a simple and non-invasive test that measures lung function: the amount (volume) and/or speed (flow) of air that can be inhaled and exhaled. The evaluated parameters are: post-bronchodilator forced expiratory volume in 1 second (FEV1) and the ratio between FEV1 and forced vital capacity (FEV1/FVC). These two measurements are used to detect COPD and assess its severity by comparison with reference values that vary with age, gender, race and height (34).

The classification of COPD ranges from mild to very severe. A normal spirometry implies a FEV1/FVC ratio superior to 70% and a FEV1 superior to 80%. COPD is diagnosed when the FEV1/FVC ratio is below 70% and its severity varies from mild to very severe depending on the reduction of post-bronchodilator FEV1 compared to the predicted in a healthy population (Table 1) (41). Radiological imaging can be used in the assessment of COPD severity since it can provide an accurate characterization of changes in the lung parenchyma and airway wall thickness, as well as exclude alternative diagnoses (42). Arterial blood gas measurement, α1-antitrypsin deficiency screening, and the six minute walk test are also useful in the diagnosis of this disease (34).

Table 1: COPD classification for patients with a FEV1/FVC ratio below 70%

Gold stage	Description	FEV1 values
GOLD 1	Mild	FEV1 ≥ 80% predicted
GOLD 2	Moderate	50% ≤ FEV1 < 80% predicted
GOLD 3	Severe	30% ≤ FEV1 < 50% predicted
GOLD 4	Very severe	FEV1 < 30% predicted

Given that no cure is currently available and there are no therapies than can halt the lung function decline, the treatment of COPD varies enormously across patients. The main goals of the treatment are to reduce symptoms and improve quality of life, reduce the frequency and severity of the exacerbations, and reduce mortality (43). One of the first steps in the treatment and management of these patients is smoking cessation. Non-pharmacological treatments include long term oxygen therapy (that can modify long term lung function decline), non-invasive ventilation, lung volume reduction surgery (for patients with emphysema), and pulmonary rehabilitation that can provide relief from dyspnoea,

increases exercise tolerance, and improves quality of life (34). Pharmacological treatments are patient-specific and can include the use of antibiotics, bronchodilators, corticosteroids, and mucolytic agents, alone or in combinations (34, 42).

1.3. Links between LC and COPD

Lung cancer and COPD are two of the major causes of mortality in the world. Apart from being caused by the same major etiological factor, tobacco smoking, epidemiological data demonstrates a clear connection between both diseases. The odds ratio (OR) value for LC increases 2.8 fold in patients with moderate to severe COPD (44). Independently of smoking history, reduced FEV1 is associated with an increased risk for developing LC: even small changes in this parameter, such as 10% differences, raise LC risk by 30-60% (45). Indeed, 50 to 70% of all LC patients have spirometric evidences of COPD (46). In addition, the presence of mild emphysema, also independent of cigarette smoking and history of other lung diseases, confers a greater risk to develop LC (47). Since only a fraction of long-term smokers develop these diseases, complex mechanisms and interactions are thought to determine an individual's risk of developing COPD, LC, or both.

At a first glance, it is difficult to envision any common mechanism between both diseases: LC is characterized by unlimited cell proliferation, sustained angiogenesis and self-sufficiency in growth factors, tissue invasion, evasion of apoptosis, and insensitivity to anti-growth signals, whereas COPD has increased levels of apoptosis, extracellular matrix degradation, limited angiogenesis, and ineffective tissue repair (48, 49). Understanding the shared mechanisms between these two dissimilar diseases is critical for developing new methods of early diagnosis and treatment of LC.

Throughout the years specific processes have been suggested to be of importance in the pathogenesis of both COPD and LC (Figure 7). Reactive oxygen species (ROS) caused by tobacco smoke are produced in excess in the setting of COPD and can ultimately produce DNA alterations that lead to LC development. The balance between cellular proliferation and cell death is also important in the lung given its contribution to the development of both diseases. Hypoxia and angiogenesis can also affect the development of COPD and LC: airflow limitation can lead hypoxia that may cause destruction of alveolar capillaries and airway enlargement, in addition to being capable of promoting, through angiogenesis, the progression and metastasis of lung tumours (46, 50). However, these processes are all influenced by and interact with a major mechanism activated by tobacco smoking that is inflammation.

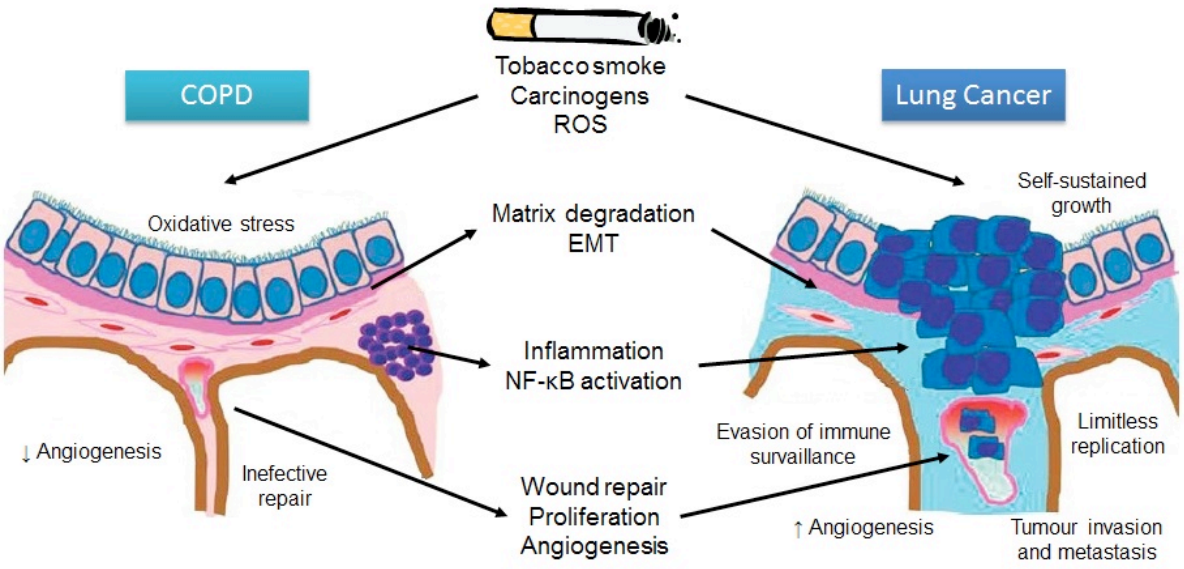


Figure 7: Mechanisms and pathways potentially involved in the development of COPD and LC, adapted from Yang et al. (2011).

Inflammation is one of the body's natural protective mechanisms against harmful stimuli. It is the result of complex interactions between immune cells and soluble factors that arises in tissues as a response to infections, trauma, or toxic substances such as the ones present in tobacco smoke (51). Its major role is to remove the source of aggression and to initiate the healing process. Nevertheless, a persistent inflammatory response in the lung can result in the development of lung diseases, such as COPD and lung tumours.

COPD was early recognized as an inflammatory disease, with profound abnormalities in the inflammatory pathways, and that airway inflammation starts at an early stage, many years prior to the onset of clinical symptoms. Tobacco smoke induces inflammation in all smokers, but those who develop COPD have a greater degree of inflammation that persists even after smoking cessation (50). Tobacco smoke is a complex mixture of thousands of chemicals generated by tobacco combustion such as carcinogenic toxins, addictive substances, and more than 10^{14} oxidants and free radicals (46). These compounds first interface with the lungs at the mucosal surfaces lining the airways. There they can damage epithelial cells and resident inflammatory cells by inducing peroxidation of lipids and other cell membrane constituents, activate oxidative-sensitive cellular pathways, induce DNA damage, and activate epithelial cell intracellular signalling cascades that lead to inflammatory gene activation (52). The release of these inflammatory mediators initiates the recruitment and activation of more pro-inflammatory cells such as neutrophils, macrophages, and lymphocytes that propagate the inflammatory response by releasing more cytokines, for example interleukin 8 (IL-8) and tumour necrosis factor α (TNF-α), and ROS, generating a vicious cycle of persistent

inflammation. Interestingly, the immune cells and mediators that drive chronic inflammation and the progression of COPD are also known to promote a tumour friendly microenvironment (53).

The observation of a relationship between cancer and inflammation is not a recent one. Almost 2000 years ago, Galenus described the similarities between inflamed tissue and cancerous tissue. Rudolph Virchow, in the year of 1863, reported the infiltration of leukocytes in malignant tissues and suggested that cancers arise in sites of chronic inflammation (54). More than one hundred years later Dvorak observed that cancer and inflammation share some developmental mechanisms and types of infiltrating immune cells, apart from classifying tumours as “wounds that do not heal” (55).

The ultimate recognition of inflammation as a major player in cancer development came with the 2011 update article on cancer hallmarks by Hanahan and Weinberg where it was classified as an enabling characteristic of tumours. The evidence gathered over the last years showed that inflammation contributes to the appearance of multiple cancer hallmark capabilities by supplying important molecules to the tumour microenvironment. Those molecules include growth factors that sustain the proliferative signalling, survival factors that limit apoptosis, pro-angiogenic factors, extracellular matrix-modifying enzymes that favour angiogenesis, invasion, and metastasis. Furthermore, inflammation manifests itself at the earliest stages of tumour progression and is capable of nurturing insipient neoplasias into developed cancers. In addition, inflammatory cells can release a number of chemicals, such as ROS, that are actively mutagenic and promote malignancy even further (14).

Studies on LC and chronic inflammation have revealed that the likelihood of developing LC was increased in patients with higher levels of C-reactive protein (CRP), an acute-phase protein that is a marker for inflammation (56). Also, the risk for developing LC was decreased among COPD patients who had been treated with inhaled corticosteroids (57). Studies on animal models have showed that, after infection by *Haemophilus influenza*, mice develop a bronchial inflammation similar to COPD and a subsequent predisposition to lung carcinogenesis (58). The constant lung injury and repair cycle triggered by chronic inflammation, such as the one that occurs in COPD, can enhance cell proliferation and genetic damage accumulation, epithelial to mesenchymal transition (EMT), and ultimately lung carcinogenesis (46).

1.4. Proteomics

In the last several years genomic and proteomic technologies have begun to uncover the molecular complexity of lung tumours by allowing a rapid and complete analysis of the genes and proteins that are expressed in the context of this disease. The use of genomic technologies such as microarrays and quantitative reverse transcription polymerase chain reaction (RT-PCR) on LC studies has allowed the separation of adenocarcinoma patients into new molecular sub-classes with different outcomes, the early identification of high risk patients, and the detection of molecular signatures associated with metastasis (59).

While genomic technologies have revealed important data regarding the biology of LC, the direct evaluation of proteins offers information that is not obtainable through the study of DNA and RNA. Proteins are crucial operators in the majority of biological systems and signalling pathways. Although they can determine the cells' phenotype, RNA levels often poorly correlate with protein expression (60).

1.4.1. Introduction to proteomics

The proteome is defined as the entire collection of proteins that is produced by a cell or tissue in a given time, making it a very dynamic entity: the type of expressed proteins, their relative abundance, and their subcellular location depend on the physiological state of a cell or tissue (61).

Proteomics, the analysis of the proteome, is performed through a variety of technologies that are continuously evolving. Traditional protein analysis technologies include approaches such as western blot, immunohistochemistry (IHC), and enzyme-linked immunosorbent assay (ELISA). Protein microarrays, which consist of large number of antibodies robotically immobilized on a slide, have a higher throughput and allow the identification of several proteins at a time in the same sample (62). Protein microarrays are rapid, automated, economical, and highly sensitive, consuming small quantities of samples and reagents. However, the previous techniques are target driven, meaning they require a previous knowledge of the proteins that are expected to appear in a given sample. In contrast to these techniques, two methods do not require a previous assumption of the identity or number of proteins in a sample and are considered classical proteomic technologies: two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and mass spectrometry (MS). The use of electrospray ionization (ESI) and liquid chromatography, second generation proteomics or shotgun proteomics, has more recently allowed the identification of complex and so far unknown proteomes.

First developed in 1975 by O'Farrel, 2D-PAGE has been the preferred method for proteomic analysis given its relative low cost and high applicability. It allows proteins to be separated in two dimensions: isoelectric point and molecular weight (63).

In the first dimension, also called isoelectric focusing, the separation of proteins occurs when electric current is applied to a protein sample in a pH gradient: proteins will move until they reach a pH that is similar to their isoelectric point. The second dimension is based on the traditional electrophoresis process in polyacrylamide gels: proteins separated in the first dimension are given negative electric charges by the use of sodium dodecyl sulfate (SDS) and migrate through the gel, when current is applied, accordingly to their molecular weight. This technique has some associated limitations that include small sensitivity to very low or very high molecular weight proteins and isoelectric points, samples with high salt contents, the under-representation of membrane proteins, and gel-to-gel variation (64, 65).

Two-dimensional difference gel electrophoresis (2D-DIGE), a variation of 2D-PAGE, allows the analysis of three samples (test, control, reference) in the same gel and the evaluation of quantitative differences among them by the use of fluorescent dyes (Cy3, Cy5, and Cy2) (64). A representation of the common 2D-PAGE and 2D-DIGE workflow is depicted in Figure 8 (66). Alternatively, SDS-PAGE alone can be used and obtained protein bands can be excised and analysed by MS. Gel-free separation techniques include liquid chromatography and high-performance liquid chromatography. Typically samples are digested and peptides are separated by this technique, with the resulting fractions being analysed by MS (67).

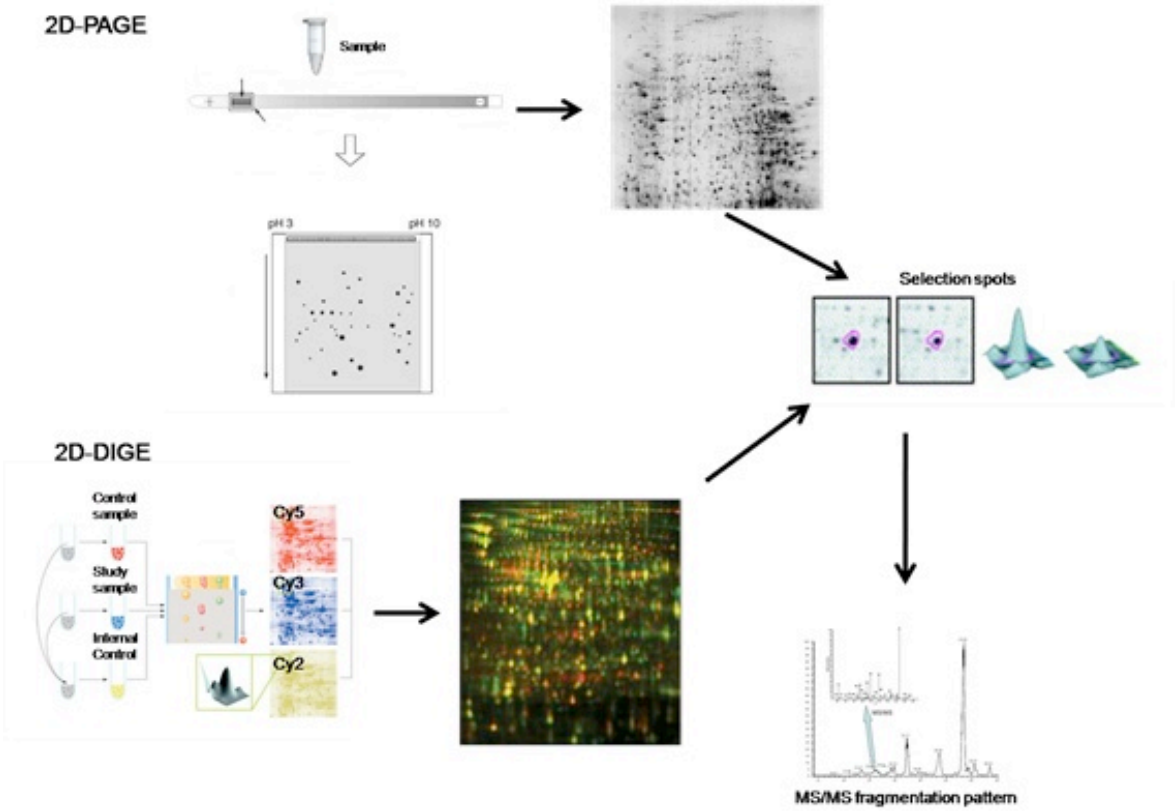


Figure 8: A typical 2D-PAGE and 2D-DIGE protocol followed by MS, from Pastor et al. (2013).

The use of labelling methods has allowed protein quantitation in a gel-free setting, as seen on Figure 9 (66). Stable isotope labelling with amino acids in cell culture (SILAC) is based on the metabolic incorporation of “heavy” and “light” forms of amino acids into the proteins of living cultured cells (68). Following trypsin digestion, peptides are analysed by MS and the light and heavy peptides appear in two distinct peaks and, by comparing the signal intensities differences, relative quantitation is achieved. Isotope-coded affinity tags (ICAT) is generally used to compare pairs of samples. Extracted proteins are labelled with the light or heavy ICAT reagent, mixed, trypsin digested, fractionated, and analysed by MS (69). Isobaric tags for relative and absolute quantitation (iTRAQ) can analyse 4 to 8 samples (4-plex, 8-plex) in the same experiment. After trypsin digestion, samples are independently labelled with the iTRAQ reagent. During MS analysis, the reporter groups of the iTRAQ reagents separate from the peptides and generate small fragments for each sample with mass-to-charge (m/z) of 114, 115, 116, and 117 for 4-plex, plus 113, 118, 119, and 121 for 8-plex. The intensity of each peak correlates with the quantity of each reporter group and thus with the quantity of the peptide (70).

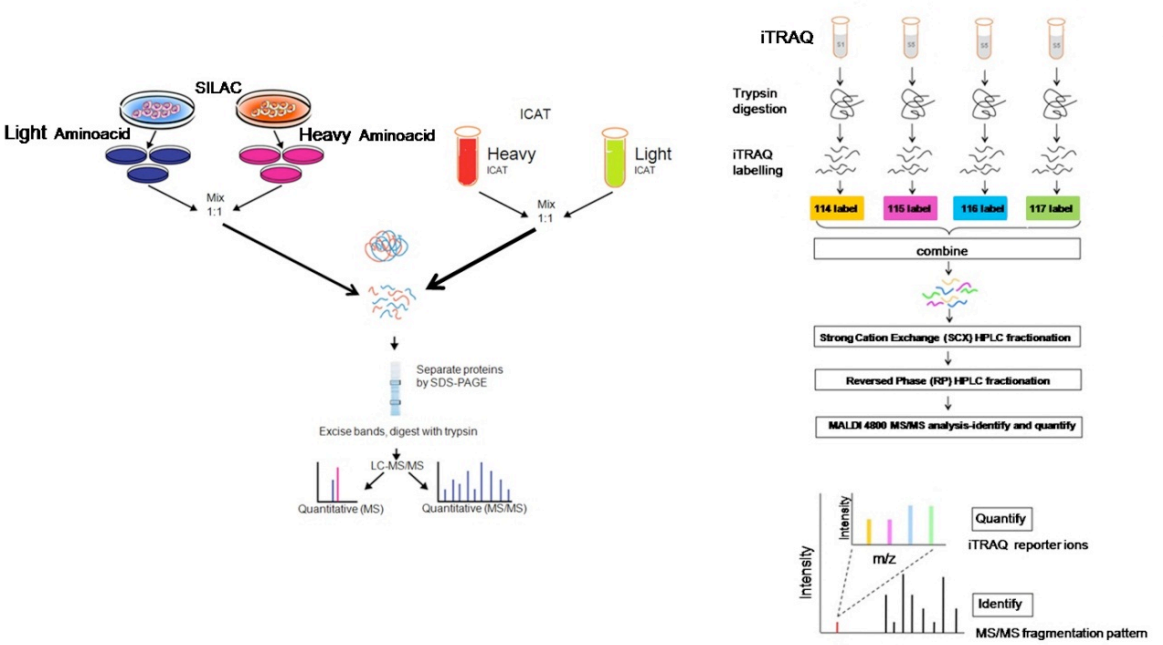


Figure 9: Gel-free quantitative approaches in proteomics, SILAC, ICAT, and iTRAQ, from Pastor et al. (2013).

The MS analysis of proteins and peptides is one of the pillars of current proteomic analysis and has revolutionized the field since the 1990s (61). It provides the accurate measurement of their molecular weight and charge (m/z ratio) that allows their identification. In order to obtain these measurements, samples are ionized by an ion source and travel through a mass analyser and then to a mass detector (71).

The most common ionization method in proteomic studies is matrix-assisted laser desorption ionization (MALDI). Proteins or peptides are mixed with an energy-absorbing matrix and spotted on a metal plate. After drying, the matrix and sample crystallize and are then hit by a laser to be ionized (64). Ionization by MALDI allows and easy and rapid data analysis, high resolution for large molecules, and sensitivity to small amounts of sample. However, small molecules may not be detected (64, 65, 71). Alternatives to MALDI include ESI, in which the protein sample is dissolved in a volatile buffer, passed through a high performance liquid chromatography column and turned into a fine aerosol by the electric current applied to a metallic capillary tube. Although ESI allows the identification of proteins in large complex samples, MALDI is still currently used given its tolerance to sample contaminants (72). Surface-enhanced laser desorption/ionization (SELDI), a variation of MALDI, uses chromatographic chip arrays that allow the analysis of a more complex sample by selectively binding subsets of proteins (64, 65).

After ionization samples pass through a mass analyser that separates ions according to their m/z ratio. The most commonly used for proteomic studies is the time-of-

flight (TOF) analyser. In this analyser, ions are accelerated by an electric field and the time that each ion takes to reach the detector is measured to determine the particles m/z ratio (71). The addition of a second mass analyser (tandem mass spectrometry) results in the identification of the amino acid sequence of peptides: the first analyser allows the m/z ratio measurement of whole peptides and the second analyser measures the ratios of the fragmented peptides. The obtained spectra are then compared against protein sequence databases to allow protein identification (65).

1.4.2. Proteomics of COPD

The progressive deterioration of pulmonary function caused by COPD and the fact that this disease is an important risk factor for the development of LC, independent of tobacco smoking, has driven researchers to discover and validate biomarkers of COPD. The use of 2D-PAGE and MALDI-TOF-MS has been the most common approach to biological COPD samples, although other techniques are used.

Several studies have evaluated tissue samples, often considered the optimal sample for studying lung diseases, by this technique revealing heat shock protein 27, cyclophilin A, matrix metallopeptidase 13, and surfactant protein A as possible tissue diagnostic markers of COPD (73-75). Plasma samples, that can reflect the systemic alterations caused by this disease, have revealed by different proteomic techniques diagnostic (aldolase, catalase, haptoglobin, among others) and prognostic (serum amyloid A) biomarkers of COPD (76-79). Induced sputum is a good sample for research and clinical diagnosis and disease monitoring, although the sampling protocol is a crucial step. Proteomic studies on COPD using this sample have identified diagnostic biomarkers, such as apolipoprotein A1, lipocalin 1, and polymeric immunoglobulin receptor, among others (80, 81).

1.4.3. Proteomics of LC

Proteomic technologies allow the analysis of thousands of cancer-related proteins and promise to improve the knowledge of LC biology and pathogenesis, and aid in the development of new early detection biomarkers and the identification of protein profiles associated with prognosis and drug resistance.

The most common proteomic workflow for LC biomarker discovery has been the use of 2D-PAGE followed by MALDI-TOF-MS analysis. Nowadays, with the development of shotgun proteomics, that allow a rapid and accurate characterization of several

thousands of proteins in a single study, a great deal of data is being generated in LC proteomic studies. However, 2D-PAGE and MALDI-TOF-MS are still widely used given its low cost and easier implementation in a laboratory setting when compared to liquid chromatography-MS approaches (82).

The proteomic studies on LC so far have used a variety of samples and techniques. When considering the studies aimed at identifying LC diagnostic biomarkers, blood based proteomics has been a major focus of research. Blood has the grand advantage of being easily obtainable and its collection is minimally invasive. Analysis of plasma or serum assumes that the perfusion of tumours or host response contributes to the modification of circulating proteins and peptides, and that tumour-derived proteins may be in circulation. However, due to the abundance of plasma proteins, depletion of these proteins is necessary to reveal the presence of less abundant ones, the ones likely to be tumour-specific markers (67). Some LC blood diagnostic biomarkers have been identified by proteomic studies: acute phase proteins, such as haptoglobin β chain and serum amyloid A (83, 84), glycosylated proteins, for example kallikrein, insulin-like growth factor-binding protein 3, prostaglandin D synthase (lipocalin-type), glycosylated forms of apolipoprotein C-III (85-87), and finally blood proteomic profiles, given that a combination of proteins is more reliable than a single marker (88-90).

Pleural effusion is the pathological accumulation of fluid that occurs in inflammatory conditions and in LC settings. In this case, pleural effusions are often drained to search for cancer cell infiltration. It has a similar protein composition to that of plasma, but its proximity to tumour cells makes it useful for lung cancer biomarker detection by proteomic techniques. A few proteomic studies have used this type of sample to search for LC diagnostic markers: they have found important proteins such as Niemann–Pick disease type C2 and membrane glycoprotein-2, among others (91, 92).

The use of tissue samples presents inherent difficulties: the lung is a heterogeneous organ composed by many different cell types (bronchial, alveolar, and inflammatory) and vascular structures, which implies that tumour tissue samples will likely have adjacent non-tumour cells, inflammatory cells, stromal components, among others, that can result in non-tumour protein contamination. However, LC proteomic studies have shown the over-expression of proteins involved in the energy-production and anti-oxidation pathways (93), of the macrophage migration inhibitory factor and cyclophilin A proteins (94), and of the pyruvate kinase M1, manganese superoxide dismutase and peroxiredoxin proteins (95).

A recent study in saliva samples identified three proteins (haptoglobin, calprotectin, and zinc-α-2-glycoprotein) capable of discriminating LC patients from healthy controls with high specificity and sensitivity (96).

In addition to LC diagnostic biomarkers, proteomic technologies have also been applied to the development of prognostic and predictive biomarkers. Some studies have identified proteins involved in cancer metabolisms, proliferation, differentiation, and apoptosis (97-100). Others have focused on the metastasis process, where proteins such as S100A11, annexins 2 and 3, cytokeratins, cathepsins D, thymosin β4 and β10, among others, have been associated with a lung cancer metastatic phenotype (101-105). Similar to what has been used for diagnostic purposes, prognosis can also be obtained based on a combination of proteins, and proteomic signal signatures have been identified that allow the discrimination between good and bad prognosis of LC patients (106, 107).

Lastly, the study of LC predictive biomarkers has focused on two major therapeutic options available for LC treatment: studies evaluating the response to traditional chemotherapy and the response to EGFR inhibitors. Regarding the traditional chemotherapy, a group of 14 proteins was identified and related to gemcitabine resistance (108), a five-peak pattern that distinguished platinum chemotherapy-resistant patients from chemotherapy-sensitive patients was observed (109), and a 13-peptide signature was able to discriminate NSCLC patients with short or long progression-free survival after treatment with cisplatin-gemcitabine in combination with bortezomib (110). When it comes to identifying patients who might benefit from treatment with EGFR inhibitors, an 8-peak profile was predictive of outcome in patients with NSCLC, after treatment with gefitinib and erlotinib (89). In conclusion, and despite some technical difficulties, several protein profiles with diagnostic, prognostic, and predictive value for LC have been identified. However, the identity of only a small proportion of the detected peaks has been achieved and for many of these identified proteins, their role in LC development is not yet known.

Considering all proteomic researches in LC, one type of sample in particular has been overlooked in these types of studies. Obtained during fiberoptic bronchoscopy, bronchoalveolar lavage fluid (BALF) is generally used to perform cell counts, cytology, and cultures, allowing the evaluation of pulmonary diseases. This biological sample contains airway cells and soluble components that can indicate possible inflammatory, infectious, or neoplastic processes that could be occurring in the lung or the alveoli. Excess BALF is normally discarded after it has been used for standard pathologic procedures, but can be used to search for biomarkers that could assist the diagnosis of pulmonary diseases (111). The use of BALF in proteomic studies has been limited to non-malignant conditions such as cystic fibrosis (112), asthma (113), acute lung injury (114), interstitial lung diseases (115), acute respiratory distress syndrome (116), among others. So far, the analysis of the proteomic profiles of BALF to diagnose LC has not been explored in depth. Moreover, to the best of our knowledge, no studies have evaluated the proteomic differences associated with COPD patients.

The analysis of the proteomic profiles of LC and COPD patients might allow us to gain more knowledge on the mechanisms of both diseases, to identify possible LC proteomic diagnostic biomarkers, as well as elucidate why some COPD patients end up developing LC.

HYPOTHESIS AND OBJECTIVES

2. HYPOTHESIS

There are differences in the protein expression (inflammation and/or other pathways) of BALF among patients with LC and/or COPD as compared to healthy individuals. These differentially expressed proteins could help to understand the molecular pathogens of LC and COPD, and may provide useful biomarker for diagnosis.

3. OBJECTIVES

3.1. Main objective

The present work focused on identifying new biomarkers as well as biological pathways involved in the development of LC and/or COPD among high-risk patients by using proteomic methodologies, focusing particularly on the inflammatory pathway.

3.2. Specific objectives

3.2.1. Identification of differently expressed proteins in BALF from patients with LC, COPD, COPD with LC, and without LC or COPD by comparative proteomic analysis.

3.2.2. Analysis of the differential expression of inflammation related proteins (cytokines and growth factors) in LC and COPD patients.

MATERIALS AND METHODS

4. MATERIALS AND METHODS

4.1. Patient selection and sample collection

From 2009 to 2011, a total of 925 patients who had required flexible bronchoscopy for diagnostic purposes at the Hospital Universitario Virgen del Rocío (Seville, Spain), were collected for this prospective study. From the original first group only 419 patients respected the selection criteria: 1) patients were under pneumologist consultation due to haemoptysis and/or a pulmonary nodule; 2) patients were requested by their physician to perform a flexible bronchoscopy for diagnostic purposes; 3) patients were smokers or ex-smokers of more than 20 pack-years; 4) patients had to be older than 40 years of age. Exclusion criteria included: 1) diagnosis of a neoplastic disease other than lung cancer; 2) active pulmonary tuberculosis; 3) previous lung resection; 4) history of drug abuse; 5) presence of other acute or chronic inflammatory diseases.

The study protocol was conducted in agreement with the Helsinki declaration and approved by the Hospital's Ethical Committee, and a written informed consent was obtained from all patients prior to their addition to the study.

Subjects were prepared with a combination of topical anaesthesia (20% benzocaine spray to the pharynx plus 2% topical lidocaine as needed) and conscious sedation using midazolam and meperidine according to institutional guidelines. All BALF samples were obtained by instillation and aspiration of 40 to 60 ml aliquots of 0.9% sterile saline in the appropriate bronchopulmonary segment. Recovered fluid was immediately passed through a 100 μm sterile nylon filter (Becton Dickinson, San Jose, CA) to remove mucus and transported on ice to the laboratory. The total volume was then centrifuged for 10 min at 1800xg and 4ºC. The supernatant was aliquoted into 2 ml tubes and frozen at –80ºC until further use.

4.2. Sample processing

4.2.1. Sample concentration

The amount of proteins in BALF samples is very small, diluted in the sterile saline used for its recovery. In order to use these samples for proteomics, they required concentration and this was accomplished by the use of a Concentrator Plus vacuum concentrator (Eppendorf, Hamburg, Germany). BALF samples were thawed on ice with a protease inhibitor cocktail kit (Pierce, Thermo Scientific, Rockford, IL, USA). Samples were then aliquoted into new tubes and placed on the vacuum concentrator. The initial volume of the samples, usually 4-8 ml, was reduced to 250-450 μl in 2-6 hours.

4.2.2. Protein quantitation

A number of colorimetric methods were tested, such as bicinchoninic acid assay (BCA from Pierce, Thermo Scientific, Rockford, IL, USA, Rockford, IL, USA), Bradford (Bio-Rad, Hercules, CA, USA) and RC-DC (Bio-Rad, Hercules, CA, USA). The RC-DC method was selected given that it was reducing agents and detergent compatible. It is based on the principle of Lowry estimation where proteins react with alkaline copper and subsequently reduce the folin reagent, leading to colour development.

Protein quantitation was assessed according to the manufacturer's instructions. Absorbance was measured at 750 nm and the linear equation created by different bovine serum albumin (BSA) dilutions allowed the determination of the protein concentration in our samples. Protein quantitation was assessed before and after sample concentration, depletion of high abundance proteins, and sample cleaning procedures, which will be described in the following corresponding sections.

4.2.3. Depletion of high abundance proteins

The presence of high abundance plasma proteins obscures the presence of low abundant ones of possible interest. Revealing these important proteins required that a depletion process be applied for the analysis of BALF proteomes.

Several commercial kits of depletion columns were tested: Proteominer (Bio-Rad, Hercules, CA, USA), ProteoPrep (Sigma-Aldrich, St. Louis, MO, USA), and SpinTrap (GE Healthcare, Waukesha, WI, USA). The latter one showed the best results in albumin and immunoglobulin G removal in our samples.

Depletion protocol followed the manufacturer's instructions. The SpinTrap columns were inverted repeatedly to resuspend the storage medium, and then centrifuged to remove it. Binding buffer was added and centrifuged out to wash the column before the addition of the concentrated BALF sample. The sample was allowed to incubate for 5 minutes to let the existent albumin and immunoglobulin G bind to the column. The column was then centrifuged and the eluate collected. Binding buffer was added twice to the column and the eluate was again collected. The depleted BALF samples were cleaned immediately or frozen at -80ºC until further use.

4.2.4. Sample cleaning

After the depletion step BALF samples required the removal of salts, thiols, denaturants, and other contaminants that could be present and interfere with the 2D-PAGE protocol. Of all available methods to clean protein samples (trichloroacetic acid

precipitation, two-step precipitation, ultra filtration, and dialysis) and taking into account their advantages and disadvantages, depleted samples were cleaned by the 2-D clean-up kit (GE Healthcare, Waukesha, WI, USA) following the manufacturer's directions. At the end, the protein pellet was rehydrated with a solution containing urea at 7M, thiourea at 2M, and 2% CHAPS, suitable for 2D-PAGE (all three reagents from GE Healthcare, Waukesha, WI, USA).

4.3. Two-dimensional gel electrophoresis

Two-dimensional gel electrophoresis is a protein analysis method that allows the separation of proteins according to two innate characteristics: isoelectric point (IP) and molecular weight (MW).

In the first dimension, also described as isoelectric focusing (IEF), proteins are separated in a gel strip that contains a predetermined pH range, allowing the proteins to separate according to their isoelectric point. In the second dimension, the strip is placed on a SDS-PAGE gel in which the focused proteins migrate according to their molecular weight. The final result is a SDS-PAGE gel where, after staining, protein spots are observed.

4.3.1. Isoelectric focusing

Treated BALF samples (n=60), equally distributed among four groups of patients (controls, LC, COPD, and LC&COPD), were independently submitted to IEF in 7 cm IPG DryStrips with a 3-11 NL pH range.

A mixture of 75 µg of protein from each sample, DeStreak rehydration solution, and 0.5% of 3-11 NL pH IPG Buffer, in a final volume of 125 µl, was prepared. This mixture was then applied to the strip holder. The strip's cover foil was carefully removed from the positive end and the strip placed, gel-side down and positive end first, onto the strip holder, distributing the solution evenly under the strip. The strip was then overlaid with cover fluid and the strip holders placed on the Ettan IPGphor II for the IEF protocol (Table 2). All reagents and instruments were from GE Healthcare (Waukesha, WI, USA).

Table 2: Isoelectric focusing protocol

Step	Voltage	Duration
1	30 V	14 hours
2	500 V	250 V/hour
3	1000 V	500 V/hour
4	5000 V	8500 V/hour
5	5000 V	30 minutes

V: volts

4.3.2. Strip equilibration

After the IEF protocol and before the SDS-PAGE protein separation, the focused IPG strips required an equilibration procedure in an appropriate buffer. Equilibration solution was prepared with double-distilled water, 6M of urea, 75 mM of 1,5 M Tris-HCl (pH 8.8) solution, 29.3% of glycerol (Sigma-Aldrich, St. Louis, MO, USA), 2% SDS (Sigma-Aldrich, St. Louis, MO, USA), and trace amounts of bromophenol blue (Bio-Rad, Hercules, CA, USA).

Focused IPG strips were equilibrated with dithiothreitol (DTT) (Bio-Rad, Hercules, CA, USA) equilibration solution (100 mg per 10 ml of equilibration solution) for 7.5 minutes (2x), to break disulphide bonds, and later with iodoacetamide (Bio-Rad, Hercules, CA, USA) equilibration solution (250 mg per 10 ml of equilibration solution) for 7.5 minutes (2x) to prevent the reformation of the disulphide bonds.

4.3.3. 2D- PAGE

The second dimension was performed in 12.5% acrylamide mini gels (1.5 mm) that were prepared for each strip and then topped with water. For one gel, 1587.5 µl of double-distilled water, 1250 µl of 1.5M Tris-HCl (pH 8.8), and 2087.5 µl of 30% acrylamide (Bio-Rad, Hercules, CA, USA) were mixed in an appropriate container. Then, 50 µl of SDS at 10% (Sigma-Aldrich, St. Louis, MO, USA), 5 µl of TEMED (Bio-Rad, Hercules, CA, USA), and 25 µl of ammonium persulfate (Bio-Rad, Hercules, CA, USA) at 10% were mixed in. The mixture was then poured into the 1.5 mm gel slides to solidify.

Electrophoresis Tris/Glycine/SDS Buffer running buffer (1x) was prepared (Bio-Rad, Hercules, CA, USA) and used to produce a 0.5% low melting agarose solution (USB, Affymetrix, Santa Clara, CA, USA) with bromophenol blue. This solution was gently

heated and placed on top of the cast gels. Equilibrated strips were quickly dipped on running buffer 1x for lubrication and then placed on the 0.5% agarose gel, with gel-side to the front, and gently pushed down until they had a good contact with the 12.5% acrylamide gel below.

Electrophoresis was performed at 20 mA per gel in mini gel electrophoresis cells (Bio-Rad, Hercules, CA, USA). Five 2D-PAGEs were performed at a time, each one with the respective sample from every group (control, COPD, LC, and LC&COPD) and the fifth gel with a 2D standard (Bio-Rad, Hercules, CA, USA) to act as an electrophoresis control.

4.3.4. Gel staining and imaging

After the second dimension, gels required staining to allow the visualization of the spots and subsequent analysis.

Several staining methods were tested, such as coomassie, silver (GE Healthcare, Waukesha, WI, USA), and SYPRO® (Bio-Rad, Hercules, CA, USA). However, considering that protein spots, after being removed from the SDS-PAGE gel, were to be submitted to MS analysis, the staining method needed to be compatible with this procedure. SYPRO® staining is MS friendly, therefore gels were SYPRO® stained according to the manufacturer's instructions. Gels were fixed for a minimum of 30 minutes with a water mixture of 10% ethanol (Panreac, Barcelona, Spain) and 7% glacial acetic acid (Sigma-Aldrich, St. Louis, MO, USA). Then, gels were covered with SYPRO® and left overnight in the dark to stain. Afterwards, gels were de-stained in the dark for 30 minutes in a water mixture of 10% ethanol and 7% glacial acetic acid. Gels were then washed in the dark with water for 10 minutes and revealed in a Typhoon 9400 (GE Healthcare, Waukesha, WI, USA). Gels were preserved until further use in a 7% glacial acetic acid solution.

A diagram of the 2D-page protocol can be found below on Figure 10.

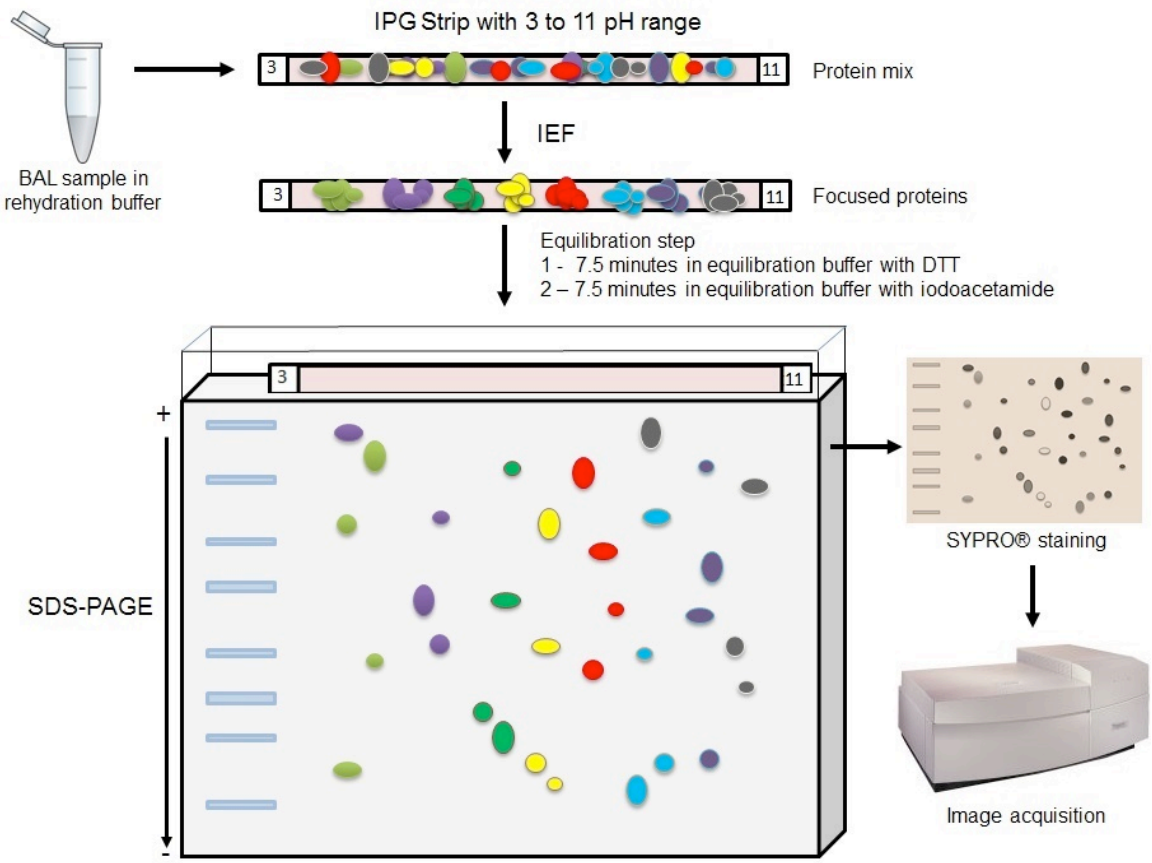


Figure 10: 2D-PAGE protocol.

4.3.5. Gel image analysis

The process of analysing 2D-PAGE images, detection of spots, spot matching, and semi-quantitative statistical analysis was performed using the Progenesis SameSpots software v4.0 (Nonlinear Dynamics, Newcastle upon Tyne, UK).

Gel images from the control, COPD, LC, and LC&COPD groups (separated in the same electrophoresis run) were uploaded to the program and submitted to image quality assessment to correct possible positional errors (Figure 11A). Gel images were then aligned to allow the matching of protein spots (Figure 11B). After computerized matching, detected spots were manually edited for greater accuracy. Gel images were analysed and the detection of the differentially expressed protein spots was performed using the INCA volume test (Figure 11C). Only proteins that were 2 fold or higher differentially expressed were considered for further analysis. Protein spots of interest were extracted from the stained gels using the ProXcision robot (Perkin Elmer, Waltham, MA, USA) and sent for MS analysis (Figure 11D).

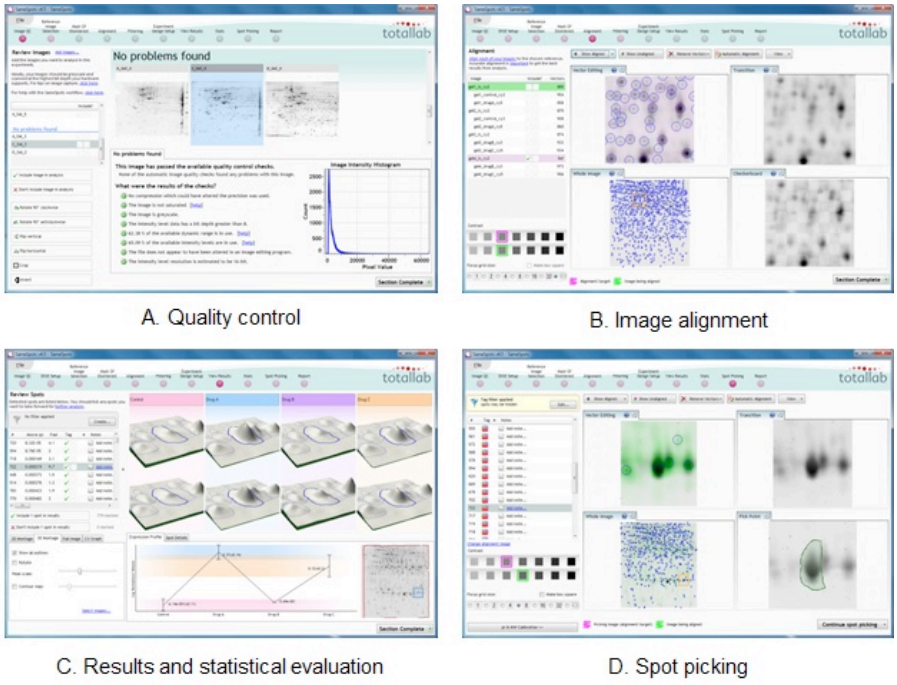


Figure 11: Gel image analysis overview. Adapted from Progenesis SameSpots v4.0

4.4. Mass spectrometry

4.4.1. MS sample preparation and analysis

Following their removal, spots were washed with water and the digestion protocol employed was the one used by Shevchenko and co-workers with minor variations (117). Gel spots were washed with 50 mM ammonium bicarbonate in 50% methanol and 70% acetonitrile and dried in a vacuum concentrator. Trypsin (Promega) was added to carry out protein digestion at 37°C overnight. Peptide extraction was carried out using 60% acetonitrile and 0.1% formic acid. An aliquot of the digestion solution was mixed with α-Cyano-3-hydroxycinnamic acid (Sigma) in 30% aqueous acetonitrile, 2-propanol 15%, and 0.1% trifluoroacetic acid. This mixture was deposited on a MALDI -TOF plate (Applied Biosystems) and left to crystalize at room temperature. After sample preparation, the MALDI plate was then placed into the mass spectrometer and irradiated. When the laser strikes the matrix crystals, the matrix crystals absorb the energy to ionize the sample. After ion creation in the MALDI ion source, they were accelerated into the TOF mass by flying through a field-free drift tube prior to impacting the detector. The total flight time of an ion traveling through the system is a combination of the flight time spent in the ion source and the flight time spent in the drift tube before hitting the detector. Larger molecules will reach the detector slower than smaller molecules. Once the ions hit the detector, spectra were recorded as the m/z of the ions and their relative intensities. MS

spectra and MS/MS data were obtained on a Voyager System DE-STR 7307 MALDI-TOF/TOF mass spectrometer (Applied Biosystems, Carlsbad, CA, USA).

4.4.2. Protein identification

Peptide matching and protein search were performed against the Swiss-Prot database using the Mascot search engine (www.matrixscience.com) with a mass tolerance over 50 parts per million. Protein scores higher than 60 indicated identity or extensive homology ($p<0.05$) and were considered significant.

From each spot, only proteins with a probability higher than 95% and at least two unique peptides matched were considered in the analysis (except for keratins which were not considered). The experimental molecular weight and isoelectric point of each identified protein were determined based on the location of the original spot in the 2D-PAGE gel by the Progenesis software.

A graphic representation of the proteomic analysis of BALF from LC and COPD patients by 2D-PAGE MALDI-TOF/TOF can be found below on Figure 12.

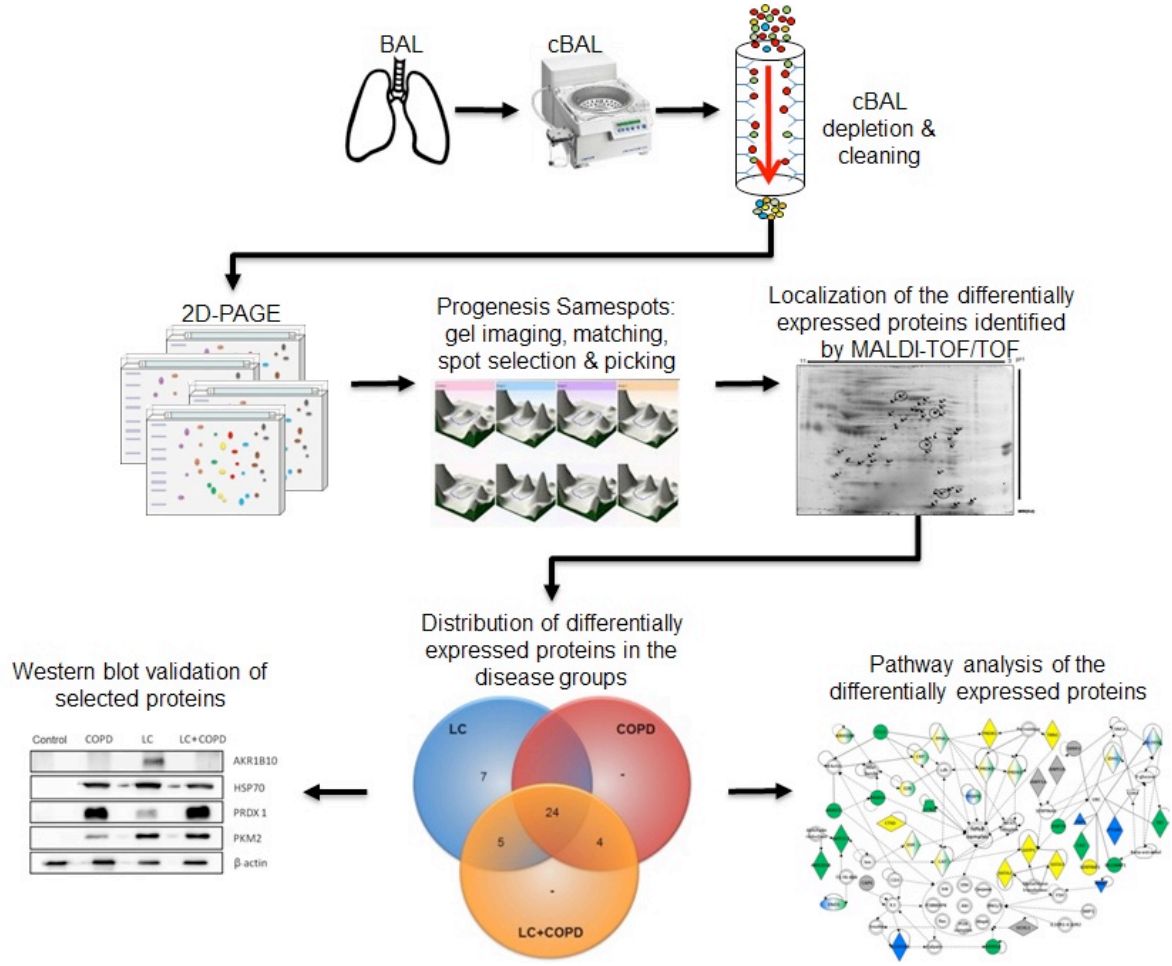


Figure 12: Workflow of the employed 2D-PAGE MALDI-TOF/TOF methodology.

4.5. Antibody arrays

Sixty BALF samples from the test cohorts were applied to the commercially available antibody array Quantibody® Human Cytokine Antibody Array 1000 (RayBiotech, Norcross, GA, USA). The human cytokine array assay, that allowed the simultaneous detection of 40 cytokines and 40 growth factors (Table 3), was performed according to the instructions provided by the manufacturer.

The glass chip was removed from the box, and let to equilibrate to room temperature inside the sealed plastic bag for 20-30 minutes. Afterwards, the slide was removed from the plastic bag, the cover film was peeled off, and the glass slide was allowed to air dry at room temperature for another 2 hours. In order to block the glass slide, sample diluent was applied to each well and left to incubate at room temperature for 30 min. The sample diluent was decanted from all wells and the cytokine/growth factor standards or BALF samples were added to each corresponding well to incubate at room temperature for 2 hours. Samples were removed from the wells and these were washed 5 times (5 minutes each) with 150 µl of 1x Wash Buffer I at room temperature with gentle shaking, completely removing the wash buffer in each wash step. Detection antibodies were reconstituted by adding 1.4 ml of sample diluent to each tube. Eighty µl of the detection antibody cocktail were added to each well and allowed to incubate at room temperature for 2 hour. Samples were again removed from the wells and these were washed 5 times with 150 µl of 1x Wash Buffer I and then 2 times with 150 µl of 1x Wash Buffer II at room temperature with gentle shaking, completely removing the wash buffer in each wash step. After a brief spin down, 1.4 ml of sample diluent were added to the Cy3 equivalent dye-conjugated streptavidin tube and mixed gently. Eighty µl of Cy3 equivalent dye-conjugated streptavidin were then applied to each well. The device was then covered with aluminium foil to avoid exposure to light and incubated at room temperature for 1 hour. Samples were removed from each well, and washed 5 times with 150 µl of 1x Wash Buffer I at room temperature with gentle shaking, completely removing the wash buffer in each wash step. The device was disassembled by pushing clips outward from the slide side and the slide was carefully removed from the gasket. The slide was then placed inside the slide Washer/Dryer (a 4-slide holder/centrifuge tube), 30 ml of 1x Wash Buffer I was added to cover the whole slide, and then gently shaken at room temperature for 15 minutes. The Wash Buffer I was removed and replaced by 30 ml of 1x Wash Buffer II and gently shaken at room temperature for 5 minutes. Water droplets were removed from the slide by centrifuging the slide Washer/Dryer at 1,000 rpm for 3 minutes without its cap. The intensity of each signal was visualized through the use of a laser scanner model

GenePix 4100 A (Molecular Devices, Sunnyvale, CA, USA), which was also used to perform data extraction.

Table 3: Cytokines and growth factors evaluated by the antibody array

Cytokine symbol	Cytokine name	Growth factor symbol	Growth factor name
BLC	B-cell lymphoma	**AR**	Androgen receptor
Eotaxin	Eotaxin	**BNDF**	Brain-derived neurotrophic factor
Eotaxin2 (CCL24)	Eotaxin 2	**bFGF**	Basic fibroblast growth factor
G-CSF	Granulocyte colony-stimulating factor	**BMP4**	Bone morphogenetic protein 4
GM-CSF	Granulocyte-macrophage colony-stimulating factor	**BMP5**	Bone morphogenetic protein 5
I309 (CCL1)	Chemokine (C-C motif) ligand 1	**BMP7**	Bone morphogenetic protein 7
ICAM-1	Inter-Cellular Adhesion Molecule 1	**B-NGF**	Beta nerve growth factor
IFN γ	Interferon-gamma	**EGF**	Epidermal growth factor
IL-1α	Interleukine-1 alpha	**EGFR**	Epidermal growth factor receptor
IL-1β	Interleukine-1 beta	**EG-VEGF**	Endocrine gland-derived vascular endothelial growth factor
IL-1Ra	Interleukin-1 receptor antagonist	**FGF4**	Fibroblast growth factor 4
IL-2	Interleukine-2	**FGF7**	Fibroblast growth factor 7
IL-4	Interleukine-4	**GDF15**	Growth differentiation factor 15
IL-5	Interleukine-5	**GDNF**	Glial cell-derived neurotrophic factor
IL-6	Interleukine-6	**GH**	Growth hormone
IL-6sR	Interleukine-6 soluble receptor	**HB-EGF**	Heparin-binding EGF-like growth factor
IL-7	Interleukine-7	**HGF**	Hepatocyte growth factor
IL-8	Interleukine-8	**IGFBP1**	Insulin-like growth factor-binding protein 1
IL-10	Interleukine-10	**IGFBP2**	Insulin-like growth factor-binding protein 2
IL-11	Interleukine-11	**IGFBP3**	Insulin-like growth factor-binding protein 3
IL-12p40	Interleukine-12 p40	**IGFBP4**	Insulin-like growth factor-binding protein 4
IL-12p70	Interleukine-12 p70	**IGFBP6**	Insulin-like growth factor-binding protein 6
IL-13	Interleukine-13	**IGF-1**	Insulin-like growth factor 1
IL-15	Interleukine-15	**Insulin**	Insulin
IL-16	Interleukine-16	**MCFR**	Macrophage chemotactic factor receptor
IL-17	Interleukine-17	**NGFR**	Nerve growth factor receptor
MCP-1 (CCL2)	Monocyte chemotactic protein-1/ Chemokine (C-C motif) ligand 2	**NT-3**	Neurotrophin 3
MCSF	Mouse stem cell factor	**NT-4**	Neurotrophin 4
MIG (CXCL9)	Monokine induced by IFN-Gamma/ Chemokine (C-X-C motif) ligand 9	**OPG**	Osteoprotegerin
MIP-1α (CCL3)	Macrophage inflammatory protein 1 alpha/ Chemokine (C-C motif) ligand 3	**PDGFAA**	Platelet-derived growth factor AA
MIP-1β (CCL4)	Macrophage inflammatory protein 1 beta/ Chemokine (C-C motif) ligand 4	**PIGF**	Phosphatidylinositol-glycan biosynthesis class F
MIP-1δ	Macrophage inflammatory protein 1 delta	**SCF**	Stem cell factor
PDGFBB	Platelet-derived growth factor BB	**SCFR**	Stem cell factor receptor
Rantes (CCL5)	Regulated upon Activation, Normal T-cell Expressed, and Secreted/ Chemokine (C-C motif) ligand 5	**TGF-α**	Transforming growth factor alpha
TIMP-1	Tissue inhibitor of metalloproteinases 1	**TGF-β1**	Transforming growth factor beta 1
TIMP-2	Tissue inhibitor of metalloproteinases 2	**TGF-β3**	Transforming growth factor beta 3
TNF-α	Tumour necrosis factor alpha	**VEGF**	Vascular endothelial growth factor
TNF-β	Tumour necrosis factor beta	**VEGFR2**	Vascular endothelial growth factor receptor 2
TNFRI	Tumour necrosis factor receptor 1	**VEGFR3**	Vascular endothelial growth factor receptor 3
TNFRII	Tumour necrosis factor receptor 2	**VEGFD**	Vascular endothelial growth factor D

4.6. Western blot

Western blot analyses were performed to allow the validation of results from the previous experiments,

The BALF samples were prepared with Laemmli buffer 4X that contained a Tris-HCl 250 mM solution with a pH of 6.8, SDS at 4% (w/v), glycerol at 60% (v/v), and 2-mercaptoethanol at 10% (v/v). They were submitted to SDS-PAGE separation after boiling for 5 min at 100ºC to denature proteins. Depending on the molecular weight of the protein, SDS-PAGE separating gels (1.5 mm) were made with 7.5, 10, and 12.5% acrylamide as previously described. Gels were covered with water and left to polymerize. As soon as the gels were set, a stacker gel with 4% acrylamide was made and placed on top of the separating gels with a comb to create wells. The electrophoresis cell was filled with 1x Tris/Glycine/SDS Buffer running buffer (Tris-Base 25mM, GIycine 192mM, SDS 0.1% (w/v)) (Bio-Rad, Hercules, CA, USA) and the electrophoresis began. The run started at a low voltage (30-50 V) until the samples had left the stacker gel. Then, the voltage was raised to 100-120 V and was kept constant until the end of the run.

Gels were then transferred to polyvinylidene fluoride (PVDF) membranes (Bio-Rad, Hercules, CA, USA) in a transfer cell with 1x Tris/Glycine transfer buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, and 20% methanol) (Bio-Rad, Hercules, CA, USA), at 4ºC for 3-4 hours. The membranes were then removed from the transfer cassettes and blocked for 1 hour in Tris-buffered saline (TBS) with 0.05% Tween-20 (Sigma-Aldrich, St. Louis, MO, USA) (0.05% TBST) and with 5% BSA (Sigma-Aldrich, St. Louis, MO, USA). Afterwards, the blots were incubated overnight at 4ºC with their corresponding primary antibody (in TBST with 1% BSA), in the appropriate proportion. The used antibodies and their respective ratios can be seen on Table 4. Membranes were then washed to remove the excess primary antibodies and the secondary antibody, in the suitable ratio, was applied to the membranes for 1 hour in TBST with 1% BSA, at room temperature. Membranes were then washed to remove the excess secondary antibodies and placed on water.

Membranes were revealed using enhanced chemiluminescence ECL according to the instructions provided by the manufacturer (GE Healthcare, Waukesha, WI, USA) and the membranes were exposed in the image analyser Mini LAS-3000 (Fujifilm, Tokyo, Japan). The relative protein levels were calculated by comparison to the amount of β-actin protein (1:1000 Abcam, Cambridge, MA, USA). The experiments were repeated three times independently. The analysis of the different expression values of the 8 proteins

obtained by western blot was performed by densitometry. The densitometry analysis of the scanned blots was done using Image J software (118). Results were normalized with the expression of the control protein (β-actin).

Table 4: Western blot validation antibodies and used ratios

Purpose	Name	Ratios
2D-PAGE validation	Anti-HSP70, Anti-AKR1B10, Anti-PRDX1 (Epitomics, Burlingame, CA, USA), Anti-PKM2, Anti-NF-κBp65, Anti-pNF-κBp65 (Cell Signaling, Beverly, MA, USA)	1/1000
Antibody array validation	Anti-IL-11, Anti-CCL1 (I309) (Abcam, Cambridge, MA, USA)	1/200

4.7. Enzyme-Linked Immunosorbent Assay

Proteins previously validated by western blot were evaluated by ELISA in a first validation cohort of 139 patients and later in a second validation cohort of 160 patients. BALF samples were applied as instructed by the manufacturer, using sandwich ELISA kits (RayBiotech, Norcross, GA, USA).

All reagents were placed at room temperature (18 - 25°C) before use. Standards were prepared at different concentrations to create a standard curve. One hundred ml of each standard and samples were added into their respective wells. All samples and standards were assayed in duplicate and the same plate was used. The ELISA plate was covered and left to incubate for 2.5 hours at room temperature with gentle shaking. The wells were washed 4 times with 1x Wash Solution by filling each well with Wash Buffer (300 µl). After the last wash, the plate was inverted and blotted against clean paper towels to remove all excess Wash Solution. The 1x prepared biotinylated antibody was added to each well (100 µl) and left to incubate for 1 hour at room temperature with gentle shaking. The wells were then washed 4 times, as previously described. Prepared Streptavidin solution (100 µl) was added to each well and incubated for 45 minutes at room temperature with gentle shaking. The wells were then washed 4 times, as previously described. Finally, 100 µl of TMB One-Step Substrate Reagent were placed on each well. Incubation was performed for 30 minutes at room temperature in the dark with gentle shaking. The Stop Solution was then added to each well and plates were read at 450 nm

immediately on an EMax Microplate Reader (Molecular Devices, Minneapolis, Minn., USA). The mean absorbance for each set of duplicate standards and samples was subtracted by the average zero standard optical density. The standard curve was created on an Excel spread sheet, with the standard concentration on the x-axis and absorbance on the y-axis. Sample concentration was calculated by using the equation provided by linear trend graph.

A graphic representation of the proteomic analysis of BALF from LC and COPD patients by antibody arrays and ELISA can be found below on Figure 13.

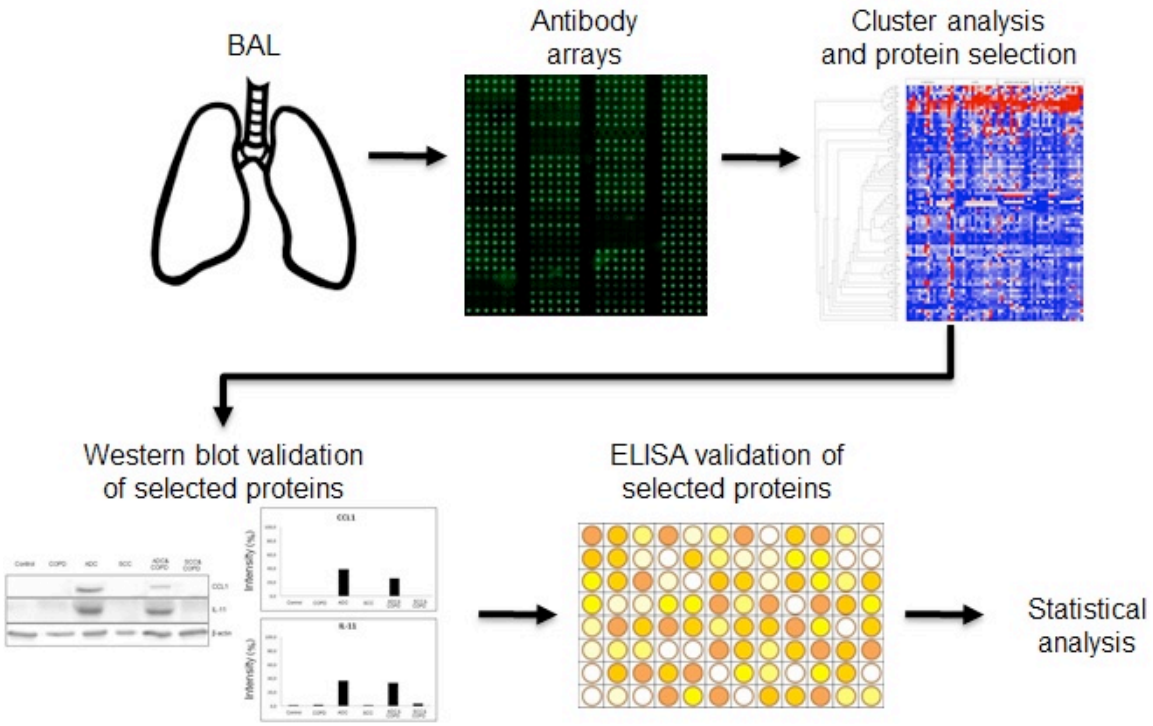


Figure 13: Workflow of the employed antibody arrays and ELISA methodology.

4.8. Bioinformatics analysis and statistical analysis

4.8.1. 2D-PAGE bioinformatics analysis

The proteins selected from the 2D-PAGE experiments and subsequently identified by MS were analysed by the Ingenuity Pathway Analysis (IPA) web-based application (www.ingenuity.com, Ingenuity® Systems, Redwood City, CA, USA). It uses a database that contains up-to-date information on thousands of genes and proteins, over one million biological interactions, and one hundred canonical pathways. This allows the identification of relationships, biological mechanisms, functions, and relevant pathways associated with the studied proteins. The IPA Core analysis allows the assessment of the pathways and biological processes most significantly altered in our dataset, the associated diseases and

disorders, the most important molecular and cellular functions, and the identification of the signalling and metabolic pathways most relevant in our data set. Only proteins identified by the IPA software were submitted to core analysis, that revealed the most important biological functions and pathways associated with our proteins, and also created interaction networks between them.

4.8.2. Antibody arrays statistical analysis

All continuous variables of patients' characteristics were expressed as a median for each variable (interquartile [IQR] range) and categorical variables as number of cases and percentage. The mean and SD of selected proteins were calculated for every patient in all study groups. The differences among groups were evaluated by the Mann–Whitney U test. Differences were considered statistically significant if their p value was ≤0.05 (*p≤0.05, **p<0.01, ***p<0.001).

Statistical analysis was performed using the Statistical Package for the Social Sciences software (SPSS 17, Chicago, Illinois, USA).

4.8.3. Antibody arrays expression data analysis

Hierarchical supervised clustering analysis was performed using the function UPGMA (Unweight Pair Group Method with Arithmetic Mean). Before statistical analysis, protein expression levels were standardized, protein by protein, across all conditions using the medians and standard deviation (SD) values. The sample conditions were clustered using euclidean distance metric tests. The results were visualized and analysed with Babelomics 4.2 (babelomics.bioinfo.cipf.es) (119). The expression level of each protein, relative to its median expression level across all conditions, was represented by a colour, with red representing expression greater than the median, blue representing expression lesser than the median, and various intermediate colour intensities representing the magnitude of variance from the median.

4.8.4. Diagnostic test validity analysis

Receiver operating characteristics curves were constructed to assess sensitivity, specificity, and respective areas under the curve (AUCs) with 95% confidence interval (CI) of possible LC biomarkers. We investigated the optimum cut-off value, which was chosen

in order to reflect the best cooperation between sensitivity, specificity, positive predictive value, negative predictive value, and likelihood ratio to predict the diagnosis of lung adenocarcinoma. To test the diagnostic accuracy of several biomarkers measured together, the chi-squared test was used.

RESULTS

5. RESULTS

5.1. First specific objective: Identification of differently expressed proteins in BALF from patients with LC, COPD, COPD with LC, and without LC or COPD by comparative proteomic analysis.

5.1.1. Selection of patients

Sixty samples from the patients in the test cohort were divided into four groups: control (without LC or COPD), COPD, LC, and LC&COPD. The principal characteristic from test cohort, were all male, with a median age of 61 (41-80) years. The majority of patients were current smokers, with similar pack-years 30 (28.6- 37.2). In the COPD group, the majority of patients had moderate and severe cases of COPD, and in the LC&COPD group, over half of patients had mild COPD. The adenocarcinoma histology was the predominant histology of LC in the LC and LC&COPD groups (Table 5). Overall patient characteristics were well matched among the four study groups.

Table 5: Characteristics of the patients analysed in the test cohort by the 2D-PAGE methodology

	Controls n=15	**COPD n=15**	**LC n=15**	**LC&COPD n=15**
Gender				
Male	100.0% (15)	100.0% (15)	100.0% (15)	100.0% (15)
Female	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
Age (years) [range]	61.3 [41.0-80.0]	61.5 [45.0-78.0]	60.7 [46.0-69.0]	60.7 [49.0-68.0]
Smoking status				
Smokers	73.3% (11)	53.3% (8)	53.3% (8)	80.0% (12)
Ex-smokers	26.7% (4)	46.7% (7)	46.7% (7)	20.0% (3)
Pack-years [range]	21.82 [18.6-25.4]	32.20 [28.6-37.2]	35.21 [31.3-41]	30.78 [27.5-33.5]
COPD				
Mild	0.0% (0)	20.0% (3)	0.0% (0)	53.3% (8)
Moderate	0.0% (0)	33.3% (5)	0.0% (0)	26.7% (4)
Severe	0.0% (0)	26.7% (4)	0.0% (0)	0.0% (0)
Very severe	0.0% (0)	20.0% (3)	0.0% (0)	20.0% (3)
LC Histology				
Adenocarcinoma	0.0% (0)	0.0% (0)	73.3% (11)	66.7% (10)
SCC	0.0% (0)	0.0% (0)	26.7% (4)	33.3% (5)

SCC: squamous cell carcinoma

5.1.2. Proteomic profiling

The identification of the differentially expressed proteins between patients with LC and COPD was accomplished by analysing BALF samples from 4 distinct groups of patients (controls, COPD, LC, and LC&COPD) by 2D-PAGE.

The software Progenesis SameSpots carried out the proteomic analysis of the 2D-PAGE gels. This program applied a t-test with a value of $p < 0.05$ to show the proteins with significant expression differences. These proteins were then later detected by MALDI-TOF/TOF and identified by comparing the obtained MS data to the one integrated in the database Swiss-Prot using the MASCOT software. Once the protein had been identified, different variables such as the isoelectric point, molecular weight, peptide matching, and sequence coverage were use to ascertain their correct identification.

A total of 123 proteins were identified across 60 gels. Out of these, 40 proteins were differentially expressed between the groups of patients with LC and/or COPD and

the control group, presenting changes in expression of over 2 fold and a p value <0.05. The spots corresponding to these 40 selected proteins can be found on Figure 14. Circled spots correspond to proteins later selected for validation by western blot. The detailed information on these proteins can be found on (Table 6).

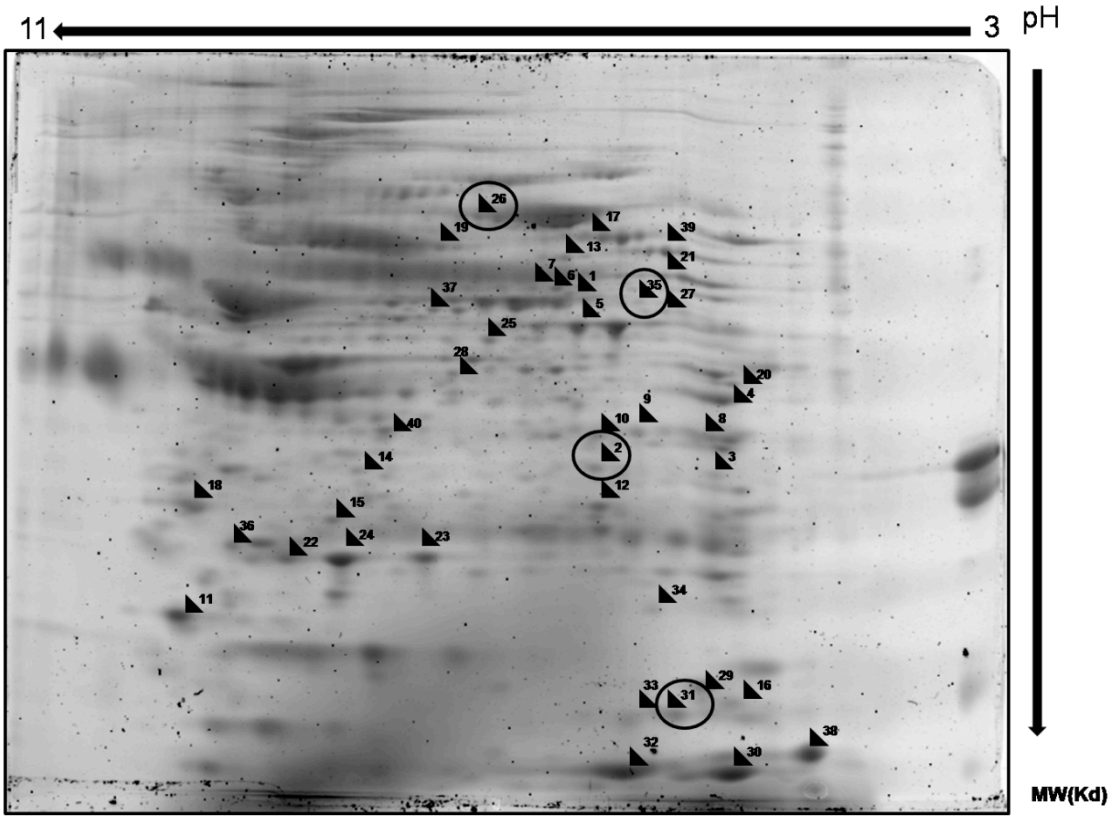


Figure 14: A 7 cm Sypro® stained 2D-PAGE gel of proteins in the 3-11 non lineal pH range of a representative patient. Circled spots correspond to proteins selected for western blot validation. Identification of protein spots can be seen on Table 6.

Table 6: Differentially expressed proteins identified by 2D-PAGE-MALDI-TOF/TOF MS

Protein name *(spot number)*	**Protein symbol**	**gi| number**	**Protein MW**	**Protein IP**	**Peptide count**	**Protein score***	**Total ion score***
Aldehyde dehydrogenase *(1)*	ALDH3A1	178375	50702.9	5.99	15	445	339
Aldo-keto reductase family 1 member B10 *(2)*	AKR1B10	4503285	37111.1	7.13	7	108	70
Aldo-keto reductase family 1 member C3 *(3)*	AKR1C3	24497583	37229.0	8.06	10	248	186
Aldolase A *(4)*	ALDOA	28614	39706.4	8.34	11	229	156
Alpha enolase *(5)*	ENO1	2661039	36627.7	6.53	9	128	70
Alpha amylase 1A precursor *(6)*	AMY1A	40254482	58415.2	6.47	18	469	326
Alpha amylase 2A *(7)*	AMY2A	224980	58443.2	6.47	15	323	221
Annexin 1 *(8)*	ANXA1	4502101	38918.1	6.57	17	495	354
Annexin 2 *(9)*	ANXA2	18645167	38779.9	7.57	20	326	159
Annexin 5 *(10)*	ANXA5	4502107	35971.4	4.94	19	472	301
Calsyphosin 2 *(11)*	CAPS2	4757908	21068.3	4.74	11	543	460
Carbonic anhydrase 1 *(12)*	CA1	4502517	28909.4	6.59	14	770	627
Catalase *(13)*	CAT	4557014	59946.8	6.90	13	310	234
Cathepsin D pro-protein *(14)*	CTSD	4503143	45036.8	6.10	8	135	95
Tubulin polymerization-promoting protein family member 3 *(15)*	TPPP3	4680715	19144.7	9.10	8	79	30
Cofilin-1 *(16)*	CFL1	5031625	18718.7	8.22	6	98	57
Complement C3 *(17)*	C3	119589476	144417.1	8.24	25	472	332
C-reactive protein *(18)*	CRP	1942435	23146.7	5.28	8	341	281
Ezrin *(19)*	EZR	46249758	69312.7	5.94	18	250	175
Fructose-1,6-biphosphatase *(20)*	FBP1	3293553	37058.9	6.60	6	124	101
Glutathione reductase *(21)*	GSR	119583848	61464.6	8.71	7	131	108
Glutathione S-transferase alpha 1 subunit *(22)*	GSTA1	163310943	25628.7	8.72	15	384	268
Glutathione S-transferase alpha 2 subunit *(23)*	GSTA2	257476	25589.6	8.81	7	105	70
Glutathione S-transferase pi 1 *(24)*	GSTP1	4504183	23569.1	5.43	10	633	541
Glycogen phosphorylase *(25)*	PYGM	225897	97386.8	6.57	17	515	436
Heat shock protein 70 *(26)*	HSP70	4529893	70280.1	5.48	12	325	269
Isocitrate dehydrogenase 1 *(27)*	IDH1	89573979	42091.0	6.19	8	62	29
Serpin peptidase inhibitor, clade B, member 1 *(28)*	SERPINB1	13489087	42828.7	5.90	18	403	274
Lipocalin 2 *(29)*	LCN2	119608155	28089.4	9.53	9	513	426
Peptidylprolyl isomerase A *(30)*	PPIA	1633054	18097.9	7.82	10	260	159
Peroxiredoxin 1 *(31)*	PRDX1	55959887	19134.7	6.41	8	170	98
Peroxiredoxin 5 *(32)*	PRDX5	6166493	22261.6	8.85	11	638	537
Peroxiredoxin 2 isoform a *(33)*	PRDX2	32189392	22049.3	5.66	12	451	325
Phosphatidylethanolamine-binding protein 4 *(34)*	PEBP4	4505621	21157.7	7.01	11	311	209
Pyruvate kinase 2 *(35)*	PKM2	35505	58411.2	7.58	22	416	249
Rho GDP dissociation inhibitor 2 *(36)*	ARHGDIB	56676393	23030.6	5.10	7	215	170
Selenium-binding protein 1 *(37)*	SELENBP1	197097476	52987.9	6.13	12	278	207
Thioredoxin *(38)*	TXN	135772	12345.0	7.93	10	241	203
Transketolase *(39)*	TKT	388891	68528.0	7.89	17	405	246
Ubiquitin carboxyl-terminal hydrolase isozyme L1 *(40)*	UCHL1	21361091	25150.6	5.33	7	126	76

MW: molecular weight; IP: isoelectric point; *: 100% confidence

The distribution of the differentially expressed proteins across the LC, COPD, and LC&COPD groups can be seen on Figure 15. The Venn diagram depicted in this figure represents the shared proteins between all the disease groups and also the unique proteins that characterize each disease group individually.

Out of the 40 differentially expressed proteins among the disease groups (compared to the control group), a total of 24 proteins were found to be over or under-expressed in the three disease groups (LC, COPD, and LC&COPD). Two proteins, CAPS2 and CFL1, were found to be under-expressed in the COPD group and over-expressed in the LC and LC&COPD groups. A set of 7 proteins were found to be differentially expressed only in the LC group AKR1B10, ALDOA, CTSD, EZR, FBP1, and TKT were over-expressed, and SELENBP1 was under-expressed in this group of patients, compared to the control group. Five proteins were observed as differentially expressed in both the LC and the LC&COPD group: ALDH3A1, AKR1C3, PYGM, PKM2, and PPIA were over-expressed in these groups compared to the control group. Catalase, PRDX1, PRDX2, and PRDX5 were found to be over-expressed in the COPD and LC&COPD groups, compared to the control group. The fold variations between the control group and the disease groups can be found on (Table 7).

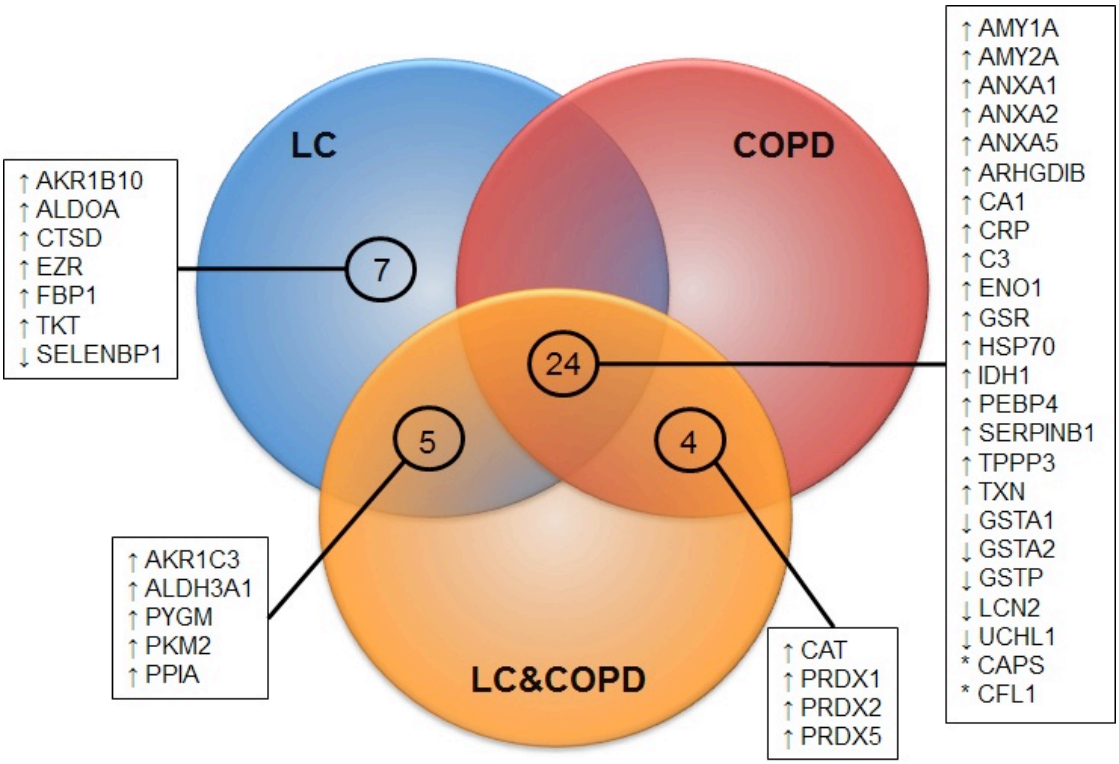


*Figure 15: Venn diagram showing the over or under-expressed proteins in each pathological group, compared to the control group. The ↑ sign precedes over-expressed proteins and the ↓ sign precedes under-expressed proteins. The * sign marks proteins under-expressed in the COPD group but over-expressed in the LC and LC&COPD groups*

Table 7: Expression fold variation of the differentially expressed proteins in the disease groups compared to the control group

Protein symbol	COPD Expression fold variation	LC Expression fold variation	LC&COPD Expression fold variation
ALDH3A1	-	↑ 2.9	↑ 2.9
AKR1B10	-	↑ 4.0	-
AKR1C3	-	↑ 2.5	↑ 2.1
ALDOA	-	↑ 2.0	-
ENO1	↑ 2.0	↑ 2.5	↑ 2.3
AMY1A	↑ 2.1	↑ 2.3	↑ 2.2
AMY2A	↑ 2.5	↑ 2.2	↑ 2.9
ANXA1	↑ 2.3	↑ 2.2	↑ 2.3
ANXA2	↑ 2.5	↑ 2.3	↑ 2.7
ANXA5	↑ 2.4	↑ 2.2	↑ 2.1
CAPS2	↓ 2.2	↑ 2.8	↑ 2.6
CA1	↑ 2.2	↑ 2.5	↑ 2.7
CAT	↑ 2.5	-	↑ 2.8
CTSD	-	↑ 3.0	-
TPPP3	↑ 2.4	↑ 2.0	↑ 2.1
CFL1	↓ 3.0	↑ 2.2	↑ 2.8
C3	↑ 2.4	↑ 3.0	↑ 2.6
CRP	↑ 3.0	↑ 3.2	↑ 2.7
EZR	-	↑ 3.0	-
FBP1	-	↑ 2.5	-
GSR	↑ 2.8	↑ 3.1	↑ 2.0
GSTA1	↓ 3.0	↓ 3.2	↓ 3.8
GSTA2	↓ 2.5	↓ 3.0	↓ 2.6
GSTP1	↓ 2.2	↓ 2.4	↓ 2.5
PYGM	-	↑ 3.0	↑ 2.1
HSP70	↑ 4.0	↑ 3.1	↑ 3.8
IDH1	↑ 2.2	↑ 2.1	↑ 2.5
SERPINB1	↑ 2.7	↑ 2.6	↑ 2.5
LCN2	↓ 3.0	↓ 3.5	↓ 2.9
PPIA	-	↑ 2.2	↑ 2.6
PRDX1	↑ 5.0	-	↑ 4.2
PRDX5	↑ 2.3	-	↑ 2.4
PRDX2	↑ 3.0	-	↑ 2.9
PEBP4	↑ 2.9	↑ 2.3	↑ 2.6
PKM2	-	↑ 3.0	↑ 3.2
ARHGDIB	↑ 2.9	↑ 2.6	↑ 2.3
SELENBP1	-	↓ 2.2	-
TXN	↑ 2.4	↑ 2.3	↑ 2.1
TKT	-	↑ 2.5	-
UCHL1	↓ 2.7	↓ 2.2	↓ 2.8

↑- Over-expression; ↓-Under-expression

5.1.3. Selection of proteins and western blot validation

Once the 2D-PAGE experiments were concluded and the MALDI-TOF/TOF identification of the differentially expressed proteins was acquired, the next step was to validate our results. This was accomplished by western blot, the preferred method to confirm protein expression levels. Considering the obtained results from the 2D-PAGE analysis and taking into account the levels of expression and score of each protein present (Tables 6 and 7), we chose for the validation process one protein representative of each disease category LC, LC and COPD, COPD, without LC and COPD. The proteins validated by western blot were AKR1B10, HSP70, PRDX1, and PKM2. Figure 16 shows the expression of these proteins across all groups and the different intensities between them.

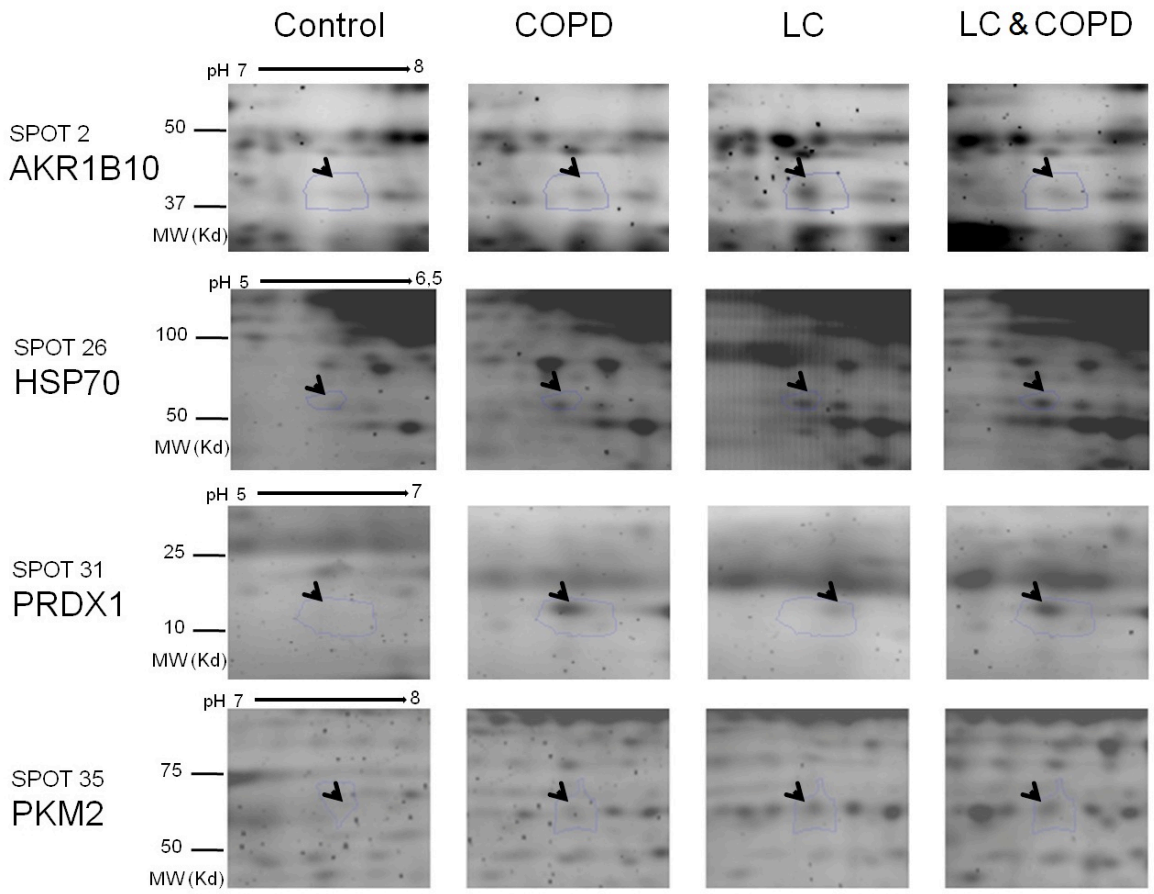


Figure 16: Different densities of spots 2, 26, 31, and 35 in each evaluated group of patients. Spots correspond to MS identified proteins ARK1B10, HSP70, PRDX1, and PKM2, respectively.

The western blot analysis of the selected proteins, together with the levels of the reference protein β-actin, are shown in Figure 17. The samples used for western blot validation were randomly selected.

Our results showed that the AKR1B10 protein had increased expression only in the LC group. On the other hand, HSP70 was over-expressed by all disease groups when compared to the control group. The PRDX1 protein showed an increased expression in the COPD and LC&COPD groups compared to the LC and control groups. Finally, PKM2 was clearly over-expressed in the BALF of the patients in the LC and LC&COPD groups compared to those in the COPD and control groups. The results of the western blot validation process confirmed those obtained by the 2D-PAGE experiments.

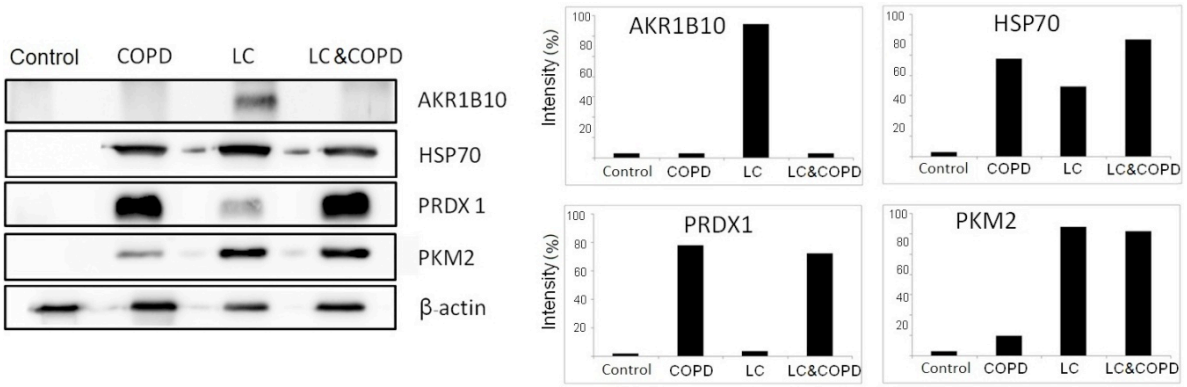


Figure 17: Western blots for AKR1B10, HSP70, PKM2, PRDX1, and β-actin in the control, COPD, LC, and LC&COPD groups. Differences in expression, normalized with the β-actin protein, are illustrated by a bar chart on the right.

5.1.4. LC and COPD pathway analysis

The connections and interactions between our 40 differentially expressed proteins and also the biological mechanisms, functions, and pathways in which they were possibly involved in were evaluated by the IPA software. In addition, the association between certain diseases and disorders and our data set were also analysed by IPA.

Core analysis was performed to include direct and indirect connections, and endogenous chemicals, and to consider only molecules and/or associations, creating an interaction network that can be seen on Figure 18. Cellular compartments as designated by IPA (gene ontology based), indicated that the majority of identified proteins were of cytoplasmatic origin.

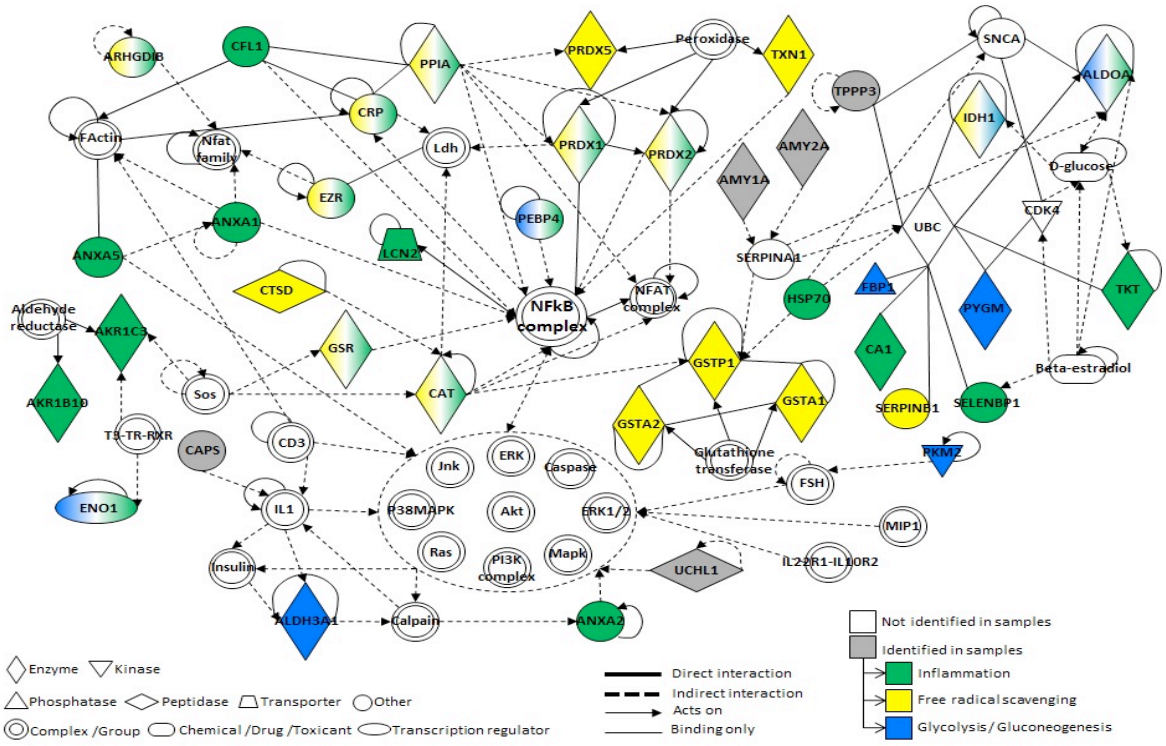


Figure 18: IPA interaction network considering the differentially expressed proteins between the control, LC, COPD, and LC&COPD groups.

The mapping of our differentially expressed proteins onto known molecular pathways and biological functions revealed three top processes occurring in the BALF samples from our four groups of patients: inflammation, free radical scavenging and oxidative stress response, and glycolysis and gluconeogenesis. Together with these findings, the recognition of the NF-kB complex as the major connection node between these three mechanisms was also accomplished.

The majority of identified proteins, a total of 22 proteins, were associated to inflammatory mechanisms (AKR1B10, AKR1C3, ALDOA, ANXA1, ANXA2, ANXA5, ARHGDIB, CA1, CAT, CFL1, CRP, ENO1, EZR, GSR, HSP70, LCN2, PEBP4, PPIA, PRDX1, PRDX2, SELENBP1, and TKT), 16 proteins were related to free radical scavenging and oxidative stress (ARHGDIB, CAT, CRP, CTSD, EZR, GSR, GSTA1, GSTA2, GSTP, IDH1, PPIA, PRDX1, PRDX2, PRDX5, SERPINB1, and TXN), and 8 proteins were found to be connected to the glycolysis and gluconeogenesis processes (ALDH3A1, ALDOA, ENO1, FBP1, IDH1, PEBP4, PKM2, and PYGM), as seen on Figure 19.

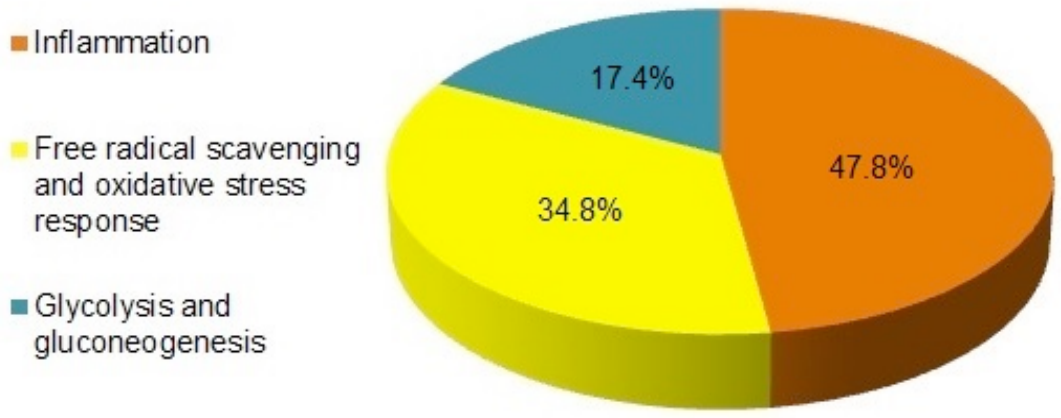


Figure 19: Top biological functions and pathways identified by IPA; inflammation related proteins (p value: $1,35*10^{-08}$–$1,42*10^{-02}$), free radical scavenging and oxidative stress response (p value: $4,93*10^{-11}$–$1,27*10^{-02}$), and glycolysis and gluconeogenesis related proteins (p value: $7,39*10^{-09}$–$1,58*10^{-02}$).

5.1.5. NF-kB functional analysis

NF-κB was the main connector between the differentially expressed proteins, even though this protein was not identified as differentially expressed among the four groups of patients analysed by 2D-PAGE.

Consequently, we analysed NF-kB in BALF samples from the control, COPD, LC, and LC&COPD groups of patients, by measuring the expression levels of the NF-κBp65 protein and its phosphorylated (activated) form, pNF-κBp65, by western blot (Figure 20).

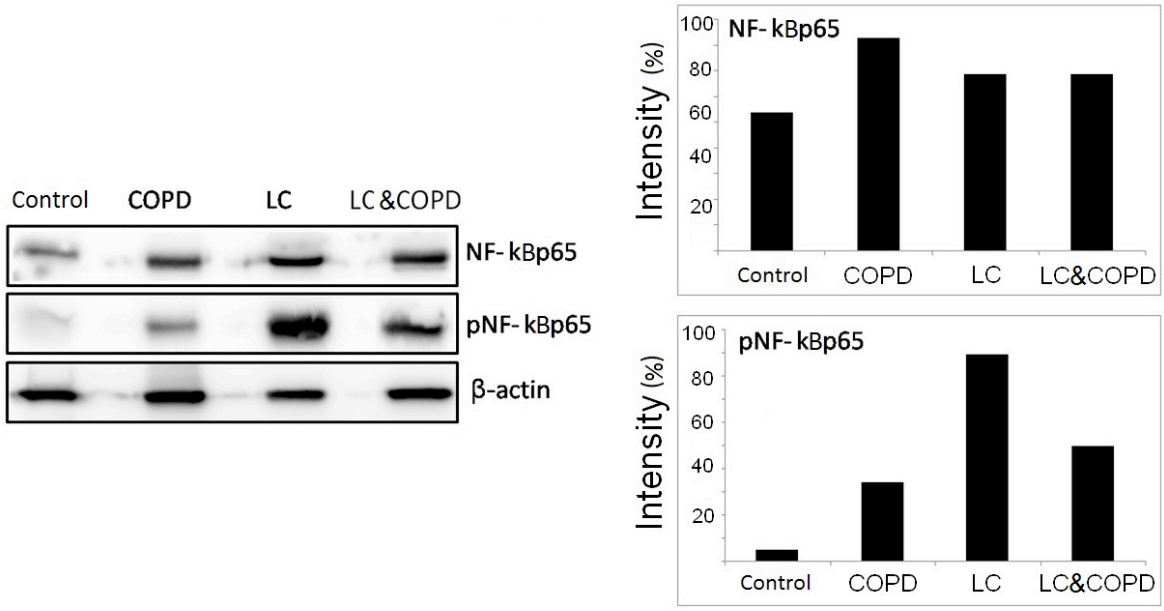


Figure 20: Western blots for NF-κBp65, p NF-κBp65, and β-actin in the control, COPD, LC, and LC&COPD groups. Differences in expression, normalized with the β-actin protein, are illustrated by a bar chart on the right.

High expression levels of NF-κBp65 were detected in all groups, with a slightly higher intensity in the COPD group. The phosphorylated form of NF-κBp65 showed higher expression levels only in the disease groups (COPD, LC, and LC&COPD), with the LC group presenting the highest intensity. These western blot analyses are compatible with the fact that activation of NF-κBp65 might be involved in the pathogenesis of both diseases.

5.2. Second specific objective: Analysis of the differential expression of inflammation related proteins (cytokines and growth factors) in LC and COPD patients.

5.2.1. Patient selection.

The antibody array methodology was applied to samples from 60 patients, whose characteristics can be found on Table 8. All selected patients were male, with median ages around 61 (41-80) years. The majority of patients were current smokers, with similar pack-years smoking exposure. Moderate and severe COPD were the most common stages in the COPD group of patients, and mild COPD was predominant in the LC&COPD group of patients. The adenocarcinoma histology was the most common among the patients in the LC group, contrary to the patients in the LC&COPD group, who had been predominantly diagnosed with the SCC histology. All other characteristics were well matched among the study groups.

Table 8: Characteristics of the discovery cohort patients for the antibody arrays methodology

	Controls n=16	**COPD n=15**	**LC n=17**	**LC&COPD n=12**
Gender				
Male	100.0% (16)	100.0% (15)	100.0% (17)	100.0% (12)
Female	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
Median age	61.3	61.5	60.7	60.7
[range]	[41.0-80.0]	[45.0-78.0]	[46.0-69.0]	[49.0-68.0]
Smoking status				
Smokers	68.8% (11)	53.3% (8)	52.9% (9)	83.3% (10)
Ex-smokers	31.2% (5)	46.7% (7)	47.1% (8)	16.7% (2)
Pack-years	38.8	50.0	58.2	52.3
[range]	[31.0-53.2]	[42.0-65.0]	[41.0-65.7]	[41.0-63.2]
COPD				
Mild	0.0% (0)	20.0% (3)	0.0% (0)	58.3% (7)
Moderate	0.0% (0)	33.3% (5)	0.0% (0)	25.0% (3)
Severe	0.0% (0)	26.7% (4)	0.0% (0)	0.0% (0)
Very severe	0.0% (0)	20.0% (3)	0.0% (0)	16.7% (2)
LC Histology				
Adenocarcinoma				
Stages I-II	0.0% (0)	0.0% (0)	11.8% (2)	0.0% (0)
Stages III-IV	0.0% (0)	0.0% (0)	58.8% (10)	33.3% (4)
SCC				
Stages I-II	0.0% (0)	0.0% (0)	5.9% (1)	8.3% (1)
Stages III-IV	0.0% (0)	0.0% (0)	23.5% (4)	58.4% (7)

SCC: squamous cell carcinoma

To validate the results obtained by the training cohort, a group of 299 patients was selected and two distinct validation cohorts were created. The first validation cohort was made up of 139 patients that were separated into four distinct groups. The patients in this cohort had similar clinical and pathological characteristics to those in the training cohorts (Table 9).

Table 9: Characteristics of the patients in the first validation cohort

	Controls n=20	**COPD n=29**	**LC n=40**	**LC&COPD n=50**
Gender				
Male	60.0% (12)	86.2% (25)	82.5% (33)	96.0% (48)
Female	40.0% (8)	13.8% (4)	17.5% (7)	4.0% (2)
Median age	52.3	65.2	61.2	64.3
[range]	[42.0-58.0]	[45.0-78.0]	[48.0-75.0]	[48.0-75.0]
Smoking status				
Smokers	100.0% (20)	51.7% (15)	50.0% (20)	52.0% (26)
Ex-smokers	0.0% (0)	48.3% (14)	50.0% (20)	48.0% (24)
Pack-years	41.8	52.8	55.0	57.0
[range]	[33.0-57.2]	[41.0-69.2]	[39.0-73.2]	[43.0-75.2]
COPD				
Mild	0.0% (0)	27.6% (8)	0.0% (0)	26.0% (13)
Moderate	0.0% (0)	51.7% (15)	0.0% (0)	50.0% (25)
Severe	0.0% (0)	20.7% (6)	0.0% (0)	24.0% (12)
Very severe	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
LC Histology				
Adenocarcinoma				
Stages I-II	0.0% (0)	0.0% (0)	10.0% (4)	10.0% (5)
Stages III-IV	0.0% (0)	0.0% (0)	27.5% (11)	36.0% (18)
SCC				
Stages I-II	0.0% (0)	0.0% (0)	25.0% (10)	30.0% (15)
Stages III-IV	0.0% (0)	0.0% (0)	37.5% (15)	24.0% (12)

SCC: squamous cell carcinoma

The characteristics of the 160 patients in the second validation cohort can be seen on Table 10. This cohort was significantly different from the first one given that it did not include a group of patients with only COPD and that the LC groups had patients diagnosed not only with adenocarcinoma and SCC, but also LCC and SCLC.

Table 10: Characteristics of the patients in the second validation cohort

	Controls n=20	**LC n=66**	**LC&COPD n=74**
Gender			
Male	60.0% (12)	83.3% (55)	94.6% (70)
Female	40.0% (8)	16.7% (11)	5.4% (4)
Median age	52.3	61.2	64.3
[range]	[42.0-58.0]	[48.0-75.0]	[48.0-75.0]
Smoking status			
Smokers	100.0% (20)	50.0% (33)	52.7% (39)
Ex-smokers	0.0% (0)	50.0% (33)	47.3% (35)
Pack-years	33.6	58.8	62.0
[range]	[30.0-51.2]	[41.0-77.2]	[43.0-79.2]
LC Histology			
Adenocarcinoma			
Stages I-II	0.0% (0)	4.5% (3)	2.7% (2)
Stages III-IV	0.0% (0)	18.2% (12)	13.6% (10)
SCC			
Stages I-II	0.0% (0)	6.1% (4)	8.1% (6)
Stages III-IV	0.0% (0)	12.1% (8)	5.4% (4)
LCC			
Stages I-II	0.0% (0)	4.5% (3)	8.1% (6)
Stages III-IV	0.0% (0)	7.6% (5)	10.8% (8)
SCLC			
Limited stage	0.0% (0)	19.7% (13)	21.6% (16)
Extensive stage	0.0% (0)	27.3% (18)	29.7% (22)

SCC: squamous cell carcinoma; LCC: large cell carcinoma; SCLC: small cell lung cancer

5.2.2. Expression of cytokines and grow factors in BALF from LC and COPD patients.

The initial protective response, elicited by the damaging complex mixture of chemicals present in tobacco smoke, can generate the release and recruitment of more pro-inflammatory immune cells and cytokines and create a chronic inflammation scenario and a tumour friendly microenvironment. So, we aimed at analysing which inflammatory proteins had altered expression in our four groups of patients. For this purpose we used

antibody arrays that measured the levels of 80 cytokines and growth factors implicated in the inflammatory process.

A total of 15 cytokines and growth factor arrays were carried out and the data analysis was performed by measuring the different expression levels between the 80 inflammatory proteins and the positive controls present in the arrays. Protein expression levels were then standardized, protein by protein, across all conditions by using the medians and SD values. Results were visualized and analysed with Babelomics 4.2 (babelomics.bioinfo.cipf.es). We created a hierarchical supervised clustering analysis of the protein arrays using the UPGMA function. The generated dendograms of the cytokines and growth factors expression levels and the groups of patients, and heatmap are represented below in Figure 21.

The original four groups of patients (control, COPD, LC, and LC&COPD) were sub-divided into six groups taking into consideration the histological sub-types of NSCLC; control, COPD, adenocarcinoma (ADC), SCC, ADC&COPD, and SCC&COPD. Analysis revealed that 20% of the evaluated proteins in the cytokine and growth factor arrays were over-expressed in the disease groups compared to the control group. These 16 proteins were MIP-1β, MIG, IGFBP1, EGF, VEGF, TNFRI and TNFRII, IL-6Sr, GDF15, IL-1Ra, MCP-1, Eotaxin2, PDGFAA, IGFBP2, IL-11, and CCL1.

A first group of proteins (MIP-1β, MIG, IGFBP2, and EGF) was highly expressed across all disease groups, compared to the control group. A second group of proteins (VEGF, TNFRI and TNFRII, IL-6Sr, GDF15, IL-1Ra, MCP-1, Eotaxin2, and PDGFAA) also had a higher expression in the disease groups, but not as high as the previous group of proteins. Interestingly, there were clear differences in protein expression between the adenocarcinoma histological subtype and the other groups. In fact, the expression of IL-11 and CCL1 was increased differentially in adenocarcinoma.

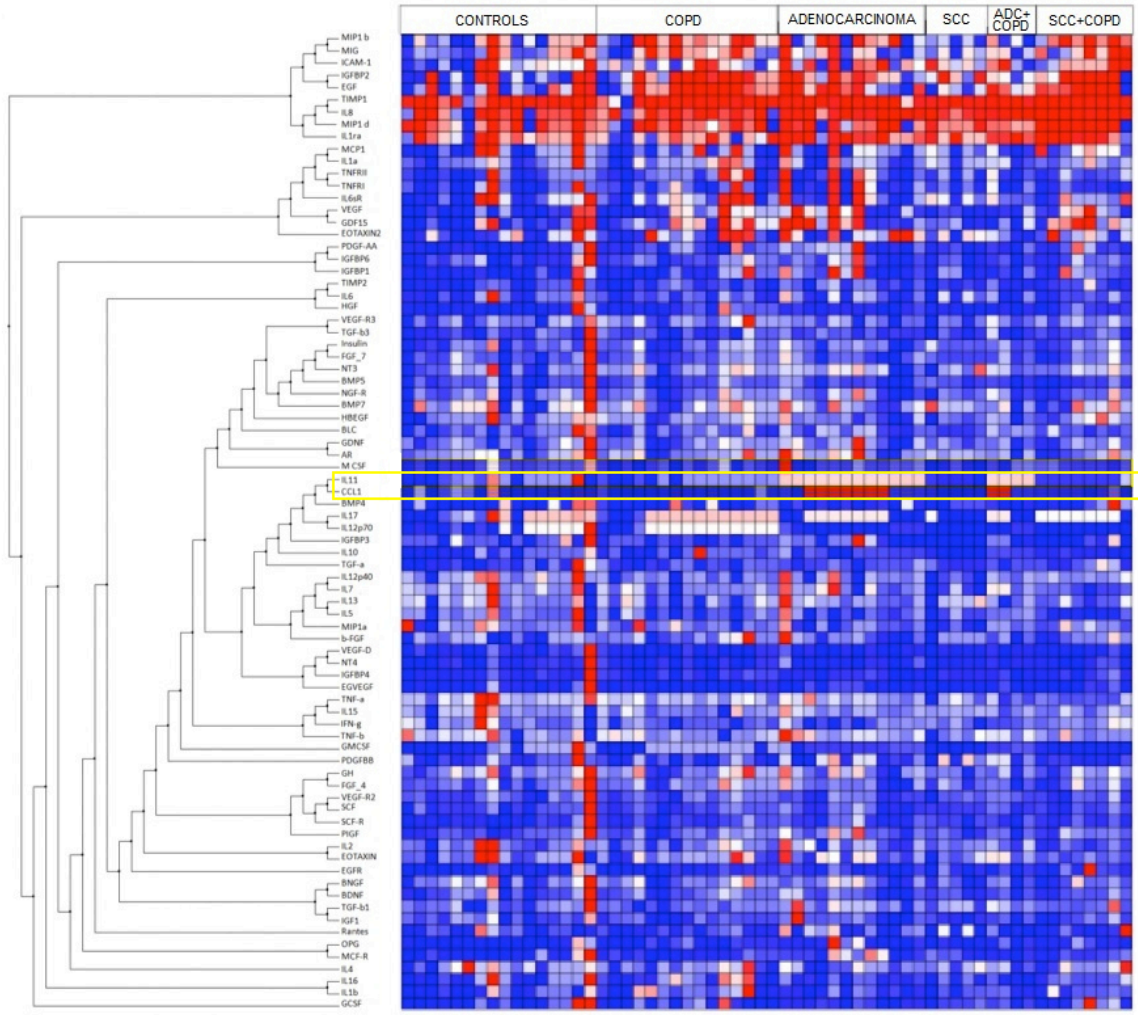


Figure 21: Heatmap of the differences in protein expression across the disease groups, compared to the control group, obtained from the inflammatory protein arrays; protein expression levels relative to its median expression level across all conditions, are represented by a colour, with red representing expression greater than the median, blue representing expression lesser than the median, and various intermediate colour intensities representing the magnitude of variance from the median.

Statistical analysis of the cytokine and growth factor arrays, seen on Figure 22, revealed significant statistical differences ($p \leq 0.05$) among groups for the majority of proteins, despite the existing dispersion between them due to patient variability.

Among the 16 inflammatory proteins that had statistically significant different expression levels when compared between the patients of the disease groups and the control group patients, IL-11 and CCL1 were observed to have a similar behaviour. These two proteins were only over-expressed in the BALF samples from patients with ADC and

patients with both ADC and COPD, and the differences of expression were statistically significant ($p<0.001$).

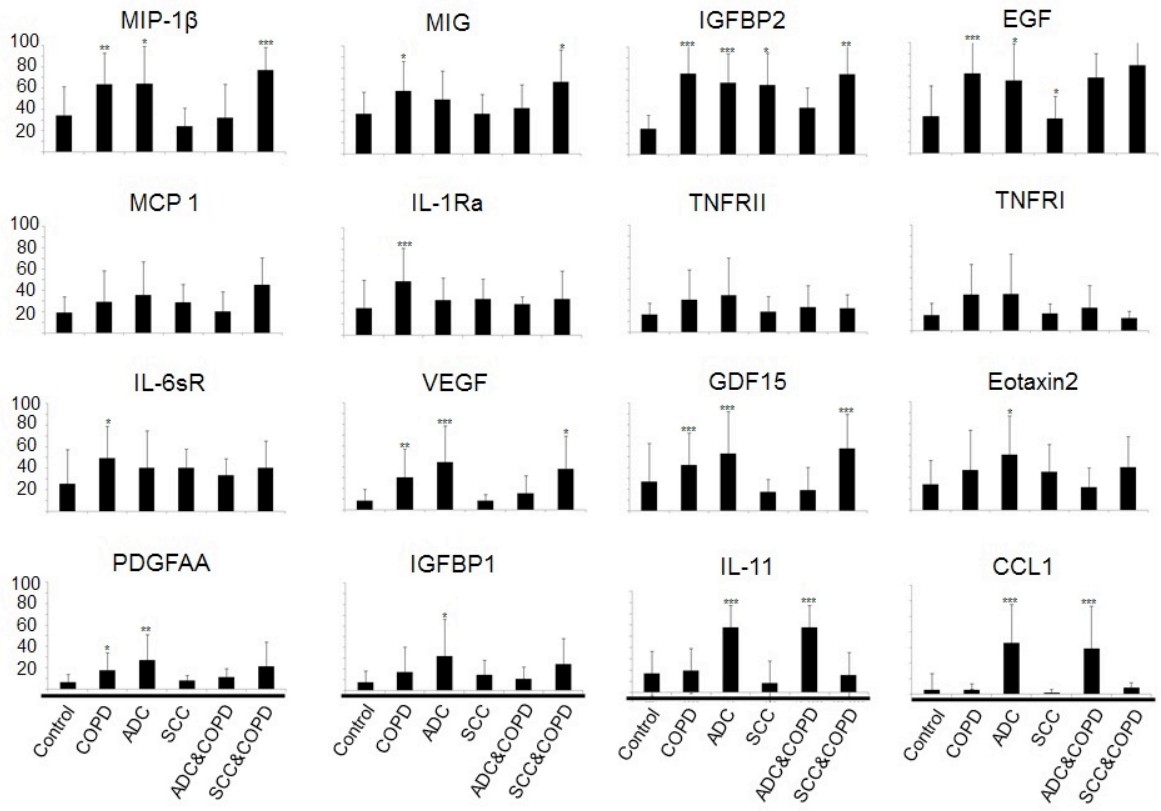


*Figure 22: Statistical analysis of the antibody arrays data (*p≤0.05, **p<0.01, ***p<0.001).*

5.2.3. Analysis of Il-11 and CCL1 as adenocarcinoma biomarkers

The confirmation of the observed altered expression of IL-11 and CCL1 was accomplished by the application of western blot to the same samples used in the protein arrays.

The IL-11 and CCL1 proteins were found to be over-expressed in the ADC and ADC&COPD groups of patients when compared to the remaining groups, confirming the results observed in the cytokine arrays. Results of the western blot analysis can be seen on Figure 23.

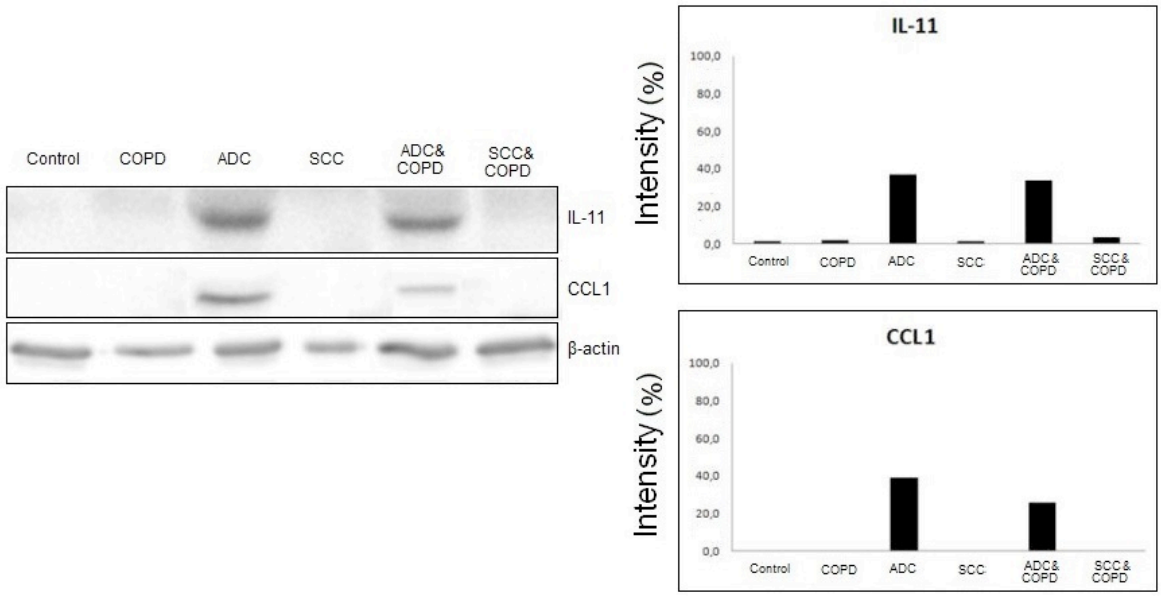


Figure 23: Western blot analysis of IL-11 and CCL1.

The reproducibility of IL-11 and CCL1 as adenocarcinoma markers was assessed by the use of the ELISA methodology on BALF samples from the first validation cohort patients (Table 9). This technology was chosen given its high sensitivity and specificity, simple and fast procedure, the fact that it requires small amounts of sample, and also because of its likelihood of implementation in a clinical setting.

Results showed that the IL-11 quantities were significantly higher in the ADC group of patients compared to the other groups of patients (median 107 pg/ml, IQR 60.5-195.5; $p<0.0001$) as well as in the ADC&COPD group (median 72 pg/ml, IQR 51-136.52; $p<0.0001$), as seen on Figure 24. The levels of CCL1 were likewise significantly elevated in the ADC group of patients (median 82.25 pg/ml, IQR 47.25-117; $p<0.0001$) and ADC&COPD group (median 62.5 pg/ml, IQR 39-85; $p<0.0001$) when compared to the remaining groups.

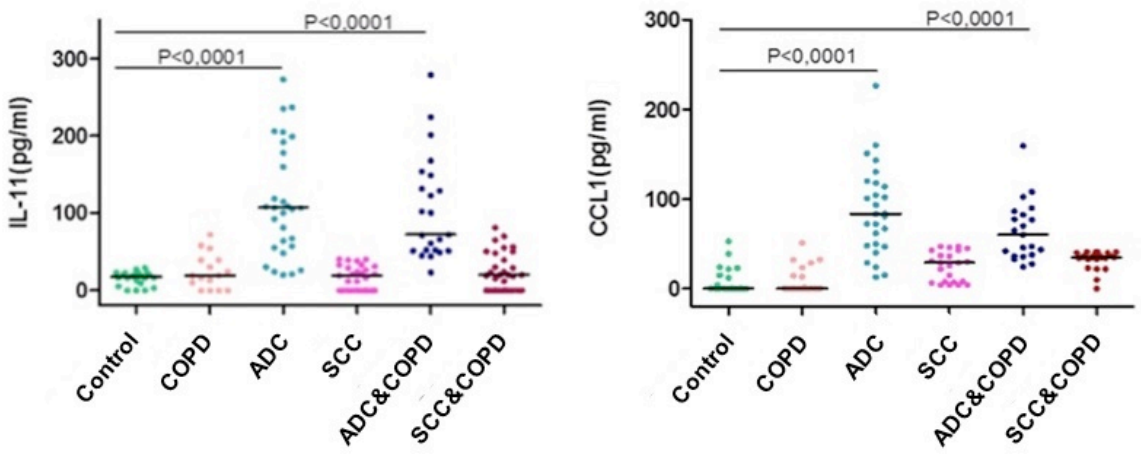


Figure 24: IL-11 and CCL1 expression levels measured by ELISA in the first validation cohort. Statistically significant expression differences (p<0.0001), obtained by the Mann-Whitney U test, were found between the ADC groups and all remaining groups, although only the differences between the control group and the ADC groups are represented.

The performance of IL-11 and CCL1 as biomarkers of lung adenocarcinoma was evaluated by receiver operating characteristic (ROC) curve analysis, a commonly used method in medicine to determine a cut-off value for a clinical test that is able to discriminate patients with and without a given condition.

In the first validation cohort the optimum diagnostic cut-off value for IL-11 was 42 pg/ml (AUC – 0.93; 95%CI: 0.90 – 0.97), with sensitivity values reaching 90% and specificity values of 88%. These results imply that a patient with adenocarcinoma will have higher levels of IL-11 than 93% of controls. They also show that 90% of patients with adenocarcinoma would be correctly identified as so by measuring the levels of IL-11 and 88% of controls would be correctly classified as such (Figure 25).

The estimated cut-off value of CCL1 was 39.5 pg/ml (AUC – 0.83; 95%CI: 0.75 – 0.90). The AUC indicates that an adenocarcinoma patient would have higher levels of CCL1 than 83% of controls. The sensitivity and specificity values for CCL1 as a biomarker of adenocarcinoma were 80% and 74%, respectively, meaning that 80% of patients with lung adenocarcinoma would be classified as so and 74% of patients without lung adenocarcinoma would be classified as controls.

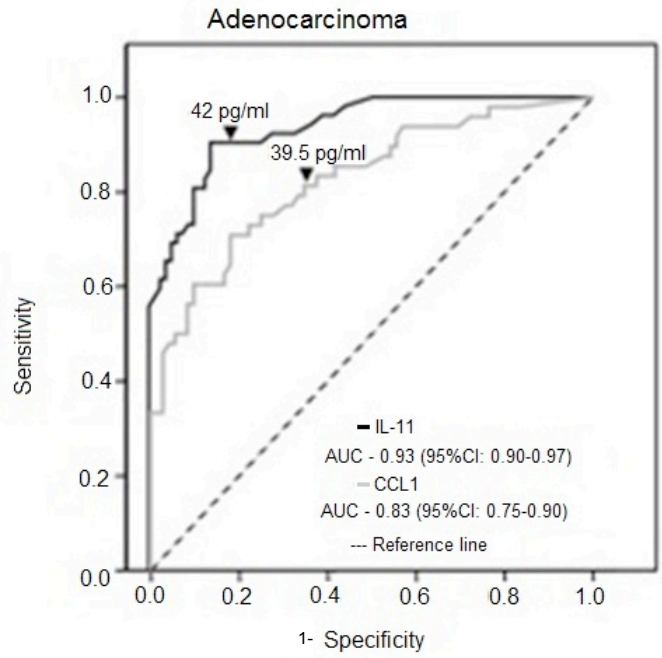


Figure 25: Diagnostic outcomes for IL-11 and CCL1 levels in BALF in the diagnosis of adenocarcinoma (versus all groups) in the first validation cohort.

Predictive values and likelihood ratios for IL-11 and CCL1 in the first validation cohort are shown in Table 11. Overall, approximately 90%, 80%, 71%, and 96.1% of patients with adenocarcinoma had positive levels of IL-11, CCL1, IL-11 and CCL1, and IL-11 and/or CCL1, respectively (Figure 26). Specificity and positive predictive value (PPV) increased when both markers were analysed together compared to measuring each one in separate (94.4% and 86% respectively). Increased values of sensitivity (94.3%) and negative predictive value (NPV) (96.4%) were observed when the evaluation of only one marker was required to positively predict the adenocarcinoma outcome. On the other hand, specificity (74,1%) and PPV (64,1%) decreased.

Table 11: Analysis of IL-11 and CCL1 levels as diagnostic markers of adenocarcinoma in BALF of patients from the first validation cohort

Adenocarcinoma vs all patients	IL-11	CCL1	IL-11 and CCL1	IL-11 and/or CCL1
AUC (95%CI)	0.93 (0.896-0.975)	0.83 (0.749-0.902)	-	-
Sensitivity (95%CI)	90.2% (79-95.7%)	80% (66.4-87.7%)	71.2% (57.7-81.7%)	94.3% (84.6-98.1%)
Specificity (95%CI)	88.7% (80.6-93.5%)	74.1% (63.9-82.2%)	94.4% (88.4-97.4%)	74.1% (65.1-81.4%)
PPV (95%CI)	80.7% (68.7-88.9%)	72.1% (59.2-73.4%)	86% (72.7-93.4%)	64.1% (53-73.9%)
NPV (95%CI)	94.5% (87.8-97.6%)	86.3% (76.6-92.4%)	87.2% (79.9-92.1%)	96.4% (89.9-98.8%)
Positive LR	7.95 (4.53-13.98)	3.02 (2.05-4.47)	12.8 (5.77-28.41)	3.64 (2.63-5.04)
Negative LR	0.11 (0.05-0.26)	0.29 (0.17-0.52)	0.31 (0.20-0.47)	0.08 (0.03-0.23)

AUC= area under curve. PPV= positive predictive value. NPV= negative predictive value. LR= likelihood ratio. The diagnostic cut-off values of IL-11 and CCL1 were 42 pg/ml and 39.5 pg/ml respectively.

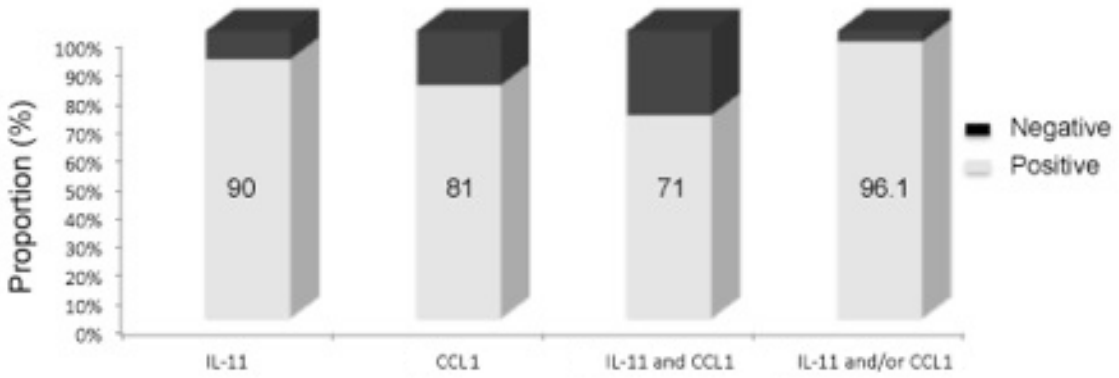


Figure 26: Proportion of adenocarcinoma patients from the first validation cohort with positive levels of IL-11, CCL1, IL-11 and CCL1, and IL-11 or CCL1.

The analysis of IL-11 and CCL1 together revealed that there was a correlation between the levels of both proteins in samples from patients with adenocarcinoma as seen on Figure 27 (Pearson r2 0.76; p=0.001). However, approximately 5.5% of patients with low levels of IL-11 were positive for CCL1, and 7.2% of patients with low levels of CCL1 were positive for IL-11.

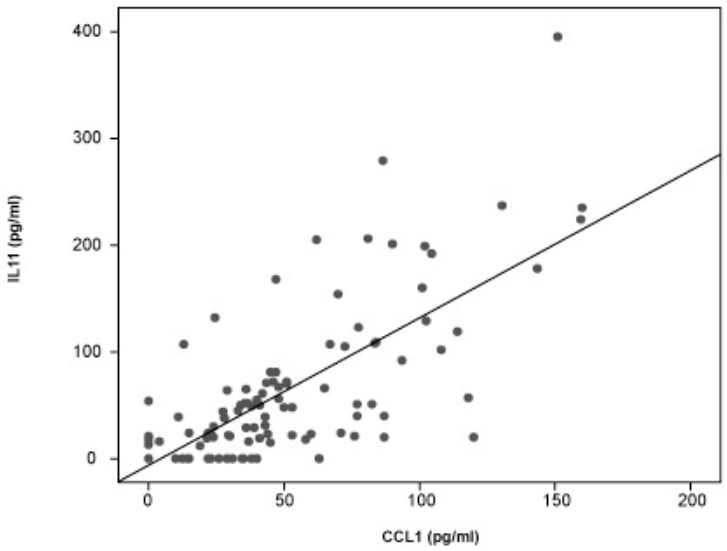


Figure 27: Correlation between IL-11 and CCL1 levels in BALF samples from the first validation cohort.

The validation of IL-11 and CCL1 as specific adenocarcinoma biomarkers was obtained by analysing their expression in the second validation cohort. The 160 patients that made up this validation cohort had distinct characteristics from the two training cohorts training and the first validation cohort and this was intentional. We decided to exclude a COPD group from this cohort because no differences were found between the adenocarcinoma and adenocarcinoma with COPD groups. On the other hand, we included other histological subgroups, such as LCC and SCLC, in order to study the performance of these proteins across all histological LC groups.

Results showed that the levels of IL-11 were significantly increased in the BALF samples of adenocarcinoma patients as compared to the control group and remaining lung cancer histological types (median 103.5 pg/ml, IQR 52.75-166; p<0.0001), as seen on Figure 28. Similarly, the levels of CCL1 were significantly higher in patients with adenocarcinoma (median 75.25 pg/ml, IQR 42.37-102.3; p<0.0001) than in the remaining groups. Among these, the BALF samples from the SCC group of patients had significantly higher levels of CCL1 than the control, LCC, and SCLC groups of patients (median 34.75, IQR 22-45; p=0.04).

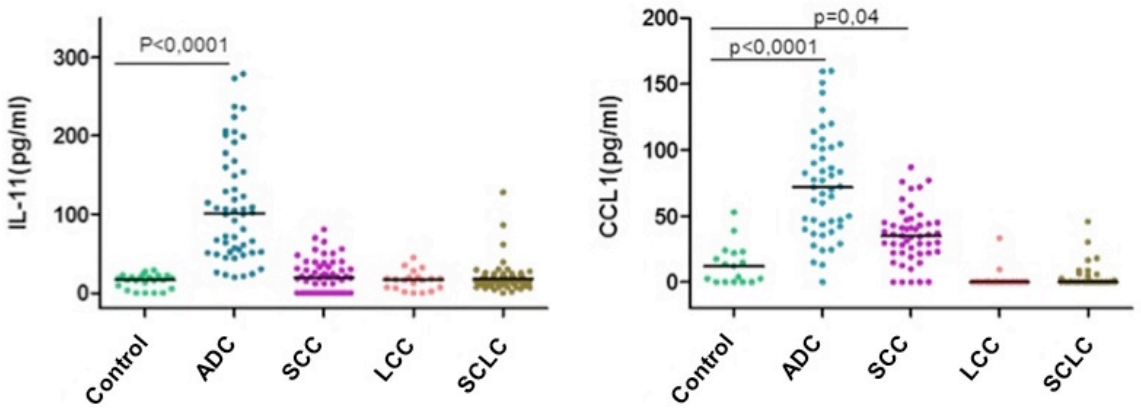


Figure 28: IL-11 and CCL1 expression levels measured by ELISA in the second validation cohort. Statistically significant expression differences (p<0.0001), obtained by the Mann-Whitney U test, were found between the ADC groups and all remaining groups.

The ROC curve analysis of the second validation cohort can be seen on Figure 29. This analysis revealed that the optimum diagnostic cut-off value for IL-11 was 29.5 pg/ml (AUC – 0.95; 95%CI: 0.92 - 0.98). The AUC associated with this cut-off indicated that a patient with lung adenocarcinoma had higher levels of IL-11 than 95% of controls. The estimated cut-off value for CCL1 was 24.25 pg/ml (AUC – 0.91; 95%CI: 0.87 - 0.96), which showed that a lung adenocarcinoma patient would have higher levels of CCL1 than 91% of patients in the control group.

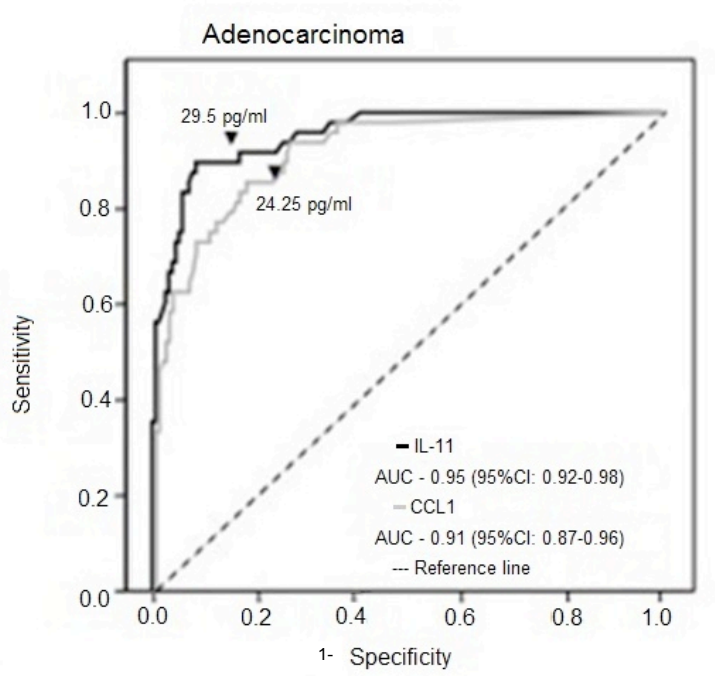


Figure 29: Diagnostic outcomes for IL-11 and CCL1 levels in BALF in the diagnosis of adenocarcinoma (versus all groups) in the second validation cohort.

In order to further validate the results from cohort validation first, the cut-off values from the first cohort were applied to the data from the second validation cohort. Predictive values and likelihood ratios for IL-11 and CCL1 in the second validation cohort can be found on Table 12.

Positive levels of IL-11, CCL1, IL-11 and CCL1, and IL-11 and/or CCL1 were detected on approximately 91%, 92%, 71%, and 92% of patients with lung adenocarcinoma (Figure 30). These percentages were similar to those obtained in the first validation cohort. Likewise, specificity and PPV increased when both markers were analysed together compared to measuring each one in separate (96.3% and 84.1% respectively). Increased values of sensitivity (92.3%) and NPV (98.1%) were observed when the evaluation of only one marker was required to positively predict the adenocarcinoma outcome, similar to what occurred with the first validation cohort. Specificity (84%) and PPV (62.5%) also increased in this case, contrary to what occurred in the first validation cohort.

Table 12: Analysis of IL-11 and CCL1 levels as diagnostic markers of adenocarcinoma in BALF of patients from the second validation cohort, with cut-off values from the first validation cohort

Adenocarcinoma vs all patients	IL-11	CCL1	IL-11 and CCL1	IL-11 and/or CCL1
AUC (95%CI)	0.95 (0.92-0.98)	0.91 (0.87-0.96)	-	-
Sensitivity (95%CI)	90.6% (79.7-95.9%)	91.7% (80.4-96.7%)	71.2% (57.7-81.7%)	92.3% (82.6-98.1%)
Specificity (95%CI)	83% (80.8-87.7%)	77.5% (71.0-82.9%)	96.3% (92.5-98.2%)	84% (78-88.5%)
PPV (95%CI)	60.8% (49.7-70.8%)	51.2% (40.8-61.4%)	84.1% (70.6-92.4%)	62.5% (51.5-72.3%)
NPV (95%CI)	96.8% (92.7-98.6%)	97.3% (93.3-99%)	92.3% (87.7-95.3%)	98.1% (94.6-99.4%)
Positive LR	5.32 (3.81-7.41)	4.08 (3.09-5.04)	19.1 (9.0-41.13)	5.88 (4.21-8.22)
Negative LR	0.11 (0.05-0.26)	0.11 (0.04-0.28)	0.3 (0.19-0.46)	0.07 (0.02-0.20)

AUC= area under curve. PPV= positive predictive value. NPV= negative predictive value. LR= likelihood ratio. The diagnostic cut-off values of IL-11 and CCL1 were 42 pg/ml and 39.5 pg/ml respectively.

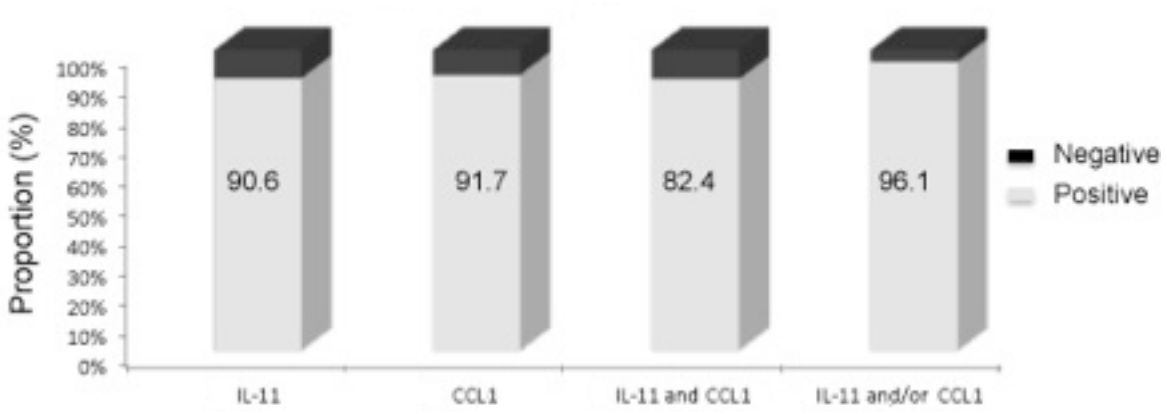


Figure 30: Proportion of adenocarcinoma patients from the second validation cohort with positive levels of IL-11, CCL1, IL-11 and CCL1, and IL-11 or CCL1.

GENERAL DISCUSSION AND MAIN CONCLUSIONS

6. GENERAL DISCUSSION AND MAIN CONCLUSIONS

Lung cancer is one of the most commonly diagnosed malignancies and kills more patients than any type of cancer in the world. Approximately 90% of all these LC types are attributable to one major factor: tobacco smoking.

The diagnosis of LC is a crucial step given its influence in the prognosis of a patient. The use of chest radiographies, bronchoscopies, among other tests, is common when searching for LC. Afterwards, available therapeutic options include surgery, radio and chemotherapy, used separately or in different combinations. However, LC is usually asymptomatic during its development and this causes the majority of patients to be diagnosed at advanced stages of the disease, limiting therapeutic options and prognosis. When a patient is diagnosed at an early stage, the 5-year survival rate lies between 43-73%. With a stage II diagnosis, 25-45% of people will live for at least 5 years. When considering more advanced stages, the outcomes change dramatically: only 7-24% of patients diagnosed with stage III will live more than 5 years after their initial diagnosis; stage IV, characterized by the presence of metastasis, has a 5-year survival rate between 2 and 13% (120).

Radiological screenings have been used for many years for the early detection of LC. Still, data from several screening trials has revealed that performing annual chest radiographies in smokers, former smokers, and non-smokers is not effective in reducing LC mortality and cannot be recommended for clinical practice (121). More recently, low dose CT scanning was associated with a significant reduction in LC mortality (20%), in one large study of high-risk individuals. However, more data are needed on the cost effectiveness of screening that takes into account the frequency of screening and both the benefits and harms (such as false positives and over diagnosis) before large-scale screening programmes are created (122).

In addition to radiological methods, the cytological analysis of certain samples is also helpful to the diagnosis of LC. However, pulmonary cytology has variable sensitivity and specificity values, and false positive diagnoses are somewhat common, giving that the cytological evaluation of the lungs has some limitations (123). The use of bronchoscopy to localize lesions and early LC presents some challenges: SCC is mostly detected in the central areas of the lungs, whereas adenocarcinomas, the most common type of LC in recent years, are mostly found in the peripheral airways of the lungs, unreachable to the majority of bronchoscopes. In conclusion, the diagnosis of LC involves a combination of the radiological and histological evaluation of a symptomatic patient or a suspicious lesion, requiring the cytological examination sputum or in bronchoscopy

obtained samples, such as biopsies and BALF (124). These tests are sometimes inconclusive, making it necessary to perform invasive surgical procedures in patients that may or may not have the disease.

The abovementioned classical methods of diagnosing LC are currently evolving to less invasive, molecular oriented approaches to the detection of this disease. The use of autofluorescence bronchoscopy, lung tissue molecular markers, blood based LC markers, airway based biomarkers, among others, are currently being evaluated for the early detection of LC. Although some recent studies have shown that these molecular biomarkers can have a positive impact in the improvement of LC diagnosis, extensive validation of the most promising molecules that arise from these studies is crucial before they can be used to screen individuals at high risk of developing LC (125-127).

Apart from being the major cause of LC development, cigarette smoke is also responsible for another common illness that affects the respiratory system, which is COPD. For a long time it was though that the link between both diseases was their aetiological factor. Nonetheless, the observation that COPD patients had high LC mortality rates challenged that concept. Despite the increased attention given to the study of the commonalities between both diseases, either by clinical studies and research, the true nature of the link that connects both diseases, remains elusive. However, as discussed below, there are key shared mechanisms that may represent links between the two diseases, and potentially diagnostic and therapeutic targets.

Some common genetic alterations play a role in the development of COPD and LC, for example, the α1 antitrypsin protein inhibits the neutrophil elastase that degrades elastin and has anti apoptotic capabilities. The α1 antitrypsin deficiency causes the development of emphysema and increases the risk of developing LC by 70% (128). Other genes that have been associated with LC and COPD include those that code proteinases, detoxifying enzymes, nicotine receptors, and inflammatory proteins (129). The enzymes that contribute to the metabolism of the compounds found in tobacco smoke can occasionally transform these molecules into more harmful substances, such as carcinogens and ROS. The action of these metabolites in the lung epithelial cells can promote the development of both LC and COPD.

Another seemingly important mechanism connecting both diseases is inflammation. Macrophage and neutrophil infiltration is a common occurrence in all smokers, but those who have COPD develop a more marked inflammatory response that correlates with disease severity (130). In addition, the cells that contribute to the development of, for example, emphysema exist both in the airways and alveolar airspaces and can therefore contribute to the development of proximal and distal LC (129). Although both diseases have chronic inflammation as a feature, the nature of the microenvironment

in COPD and LC is different and contributes to the distinct behaviours that characterize both diseases. The COPD microenvironment is more cytotoxic, genotoxic, and has more matrix degrading capabilities than the LC microenvironment, which is more angiogenic and growth promoting (129).

There are still many unanswered questions concerning LC and COPD and the links between both diseases. In addition, a pressing need exists to discover new molecular biomarkers that can aid in the early detection of LC. This latter area of research has been granted a major ally with the development and the improvement of certain analytical technologies, such as genomics and proteomics. Both technologies were designed to perform the rapid and complete analysis of the genes and proteins that are expressed in a given cell or tissue in a given time. Despite the better understanding of the genetic alterations present in LC given by genomic approaches, such as mutations and polymorphism in key genes, the importance of the evaluation of proteins cannot be overlooked.

Proteomics complements the genomic-based approaches by providing additional information that cannot be obtained by the study of genes alone. One of the most important advantages of the proteomic approach in cancer research is based on the fact that there is generally a poor correlation between the transcriptional levels of many genes and the relative abundance of the corresponding protein. Furthermore, due to differential splicing, one gene can encode several protein variants with distinct properties (131). In addition to relative abundance, the activity of a given protein is not a direct consequence of its genetic expression, but the result of post-translational modifications such as phosphorylation, glycosylation, methylation, ubiquitination, cleavage, and acetylation (132). Despite the many advantages when compared to other approaches, the use of proteomic technologies has to be applied carefully in order to provide relevant biological results and to be successfully translated into the clinical practice. Challenges associated with proteomics include the heterogeneity of cells that exist in the lungs, which can be malignant or normal, the broad dynamic range of protein abundances in proximal body fluids and serum that difficult the identification of biomarkers in those types of samples, and finally the need for extensive validation of the molecular changes in independent cohorts of patients using different techniques (131).

Proteomic methodologies can be applied to the discovery of new LC biomarkers as well as aid with the prognosis, prediction and monitoring of a patient response to a given therapeutic intervention. Such markers of disease can be released directly from cancer cells or can be attributable to the host's response against the malignancy itself. The detection of these candidate biomarkers can be performed in a variety of body fluids collected by a non-invasive method. Biomarkers obtained by these methods are usually

more sensitive, more specific, and more easily detected than invasive biomarkers (obtained by procedures such as surgery). In addition, they are associated with less anxiety and discomfort to the patients. Samples such as blood, urine, serum, and saliva are major sources of non-invasive biomarkers and commonly used in cancer research. Although more invasive than the previous samples, BALF is a great source of possible biomarkers of lung diseases given that it is thought to reflect most faithfully the protein and cell composition of the pulmonary airways among the non-invasive techniques (133). The present work aimed at evaluating this particular type of sample by different proteomic techniques with the aim of identifying proteins that could be used as biomarkers for the diagnosis of LC, something that to the best of our knowledge has not been done before.

The first proteomic technique used in this study was 2D-PAGE. This tool has been the driving force behind the development of proteomics and protein analysis throughout the years. This technique has been extensively tested and compared between different laboratories throughout the years makes it one of the most robust proteomic methods.

Our results showed that the protein composition of BALF was different among the four groups of studied patients, especially when we compared the control group to the disease groups. A total of 40 were differentially expressed among the four groups of patients and the expression changes were superior to 2 fold (p<0.05). The distribution of these differentially expressed proteins allowed us to determine the specific proteomic profiles of the various groups of patients.

Seventeen proteins were over-expressed in all the patients from the disease groups (LC, COPD, and LC&COPD) compared to those in the control group. These proteins were AMY1A, AMY2A, ANXA1, ANXA2, ANXA5, ARHGDIB, CA1, CRP, C3, ENO1, GSR, HSP70, IDH1, PEBP4, SERPINB1, TPPP3, and TXN. The expression of HSP70, which in the present study had an average over-expression of 3.6 fold across the three disease groups, was also found significantly elevated in serum samples from COPD patients, when compared to controls (134). It has also been linked to the number of cigarettes smoked daily in a study of lung adenocarcinoma in smokers (135). The use of proteomic approaches has revealed the association between the over-expression of HSP70 and the differentiation level and/or aggressiveness of several types of cancer, such as gastric adenocarcinomas, hepatocarcinomas, and oesophageal cancer (136-138). Furthermore, various oncoproteomic studies have correlated the elevated levels of HSP70 with therapeutic resistance (139-142).

Another important protein in this group is CA1. This protein is a marker of cellular hypoxia, which is a natural phenotype of solid tumours. Several studies have investigated the potential of carbonic anhydrases as cancer diagnostic, prognostic, and therapeutic markers (143). Recently, the over-expression of CA1 was reported in carcinoma cell lines

from multiple organs, such as lung, breast, cervix, bladder, oesophagus, colorectal, kidney, and head and neck (144). PEBP4 which plays a role in the inhibition of the mitogen-activated protein kinase (MAPK) signalling pathway but is also involved in the inhibition of the c-Jun N-terminal kinases (JNK) pathway that promotes the activation of protein kinase B (AKT), was recently associated with an increased invasion and metastasis in NSCLC and colorectal tumours (145-147).

The IDH1 enzyme is essential for cell metabolism and energy production, given its role in the Krebs cycle. In the present study, the IDH1 protein had an average over-expression of 2.3 fold in the disease groups. The oncogenic potential of this enzyme has been studied in acute myeloid leukaemia (148). An also recent report has observed the over-expression of IDH1 in more than 70% of NSCLC tumours. It was also correlated with lower 5-year survival rates and considered an independent unfavourable prognostic marker for overall survival of NSCLC patients in a multivariate analysis (149). In addition, IDH1 plasma levels were considered a biomarker for lung adenocarcinoma with relatively high sensitivity and specificity (150).

The next group of differentially expressed proteins appeared in the LC and LC&COPD groups of patients. In the LC group of patients the AKR1B10, ALDOA, CTSD, EZR, FBP1, and TKT proteins were over-expressed when compared to the control group and the SELENBP1 protein was under-expressed in this group of patients compared to those in the control group. A number of studies have associated the presence of CTSD, Ezrin, and SELENBP1 with increased cancer growth, invasion, and metastasis in various types of tumours (151-153). Furthermore, three different studies have highlighted the importance and potential role of these proteins as prognostic biomarkers of lung tumours (154-156). Another over-expressed protein in our results was the AKR1B10 protein. In other studies, this over-expression was associated in most cases with smoking, suggesting a possible involvement of this enzyme in tobacco-induced LC (157-159). In fact, AKR1B10 has a high catalytic efficiency for the reduction of retinoids, and retinoic acid deficiency has been linked to airway epithelial squamous metaplasia and epithelial-to-mesenchymal transition (160). In addition, AKR1B10 has also been recently involved in the resistance to different chemotherapeutic drugs, such as cisplatin (161).

Apart from these differentially expressed proteins, the LC group of patients together with the LC&COPD group of patients, had an over-expression of the following proteins when compared to the control group; ALDH3A1, AKR1C3, PYGM, PKM2, and PPIA. The ALDH3A1 protein is a phase II drug-metabolizing enzyme that is highly expressed in the lung, stomach, keratinocytes, and cornea. Cytosolic ALDH3A1, which is induced by polycyclic aromatic hydrocarbons or chlorinated compounds (present in tobacco smoke), is thought to play an important role in alveolar pneumocyte physiology,

and increases during carcinogenesis (162). On the other hand, the TKT, FBP1, ALDOA, PYGM, and PKM2 proteins are involved in the regulation of glycolysis and are probably contributing to the Warburg effect that leads to a state that has been termed "aerobic glycolysis". Previous studies have showed that the PKM2 protein has a role in the production of lactic acid through glycolysis rather than producing energy through mitochondrial oxidative phosphorylation. This may help tumour cells to survive in low glucose and low oxygen environments, and facilitate tumour invasion (163, 164).

Lastly, a total of four proteins were found to be over-expressed in the COPD and LC&COPD groups of patients: CAT, PRDX1, PRDX2, and PRDX5. These proteins are enzymatic antioxidants and represent the first line of defence against oxidative stress. The evaluation of the role of these proteins and the development of COPD has generated some conflicting results. Some studies have observed a down-regulation of enzymatic antioxidants, such as peroxiredoxins and catalase, in COPD patients (165-167), while others claim their up-regulation in the same conditions (168). Other studies on LC have observed that the up-regulation of different peroxiredoxins correlates with a poor prognosis (95, 169, 170).

The identified 40 differentially expressed proteins in our groups of patients were submitted to an additional analysis to ascertain the common biological functions they might share, their presence in known pathways, and any possible connection between them. Our results showed three major pathways: inflammation, free radical scavenging, and glycolysis and gluconeogenesis.

Altered glucose metabolism has been observed in cancer cells. To support continuous cell growth and proliferation, these cells exhibit high rates of aerobic glycolysis, a phenomenon known as the Warburg effect (171). Some oncogenes and transcription factors that are frequently altered in cancer are also responsible for the increased expression of glycolytic enzymes and glucose transporters, altering the glucose metabolism in cancer cells (172).

Free radicals such as ROS are involved in a variety of diseases, degenerative changes, and tissue degradation, a hallmark of carcinogenesis (173). In a metabolically active cell the redox system pathways maintain the balance between oxidant and antioxidant factors, by regulating the activation of specific transcription factors and the production of substances that neutralize oxidants (174, 175). However, in cancer settings, alterations in these redox pathways occur and the cell is no longer able to produce antioxidant substances to adjust the balance between oxidant and antioxidant factors. Alterations in the physiological pathways involved in the regulation of the redox system have been identified in tumours (176). Several proteins detected in our samples are involved in this system; such as the PPIA protein is secreted in response to ROS from

vascular smooth muscle cells (177). It is a chaperone protein that has several functions including protein trafficking, such as the nuclear translocation of extracellular-signal-regulated kinases 1/2 (ERK1/2) (178) and apoptosis-inducing factor (AIF) (179). There is also evidence that PPIA might be a valuable biomarker. Up-regulation of PPIA in small cell lung cancer, pancreatic cancer, breast cancer, colorectal cancer, squamous cell carcinoma, and melanoma has been reported (180). In addition, knockdown of PPIA *in vivo* was found to correlate with slower growth, decreased proliferation, and a greater degree of apoptosis of lung tumours (181). Another over-expressed protein in our results was TXN that is a potent growth and cell survival factor, which activates specific transcription factors such as NF-kB, p53, hypoxia-inducible factor a (HIFa), and activator protein 1 (AP-1). TXN regulates the production of substances that protect cells from oxidative stress induced by oxygen free radicals (182, 183).

The final mechanism identified was inflammation, one of the body's natural mechanisms against harmful stimuli. However, the components present in tobacco smoke can induce an inflammatory response of the lungs that can become persistent, promoting the development of COPD and a tumour friendly microenvironment. Inflammation is currently considered an enabling characteristic of tumours, given that it contributes to the appearance of multiple cancer hallmark capabilities by supplying important molecules to the tumour microenvironment.

A total of 22 were associated with the inflammatory pathway. The ANXA2 protein is highly expressed in NSCLC and is positively correlated with a poor prognosis, apart from promoting cell proliferation in LC (184). The expression levels of the CFL1 gene have been considered prognostic and drug resistance markers of NSCLC (185). Elevated levels of CRP, a systemic marker of chronic inflammation, were associated with subsequently increased LC risk (186). The LCN2 protein, a member of the lipocalin family that transports small lipophilic ligands, has gained recent attention as both a potential biomarker and a modulator of human cancers. High levels of LCN2 are common at sites of inflammation and its over-expression reduced the apoptosis of a LC cell line (187).

The principal core of all identified pathways was NF-κB. This transcription factor is known to regulate genes that express cytokines, adhesion molecules, angiogenic factors, anti-apoptotic factors, and matrix metalloproteinases (MMPs). These proteins are involved in distinct aspects of cellular physiology, such as inflammation, cell survival and proliferation (188).

Nuclear factor-kappa B represents a family of transcription factors that include five members in mammalian cells: p65 (RelA), RelB, c-Rel, p50/p105 (NF-kappaB1), and p52 (NF-kappaB2). Most commonly, NF-kB exists as a p65/p50 heterodimer that is retained in its inactive state by its association with IkBα, an inhibitory protein. NF-kB is activated by a

wide variety of stimuli that phosphorylate IkBα, dissociating it from the p65/p50 heterodimers. These stimuli can be for example DNA damage or the presence of certain cytokines. Then, the NF-kB protein, free from its inhibitory regulator, is transported to the nucleus, where it regulates transcription (189).

The NF-kB protein is thought to play a role in the multiple steps of carcinogenesis. In the initiation process, the multiple irritants present in tobacco smoke can cause DNA damage in lung epithelial cells via the generation of ROS. When cells fail to repair the damaged DNA, they are subjected to apoptosis. However, these same environmental risk factors can simultaneously induce chronic inflammation in lung tissue, which is accompanied by the activation of cell survival signalling pathways. As a key player of the inflammatory process, NF-kB is activated in these conditions and can induce autocrine production of the inflammatory cytokine IL-6 and activation of the transcription factor STAT3, causing spontaneous lung cancer *in vivo* (190). The NF-kB protein also plays a role in promoting LC by activating the transcription of several growth factors necessary for angiogenesis (188).

The importance of the inflammatory process observed in our results, either by the amount of inflammation related proteins identified or by the presence of NF-kB as the major connector between the most of identified proteins, encouraged further investigation. It has been suggested that NF-kB promotes LC mainly by promoting the secretion of inflammatory cytokines that establish a cancer-prone inflammatory microenvironment (191). The NF-κB complex induces cytokines that regulate the immune response (such as TNFα, IL-1, IL-6 and IL-8), as well as adhesion molecules, which lead to the recruitment of more leukocytes to sites of inflammation (192). The better understanding of the mechanisms by which chronic inflammation, such as the one occurring in COPD patients, leads to the initiation of LC may identify early diagnostic biomarkers and screening tests, in addition to identifying more effective therapeutic strategies and targets.

In order to search for proteomic markers involved in inflammation and in the pathogenesis of the two most prevalent and devastating smoking associated respiratory diseases, BALF samples from 60 patients with LC and/or COPD and a control group were applied to an inflammation antibody array. With this strategy we were able to identity 16 proteins differentially expressed in BALF fluid in patients with LC, COPD, or both diseases as compared to healthy controls. Remarkably two proteins, IL-11 and CCL1, were almost selectively expressed in adenocarcinoma BALFs, regardless of the presence or absence of COPD. These results were fully validated by western blotting of the same samples and individually validated by ELISA in two additional patient cohorts. ROC curves showed the optimum diagnostic cut-off levels in BALF were 42 pg/ml for IL-11 and 39·5 pg/ml for CCL1. The diagnostic accuracy for adenocarcinoma was confirmed for each biomarker.

Even though there was a positive correlation between the BALF levels of IL-11 and CCL1, the measurement of both proteins seemed to optimize the sensitivity and specificity of the diagnostic to level of 90% and 89%, respectively. A recent study analysed the so far best known LC tumour markers: squamous cell carcinoma antigen (SCCA), carcinoembryonic antigen (CEA), cytokeratin 19 fragment antigen 21-1 (CYFRA21-1), and neuron specific enolase (NSE) (193). These markers have been extensively analysed and their importance in the detection, prognosis and follow-up of LC has been revealed (194-199). However, the ROC curve analysis of these markers revealed AUC between 0.601 and 0.726 and low values of sensitivity (below 30%). Interestingly, both proteins were similarly predictive of adenocarcinoma, regardless the presence of concurrent COPD. They were, however, not useful for other histological subtypes such as SCC. Studies in human tumours have suggested the existence of specific molecular networks and subtype-specific differences between lung adenocarcinoma and SCC subtypes. Another study focusing on human tumours has suggested that the genes expressed by SCC present higher levels of methylation than the genes expressed by adenocarcinomas (200), probably accounting for the differential expression of 178 proteins between the two histological subtypes reported in a recent study (201). Furthermore, it has been suggested the existence of specific molecular networks and subtype-specific differences between lung adenocarcinoma and SCC subtypes, mostly found in cell cycle, DNA repair, and metabolic pathways (202). Our findings firmly suggest differences at inflammatory networks between both histological subtypes.

Interleukin-11 is a member of the IL6-type cytokines. Epithelial cells, fibroblasts, and airway smooth muscle cells are known to produce IL-11 in response to certain stimuli. Functions of this cytokine include the stimulation of hematopoiesis, thrombopoiesis, and megakaryocytopoiesis, the regulation of macrophage differentiation, the mucosal protection after chemotherapy and radiation therapy, among others (203). The cellular action of IL-11 is mediated by its specific receptor (IL-11Rα), also a receptor for IL-6. Following the connection between ligand and receptor, gp130 proteins are activated and phosphorylate the janus kinase (JAK) family of tyrosine kinases. These in turn activate STAT3 and STAT1, resulting in their subsequent translocation to the nucleus. Here they bind DNA in a sequence-specific context and elicit transcriptional activation of target genes that play a role in multiple cellular mechanisms such as proliferation, survival, angiogenesis, and inflammation (204). Regarding STAT3, its aberrant or persistent activation has been observed in human cancers and is associated with poor outcome (205). Apart from STAT3, the expression of IL-11 in cancer cells is also regulated by another transcription factor: AP-1 (206, 207). The AP-1 proteins are composed of homo or heterodimers of Jun or Fos family member. Most AP-1 proteins may be associated with

oncogene-induced transformation and are up-regulated by oncogenic Ras (208-210). A recent cell lines study suggested that both phosphatidylinositide 3-kinase (PI3K) and Raf pathways are necessary for the expression of IL-11, and that JNK, p38, and Stat3 also contribute to oncogenic Ras-induced IL-11 expression (211).

The expression of IL-11 and its receptor has been described in several human malignancies, including ovarian, prostate, breast, colorectal, and gastric carcinoma (212-216). The regulation of IL-11 expression and its contribution to tumour development is still not fully understood. A recent study has demonstrated the action of IL-11 in the development of gastrointestinal cancers, casting some light on the importance of this protein in the development of inflammation-associated neoplasias. Putoczki and co-workers revealed that increased expression of IL-11 was associated with excessive STAT3 activation in gastrointestinal cancers and that this excess resulted in epithelial tumorogenesis and submucosa invasion. The possible use of IL-11 as a therapeutic agent was assessed and the pharmacologic inhibition of this protein alleviated STAT3 activation, supressed tumour cell proliferation, and reduced the invasive capacity and growth of gastrointestinal tumours. In addition, they observed that gastric tumours arising from the glandular epithelium might be biased toward IL-11 responsiveness (217). Another study has detected IL-11 mRNA and protein after incubating human cancer cell lines under hypoxic conditions and has discovered that HIF-1 is involved in the regulation of IL-11 expression (218). Interestingly, it has been recently shown that IL-11 increases HIF-1α nuclear translocation and transcriptional activity, suggesting the existence of a positive feedback loop between IL-11 and HIF (219).

In the lungs, IL-11 is produced by a variety of structural cells and eosinophils in response to a variety of stimuli, including TGF-β, ROS, and viruses. The targeted expression of this cytokine in the lungs of mice resulted in airway remodelling and inflammatory cell infiltration (220). Furthermore, IL-11 was found to be selectively expressed in eosinophils and epithelial cells in patients with moderate and severe asthma where expression correlates directly with disease severity and inversely with FEV1 (221). To the best of our knowledge, the expression of IL-11 in the context of LC has not been evaluated so far. On the other hand, its close relative IL-6 has been analysed in LC patients. Yeh and co-workers also found higher expression of IL-6 in the pleural fluids of patients with lung adenocarcinoma (222). In a study with lung adenocarcinoma, STAT3 activation in mutant EGFR–expressing cell lines occurred, in part, through mutant EGFR–mediated transcriptional up-regulation of the IL-6 gene, which, in turn, activates the gp130/JAK signalling pathway (223). A similar situation might be also occurring with IL-11, given the existing overlap between IL-11 and IL-6 targets. In addition, STAT3 and its downstream genes have been showed to serve as biomarkers for lung adenocarcinoma

and COPD diagnosis and prognosis in mice and humans (224). In conclusion, the local expression of IL-11 in the lungs and the consequent activation by it of STAT3, might contribute to the development of an inflammatory microenvironment, a positive feedback loop, and ultimately to the development of a carcinogenic process similar to what has been proven to occur in other types of cancers.

The other protein that predicted lung adenocarcinoma in our BALF samples was CCL1. It is a chemokine produced by activated T lymphocytes, monocytes, and mast cells, which binds selectively to the CCR8 receptor. Unfortunately, insights to its function are less than the ones referring to IL-11.

CCL1 has been known to act as a potent monocyte and lymphocyte chemoattractant and is thought to play a major role in inflammatory processes (225). The CCL1-CCR8 axis may play a role in the initiation and amplification of atopic skin inflammation (226). This chemokine was also shown to bind to endothelial cells and to stimulate differentiation, chemotaxis, and invasion, indicating a possible role for this chemokine in angiogenesis (227). The role of this protein in cancer has been described in leukaemia, where it is ascribed to activate the anti-apoptotic function of the RAS/MAPK pathway (228, 229). More recently, a report by Das and co-workers revealed that tumour cell migration to lymphatic endothelial cells in vitro was inhibited by the blocking of CCR8 or CCL1 and that the pro-inflammatory mediators TNF, IL-1β, and lipopolysaccharides (LPS) increased CCL1 production by lymphatic endothelial cells and tumour cell migration to these same cells (230). Furthermore, the CCL1/CCR8 axis was identified as a component of cancer-related inflammation and may contribute to immune evasion in human urothelial and renal cancers (231).

The role of CCL1 in lung diseases has been evaluated in a few studies. Measurements of CCL1 using ELISA showed that levels of this cytokine were significantly elevated in BALF from asthmatics compared with normal individuals (232). A study evaluating the effects of specific tobacco smoke components revealed that polycyclic aromatic hydrocarbons, well known to regulate expression of pro-inflammatory cytokines such as IL-1β and TNF-α, markedly increased mRNA expression and secretion of CCL1 in primary human macrophage cultures (233). When evaluating COPD, Takabatake and co-workers identified polymorphic variants of the CCL1 gene that were associated with susceptibility to acute exacerbations, a major cause of morbidity and mortality in patients with this disease (234). Another study revealed the presence of CCR8 on inflammatory macrophages in human COPD lung tissue and that CCL1 induced the production of pro-inflammatory cytokines TNF-α and IL-6 (235).

Overall, the existing data on the involvement of IL-11 and CCL1 in lung carcinogenesis is not prominent. However, the evidence provided by the present study is

encouraging. The present work, through the application of a 2D-PAGE MALDI-TOF/TOF methodology, identified distinct proteomic profiles characteristic of LC, COPD and LC&COPD, which were later validated. The bioinformatics analysis of the 40 differentially expressed proteins identified in BALF suggested that LC and COPD shared some pathogenic pathways such as inflammation, free radical scavenging and oxidative stress response, and glycolysis and gluconeogenesis. In addition, IPA analysis exposed a major connector of these pathogenic mechanisms: the NF-kB transcription factor. The further exploration of the inflammatory process in our patients revealed that the determination of IL-11 and CCL1 levels in BALF samples by a simple ELISA assay, could be used to improve the diagnosis of lung adenocarcinoma in smokers, regardless the presence or absence of COPD. The application of this assay to routine clinical practice could in the future improve the early detection of lung adenocarcinoma, the currently most common form of NSCLC, and possibly contribute to an increase in the 5-year survival rate associated with this disease.

FUTURE PERSPECTIVES

7. FUTURE PERSPECTIVES

The search for molecular biomarkers in easily accessible samples is of vital interest in LC research. The diagnosis of LC is currently commonly obtained in advanced stages of the disease and is unfortunately associated with poor survival and a lack of therapeutic options.

Our results revealed the existence of different proteomic profiles between patients with LC and COPD. The proteins included in these signatures deserve further validation and investigation as potential biomarkers for early diagnosis and prognosis, and possibly as therapeutic targets. The exploration of the inflammatory process in the samples from our patients revealed the potential of two proteins as LC biomarkers. Higher levels of IL-11 and CCL1 in BALF from smokers were associated with the presence of adenocarcinoma of the lung.

Future studies aiming at validating the findings in the present work should include a larger sample size, both in LC cases and controls. Further validation of the diagnostic performance of IL-11 and CCL1 in more accessible fluids such as plasma would be desirable and more practical for eventual screening programs or routine diagnosis. In addition, the evaluation of IL-11 and CCL1 as prognostic and therapeutic markers should also be addressed in future investigations. Furthermore, the performance of the IL-11 and CCL1 proteins on the differential diagnosis of pleural effusions (between lung adenocarcinoma, mesothelioma, adenocarcinoma of other origins, and non tumoral) would also be useful. The association of these two proteins with the development of other types of inflammation-related cancers should also be of interest.

Finally, the functional role of IL-11 and CCL1 in the carcinogenic process should also be assessed. Studies aiming at evaluating the capabilities of these two proteins in different cell lines and animal models could shed some light on their action in the lung carcinogenic process and also uncover new possible targeted therapies.

REFERENCES

8. REFERENCES

1. WHO. Global burden of disease: 2004 update. Geneva: World Health Organization; 2008.

2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011;61(2):69-90.

3. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3(11):e442.

4. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010;127(12):2893-917.

5. Alberg AJ, Ford JG, Samet JM. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007;132(3 Suppl):29S-55S.

6. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. Br Med J. 1950;2(4682):739-48.

7. Mills CA, Porter MM. Tobacco smoking habits and cancer of the mouth and respiratory system. Cancer Res. 1950;10(9):539-42.

8. Dresler C. The Tobacco Epidemic. In: Pass HI, Carbone DP, Johnson DH, Minna JD, Scagliotti GV, Turrisi AT, editors. Principles & Practice of Lung Cancer The Official Reference Text of the IASLC. 1. Philadelphia: Wolters Kluwer - Lippincott Williams & Wilkins; 2010. p. 23-32.

9. Alberg AJ, Samet JM. Epidemiology of lung cancer. Chest. 2003;123(1 Suppl):21S-49S.

10. Sasco AJ, Secretan MB, Straif K. Tobacco smoking and cancer: a brief review of recent epidemiological evidence. Lung Cancer. 2004;45 Suppl 2:S3-9.

11. Haugen A, Mollerup S. Etiology of lung cancer. In: Hansen H, editor. Textbook of Lung Cancer. 2 ed. London: Informa UK Ltd; 2008. p. 1-9.

12. IARC. World Cancer Report 2008. Geneva: World Health Organization; 2008.

13. Mattson ME, Pollack ES, Cullen JW. What are the odds that smoking will kill you? Am J Public Health. 1987;77(4):425-31.

14. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

15. Curado M, Edwards B, Shin H, Storm H, Ferlay J, Heanue M, et al., editors. Cancer incidence in five continents. Volume IX. 2008/01/01 ed2007.

16. Brambilla E, Lantejoul S. Histopathology of lung tumors. In: Hansen H, editor. Textbook of Lung Cancer. 2 ed. London: Informa UK Ltd; 2008.

17. Jackman DM, Johnson BE. Small-cell lung cancer. Lancet. 2005;366(9494):1385-96.

18. Lee CM, Sause WT. Treatment of SCLC: radiotherapy. In: Hansen H, editor. Textbook of Lung Cancer. 2 ed. London: Informa UK Ltd; 2008.

19. Brambilla E, Gazdar A. Pathogenesis of lung cancer signalling pathways: roadmap for therapies. Eur Respir J. 2009;33(6):1485-97.

20. Brambilla E, Travis WD, Colby TV, Corrin B, Shimosato Y. The new World Health Organization classification of lung tumours. Eur Respir J. 2001;18(6):1059-68.

21. Saqi A, Vazquez MF. Fine-Needle Aspiration Cytology of Benign and Malignant Tumors of the Lung. In: Pass HI, Carbone DP, Johnson DH, Minna JD, Scagliotti GV, Turrisi AT, editors. Principles & Pratice of Lung Cancer The Official Reference Text of the IASLC. 1. Philadelphia: Wolters Kluwer - Lippincott Williams & Wilkins; 2010. p. 241-56.

22. Colby TV, Noguchi M, Henschke C, Vazquez MF, Geisinger K, Yokose T, et al. Adenocarcinoma. In: Travis WD, Brambilla E, Müller-Hermelink HK, Harris CC, editors. Pathology and Genetics of Tumours of the Lung, Pleura,Thymus and Heart. Lyon: International Agency for Research on Cancer; 2004. p. 35-44.

23. Hammar SP, Brambilla C, Pugatch B, Geisinger K, Fernandez EA, Vogt P, et al. Squamous cell carcinoma. In: Travis WD, Brambilla E, Müller-Hermelink HK, Harris CC, editors. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart. Lyon: International Agency for Research on Cancer; 2004. p. 26-30.

24. Brambilla E, Pugatch B, Geisinger K, Gal A, Sheppard MN, Guinee DG, et al. Large cell carcinoma. In: Travis WD, Brambilla E, Müller-Hermelink HK, Harris CC, editors. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart. Lyon: International Agency for Research on Cancer; 2004. p. 45-50.

25. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc. 2008;83(5):584-94.

26. Detterbeck FC, Boffa DJ, Tanoue LT. The new lung cancer staging system. Chest. 2009;136(1):260-71.

27. UyBico SJ, Wu CC, Suh RD, Le NH, Brown K, Krishnam MS. Lung cancer staging essentials: the new TNM staging system and potential imaging pitfalls. Radiographics. 2010;30(5):1163-81.

28. Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest. 1997;111(6):1710-7.

29. Dusmet M, Goldstraw P. Staging, classification, and prognosis. In: Hansen H, editor. Textbook of Lung Cancer. London: Informa UK Ltd; 2008. p. 97-122.

30. Romney BM, Austin JH. Plain film evaluation of carcinoma of the lung. Semin Roentgenol. 1990;25(1):45-63.

31. Wendland MM, Sause WT. Treatment of NSCLC: radiotherapy. In: Hansen H, editor. Textbook of Lung Cancer. London: Informa UK Ltd; 2008. p. 136-46.

32. Pallis AG, Agelaki S, Georgoulias V. Treatment of NSCLC: chemotherapy. In: Hansen H, editor. Textbook of Lung Cancer. London: Informa UK Ltd; 2008. p. 137-69.

33. Pal SK, Figlin RA, Reckamp K. Targeted therapies for non-small cell lung cancer: an evolving landscape. Mol Cancer Ther. 2010;9(7):1931-44.

34. Disease GIfCOL. Global Strategy for the Diagnosis, Management and Prevention of COPD. Updated 2013: Global Initiative for Chronic Obstructive Lung Disease, Inc.; 2013.

35. Macnee W. Pathogenesis of chronic obstructive pulmonary disease. Clin Chest Med. 2007;28(3):479-513, v.

36. Kim V, Rogers TJ, Criner GJ. New concepts in the pathobiology of chronic obstructive pulmonary disease. Proc Am Thorac Soc. 2008;5(4):478-85.

37. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet. 2007;370(9589):765-73.

38. Viegi G, Pistelli F, Sherrill DL, Maio S, Baldacci S, Carrozzi L. Definition, epidemiology and natural history of COPD. Eur Respir J. 2007;30(5):993-1013.

39. Franchi M, editor. EFA Book on Chronic Obstructive Pulmonary Disease in Europe. Sharing and Caring. Brussels: European Federation of Allergy and Airways Diseases Patients Associations; 2010.

40. Pauwels RA, Rabe KF. Burden and clinical features of chronic obstructive pulmonary disease (COPD). Lancet. 2004;364(9434):613-20.

41. Molfino NA, Jeffery PK. Chronic obstructive pulmonary disease: histopathology, inflammation and potential therapies. Pulm Pharmacol Ther. 2007;20(5):462-72.

42. Cazzola M, Donner CF, Hanania NA. One hundred years of chronic obstructive pulmonary disease (COPD). Respir Med. 2007;101(6):1049-65.

43. Cosio BG, Agusti A. Update in chronic obstructive pulmonary disease 2009. Am J Respir Crit Care Med. 2010;181(7):655-60.

44. Mannino DM, Aguayo SM, Petty TL, Redd SC. Low lung function and incident lung cancer in the United States: data From the First National Health and Nutrition Examination Survey follow-up. Arch Intern Med. 2003;163(12):1475-80.

45. Wasswa-Kintu S, Gan WQ, Man SF, Pare PD, Sin DD. Relationship between reduced forced expiratory volume in one second and the risk of lung cancer: a systematic review and meta-analysis. Thorax. 2005;60(7):570-5.

46. Yao H, Rahman I. Current concepts on the role of inflammation in COPD and lung cancer. Curr Opin Pharmacol. 2009;9(4):375-83.

47. de Torres JP, Bastarrika G, Wisnivesky JP, Alcaide AB, Campo A, Seijo LM, et al. Assessing the relationship between lung cancer risk and emphysema detected on low-dose CT of the chest. Chest. 2007;132(6):1932-8.

48. Punturieri A, Szabo E, Croxton TL, Shapiro SD, Dubinett SM. Lung cancer and chronic obstructive pulmonary disease: needs and opportunities for integrated research. J Natl Cancer Inst. 2009;101(8):554-9.

49. Brody JS, Spira A. State of the art. Chronic obstructive pulmonary disease, inflammation, and lung cancer. Proc Am Thorac Soc. 2006;3(6):535-7.

50. Lee G, Walser TC, Dubinett SM. Chronic inflammation, chronic obstructive pulmonary disease, and lung cancer. Curr Opin Pulm Med. 2009;15(4):303-7.

51. Nathan C. Points of control in inflammation. Nature. 2002;420(6917):846-52.

52. Lee J, Taneja V, Vassallo R. Cigarette smoking and inflammation: cellular and molecular mechanisms. J Dent Res. 2012;91(2):142-9.

53. Yang IA, Relan V, Wright CM, Davidson MR, Sriram KB, Savarimuthu Francis SM, et al. Common pathogenic mechanisms and pathways in the development of COPD and lung cancer. Expert Opin Ther Targets. 2011;15(4):439-56.

54. Trinchieri G. Inflammation in cancer: a therapeutic target? Oncology (Williston Park). 2011;25(5):418-20.

55. Dvorak HF. Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. N Engl J Med. 1986;315(26):1650-9.

56. Siemes C, Visser LE, Coebergh JW, Splinter TA, Witteman JC, Uitterlinden AG, et al. C-reactive protein levels, variation in the C-reactive protein gene, and cancer risk: the Rotterdam Study. J Clin Oncol. 2006;24(33):5216-22.

57. Parimon T, Chien JW, Bryson CL, McDonell MB, Udris EM, Au DH. Inhaled corticosteroids and risk of lung cancer among patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2007;175(7):712-9.

58. Moghaddam SJ, Li H, Cho SN, Dishop MK, Wistuba, II, Ji L, et al. Promotion of lung carcinogenesis by chronic obstructive pulmonary disease-like airway inflammation in a K-ras-induced mouse model. Am J Respir Cell Mol Biol. 2009;40(4):443-53.

59. Agullo-Ortuno MT, Lopez-Rios F, Paz-Ares L. Lung cancer genomic signatures. J Thorac Oncol. 2010;5(10):1673-91.

60. Ocak S, Chaurand P, Massion PP. Mass spectrometry-based proteomic profiling of lung cancer. Proc Am Thorac Soc. 2009;6(2):159-70.

61. Aebersold R, Goodlett DR. Mass spectrometry in proteomics. Chem Rev. 2001;101(2):269-95.

62. Dhamoon AS, Kohn EC, Azad NS. The ongoing evolution of proteomics in malignancy. Drug Discov Today. 2007;12(17-18):700-8.

63. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. J Biol Chem. 1975;250(10):4007-21.

64. Conrad DH, Goyette J, Thomas PS. Proteomics as a method for early detection of cancer: a review of proteomics, exhaled breath condensate, and lung cancer screening. J Gen Intern Med. 2008;23 Suppl 1:78-84.

65. Hirsch J, Hansen KC, Burlingame AL, Matthay MA. Proteomics: current techniques and potential applications to lung disease. Am J Physiol Lung Cell Mol Physiol. 2004;287(1):L1-23.

66. Pastor MD, Nogal A, Molina-Pinelo S, Paz-Ares L, Carnero A. Oncoproteomic Approaches in Lung Cancer Research. In: López-Camarillo Cs, Aréchaga-Ocampo E, editors. Oncogenomics and Cancer Proteomics - Novel Approaches in Biomarkers Discovery and Therapeutic Targets in Cancer: InTech; 2013. p. 169-84.

67. Massion PP, Caprioli RM. Proteomic strategies for the characterization and the early detection of lung cancer. J Thorac Oncol. 2006;1(9):1027-39.

68. Mann M. Functional and quantitative proteomics using SILAC. Nat Rev Mol Cell Biol. 2006;7(12):952-8.

69. Graves PR, Haystead TA. Molecular biologist's guide to proteomics. Microbiol Mol Biol Rev. 2002;66(1):39-63; table of contents.

70. DeSouza L, Diehl G, Rodrigues MJ, Guo J, Romaschin AD, Colgan TJ, et al. Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. J Proteome Res. 2005;4(2):377-86.

71. Bowler RP, Ellison MC, Reisdorph N. Proteomics in pulmonary medicine. Chest. 2006;130(2):567-74.

72. Domon B, Aebersold R. Mass spectrometry and protein analysis. Science. 2006;312(5771):212-7.

73. Hu R, Ouyang Q, Dai A, Tan S, Xiao Z, Tang C. Heat shock protein 27 and cyclophilin A associate with the pathogenesis of COPD. Respirology. 2011;16(6):983-93.

74. Lee EJ, In KH, Kim JH, Lee SY, Shin C, Shim JJ, et al. Proteomic analysis in lung tissue of smokers and COPD patients. Chest. 2009;135(2):344-52.

75. Ohlmeier S, Vuolanto M, Toljamo T, Vuopala K, Salmenkivi K, Myllarniemi M, et al. Proteomics of human lung tissue identifies surfactant protein A as a marker of chronic obstructive pulmonary disease. J Proteome Res. 2008;7(12):5125-32.

76. Alexandre BM, Charro N, Blonder J, Lopes C, Azevedo P, Bugalho de Almeida A, et al. Profiling the erythrocyte membrane proteome isolated from patients diagnosed with chronic obstructive pulmonary disease. J Proteomics. 2012.

77. Bowler RP, Canham ME, Ellison MC. Surface enhanced laser desorption/ionization (SELDI) time-of-flight mass spectrometry to identify patients with chronic obstructive pulmonary disease. COPD. 2006;3(1):41-50.

78. Bozinovski S, Hutchinson A, Thompson M, Macgregor L, Black J, Giannakis E, et al. Serum amyloid a is a biomarker of acute exacerbations of chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2008;177(3):269-78.

79. Verrills NM, Irwin JA, He XY, Wood LG, Powell H, Simpson JL, et al. Identification of novel diagnostic biomarkers for asthma and chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2011;183(12):1633-43.

80. Nicholas B, Djukanovic R. Induced sputum: a window to lung pathology. Biochem Soc Trans. 2009;37(Pt 4):868-72.

81. Ohlmeier S, Mazur W, Linja-Aho A, Louhelainen N, Ronty M, Toljamo T, et al. Sputum proteomics identifies elevated PIGR levels in smokers and mild-to-moderate COPD. J Proteome Res. 2012;11(2):599-608.

82. Hassanein M, Rahman JS, Chaurand P, Massion PP. Advances in proteomic strategies toward the early detection of lung cancer. Proc Am Thorac Soc. 2011;8(2):183-8.

83. Kang SM, Sung HJ, Ahn JM, Park JY, Lee SY, Park CS, et al. The Haptoglobin beta chain as a supportive biomarker for human lung cancers. Molecular BioSystems. 2011;7(4):1167-75.

84. Sung HJ, Ahn JM, Yoon YH, Rhim TY, Park CS, Park JY, et al. Identification and validation of SAA as a potential lung cancer biomarker and its involvement in metastatic pathogenesis of lung cancer. Journal of Proteome Research. 2011;10(3):1383-95.

85. Heo SH, Lee SJ, Ryoo HM, Park JY, Cho JY. Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC-MS/MS. Proteomics. 2007;7(23):4292-302.

86. Ueda K, Katagiri T, Shimada T, Irie S, Sato TA, Nakamura Y, et al. Comparative profiling of serum glycoproteome by sequential purification of glycoproteins and 2-nitrobenzensulfenyl (NBS) stable isotope labeling: a new approach for the novel biomarker discovery for cancer. Journal of Proteome Research. 2007;6(9):3475-83.

87. Zeng X, Hood BL, Sun M, Conrads TP, Day RS, Weissfeld JL, et al. Lung cancer serum biomarker discovery using glycoprotein capture and liquid chromatography mass spectrometry. J Proteome Res. 2010;9(12):6440-9.

88. Patz EF, Jr., Campa MJ, Gottlin EB, Kusmartseva I, Guan XR, Herndon JE, 2nd. Panel of serum biomarkers for the diagnosis of lung cancer. Journal of Clinical Oncology. 2007;25(35):5578-83.

89. Taguchi F, Solomon B, Gregorc V, Roder H, Gray R, Kasahara K, et al. Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. Journal of the National Cancer Institute. 2007;99(11):838-46.

90. Han KQ, Huang G, Gao CF, Wang XL, Ma B, Sun LQ, et al. Identification of lung cancer patients by serum protein profiling using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry. Am J Clin Oncol. 2008;31(2):133-9.

91. Pernemalm M, De Petris L, Eriksson H, Branden E, Koyi H, Kanter L, et al. Use of narrow-range peptide IEF to improve detection of lung adenocarcinoma markers in plasma and pleural effusion. Proteomics. 2009;9(13):3414-24.

92. Soltermann A, Ossola R, Kilgus-Hawelski S, von Eckardstein A, Suter T, Aebersold R, et al. N-glycoprotein profiling of lung adenocarcinoma pleural effusions by shotgun proteomics. Cancer. 2008;114(2):124-33.

93. Chen G, Gharib TG, Huang CC, Thomas DG, Shedden KA, Taylor JM, et al. Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. Clinical Cancer Research. 2002;8(7):2298-305.

94. Campa MJ, Wang MZ, Howard B, Fitzgerald MC, Patz EF, Jr. Protein expression profiling identifies macrophage migration inhibitory factor and cyclophilin a as potential molecular targets in non-small cell lung cancer. Cancer Res. 2003;63(7):1652-6.

95. Alfonso P, Catala M, Rico-Morales ML, Durante-Rodriguez G, Moro-Rodriguez E, Fernandez-Garcia H, et al. Proteomic analysis of lung biopsies: Differential protein expression profile between peritumoral and tumoral tissue. Proteomics. 2004;4(2):442-7.

96. Xiao H, Zhang L, Zhou H, Lee JM, Garon EB, Wong DT. Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. Mol Cell Proteomics. 2012;11(2):M111 012112.

97. Chen G, Gharib TG, Wang H, Huang CC, Kuick R, Thomas DG, et al. Protein profiles associated with survival in lung adenocarcinoma. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(23):13537-42.

98. Chen G, Wang H, Gharib TG, Huang CC, Thomas DG, Shedden KA, et al. Overexpression of oncoprotein 18 correlates with poor differentiation in lung adenocarcinomas. Mol Cell Proteomics. 2003;2(2):107-16.

99. De Petris L, Orre LM, Kanter L, Pernemalm M, Koyi H, Lewensohn R, et al. Tumor expression of S100A6 correlates with survival of patients with stage I non-small-cell lung cancer. Lung Cancer. 2009;63(3):410-7.

100. Tang SJ, Ho MY, Cho HC, Lin YC, Sun GH, Chi KH, et al. Phosphoglycerate kinase 1-overexpressing lung cancer cells reduce cyclooxygenase 2 expression and promote anti-tumor immunity in vivo. International Journal of Cancer. 2008;123(12):2840-8.

101. Tian T, Hao J, Xu A, Luo C, Liu C, Huang L, et al. Determination of metastasis-associated proteins in non-small cell lung cancer by comparative proteomic analysis. Cancer Science. 2007;98(8):1265-74.

102. De Petris L, Branden E, Herrmann R, Sanchez BC, Koyi H, Linderholm B, et al. Diagnostic and prognostic role of plasma levels of two forms of cytokeratin 18 in patients with non-small-cell lung cancer. European Journal of Cancer. 2011;47(1):131-7.

103. Gharib TG, Chen G, Wang H, Huang CC, Prescott MS, Shedden K, et al. Proteomic analysis of cytokeratin isoforms uncovers association with survival in lung adenocarcinoma. Neoplasia. 2002;4(5):440-8.

104. Liu YF, Xiao ZQ, Li MX, Li MY, Zhang PF, Li C, et al. Quantitative proteome analysis reveals annexin A3 as a novel biomarker in lung adenocarcinoma. Journal of Pathology. 2009;217(1):54-64.

105. Xu BJ, Gonzalez AL, Kikuchi T, Yanagisawa K, Massion PP, Wu H, et al. MALDI-MS derived prognostic protein markers for resected non-small cell lung cancer. Proteomics - Clinical Applications. 2008;2(10-11):1508-17.

106. Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, et al. Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet. 2003;362(9382):433-9.

107. Yanagisawa K, Tomida S, Shimada Y, Yatabe Y, Mitsudomi T, Takahashi T. A 25-signal proteomic signature and outcome for patients with resected non-small-cell lung cancer. J Natl Cancer Inst. 2007;99(11):858-67.

108. Qu Y, Yang Y, Liu B, Xiao W. Comparative proteomic profiling identified sorcin being associated with gemcitabine resistance in non-small cell lung cancer. Med Oncol. 2010;27(4):1303-8.

109. Han M, Liu Q, Yu J, Zheng S. Identification of candidate molecular markers predicting chemotherapy resistance in non-small cell lung cancer. Clin Chem Lab Med. 2010;48(6):863-7.

110. Voortman J, Pham TV, Knol JC, Giaccone G, Jimenez CR. Prediction of outcome of non-small cell lung cancer patients treated with chemotherapy and bortezomib by time-course MALDI-TOF-MS serum peptide profiling. Proteome Science. 2009;7:34.

111. Oumeraci T, Schmidt B, Wolf T, Zapatka M, Pich A, Brors B, et al. Bronchoalveolar lavage fluid of lung cancer patients: mapping the uncharted waters using proteomics technology. Lung Cancer. 2011;72(1):136-8.

112. von Bredow C, Birrer P, Griese M. Surfactant protein A and other bronchoalveolar lavage fluid proteins are altered in cystic fibrosis. Eur Respir J. 2001;17(4):716-22.

113. Wu J, Kobayashi M, Sousa EA, Liu W, Cai J, Goldman SJ, et al. Differential proteomic analysis of bronchoalveolar lavage fluid in asthmatics following segmental antigen challenge. Mol Cell Proteomics. 2005;4(9):1251-64.

114. Bowler RP, Duda B, Chan ED, Enghild JJ, Ware LB, Matthay MA, et al. Proteomic analysis of pulmonary edema fluid and plasma in patients with acute lung injury. Am J Physiol Lung Cell Mol Physiol. 2004;286(6):L1095-104.

115. Magi B, Bargagli E, Bini L, Rottoli P. Proteome analysis of bronchoalveolar lavage in lung diseases. Proteomics. 2006;6(23):6354-69.

116. Schnapp LM, Donohoe S, Chen J, Sunde DA, Kelly PM, Ruzinski J, et al. Mining the acute respiratory distress syndrome proteome: identification of the insulin-like growth factor (IGF)/IGF-binding protein-3 pathway in acute lung injury. Am J Pathol. 2006;169(1):86-95.

117. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem. 1996;68(5):850-8.

118. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5.

119. Medina I, Carbonell J, Pulido L, Madeira SC, Goetz S, Conesa A, et al. Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. Nucleic Acids Res. 2010;38(Web Server issue):W210-3.

120. Vallieres E, Shepherd FA, Crowley J, Van Houtte P, Postmus PE, Carney D, et al. The IASLC Lung Cancer Staging Project: proposals regarding the relevance of TNM in the pathologic staging of small cell lung cancer in the forthcoming (seventh) edition of the TNM classification for lung cancer. J Thorac Oncol. 2009;4(9):1049-59.

121. Manser R, Lethaby A, Irving LB, Stone C, Byrnes G, Abramson MJ, et al. Screening for lung cancer. Cochrane Database Syst Rev. 2013;6:CD001991.

122. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395-409.

123. Jay SJ, Wehr K, Nicholson DP, Smith AL. Diagnostic sensitivity and specificity of pulmonary cytology: comparison of techniques used in conjunction with flexible fiber optic bronchoscopy. Acta Cytol. 1980;24(4):304-12.

124. Li QK, Shah P, Li Y, Aiyetan PO, Chen J, Yung R, et al. Glycoproteomic analysis of bronchoalveolar lavage (BAL) fluid identifies tumor-associated glycoproteins from lung adenocarcinoma. J Proteome Res. 2013;12(8):3689-96.

125. Bigbee WL, Gopalakrishnan V, Weissfeld JL, Wilson DO, Dacic S, Lokshin AE, et al. A multiplexed serum biomarker immunoassay panel discriminates clinical lung cancer patients from high-risk individuals found to be cancer-free by CT screening. J Thorac Oncol. 2012;7(4):698-708.

126. Kikuchi T, Hassanein M, Amann JM, Liu Q, Slebos RJ, Rahman SM, et al. In-depth proteomic analysis of nonsmall cell lung cancer to discover molecular targets and candidate biomarkers. Mol Cell Proteomics. 2012;11(10):916-32.

127. Nolen BM, Langmead CJ, Choi S, Lomakin A, Marrangoni A, Bigbee WL, et al. Serum biomarker profiles as diagnostic tools in lung cancer. Cancer Biomark. 2011;10(1):3-12.

128. Yang P, Sun Z, Krowka MJ, Aubry MC, Bamlet WR, Wampfler JA, et al. Alpha1-antitrypsin deficiency carriers, tobacco smoke, chronic obstructive pulmonary disease, and lung cancer risk. Arch Intern Med. 2008;168(10):1097-103.

129. Houghton AM. Mechanistic links between COPD and lung cancer. Nat Rev Cancer. 2013;13(4):233-45.

130. Hogg JC, Chu F, Utokaparch S, Woods R, Elliott WM, Buzatu L, et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. N Engl J Med. 2004;350(26):2645-53.

131. Martinkova J, Gadher SJ, Hajduch M, Kovarova H. Challenges in cancer research and multifaceted approaches for cancer biomarker quest. FEBS Lett. 2009;583(11):1772-84.

132. Indovina P, Marcelli E, Maranta P, Tarro G. Lung cancer proteomics: recent advances in biomarker discovery. Int J Proteomics. 2011;2011:726869.

133. Miller I, Eberini I, Gianazza E. Proteomics of lung physiopathology. Proteomics. 2008;8(23-24):5053-73.

134. Hacker S, Lambers C, Hoetzenecker K, Pollreisz A, Aigner C, Lichtenauer M, et al. Elevated HSP27, HSP70 and HSP90 alpha in chronic obstructive pulmonary disease: markers for immune activation and tissue destruction. Clin Lab. 2009;55(1-2):31-40.

135. Volm M, Mattern J, Stammler G. Up-regulation of heat shock protein 70 in adenocarcinomas of the lung in smokers. Anticancer Res. 1995;15(6B):2607-9.

136. Yoshihara T, Kadota Y, Yoshimura Y, Tatano Y, Takeuchi N, Okitsu H, et al. Proteomic alteration in gastic adenocarcinomas from Japanese patients. Mol Cancer. 2006;5:75.

137. Lee IN, Chen CH, Sheu JC, Lee HS, Huang GT, Yu CY, et al. Identification of human hepatocellular carcinoma-related biomarkers by two-dimensional difference gel electrophoresis and mass spectrometry. J Proteome Res. 2005;4(6):2062-9.

138. Jazii FR, Najafi Z, Malekzadeh R, Conrads TP, Ziaee AA, Abnet C, et al. Identification of squamous cell carcinoma associated proteins by proteomics and loss of beta tropomyosin expression in esophageal cancer. World J Gastroenterol. 2006;12(44):7104-12.

139. Allal AS, Kahne T, Reverdin AK, Lippert H, Schlegel W, Reymond MA. Radioresistance-related proteins in rectal cancer. Proteomics. 2004;4(8):2261-9.

140. Bottoni P, Giardina B, Scatena R. Proteomic profiling of heat shock proteins: An emerging molecular approach with direct pathophysiological and clinical implications. Proteomics Clin Appl. 2009;3(6):636-53.

141. Castagna A, Antonioli P, Astner H, Hamdan M, Righetti SC, Perego P, et al. A proteomic approach to cisplatin resistance in the cervix squamous cell carcinoma cell line A431. Proteomics. 2004;4(10):3246-67.

142. Smith L, Welham KJ, Watson MB, Drew PJ, Lind MJ, Cawkwell L. The proteomic analysis of cisplatin resistance in breast cancer cells. Oncol Res. 2007;16(11):497-506.

143. Potter CP, Harris AL. Diagnostic, prognostic and therapeutic implications of carbonic anhydrases in cancer. Br J Cancer. 2003;89(1):2-7.

144. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nat Rev Cancer. 2011;11(2):85-95.

145. Yu GP, Chen GQ, Wu S, Shen K, Ji Y. The expression of PEBP4 protein in lung squamous cell carcinoma. Tumour Biol. 2011;32(6):1257-63.

146. Yu GP, Huang B, Chen GQ, Wu S, Ji Y, Shen ZY. PEBP4 gene expression and its significance in invasion and metastasis of non-small cell lung cancer. Tumour Biol. 2012;33(1):223-8.

147. Liu H, Kong Q, Li B, He Y, Li P, Jia B. Expression of PEBP4 protein correlates with the invasion and metastasis of colorectal cancer. Tumour Biol. 2012;33(1):267-73.

148. Chotirat S, Thongnoppakhun W, Promsuwicha O, Boonthimat C, Auewarakul CU. Molecular alterations of isocitrate dehydrogenase 1 and 2 (IDH1 and IDH2) metabolic genes and additional genetic mutations in newly diagnosed acute myeloid leukemia patients. J Hematol Oncol. 2012;5:5.

149. Tan F, Jiang Y, Sun N, Chen Z, Lv Y, Shao K, et al. Identification of isocitrate dehydrogenase 1 as a potential diagnostic and prognostic biomarker for non-small cell lung cancer by proteomic analysis. Mol Cell Proteomics. 2012;11(2):M111 008821.

150. Sun N, Chen Z, Tan F, Zhang B, Yao R, Zhou C, et al. Isocitrate dehydrogenase 1 is a novel plasma biomarker for the diagnosis of non-small cell lung cancer. Clin Cancer Res. 2013;19(18):5136-45.

151. Garcia M, Derocq D, Pujol P, Rochefort H. Overexpression of transfected cathepsin D in transformed cells increases their malignant phenotype and metastatic potency. Oncogene. 1990;5(12):1809-14.

152. Kim H, Kang HJ, You KT, Kim SH, Lee KY, Kim TI, et al. Suppression of human selenium-binding protein 1 is a late event in colorectal carcinogenesis and is associated with poor survival. Proteomics. 2006;6(11):3466-76.

153. Zeng H, Xu L, Xiao D, Zhang H, Wu X, Zheng R, et al. Altered expression of ezrin in esophageal squamous cell carcinoma. J Histochem Cytochem. 2006;54(8):889-96.

154. Wang Z, Zhao X. [Expression and prognostic relation of cathepsin D in non-small cell lung cancer tissues and lymph nodes]. Zhonghua Jie He He Hu Xi Za Zhi. 1998;21(3):164-6.

155. Chen G, Wang H, Miller CT, Thomas DG, Gharib TG, Misek DE, et al. Reduced selenium-binding protein 1 expression is associated with poor outcome in lung adenocarcinomas. J Pathol. 2004;202(3):321-9.

156. Zhang XQ, Chen GP, Wu T, Yan JP, Zhou JY. Expression and clinical significance of ezrin in non--small-cell lung cancer. Clin Lung Cancer. 2012;13(3):196-204.

157. Li CP, Goto A, Watanabe A, Murata K, Ota S, Niki T, et al. AKR1B10 in usual interstitial pneumonia: expression in squamous metaplasia in association with smoking and lung cancer. Pathol Res Pract. 2008;204(5):295-304.

158. Kang MW, Lee ES, Yoon SY, Jo J, Lee J, Kim HK, et al. AKR1B10 is associated with smoking and smoking-related non-small-cell lung cancer. J Int Med Res. 2011;39(1):78-85.

159. Ruiz FX, Gallego O, Ardevol A, Moro A, Dominguez M, Alvarez S, et al. Aldo-keto reductases from the AKR1B subfamily: retinoid specificity and control of cellular retinoic acid levels. Chem Biol Interact. 2009;178(1-3):171-7.

160. Tang XH, Gudas LJ. Retinoids, retinoic acid receptors, and cancer. Annu Rev Pathol. 2011;6:345-64.

161. Matsunaga T, Wada Y, Endo S, Soda M, El-Kabbani O, Hara A. Aldo-Keto Reductase 1B10 and Its Role in Proliferation Capacity of Drug-Resistant Cancers. Front Pharmacol. 2012;3:5.

162. Muzio G, Maggiora M, Paiuzzi E, Oraldi M, Canuto RA. Aldehyde dehydrogenases and cell proliferation. Free radical biology & medicine. 2012;52(4):735-46.

163. Eigenbrodt E, Kallinowski F, Ott M, Mazurek S, Vaupel P. Pyruvate kinase and the interaction of amino acid and carbohydrate metabolism in solid tumors. Anticancer Res. 1998;18(5A):3267-74.

164. Christofk HR, Vander Heiden MG, Harris MH, Ramanathan A, Gerszten RE, Wei R, et al. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. Nature. 2008;452(7184):230-3.

165. Bandow JE, Baker JD, Berth M, Painter C, Sepulveda OJ, Clark KA, et al. Improved image analysis workflow for 2-D gels enables large-scale 2-D gel-based proteomics studies--COPD biomarker discovery study. Proteomics. 2008;8(15):3030-41.

166. Comandini A, Marzano V, Curradi G, Federici G, Urbani A, Saltini C. Markers of anti-oxidant response in tobacco smoke exposed subjects: a data-mining review. Pulm Pharmacol Ther. 2010;23(6):482-92.

167. Alexandre BM, Charro N, Blonder J, Lopes C, Azevedo P, Bugalho de Almeida A, et al. Profiling the erythrocyte membrane proteome isolated from patients diagnosed with chronic obstructive pulmonary disease. J Proteomics. 2012;76 Spec No.:259-69.

168. Pierrou S, Broberg P, O'Donnell RA, Pawlowski K, Virtala R, Lindqvist E, et al. Expression of genes involved in oxidative stress responses in airway epithelial cells of smokers with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2007;175(6):577-86.

169. Lehtonen ST, Svensk AM, Soini Y, Paakko P, Hirvikoski P, Kang SW, et al. Peroxiredoxins, a novel protein family in lung cancer. Int J Cancer. 2004;111(4):514-21.

170. Kim JH, Bogner PN, Ramnath N, Park Y, Yu J, Park YM. Elevated peroxiredoxin 1, but not NF-E2-related factor 2, is an independent prognostic factor for disease recurrence and reduced survival in stage I non-small cell lung cancer. Clin Cancer Res. 2007;13(13):3875-82.

171. Warburg O. On the origin of cancer cells. Science. 1956;123(3191):309-14.

172. Ros S, Schulze A. Balancing glycolytic flux: the role of 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatases in cancer metabolism. Cancer & metabolism. 2013;1(1):8.

173. Luo C, Urgard E, Vooder T, Metspalu A. The role of COX-2 and Nrf2/ARE in anti-inflammation and antioxidative stress: Aging and anti-aging. Med Hypotheses. 2011;77(2):174-8.

174. Holmgren A, Lu J. Thioredoxin and thioredoxin reductase: current research with special reference to human disease. Biochem Biophys Res Commun. 2010;396(1):120-4.

175. Kakolyris S, Giatromanolaki A, Koukourakis M, Powis G, Souglakos J, Sivridis E, et al. Thioredoxin expression is associated with lymph node status and prognosis in early operable non-small cell lung cancer. Clin Cancer Res. 2001;7(10):3087-91.

176. Kim SJ, Miyoshi Y, Taguchi T, Tamaki Y, Nakamura H, Yodoi J, et al. High thioredoxin expression is associated with resistance to docetaxel in primary breast cancer. Clin Cancer Res. 2005;11(23):8425-30.

177. Pan H, Luo C, Li R, Qiao A, Zhang L, Mines M, et al. Cyclophilin A is required for CXCR4-mediated nuclear export of heterogeneous nuclear ribonucleoprotein A2, activation and nuclear translocation of ERK1/2, and chemotactic cell migration. J Biol Chem. 2008;283(1):623-37.

178. Zhu C, Wang X, Deinum J, Huang Z, Gao J, Modjtahedi N, et al. Cyclophilin A participates in the nuclear translocation of apoptosis-inducing factor in neurons after cerebral hypoxia-ischemia. J Exp Med. 2007;204(8):1741-8.

179. Peng Y, Li X, Wu M, Yang J, Liu M, Zhang W, et al. New prognosis biomarkers identified by dynamic proteomic analysis of colorectal cancer. Mol Biosyst. 2012;8(11):3077-88.

180. Lee J. Role of cyclophilin a during oncogenesis. Archives of pharmacal research. 2010;33(2):181-7.

181. Howard BA, Furumai R, Campa MJ, Rabbani ZN, Vujaskovic Z, Wang XF, et al. Stable RNA interference-mediated suppression of cyclophilin A diminishes non-small-cell lung tumor growth in vivo. Cancer research. 2005;65(19):8853-60.

182. Ceccarelli J, Delfino L, Zappia E, Castellani P, Borghi M, Ferrini S, et al. The redox state of the lung cancer microenvironment depends on the levels of thioredoxin expressed by tumor cells and affects tumor progression and response to prooxidants. Int J Cancer. 2008;123(8):1770-8.

183. Jones DT, Pugh CW, Wigfield S, Stevens MF, Harris AL. Novel thioredoxin inhibitors paradoxically increase hypoxia-inducible factor-alpha expression but decrease functional transcriptional activity, DNA binding, and degradation. Clin Cancer Res. 2006;12(18):5384-94.

184. Wang CY, Chen CL, Tseng YL, Fang YT, Lin YS, Su WC, et al. Annexin A2 silencing induces G2 arrest of non-small cell lung cancer cells through p53-dependent and -independent mechanisms. The Journal of biological chemistry. 2012;287(39):32512-24.

185. Castro MA, Dal-Pizzol F, Zdanov S, Soares M, Muller CB, Lopes FM, et al. CFL1 expression levels as a prognostic and drug resistance marker in nonsmall cell lung cancer. Cancer. 2010;116(15):3645-55.

186. Chaturvedi AK, Caporaso NE, Katki HA, Wong HL, Chatterjee N, Pine SR, et al. C-reactive protein and risk of lung cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(16):2719-26.

187. Yang J, Moses MA. Lipocalin 2: a multifaceted modulator of human cancer. Cell cycle. 2009;8(15):2347-52.

188. Cai Z, Tchou-Wong KM, Rom WN. NF-kappaB in lung tumorigenesis. Cancers. 2011;3(4):4258-68.

189. Tang X, Liu D, Shishodia S, Ozburn N, Behrens C, Lee JJ, et al. Nuclear factor-kappaB (NF-kappaB) is frequently expressed in lung cancer and preneoplastic lesions. Cancer. 2006;107(11):2637-46.

190. Dougan M, Li D, Neuberg D, Mihm M, Googe P, Wong KK, et al. A dual role for the immune response in a mouse model of inflammation-associated lung cancer. The Journal of clinical investigation. 2011;121(6):2436-46.

191. Wong KK, Jacks T, Dranoff G. NF-kappaB fans the flames of lung carcinogenesis. Cancer Prev Res (Phila). 2010;3(4):403-5.

192. Hoesel B, Schmid JA. The complexity of NF-kappaB signaling in inflammation and cancer. Molecular cancer. 2013;12:86.

193. Chu XY, Hou XB, Song WA, Xue ZQ, Wang B, Zhang LB. Diagnostic values of SCC, CEA, Cyfra21-1 and NSE for lung cancer in patients with suspicious pulmonary masses: a single center analysis. Cancer biology & therapy. 2011;11(12):995-1000.

194. Vinolas N, Molina R, Galan MC, Casas F, Callejas MA, Filella X, et al. Tumor markers in response monitoring and prognosis of non-small cell lung cancer: preliminary report. Anticancer research. 1998;18(1B):631-4.

195. Sanchez De Cos J, Masa F, de la Cruz JL, Disdier C, Vergara C. Squamous cell carcinoma antigen (SCC Ag) in the diagnosis and prognosis of lung cancer. Chest. 1994;105(3):773-6.

196. Niewoehner DE, Rubins JB. Clinical utility of tumor markers in the management of non-small cell lung cancer. Methods in molecular medicine. 2003;75:135-41.

197. Tomita M, Shimizu T, Ayabe T, Yonei A, Onitsuka T. Prognostic significance of tumour marker index based on preoperative CEA and CYFRA 21-1 in non-small cell lung cancer. Anticancer research. 2010;30(7):3099-102.

198. Trape J, Buxo J, Perez de Olaguer J, Vidal C. Tumor markers as prognostic factors in treated non-small cell lung cancer. Anticancer research. 2003;23(5b):4277-81.

199. Holdenrieder S, Nagel D, Stieber P. Estimation of prognosis by circulating biomarkers in patients with non-small cell lung cancer. Cancer biomarkers : section A of Disease markers. 2010;6(3-4):179-90.

200. Ji M, Zhang Y, Shi B, Hou P. Association of promoter methylation with histologic type and pleural indentation in non-small cell lung cancer (NSCLC). Diagn Pathol. 2011;6:48.

201. Wei Y, Tong J, Taylor P, Strumpf D, Ignatchenko V, Pham NA, et al. Primary tumor xenografts of human lung adeno and squamous cell carcinoma express distinct proteomic signatures. J Proteome Res. 2011;10(1):161-74.

202. Daraselia N, Wang Y, Budoff A, Lituev A, Potapova O, Vansant G, et al. Molecular signature and pathway analysis of human primary squamous and adenocarcinoma lung cancers. Am J Cancer Res. 2012;2(1):93-103.

203. Lee CG, Hartl D, Matsuura H, Dunlop FM, Scotney PD, Fabri LJ, et al. Endogenous IL-11 signaling is essential in Th2- and IL-13-induced inflammation and mucus production. American journal of respiratory cell and molecular biology. 2008;39(6):739-46.

204. Putoczki T, Ernst M. More than a sidekick: the IL-6 family cytokine IL-11 links inflammation to cancer. Journal of leukocyte biology. 2010;88(6):1109-17.

205. Bromberg J. Stat proteins and oncogenesis. The Journal of clinical investigation. 2002;109(9):1139-42.

206. Tang W, Yang L, Yang YC, Leng SX, Elias JA. Transforming growth factor-beta stimulates interleukin-11 transcription via complex activating protein-1-dependent pathways. J Biol Chem. 1998;273(10):5506-13.

207. Ernst M, Najdovska M, Grail D, Lundgren-May T, Buchert M, Tye H, et al. STAT3 and STAT1 mediate IL-11-dependent and inflammation-associated gastric tumorigenesis in gp130 receptor mutant mice. J Clin Invest. 2008;118(5):1727-38.

208. Mechta F, Lallemand D, Pfarr CM, Yaniv M. Transformation by ras modifies AP1 composition and activity. Oncogene. 1997;14(7):837-47.

209. Casalino L, De Cesare D, Verde P. Accumulation of Fra-1 in ras-transformed cells depends on both transcriptional autoregulation and MEK-dependent posttranslational stabilization. Mol Cell Biol. 2003;23(12):4401-15.

210. Verde P, Casalino L, Talotta F, Yaniv M, Weitzman JB. Deciphering AP-1 function in tumorigenesis: fra-ternizing on target promoters. Cell Cycle. 2007;6(21):2633-9.

211. Shin SY, Choi C, Lee HG, Lim Y, Lee YH. Transcriptional regulation of the interleukin-11 gene by oncogenic Ras. Carcinogenesis. 2012;33(12):2467-76.

212. Campbell CL, Guardiani R, Ollari C, Nelson BE, Quesenberry PJ, Savarese TM. Interleukin-11 receptor expression in primary ovarian carcinomas. Gynecol Oncol. 2001;80(2):121-7.

213. Campbell CL, Jiang Z, Savarese DM, Savarese TM. Increased expression of the interleukin-11 receptor and evidence of STAT3 activation in prostate carcinoma. Am J Pathol. 2001;158(1):25-32.

214. Hanavadi S, Martin TA, Watkins G, Mansel RE, Jiang WG. Expression of interleukin 11 and its receptor and their prognostic value in human breast cancer. Ann Surg Oncol. 2006;13(6):802-8.

215. Yamazumi K, Nakayama T, Kusaba T, Wen CY, Yoshizaki A, Yakata Y, et al. Expression of interleukin-11 and interleukin-11 receptor alpha in human colorectal

adenocarcinoma; immunohistochemical analyses and correlation with clinicopathological factors. World J Gastroenterol. 2006;12(2):317-21.

216. Nakayama T, Yoshizaki A, Izumida S, Suehiro T, Miura S, Uemura T, et al. Expression of interleukin-11 (IL-11) and IL-11 receptor alpha in human gastric carcinoma and IL-11 upregulates the invasive activity of human gastric carcinoma cells. Int J Oncol. 2007;30(4):825-33.

217. Putoczki TL, Thiem S, Loving A, Busuttil RA, Wilson NJ, Ziegler PK, et al. Interleukin-11 is the dominant IL-6 family cytokine during gastrointestinal tumorigenesis and can be targeted therapeutically. Cancer cell. 2013;24(2):257-71.

218. Onnis B, Fer N, Rapisarda A, Perez VS, Melillo G. Autocrine production of IL-11 mediates tumorigenicity in hypoxic cancer cells. The Journal of clinical investigation. 2013;123(4):1615-29.

219. Lee HT, Park SW, Kim M, Ham A, Anderson LJ, Brown KM, et al. Interleukin-11 protects against renal ischemia and reperfusion injury. American journal of physiology Renal physiology. 2012;303(8):F1216-24.

220. Tang W, Geba GP, Zheng T, Ray P, Homer RJ, Kuhn C, 3rd, et al. Targeted expression of IL-11 in the murine airway causes lymphocytic inflammation, bronchial remodeling, and airways obstruction. The Journal of clinical investigation. 1996;98(12):2845-53.

221. Zheng T, Zhu Z, Wang J, Homer RJ, Elias JA. IL-11: insights in asthma from overexpression transgenic modeling. The Journal of allergy and clinical immunology. 2001;108(4):489-96.

222. Yeh HH, Lai WW, Chen HH, Liu HS, Su WC. Autocrine IL-6-induced Stat3 activation contributes to the pathogenesis of lung adenocarcinoma and malignant pleural effusion. Oncogene. 2006;25(31):4300-9.

223. Gao SP, Mark KG, Leslie K, Pao W, Motoi N, Gerald WL, et al. Mutations in the EGFR kinase domain mediate STAT3 activation via IL-6 production in human lung adenocarcinomas. The Journal of clinical investigation. 2007;117(12):3846-56.

224. Qu P, Roberts J, Li Y, Albrecht M, Cummings OW, Eble JN, et al. Stat3 downstream genes serve as biomarkers in human lung carcinomas and chronic obstructive pulmonary disease. Lung cancer. 2009;63(3):341-7.

225. Harpel PC, Haque NS. Chemokine receptor-8: potential role in atherogenesis. Isr Med Assoc J. 2002;4(11):1025-7.

226. Gombert M, Dieu-Nosjean MC, Winterberg F, Bunemann E, Kubitza RC, Da Cunha L, et al. CCL1-CCR8 interactions: an axis mediating the recruitment of T cells and Langerhans-type dendritic cells to sites of atopic skin inflammation. Journal of immunology. 2005;174(8):5082-91.

227. Bernardini G, Spinetti G, Ribatti D, Camarda G, Morbidelli L, Ziche M, et al. I-309 binds to and activates endothelial cell functions and acts as an angiogenic molecule in vivo. Blood. 2000;96(13):4039-45.

228. Ruckes T, Saul D, Van Snick J, Hermine O, Grassmann R. Autocrine antiapoptotic stimulation of cultured adult T-cell leukemia cells by overexpression of the chemokine I-309. Blood. 2001;98(4):1150-9.

229. Louahed J, Struyf S, Demoulin JB, Parmentier M, Van Snick J, Van Damme J, et al. CCR8-dependent activation of the RAS/MAPK pathway mediates anti-apoptotic activity of I-309/ CCL1 and vMIP-I. European journal of immunology. 2003;33(2):494-501.

230. Das S, Sarrou E, Podgrabinska S, Cassella M, Mungamuri SK, Feirt N, et al. Tumor cell entry into the lymph node is controlled by CCL1 chemokine expressed by lymph node lymphatic sinuses. The Journal of experimental medicine. 2013;210(8):1509-28.

231. Eruslanov E, Stoffs T, Kim WJ, Daurkin I, Gilbert SM, Su LM, et al. Expansion of CCR8(+) inflammatory myeloid cells in cancer patients with urothelial and renal carcinomas. Clinical cancer research : an official journal of the American Association for Cancer Research. 2013;19(7):1670-80.

232. Montes-Vizuet R, Vega-Miranda A, Valencia-Maqueda E, Negrete-Garcia MC, Velasquez JR, Teran LM. CC chemokine ligand 1 is released into the airways of atopic asthmatics. The European respiratory journal. 2006;28(1):59-67.

233. N'Diaye M, Le Ferrec E, Lagadic-Gossmann D, Corre S, Gilot D, Lecureur V, et al. Aryl hydrocarbon receptor- and calcium-dependent induction of the chemokine CCL1 by the environmental contaminant benzo[a]pyrene. The Journal of biological chemistry. 2006;281(29):19906-15.

234. Takabatake N, Shibata Y, Abe S, Wada T, Machiya J, Igarashi A, et al. A single nucleotide polymorphism in the CCL1 gene predicts acute exacerbations in chronic obstructive pulmonary disease. American journal of respiratory and critical care medicine. 2006;174(8):875-85.

235. Reimer MK, Brange C, Rosendahl A. CCR8 signaling influences Toll-like receptor 4 responses in human macrophages in inflammatory diseases. Clinical and vaccine immunology : CVI. 2011;18(12):2050-9.

ARTICLES

Clin Transl Oncol (2013) 15:671–682
DOI 10.1007/s12094-013-1034-0

ADVANCES IN TRANSLATIONAL ONCOLOGY

Proteomic biomarkers in lung cancer

M. D. Pastor · A. Nogal · S. Molina-Pinelo · A. Carnero · L. Paz-Ares

Received: 14 February 2013 / Accepted: 25 March 2013 / Published online: 20 April 2013

Abstract The correct understanding of tumour development relies on the comprehensive study of proteins. They are the main orchestrators of vital processes, such as signalling pathways, which drive the carcinogenic process. Proteomic technologies can be applied to cancer research to detect differential protein expression and to assess different responses to treatment. Lung cancer is the number one cause of cancer-related death in the world. Mostly diagnosed at late stages of the disease, lung cancer has one of the lowest 5-year survival rates at 15 %. The use of different proteomic techniques such as two-dimensional gel electrophoresis (2D-PAGE), isotope labelling (ICAT, SILAC, iTRAQ) and mass spectrometry may yield new knowledge on the underlying biology of lung cancer and also allow the development of new early detection tests and the identification of changes in the cancer protein network that are associated with prognosis and drug resistance.

Keywords Lung cancer · Proteomics · Diagnostic biomarkers · Prognostic biomarkers · Predictive biomarkers

M. D. Pastor and A. Nogal contributed equally to the present work.

M. D. Pastor · A. Nogal · S. Molina-Pinelo · A. Carnero · L. Paz-Ares (✉)
Servicio de Oncología Médica, Instituto de Biomedicina de Sevilla (IBIS), University Hospital Virgen del Rocío, Av. Manuel Siurot s/n, 41013 Seville, Spain
e-mail: luis.pazares.sspa@juntadeandalucia.es

A. Nogal
Instituto de Ciências Biomédicas Abel Salazar, University of Porto, Porto, Portugal

Lung cancer overview

Lung cancer was a rare disease in the beginning of the 20th century, but environmental and lifestyle changes have prompted it to the top of the world cancer incidence and mortality rates. European estimates for 2012 indicate that over 262,000 people will die from lung cancer [1]. The major contributor to the development of this disease is tobacco smoking. Approximately 85–90 % of all lung cancer cases are attributed to this habit and smokers have a 15–30-fold increase risk to develop this disease [2]. Tobacco smoking allows over 4,000 compounds, 66 of which are known carcinogens, to enter the airways and cause damage that can lead to the development of lung cancer. Individual variation in the susceptibility to carcinogens and the presence of previous lung diseases also play a part in the pathogenesis of this disease [2]. Other risk factors for lung cancer development include passive smoking, indoor radon, heavy metals, outdoor pollution, asbestos, and fumes from cooking stoves and fires [3].

Current lung cancer classification makes use of histological data. The majority of these tumours are firstly divided into two categories: small cell lung cancer (SCLC) that comprises approximately 15 % of all lung cancer cases and non-small cell lung cancer (NSCLC). This latter group, with 85 % of cases, is sub-divided into adenocarcinoma, squamous cell carcinoma (SCC) and large cell carcinoma (LCC). The SCC type used to be the most common, but in the past years adenocarcinoma has become the most prevalent form of NSCLC [3]. This useful histological classification can every so often be inadequate since it does not take into consideration the heterogeneity of lung tumours. Important molecular and genetic alterations between SCLC and NSCLC or between the NSCLC sub-types could, for example, imply different treatments and outcomes for lung

cancer patients [4]. The genetic instability that characterises lung cancer can occur in the chromosomes, nucleotides, and in the transcriptome. These alterations arise in proto-oncogenes, tumour suppressor genes, DNA repair genes, etc. For example, the activation of the TERT gene, that codifies the catalytic sub-unit of telomerase, is present in 80 % of NSCLC and almost all SCLC cases [5]. Mutations of the genome guardian gene TP53 are detected in 50–70 % of NSCLC and over 75 % of SCLC [6]. Amplifications of the EGFR and PI3KCA genes are more frequent in SCC and mutations of the KRAS and LKB1 genes are more common in adenocarcinomas. In addition, more frequent in adenocarcinomas is the ALK-EML4 gene fusion [6], although a number of new gene mutations in lung cancer are being identified [7]. The discovery and further research of these alterations have allowed the development of specific drugs, such as kinase inhibitors and monoclonal antibodies, which improved survival of lung cancer patients, highlighting the importance of the molecular characterisation of lung tumours to improve detection, treatment and better outcome of these patients. Regardless of the advances that have been made, a grand majority of lung cancers is still diagnosed at advanced disease stages when curative therapeutic options are no longer successful. This fact is demonstrated by one of the lowest 5-year survival rates among cancers with 15 %, making it necessary to uncover new diagnostic biomarkers and therapeutic targets that could assist in the early diagnosis and treatment of lung cancer [3].

In the last several years, genomic and proteomic technologies have begun to reveal the molecular complexity of lung tumours by allowing a rapid and complete analysis of the genes and proteins that are expressed in the context of this disease. The use of genomic technologies such as microarrays and quantitative RT-PCR on lung cancer has generated a great deal of knowledge. It has allowed the separation of adenocarcinoma patients into new molecular sub-classes with different outcomes, the early identification of high risk patients, and the detection of molecular signatures associated with metastasis, among many more discoveries that improve patient management and treatment [4]. While genomics has provided a great deal of information on lung cancer biology, the proteomic approach opens a new window to the better understanding of this disease, given that the direct evaluation of the expressed proteins offers information that is not obtainable through the study of DNA and RNA. The phenotype of a cell is determined by proteins and RNA levels are often poorly correlated with protein expression [8]. Furthermore, common post-translational modifications, such as phosphorylation and glycosylation, go undetected in genomic studies in spite of their importance in the modification of protein function and in the phenotype of a cell or tissue [8].

In conclusion, the new proteomic technologies that allow the analysis of thousands of cancer-related proteins will improve the knowledge of lung cancer biology and pathogenesis, and aid in the development of new early detection biomarkers and the identification of protein profiles associated with prognosis and drug resistance. This review describes the current proteomic approaches to the study of lung cancer and also the so far discovered proteomic biomarkers for this disease.

Proteomics in lung cancer: technologies and samples

The major proteomic approaches typically involve electrophoresis and/or chromatography combined with chemical or metabolic labelling and mass spectrometry (MS) analysis. It is important to recognise that all of them have advantages and disadvantages and that, when choosing a proteomic method, one must consider the most suitable for the sample and hypothesis at study.

Gel-based proteomics

Two-dimensional gel electrophoresis (2D-PAGE) has been the preferred method for proteomic analysis given its relative low cost and high applicability. It allows proteins to be separated by two dimensions. Briefly, in the first dimension (isoelectric focusing) solubilised, denatured proteins are separated by their isoelectric point in a polyacrylamide gel strip with a specific pH range. Then, the focused strip is placed on a sodium dodecyl sulphate polyacrylamide gel (SDS-PAGE) to allow the separation of the focused proteins by molecular weight. Proteins generate spots than can be identified by comparison with gel maps in databases, by western blot, or by MS. This technique can separate complex samples with good resolution and allows the comparison of multiple gels. Nonetheless, limitations include a small sensitivity to very low or very high molecular weight proteins and isoelectric points, samples with high salt contents, the under-representation of membrane proteins, and gel-to-gel variation [9, 10]. A variation of 2D-PAGE is two-dimensional difference gel electrophoresis (2D-DIGE). It allows the marking of proteins from a test, control, and reference sample with fluorescent dyes (Cy3, Cy5, and Cy2) that are mixed together before 2D-PAGE. Although the disadvantages when compared to 2D-PAGE remain the same, advantages include the analysis of different samples in the same gel, the evaluation of quantitative differences in the same gel, and the use of the reference sample that can reduce gel-to-gel variation [10]. At the final stage of these two methods, protein spots are digested with trypsin prior to their MS analysis. The use of two-dimensional alternative separation

methods such as liquid chromatography (LC) has become popular given its ability to separate and allow the prior MS identification of complex samples. Typically samples are digested and the resulting peptides are separated by LC, with the resulting fractions being analysed by MS [11]. Multidimensional protein identification technology (MudPIT) uses high-performance LC to separate peptides online with the ion source of the mass spectrometer, allowing the separation and identification of peptides [12].

Isotope labelling proteomics

Developed to allow protein quantitation in a gel-free setting, isotope labelling techniques involve the metabolic incorporation of the labels in vivo in cell cultures or the global labelling of samples before or after protein digestion.

Stable isotope labelling with amino acids in cell culture (SILAC) is, as the name states, based on the metabolic incorporation of "heavy" and "light" forms of amino acids into the proteins of living cultured cells [13]. Typically, heavy (13C or 15N) arginine or lysine are used in the culture medium of a cell culture, while the other cell culture is supplied with regular amino acids, allowing these amino acids to be incorporated into the newly synthesised proteins. Following trypsin digestion, peptides are analysed by MS and the light and heavy peptides appear in two distinct peaks and, by comparing the signal intensities differences, relative quantitation is achieved. Isotope-coded affinity tags (ICAT) are generally used to compare pairs of samples. Extracted proteins are labelled with the light or heavy ICAT reagent, mixed, trypsin digested, fractionated, and analysed by MS [14]. Isobaric tags for relative and absolute quantitation (iTRAQ) can analyse 4–8 samples (4-plex, 8-plex) in the same experiment. After trypsin digestion, samples are independently labelled with the iTRAQ reagent. During MS analysis, the reporter groups of the iTRAQ reagents separate from the peptides and generate small fragments for each sample with mass-to-charge (m/z) of 114, 115, 116, and 117 for 4-plex, plus 113, 118, 119, and 121 for 8-plex. The intensity of each peak correlates with the quantity of each reporter group and thus with the quantity of the peptide [15].

MS based proteomics

The MS analysis of proteins and peptides provides the accurate measurement of their molecular weight and charge (m/z ratio) that allows their identification. In order to obtain these measurements, samples are ionised by the ion source and travel through a mass analyser and then to a mass detector [16]. Two ionisation methods are the most commonly used: matrix-assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI). In the first one, proteins or peptides are mixed with an energy-absorbing matrix, dried on a metal plate, and then blasted by a laser that ionises the sample [10]. With ESI, the sample is introduced in a liquid form and pumped through a capillary tube at high voltages, resulting in a fine spray [16]. Advantages of MALDI include easy and rapid data analysis, high resolution for large molecules, and sensitivity to small amounts of sample. However, small molecules may not be detected [9, 10, 16]. Ionisation by ESI allows the interface with LC, has a large mass range, however, it is more susceptible to contamination and requires more maintenance [9]. A variation of the MALDI method is the surface-enhanced laser desorption/ionisation (SELDI) method where chromatographic chip arrays selectively bind subsets of proteins from a complex sample. It allows a wider proteome coverage and quantitative information on a large number of proteins. Nonetheless, it is technically demanding and most samples require trypsin digestion prior to their inclusion in the chips [10]. After this step, ionised samples pass through a mass analyser that separate ions according to their m/z ratio and the most commonly used are time-of-flight (TOF), Fourier transform, and ion traps (quadrupole -Q, linear quadrupole -LTQ, Orbitrap) [16]. The addition of a second mass analyser (tandem mass spectrometry or MS/MS) results in the identification of the amino acid sequence of peptides. The obtained data is then compared against protein sequence databases to allow protein identification [9].

Samples for lung cancer proteomic studies

The lung is a heterogeneous organ composed by many different cell types (bronchial, alveolar, inflammatory) and vascular structures. Its main function is to perform gas exchanges between the atmosphere and the bloodstream. When studying lung cancer with proteomic tools, several different samples can be used: lung tissue, blood (serum or plasma), pleural effusions, and also urine [17].

Lung cancer tissue samples, either fresh-frozen or paraffin-embedded, are the ultimate sample for any type of cancer research. Nonetheless, the presence of adjacent non-tumour cells, inflammatory cells and stromal components, among others, can result in non-tumour protein contamination. Using laser micro-dissection or a larger number of samples, this issue might be overcome. In addition, tissue samples from early disease stages are not easily obtained. The accessibility of blood makes it a great sample for lung cancer proteomic studies. Analysis of plasma or serum assumes that the perfusion of tumours or host response contributes to the modification of circulating proteins and peptides, and that tumour-derived proteins may be in circulation. However, due to the abundance of plasma

proteins, depletion of these proteins is necessary to reveal the presence of less abundant ones, the ones likely to be tumour-specific markers [11]. Pleural effusion is the pathological accumulation of fluid that occurs in inflammatory conditions and in the lung cancer setting. In this case, pleural effusions are often drained to search for cancer cell infiltration. It has a similar protein composition to that of plasma, but its proximity to tumour cells makes it useful for lung cancer biomarker detection by proteomic techniques. Bronchoalveolar lavage fluid and induced sputum are also good samples, closely reflecting the content of the lungs, nonetheless, their use has been so far limited to the study of non-malignant conditions [18, 19]. More recently, urine has been receiving attention as a potential sample, given its non-invasive collection and also because urine protein composition can demonstrate changes in the circulatory system. The urinary proteome, usually applied to proteomic studies of the kidney and urinary tract related diseases, can also be used in non-urogenital diseases such as lung cancer [20] In addition, one proteomic study making use of saliva was also conducted [21].

Proteomic biomarkers in lung cancer

This section describes the potential diagnostic, prognostic, and predictive biomarkers revealed by proteomic techniques using biological lung cancer samples and that were subsequently validated by western blot, immunohistochemistry, or enzyme-linked immunosorbent assay (ELISA), or using another group of samples.

Diagnostic

The identification of a lung cancer diagnostic biomarker requires that the protein in question must be specific and directly correlated with the presence of the disease. The majority of studies have performed a comparison between the protein profiles of tumour samples and normal lung samples. We have divided the performed proteomic studies according to the histological type of lung cancer used for the study.

The following proteomic studies did not discriminate between the different lung cancer histological subtypes, giving a broad idea of the main molecular alterations that occur between the patients with lung cancer and healthy subjects. Proteomic studies involving tissue samples from lung cancer patients have identified a great number of differentially expressed proteins when compared to control samples. The macrophage migration inhibitory factor (MIF), cyclophilin A (CyP-A) [22], CD98, fascin, polymeric immunoglobulin receptor/secretory (PIGR), 14-3-3 η [23], thymosin β4 (TMSB4), ubiquitin, acyl-coA binding protein (ACBP), cystatin A (CSTA), and cytochrome C [24] were identified as lung cancer diagnostic tissue biomarkers. The MIF and 14-3-3 η proteins have been known to have an important role in the carcinogenic process: the first one is a pro-inflammatory cytokine and its expression correlates with the levels of several angiogenic chemokines and a higher vessel density in lung cancer tissue, and the 14-3-3 η protein regulates many cellular processes that are important in cancer biology [23, 25]. The use of tissue and serum from lung cancer, other cancers and healthy individuals, allowed the identification of ubiquitin carboxy-terminal hydrolase L1 (UCH-L1) as a possible diagnostic biomarker for lung cancer [26].

Lung cancer proteomic studies that made use of serum and plasma samples have identified different sets of proteins such as leucine-rich alpha-2-glycoprotein (LRG1), inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4), plasma retinol-binding protein (plasmaRBP), complement component C4 and C3, prothrombin [27] haptoglobin (HP) [27–29], carcinoembryonic antigen (CEA), RBP, α-1-antitrypsin (A1AT), transferrin (TF), squamous cell carcinoma antigen 1 (SCCA1) [29], chaperonin, haemoglobin [30], and serum amyloid A (SAA) [30–32] as over-expressed and characteristic of lung cancer samples. The most often identified proteins were acute phase reactant proteins, which are produced in response to inflammation, and can, therefore, have altered levels in the tumour microenvironment. Haptoglobin is an acute phase protein that has angiogenic and antioxidant properties, and has an important role in cell migration [28]. The SAA protein, also an acute phase protein, is secreted into circulation in a number of inflammatory and malignant conditions. It is involved in the transport of cholesterol to the liver, the recruitment of immune cells, and the induction extracellular matrix degrading enzymes [31]. Decreased levels of transthyretin (TTR), identified in serum by SELDI–TOF–MS, were detected in lung cancer samples in comparison to benign lung diseases and healthy subjects [33].

Other sample types, such as pleural effusions and saliva have been used with the purpose of finding diagnostic lung cancer biomarkers. A 1D-PAGE–LC–ESI–MS/MS methodology in pleural effusions from lung cancer patients and non-malignant conditions, identified higher alpha-2-HSglycoprotein (AHSG) and insulin-like growth factor-binding protein 2 (IGFBP2) levels in malignant pleural effusions compared to non-malignant ones [34]. A recent study using 2D-DIGE, MALDI–TOF-MS, and LC–MS/MS identified HP, calprotectin, and zinc-a-2-glycoprotein (AZGP1) as capable of discriminating lung cancer from healthy controls with high specificity and sensitivity [21].

Several studies on SCLC samples have revealed diagnostic biomarkers for this type of lung cancer that could be helpful in the detection of this aggressive malignancy.

Using distinct proteomic techniques on tumour tissue samples, several proteins were identified: β-tubulin, heat shock protein 73 (HSP73), HSP90, lamin B, proliferating cell nuclear antigen (PCNA) [35], coactosin-like protein-1 (COTL-1), gamma actin (ACTG1), alpha tubulin, laminin B (LAMB1), UCH-L1, ubiquitin-conjugating enzyme E2 (UBE2), carbonic anhydrase 1 (CA1) [36], and stathmin (STMN1) [37] were found to be over-expressed in SCLC tissue samples, while calcyclin (S100A6) was under-expressed when compared to adjacent normal tissue [38]. Lastly, HP was identified as a marker for SCLC in serum samples [39]. Some of the identified proteins, such as HSPs and CA, have an important role in cancer development. Heat shock proteins are known for stabilising many oncogenic and growth promoting proteins in cancer cells [40]. Carbonic anhydrases have an important role in maintaining the pH homeostasis. Extracellular acidification has been associated with many tumour progression effects and CA inhibitors may suppress the invasion capacity of malignant cell lines [41].

Similarly, the use of NSCLC samples has also revealed interesting diagnostic biomarkers. These could be used to improve detection of the most common form of lung cancer. Studies that do not differentiate the histological subtypes of NSCLC have identified alpha enolase (ENO1) [42], peroxiredoxin 1 (PRDXI), PRDX3, thioredoxin (TXN) [43], coatomer protein complex subunit gamma (COPG), PRDX4, thymopoietin (TMPO) [44], MIF [45], TMSB4, TMSB10, ribosomal protein L39 and S30, S100A6, and histone H2A.2 [46] as over-expressed in NSCLC tissue samples. In the same samples, the proteins transgelin (TAGLN), HSP20-like protein, CA, selenium binding protein 1 (SELENBP1), alloalbumin venezia [42], and polymerase I and transcript release factor (PTRF or cavin-1) [45] were found to be under-expressed in NSCLC. Peroxiredoxins are a group of antioxidant enzymes that also control cytokine-induced peroxide levels. Their up-regulation in lung cancer may compensate the increased ROS formation owing to elevated mitochondrial respiration and the higher energy requirements of proliferating lung cancers [42]. The use of serum samples in NSCLC proteomic studies revealed the over-expression of SAA and HP, compared to healthy individuals [47, 48]. A 2D-DIGE–MALDI–TOF–MS study on pleural effusion and serum from NSCLC and benign lung diseases patients revealed that pigment epithelium-derived factor (PEDF), gelsolin, and metalloproteinase inhibitor 2 (TIMP2) were over-expressed, and that S100A8 and S100A9 were under-expressed in cancer samples [49]. The analysis of urine and also tissue from lung cancer patients and healthy subjects by 1D-PAGE-nano HPLC–MS/MS found that LRG1 was expressed at higher levels in urinary exosomes and lung tissue of NSCLC patients [50]. The proteomic analysis of the different histological subtypes of NSCLC has focused mainly on adenocarcinoma and SCC.

The study of adenocarcinoma compared to healthy controls or other lung diseases might reveal important markers for this heterogeneous subtype and currently the most commonly diagnosed NSCLC. The proteomic evaluation of adenocarcinoma tissue revealed several over-expressed proteins, such as PRDX4, ATP synthase subunit delta (ATP5D), beta 1,4-galactosyltransferase, cytosolic inorganic pyrophosphatase, glutathione S-transferase Mu 4 (GSTM4), glucose-related Mr 58000 protein, propyl-4-hydrolase beta subunit (P4HB), triose-phosphate isomerase (TIM), UCH-L1 [51], CyP-A, TAGLN, TAGLN2 [52], 14-3-3 σ (stratifin—SNF), which has been directly implicated in the aetiology of human cancer, annexin 1 (ANXA1), an anti-inflammatory response protein, manganese superoxide dismutase (SOD2) [53], cofilin-1 (CFL1), and pyruvate kinase M2 (PKM2) [54]. The over-expression of immunoglobulin lambda chain, transthyretin (TTR) monomer, HP, and two isoforms of the SAA protein, and the down regulation of a fragment of apolipoprotein A-I (APO-AI) were detected in serum samples from lung cancer adenocarcinoma subjects [55]. Pernemalm and colleagues used iTRAQ-LC–MS/MS to analyse pleural effusion and plasma samples from adenocarcinoma patients and inflammatory pleuritis patients. This strategy revealed the under-expression of α-2-macroglobulin (A2M), tetranectin (CLEC3B), cystatin-C precursor, EGF-containing fibulin-like extracellular matrix protein 1 (EFEMP1), gelsolin, and vascular cell adhesion protein 1 (VCAM1), and the over-expression of keratin type II cytoskeletal 8 (CK8), A1AT, and epididymal secretory protein E1 (NPC2) in lung cancer samples [56]. Urine samples from adenocarcinoma and normal individuals were compared by 1D-PAGE–HPLC–MS/MS revealing clusterin (CLU), kallikrein 1 (KLK1), gelsolin, LRG1, and α-1-antichymotrypsin (AACT) as proteins of importance to adenocarcinoma diagnosis [20].

The proteomic evaluation of SCC patients could provide useful biomarkers for the second most commonly diagnosed NSCLC subtype. Considering studies assessing tissue samples from SCC patients, several proteins were found to be over-expressed. They were ENO1, pre-B cell-enhancing factor precursor (PBEF), TIM, phosphoglycerate mutase 1 (PGAM1), fructose-biphosphate aldolase A (ALDOA), guanine nucleotide-binding protein subunit beta-like protein (GBN1L) [57] and HSPB1, that prevents stress-induced cellular damage and is elevated in multiple cancer types, S100A8 and S100A9 [58] and HSP60, that plays an important role in regulating cell differentiation and proliferation, ANXA5 and ANXA6, ACTG1 [59] and isocitrate dehydrogenase 1 (IDH1), which play a role in the antioxidant system by producing NADPH, SOD2, SNF,

PRDX2, and receptor of activated protein kinase 1 [60]. Five proteins were found to be under-expressed in the same samples: creatine kinase B (CKB), SCCA1 as down-regulated [58], cathepsin D preprotein (CTSD), ferritin heavy chain (FTH1), and ANXA3 (under-expressed) [59]. The identified annexins in these proteomic studies have a potential role in cellular signal transduction, exocytosis, inflammation, coagulation, cellular growth and differentiation. Using tissue and serum from SCC and healthy individuals, Yang and co-workers detected an increase in autoantibodies against TIM and SOD2 that could aid in the diagnosis of this type of lung cancer [61]. Finally, by comparing serum from SCC patients and controls, five over-expressed (HP, APO-A4, complement component C3c, SAA, and Ras-related protein 7B—RAB7B) and five under-expressed (AHSG, hemopexin precursor—HPX, proapolipoprotein, antithrombin III—SERPINC1, and CLU) proteins were identified and were able to distinguish SCC patients [62].

A summary of the most recurrently identified diagnostic biomarkers on lung cancer tissue samples and in fluid samples can be found on Tables 1 and 2, respectively.

Prognostic

When the expression levels of a protein correlate with the natural history of the disease they are considered to have prognostic value. The study of prognostic biomarkers in lung cancer has been made by correlating the expression of a molecule to the patient survival. An alternative approach is to compare groups of patients with different clinical stages of disease, based on the assumption that a more advanced tumour is more aggressive and may express proteins that drive the metastatic process. The published proteomic studies that have focused on finding prognostic biomarkers for lung cancer have only made use of samples from NSCLC patients as a whole or in separate as adenocarcinoma or SCC patients.

The grand majority of studies using samples from NSCLC patients were performed using tissue. The proteins small ubiquitin-related modifier-2 protein (SUMO-2), TMSB4, and ubiquitin were identified in a 15 MS peak profile that distinguished between patients with resected NSCLC who had poor or good prognosis [63]. Using a similar approach, 25 proteins (not all identified, but including TMSB4, TMSB10, ribosomal protein L39 and S30, S100A6, and histone H2A.2) were associated with survival among NSCLC patients and appeared to distinguish those with poor prognosis from those with good prognosis [46]. In a different study, the levels of TMSB4, TMSB10, and calmodulin were also associated with patient survival. In addition, low levels of CFL1 were associated with better outcome for patients with negative lymph nodes and high levels of CFL1 with better outcome for those with positive lymph nodes [64]. TMSB4 is a regulator of actin polymerisation whose over-expression seems to stimulate lung tumour metastasis [64]. Improved survival of NSCLC was additionally associated with the levels of S100A6 [65]. The S100 proteins are involved in the regulation of a number of cellular processes such as cell cycle progression, differentiation, and also inflammation.

Proteomic studies of adenocarcinoma tissue samples identified CFL1, an actin-modulating protein that is increased in some invasive cancers, and PKM2, a key glycolytic enzyme during tumorigenesis, as associated with poor prognosis in adenocarcinoma patients [54]. The comparison between adenocarcinoma and adjacent lung tissue revealed that specific isoforms of CK7, 8, 18, and 19 were associated with patient survival [66]. Cytoskeletal reorganization is a central process regulating cell migration and metastasis, and CKs, a family of cytoskeletal intermediate filaments, have been suggested to play a role in carcinogenesis, by promoting cellular architecture reorganisation during tumour development and progression. Another adenocarcinoma study associated the expression of phosphoglycerate kinase 1 (PGK1), HSP70, CK19, PGAM1, and G protein-coupled receptor 4 (GRK4) with poor survival [67]. Several annexins (ANXA1, ANXA2, and ANXA3) were associated with advanced clinical stage, by presenting higher expression levels in lymph node metastatic tissue [68, 69]. The deregulation of annexins has been reported in numerous cancers and is thought to influence the patterns of cellular behaviour, such as cell proliferation, motility, invasiveness and signalling pathways. Finally, Maeda and colleagues revealed that the absence of expression of LC–MS/MS identified myosin IIA and vimentin, two proteins associated with the cytoskeleton and cell motility, correlated with good prognosis in adenocarcinoma patients who did not receive post-operative adjuvant chemotherapy, indicating that these patients do not require such treatment [70].

Considering SCC proteomic studies searching for prognostic biomarkers, the previously identified IDH1 protein was associated with poor overall survival [60]. The evaluation of micro-dissected primary SCC and matched lymph node metastatic tissues revealed the under-expression of SNF in lymph node metastatic tumour versus primary SCC [71]. Yao and co-workers compared tissue from SCC with and without lymph node metastasis. The ANXA2, HSP27, and CK17 proteins were over-expressed in metastatic SCC, and SNF was under-expressed in these tissues [72]. A summary of these potential prognostic biomarkers can be found on Table 3.

Table 1 Candidate proteomic diagnostic biomarkers for lung cancer identified in tissue samples

Diagnostic biomarker	Lung cancer type	Proteomic technique
S100		
A6	SCC	MALDI–TOF–MS [38]
	NSCLC	MALDI–MS [46]
A8	SCC	iTRAQ, LC-Q-MS/MS [58]
A9	SCC	iTRAQ, LC-Q-MS/MS [58]
SCCA1 (squamous cell carcinoma antigen 1)	SCC	iTRAQ, LC-Q-MS/MS [58]
Cystatin A	Lung cancer	MALDI–TOF–MS [24]
Tubulin (α, β)	SCLC	2D-PAGE; 2D-PAGE, MALDI–TOF–MS [35, 36]
UCH-L1 (ubiquitin carboxy-terminal hydrolase L1)	Lung cancer	2D-PAGE, MALDI–TOF–MS [26]
	SCLC	2D-PAGE, MALDI–TOF–MS [36]
	Adenocarcinoma	2D-PAGE, MALDI–MS [51]
Peroxiredoxin		
PRDX4	NSCLC, Adenocarcinoma	1D-LC–ESI–MS/MS [44]; 2D-PAGE, MALDI–MS [51]
PRDX1	NSCLC	2D-PAGE [43]
PRDX2	SCC	2D-DIGE, MALDI–TOF–TOF [60]
PRDX3	NSCLC	2D-PAGE [43]
PRDX6	SCC	2D-PAGE, MALDI–TOF–MS [61]
TIM (triose-phosphate isomerase)	Adenocarcinoma	2D-PAGE, MALDI–MS [51]
	SCC	2D-PAGE, MALDI–TOF–MS [57, 61]
MIF (macrophage migration inhibitory factor)	Lung cancer	MALDI–TOF–MS [22]
	NSCLC	MALDI–TOF–MS; IMAC, LC–MS/MS [24, 45]
CyP-A (Cyclophilin A)	Lung cancer	MALDI–TOF–MS [22]
	Adenocarcinoma	2D-DIGE, LC–MS/MS [52]
TAGLN (Transgelin)	NSCLC	2D-PAGE, MALDI–TOF–MS [42]
	Adenocarcinoma	2D-DIGE, LC–MS/MS [52]
CA (Carbonic anhydrase)	SCLC	2D-PAGE, MALDI–TOF–MS [36]
	NSCLC	2D-PAGE, MALDI–TOF–MS [42]
ENO1 (Alpha enolase)	NSCLC	2D-PAGE, MALDI–TOF–MS [42]
	SCC	2D-PAGE, MALDI–TOF–MS [57, 61]
14-3-3		
η	Lung cancer	1D-PAGE, nanoESI–MS/MS [23]
σ (SNF)	Adenocarcinoma	2D-PAGE, MALDI–TOF–MS, Q-TOF–MS/MS [53]
	SCC	2D-DIGE, MALDI–TOF–TOF [60]
SOD2 (Superoxide dismutase 2)	Adenocarcinoma	2D-PAGE, MALDI–TOF–MS, Q-TOF–MS/MS [53]
	SCC	2D-DIGE, MALDI–TOF–TOF; 2D-PAGE, MALDI–TOF–MS [60, 61]
Heat shock protein		
HSP73	SCLC	2D-PAGE [35]
HSP90	SCLC	2D-PAGE [35]
HSP20-like	NSCLC	2D-PAGE, MALDI–TOF–MS [42]
HSP70	SCC	2D-PAGE, MALDI–TOF–MS [61]
HSP60	SCC	2D-PAGE, MALDI–TOF–MS [59, 61]
HSP27	SCC	iTRAQ, LC-Q-MS/MS [58]
Annexin		
ANXA1	Adenocarcinoma	2D-PAGE, MALDI–TOF–MS, Q-TOF–MS/MS [53]
ANXA2	SCC	2D-PAGE, MALDI–TOF–MS [61]
ANXA3	SCC	2D-PAGE, MALDI–TOF–MS [59]
ANXA5	SCC	2D-PAGE, MALDI–TOF–MS [59]

Table 1 continued

Diagnostic biomarker	Lung cancer type	Proteomic technique
ANXA6	SCC	2D-PAGE, MALDI–TOF–MS [59]
ACTG1 (Gamma actin)	SCLC	2D-PAGE, MALDI–TOF–MS [36]
	SCC	2D-PAGE, MALDI–TOF–MS [59]
Thymosin		
β4	Lung cancer	MALDI–TOF–MS [24]
	NSCLC	MALDI–MS [46]
β10	NSCLC	MALDI–MS [46]
CFL1 (Cofilin-1)	Adenocarcinoma	2D-PAGE, ESI-Q-TOF–MS/MS [54]

NSCLC non-small cell lung cancer, *SCLC* small cell lung cancer, *SCC* squamous cell carcinoma

Table 2 Candidate proteomic diagnostic biomarkers for lung cancer identified in fluid samples

Diagnostic biomarker	Sample type	Lung cancer type	Proteomic technique
CLU (Clusterin)	Urine	Adenocarcinoma	1D-PAGE, HPLC–MS/MS [20]
	Serum	SCC	2D-DIGE, MALDI–TOF–MS [62]
Gelsolin	Urine	Adenocarcinoma	1D-PAGE, HPLC–MS/MS [20]
	Serum	NSCLC	2D-DIGE, MALDI–TOF–MS [49]
	Pleural effusion, plasma	Adenocarcinoma	iTRAQ, LC–MS/MS [56]
LRG1 (leucine-rich alpha-2-glycoprotein)	Urine	Adenocarcinoma	1D-PAGE, HPLC–MS/MS [20]
	Serum	Lung cancer	2D-DIGE, LC–MS/MS [27]
	Urine	NSCLC	1D-PAGE, nanoHPLC-MS/MS [50]
SAA (Serum amyloid A)	Serum	Lung cancer	SELDI–TOF–MS; MALDI–MS [30, 32]
		NSCLC	MALDI–TOF–MS [48]
		Adenocarcinoma	2D-PAGE, MALDI–TOF–PMF [55]
		SCC	2D-DIGE, MALDI–TOF–MS [62]
	Pleural effusion, serum	Lung cancer	1D-PAGE, LC–ESI–MS/MS [31]
HP (haptoglobin)	Serum	Lung cancer	2D-DIGE, LC–MS/MS; 2D-DIGE, MALDI–TOF–MS [27, 29]
		SCLC	SDS–PAGE, MALDI–TOF–MS [39]
		NSCLC	2D-DIGE, MALDI–TOF/TOF [47]
		Adenocarcinoma	2D-PAGE, MALDI–TOF–PMF [55]
		SCC	2D-DIGE, MALDI–TOF–MS [62]
	Plasma, serum	Lung cancer	1D-PAGE, LC–ESI–MS/MS [28]
	Saliva	Lung cancer	2D-DIGE, MALDI–TOF–MS, LC–MS/MS [21]
TTR (Transthyretin)	Serum	Lung cancer	SELDI–TOF–MS [33]
		Adenocarcinoma	2D-PAGE, MALDI–TOF–PMF [55]
Complement component			
C4	Serum	Lung cancer	2D-DIGE, LC–MS/MS [27]
C3, C3c		Lung cancer, SCC	2D-DIGE, LC–MS/MS [27]; 2D-DIGE, MALDI–TOF–MS [62]
A1AT (α-1-antitrypsin)	Serum	Lung cancer	2D-DIGE, MALDI–TOF–MS [29]
	Pleural effusion, plasma	Adenocarcinoma	iTRAQ, LC–MS/MS [56]

Table 2 continued

Diagnostic biomarker	Sample type	Lung cancer type	Proteomic technique
Apolipoprotein			
A1	Serum	Adenocarcinoma	2D-PAGE, MALDI–TOF–PMF [55]
A4		SCC	2D-DIGE, MALDI–TOF–MS [62]
CK8 (Cytokeratin 8)	Pleural effusion, plasma	Adenocarcinoma	iTRAQ, LC–MS/MS [56]
AHSG (alpha-2-HSglycoprotein)	Pleural effusion	Lung cancer	1D-PAGE, LC–ESI–MS/MS [34]
	Serum	SCC	2D-DIGE, MALDI–TOF–MS [62]
S100A8 and S100A9	Pleural effusion, serum	NSCLC	2D-DIGE, MALDI–TOF–MS [49]
SCCA1 (Squamous cell carcinoma antigen 1)	Serum	Lung cancer	2D-DIGE, MALDI–TOF–MS [29]
Cystatin C3	Pleural effusion	Lung cancer	1D-PAGE, LC–ESI–MS/MS [34]
	Pleural effusion, plasma	Adenocarcinoma	iTRAQ, LC–MS/MS [56]

NSCLC non-small cell lung cancer, *SCLC* small cell lung cancer, *SCC* squamous cell carcinoma

Table 3 Candidate proteomic prognostic biomarkers for lung cancer

Prognostic biomarker	Sample type	Lung cancer type	Proteomic technique
S100A6	Tissue	NSCLC	MALDI–MS [46]
	Tissue, plasma, pleural effusion	NSCLC	SELDI–TOF–MS [65]
Thymosin			
β4	Tissue	NSCLC	MALDI–MS; MALDI–TOF–MS [63, 64]
β10		NSCLC	MALDI–MS; MALDI–TOF–MS [46, 64]
CFL1 (Cofilin-1)	Tissue	NSCLC	MALDI–TOF–MS [64]
		Adenocarcinoma	2D-PAGE, ESI-Q-TOF–MS/MS [54]
IDH1 (isocitrate dehydrogenase 1)	Tissue	SCC	2D-DIGE, MALDI–TOF–TOF [60]
Cytokeratin			
CK7, CK8, CK18	Tissue	Adenocarcinoma	2D-PAGE, MALDI–TOF–MS [66]
CK19		Adenocarcinoma	2D-PAGE, MALDI–TOF–MS; 2D-PAGE [66, 67]
		SCC	2D-DIGE, MALDI–TOF–PMF [72]
Heat shock protein			
HSP70	Tissue	Adenocarcinoma	2D-PAGE, MALDI–MS [51]
HSP27		SCC	2D-DIGE, MALDI–TOF–PMF [72]
Annexin			
ANXA1	Tissue	Adenocarcinoma	2D-DIGE, MALDI–TOF–PMF [69]
ANXA2		Adenocarcinoma	2D-DIGE, MALDI–TOF–PMF [69]
		SCC	2D-DIGE, MALDI–TOF–PMF [72]
ANXA3		Adenocarcinoma	2D-DIGE, MALDI–TOF–MS; 2D-DIGE, MALDI–TOF–PMF [68, 69]
Myosin IIA	Tissue	Adenocarcinoma	LC–MS/MS [70]
Vimentin			
SNF (14-3-3 σ)	Tissue	SCC	2D-PAGE, MALDI–TOF–MS; 2D-DIGE, MALDI–TOF–PMF [71, 72]

NSCLC non-small cell lung cancer, *SCC* squamous cell carcinoma

Table 4 Candidate proteomic predictive biomarkers for lung cancer

Predictive biomarker	Sample type	Lung cancer type	Proteomic technique
H-FABP (fatty acid-binding protein heart)	Tissue	Adenocarcinoma	2D-DIGE, LC–MS/MS [73]
8-peak signature (VeriStrat®)	Serum	NSCLC	MALDI–MS [74]

NSCLC non-small cell lung cancer

Predictive

A good predictive biomarker can anticipate the efficacy of a specific treatment. It aims at individualising therapies in lung cancer and studies have been based on studying clinical samples from responding and non-responding patients and then validating results on selected cohorts. Proteomic studies developed with this purpose have focused on the response to EGFR inhibitors.

Okano and co-workers analysed lung adenocarcinoma tissue from patients who showed a different response to gefitinib treatment by 2D-DIGE–LC–MS/MS. High levels of fatty acid-binding protein heart (H-FABP) were associated with partial and complete responses. This protein participates in intracellular lipid transport, storage, and metabolism, in addition to having a possible role in cancer biology [73]. In addition to this study, a serum MALDI–MS study conducted by Taguchi and co-workers in NSCLC patients treated with gefitinib and erlotinib revealed an 8-peak profile predictive of outcome [74]. This 8-peak signature was commercially launched as a commercial product (Veristrat®, Biodesix, Broom field, CO, USA) and has been used to discriminate NSCLC patients treated with combinations of erlotinib, gefitinib, bevacizumab, or cetuximab with good and poor prognosis in a number of studies [75–82]. The SAA1 protein was recently identified as part of this proteomic signature [82]. A summary of the abovementioned predictive biomarkers can be found on Table 4.

Conclusions

The detection of new biomarkers is of foremost importance in lung cancer research. The vast majority of lung cancers are still diagnosed at advanced stages of the disease, reducing in some cases the available therapeutic options to only palliative care, resulting in one of the lowest 5-year survival rates among cancers.

The proteomic field is expanding and the development of high-throughput technologies has allowed the identification of several protein profiles with diagnostic, prognostic, and predictive value for lung cancer. Nonetheless, it is necessary to take into account several key points order to consider this proteins as biomarkers. First, an extensive clinical validation with strict statistical criteria is required to evaluate these profiles, to reduce the occurrence of false positive results. Second, large cohorts of carefully selected patients to determine the actual usefulness of the biomarkers are necessary. Third, it would be interesting to study the reproducibility of the obtained protein profiles and this requires a validation in different biological samples and different laboratories to prove their sensitivity. Finally, it is necessary to address the problems of heterogeneity to the lung cancer patients. Studies considering the TNM classification, tumour grade, and demographic characteristics of the population would be required.

In summary, proteomics technologies are one of the most powerful tools to expand the repertoire of known biomarkers for lung cancer early diagnosis, prognosis, and prediction of response to therapy. Moreover, many of the identified candidate biomarkers could serve as targets for more rationale and effective therapeutic strategies due to their role in tumorigenesis and cancer progression.

Conflict of interest The authors declare to have no conflict of interest.

References

1. Malvezzi M, Bertuccio P, Levi F, La Vecchia C, Negri E. European cancer mortality predictions for the year 2012. Ann Oncol. 2012;23(4):1044–52.
2. Alberg AJ, Samet JM. Epidemiology of lung cancer. Chest. 2003;123(1 Suppl):21S-49S.
3. Alberg AJ, Ford JG, Samet JM. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edn). Chest. 2007;132(3 Suppl):29S–55S.
4. Agullo-Ortuno MT, Lopez-Rios F, Paz-Ares L. Lung cancer genomic signatures. J Thorac Oncol. 2010;5(10):1673–91.
5. Meyerson M, Franklin WA, Kelley MJ. Molecular classification and molecular genetics of human lung cancers. Semin Oncol. 2004;1(Suppl 1):4–19.
6. Herbst RS, Heymach JV, Lippman SM. Lung cancer. N Engl J Med. 2008;359(13):1367–80.
7. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell. 2012;150(6):1121–34.
8. Ocak S, Chaurand P, Massion PP. Mass spectrometry-based proteomic profiling of lung cancer. Proc Am Thorac Soc. 2009;6(2):159–70.
9. Hirsch J, Hansen KC, Burlingame AL, Matthay MA. Proteomics: current techniques and potential applications to lung disease. Am J Physiol Lung Cell Mol Physiol. 2004;287(1):L1–23.
10. Conrad DH, Goyette J, Thomas PS. Proteomics as a method for early detection of cancer: a review of proteomics, exhaled breath condensate, and lung cancer screening. J Gen Intern Med. 2008;23(Suppl 1):78–84.
11. Massion PP, Caprioli RM. Proteomic strategies for the characterization and the early detection of lung cancer. J Thorac Oncol. 2006;1(9):1027–39.
12. Guerrera IC, Keep NH, Godovac-Zimmermann J. Proteomics study reveals cross-talk between Rho guanidine nucleotide dissociation inhibitor 1 post-translational modifications in epidermal growth factor stimulated fibroblasts. J Proteome Res. 2007;6(7):2623–30.
13. Mann M. Functional and quantitative proteomics using SILAC. Nat Rev Mol Cell Biol. 2006;7(12):952–8.

14. Graves PR, Haystead TA. Molecular biologist's guide to proteomics. Microbiol Mol Biol Rev. 2002;66(1):39–63; table of contents.
15. DeSouza L, Diehl G, Rodrigues MJ, Guo J, Romaschin AD, Colgan TJ, et al. Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. J Proteome Res. 2005;4(2):377–86.
16. Bowler RP, Ellison MC, Reisdorph N. Proteomics in pulmonary medicine. Chest. 2006 Aug;130(2):567–74.
17. De Petris L, Pernemalm M, Elmberger G, Bergman P, Orre L, Lewensohn R, et al. A novel method for sample preparation of fresh lung cancer tissue for proteomics analysis by tumor cell enrichment and removal of blood contaminants. Proteome Science. 2010;8:9.
18. Magi B, Bargagli E, Bini L, Rottoli P. Proteome analysis of bronchoalveolar lavage in lung diseases. Proteomics. 2006;6(23):6354–69.
19. Gray RD, MacGregor G, Noble D, Imrie M, Dewar M, Boyd AC, et al. Sputum proteomics in inflammatory and suppurative respiratory diseases. Am J Respir Crit Care Med. 2008;178(5):444–52.
20. Zhang Y, Li Y, Qiu F, Qiu Z. Comparative analysis of the human urinary proteome by 1D SDS-PAGE and chip-HPLC-MS/MS identification of the AACT putative urinary biomarker. J Chromatogr B Analyt Technol Biomed Life Sci. 2010;878(32):3395–401.
21. Xiao H, Zhang L, Zhou H, Lee JM, Garon EB, Wong DT. Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. Mol Cell Proteomics. 2012;11(2):M111 012112.
22. Campa MJ, Wang MZ, Howard B, Fitzgerald MC, Patz EF, Jr. Protein expression profiling identifies macrophage migration inhibitory factor and cyclophilin a as potential molecular targets in non-small cell lung cancer. Cancer Res. 2003;63(7):1652–6.
23. Xiao T, Ying W, Li L, Hu Z, Ma Y, Jiao L, et al. An approach to studying lung cancer-related proteins in human blood. Mol Cell Proteomics. 2005;4(10): 1480–6.
24. Rahman SM, Gonzalez AL, Li M, Seeley EH, Zimmerman LJ, Zhang XJ, et al. Lung cancer diagnosis from proteomic analysis of preinvasive lesions. Cancer Res. 2011;71(8):3009–17.
25. White ES, Flaherty KR, Carskadon S, Brant A, Iannettoni MD, Yee J, et al. Macrophage migration inhibitory factor and CXC chemokine expression in non-small cell lung cancer: role in angiogenesis and prognosis. Clin Cancer Res. 2003;9(2):853–60.
26. Brichory F, Beer D, Le Naour F, Giordano T, Hanash S. Proteomics-based identification of protein gene product 9.5 as a tumor antigen that induces a humoral immune response in lung cancer. Cancer Res. 2001;61(21):7908–12.
27. Okano T, Kondo T, Kakisaka T, Fujii K, Yamada M, Kato H, et al. Plasma proteomics of lung cancer by a linkage of multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis. Proteomics. 2006;6(13):3938–48.
28. Kang SM, Sung HJ, Ahn JM, Park JY, Lee SY, Park CS, et al. The Haptoglobin beta chain as a supportive biomarker for human lung cancers. Mol Biosyst. 2011;7(4):1167–75.
29. Patz EF, Jr, Campa MJ, Gottlin EB, Kusmartseva I, Guan XR, Herndon JE, 2nd. Panel of serum biomarkers for the diagnosis of lung cancer. J Clin Oncol. 2007;25(35):5578–83.
30. Song QB, Hu WG, Wang P, Yao Y, Zeng HZ. Identification of serum biomarkers for lung cancer using magnetic bead-based SELDI–TOF–MS. Acta Pharmacol Sin. 2011;32(12):1537–42.
31. Sung HJ, Ahn JM, Yoon YH, Rhim TY, Park CS, Park JY, et al. Identification and validation of SAA as a potential lung cancer biomarker and its involvement in metastatic pathogenesis of lung cancer. J Proteome Res. 2011;10(3):1383–95.
32. Yildiz PB, Shyr Y, Rahman JS, Wardwell NR, Zimmerman LJ, Shakhtour B, et al. Diagnostic accuracy of MALDI mass spectrometric analysis of unfractionated serum in lung cancer. J Thorac Oncol. 2007;2(10):893–901.
33. Liu L, Liu J, Dai S, Wang X, Wu S, Wang J, et al. Reduced transthyretin expression in sera of lung cancer. Cancer Sci. 2007;98(10):1617–24.
34. Yu CJ, Wang CL, Wang CI, Chen CD, Dan YM, Wu CC, et al. Comprehensive proteome analysis of malignant pleural effusion for lung cancer biomarker discovery by using multidimensional protein identification technology. J Proteome Res. 2011;10(10):4671–82.
35. Okuzawa K, Franzen B, Lindholm J, Linder S, Hirano T, Bergman T, et al. Characterization of gene expression in clinical lung cancer materials by two-dimensional polyacrylamide gel electrophoresis. Electrophoresis. 1994;15(3–4): 382–90.
36. Jeong HC, Kim GI, Cho SH, Lee KH, Ko JJ, Yang JH, et al. Proteomic analysis of human small cell lung cancer tissues: up-regulation of coactosin-like protein-1. J Proteome Res. 2011;10(1):269–76.
37. Tanca A, Addis MF, Pagnozzi D, Cossu-Rocca P, Tonelli R, Falchi G, et al. Proteomic analysis of formalin-fixed, paraffin-embedded lung neuroendocrine tumor samples from hospital archives. J Proteomics. 2011;74(3):359–70.
38. Lee HS, Park JW, Chertov O, Colantonio S, Simpson JT, Fivash MJ, et al. Matrix-assisted laser desorption/ionization mass spectrometry reveals decreased calcylcin expression in small cell lung cancer. Pathol Int. 2012; 62(1):28–35.
39. Bharti A, Ma PC, Maulik G, Singh R, Khan E, Skarin AT, et al. Haptoglobin alpha-subunit and hepatocyte growth factor can potentially serve as serum tumor biomarkers in small cell lung cancer. Anticancer Res. 2004;24(2C):1031–8.
40. Trepel J, Mollapour M, Giaccone G, Neckers L. Targeting the dynamic HSP90 complex in cancer. Nat Rev Cancer. 2010;10(8):537–49.
41. Pastorekova S, Parkkila S, Pastorek J, Supuran CT. Carbonic anhydrases: current state of the art, therapeutic applications and future prospects. J Enzyme Inhib Med Chem. 2004;19(3):199–229.
42. Li LS, Kim H, Rhee H, Kim SH, Shin DH, Chung KY, et al. Proteomic analysis distinguishes basaloid carcinoma as a distinct subtype of nonsmall cell lung carcinoma. Proteomics. 2004;4(11):3394–400.
43. Park JH, Kim YS, Lee HL, Shim JY, Lee KS, Oh YJ, et al. Expression of peroxiredoxin and thioredoxin in human lung cancer and paired normal lung. Respirology. 2006;11(3):269–75.
44. Park HJ, Kim BG, Lee SJ, Heo SH, Kim JY, Kwon TH, et al. Proteomic profiling of endothelial cells in human lung cancer. J Proteome Res. 2008;7(3):1138–50.
45. Gamez-Pozo A, Sanchez-Navarro I, Calvo E, Agullo-Ortuno MT, Lopez-Vacas R, Diaz E, et al. PTRF/cavin-1 and MIF proteins are identified as non-small cell lung cancer biomarkers by label-free proteomics. PLoS One. 2012;7(3):e33752.
46. Yanagisawa K, Tomida S, Shimada Y, Yatabe Y, Mitsudomi T, Takahashi T. A 25-signal proteomic signature and outcome for patients with resected non-small-cell lung cancer. J Natl Cancer Inst. 2007;99(11):858–67.
47. Hoagland LFt, Campa MJ, Gottlin EB, Herndon JE, 2nd, Patz EF, Jr. Haptoglobin and posttranslational glycan-modified derivatives as serum biomarkers for the diagnosis of nonsmall cell lung cancer. Cancer. 2007;110(10):2260–8.
48. Howard BA, Wang MZ, Campa MJ, Corro C, Fitzgerald MC, Patz EF, Jr. Identification and validation of a potential lung cancer serum biomarker detected by matrix-assisted laser desorption/ionization-time of flight spectra analysis. Proteomics. 2003;3(9):1720–4.
49. Rodriguez-Pineiro AM, Blanco-Prieto S, Sanchez-Otero N, Rodriguez-Berrocal FJ, de la Cadena MP. On the identification of biomarkers for non-small cell lung cancer in serum and pleural effusion. J Proteomics. 2010;73(8):1511–22.
50. Li Y, Zhang Y, Qiu F, Qiu Z. Proteomic identification of exosomal LRG1: a potential urinary biomarker for detecting NSCLC. Electrophoresis. 2011;32(15):1976–83.
51. Chen G, Gharib TG, Huang CC, Thomas DG, Shedden KA, Taylor JM, et al. Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. Clin Cancer Res. 2002;8(7):2298–305.
52. Rho JH, Roehrl MH, Wang JY. Tissue proteomics reveals differential and compartment-specific expression of the homologs transgelin and transgelin-2 in lung adenocarcinoma and its stroma. J Proteome Res. 2009;8(12):5610–8.
53. Xiao G, Lu Q, Li C, Wang W, Chen Y, Xiao Z. Comparative proteome analysis of human adenocarcinoma. Med Oncol. 2010;27(2):346–56.
54. Peng XC, Gong FM, Zhao YW, Zhou LX, Xie YW, Liao HL, et al. Comparative proteomic approach identifies PKM2 and cofilin-1 as potential diagnostic, prognostic and therapeutic targets for pulmonary adenocarcinoma. PLoS One. 2011;6(11):e27309.
55. Maciel CM, Junqueira M, Paschoal ME, Kawamura MT, Duarte RL, Carvalho Mda G, et al. Differential proteomic serum pattern of low molecular weight proteins expressed by adenocarcinoma lung cancer patients. J Exp Ther Oncol. 2005;5(1):31–8.
56. Pernemalm M, De Petris L, Eriksson H, Branden E, Koyi H, Kanter L, et al. Use of narrow-range peptide IEF to improve detection of lung adenocarcinoma markers in plasma and pleural effusion. Proteomics. 2009;9(13):3414–24.
57. Li C, Xiao Z, Chen Z, Zhang X, Li J, Wu X, et al. Proteome analysis of human lung squamous carcinoma. Proteomics. 2006;6(2):547–58.
58. Xu Y, Cao LQ, Jin LY, Chen ZC, Zeng GQ, Tang CE, et al. Quantitative proteomic study of human lung squamous carcinoma and normal bronchial epithelial acquired by laser capture microdissection. J Biomed Biotechnol. 2012;2012:510418.
59. Li B, Chang J, Chu Y, Kang H, Yang J, Jiang J, et al. Membrane proteomic analysis comparing squamous cell lung cancer tissue and tumour-adjacent normal tissue. Cancer Lett. 2012;319(1):118–24.
60. Tan F, Jiang Y, Sun N, Chen Z, Lv Y, Shao K, et al. Identification of isocitrate dehydrogenase 1 as a potential diagnostic and prognostic biomarker for non-small cell lung cancer by proteomic analysis. Mol Cell Proteomics. 2012;11(2):M111 008821.
61. Yang F, Xiao ZQ, Zhang XZ, Li C, Zhang PF, Li MY, et al. Identification of tumor antigens in human lung squamous carcinoma by serological proteome analysis. J Proteome Res. 2007;6(2):751–8.
62. Dowling P, O'Driscoll L, Meleady P, Henry M, Roy S, Ballot J, et al. 2-D difference gel electrophoresis of the lung squamous cell carcinoma versus normal sera demonstrates consistent alterations in the levels of ten specific proteins. Electrophoresis. 2007;28(23):4302–10.
63. Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, et al. Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet. 2003;362(9382):433–9.
64. Xu BJ, Gonzalez AL, Kikuchi T, Yanagisawa K, Massion PP, Wu H, et al. MALDI–MS derived prognostic protein markers for resected non-small cell lung cancer. Proteomics Clin Appl. 2008;2(10–11):1508–17.

65. De Petris L, Orre LM, Kanter L, Pernemalm M, Koyi H, Lewensohn R, et al. Tumor expression of S100A6 correlates with survival of patients with stage I non-small-cell lung cancer. Lung Cancer. 2009;63(3):410–7.
66. Gharib TG, Chen G, Wang H, Huang CC, Prescott MS, Shedden K, et al. Proteomic analysis of cytokeratin isoforms uncovers association with survival in lung adenocarcinoma. Neoplasia. 2002;4(5):440–8.
67. Chen G, Gharib TG, Wang H, Huang CC, Kuick R, Thomas DG, et al. Protein profiles associated with survival in lung adenocarcinoma. Proc Natl Acad Sci U S A. 2003;100(23):13537–42.
68. Liu YF, Xiao ZQ, Li MX, Li MY, Zhang PF, Li C, et al. Quantitative proteome analysis reveals annexin A3 as a novel biomarker in lung adenocarcinoma. J Pathol. 2009;217(1):54–64.
69. Liu YF, Chen YH, Li MY, Zhang PF, Peng F, Li GQ, et al. Quantitative proteomic analysis identifying three annexins as lymph node metastasis-related proteins in lung adenocarcinoma. Med Oncol. 2012;29(1):174–84.
70. Maeda J, Hirano T, Ogiwara A, Akimoto S, Kawakami T, Fukui Y, et al. Proteomic analysis of stage I primary lung adenocarcinoma aimed at individualisation of postoperative therapy. Br J Cancer. 2008;98(3):596–603.
71. Li DJ, Deng G, Xiao ZQ, Yao HX, Li C, Peng F, et al. Identificating 14-3-3 sigma as a lymph node metastasis-related protein in human lung squamous carcinoma. Cancer Lett. 2009;279(1):65–73.
72. Yao H, Zhang Z, Xiao Z, Chen Y, Li C, Zhang P, et al. Identification of metastasis associated proteins in human lung squamous carcinoma using two-dimensional difference gel electrophoresis and laser capture microdissection. Lung Cancer. 2009;65(1):41–8.
73. Okano T, Kondo T, Fujii K, Nishimura T, Takano T, Ohe Y, et al. Proteomic signature corresponding to the response to gefitinib (Iressa, ZD1839), an epidermal growth factor receptor tyrosine kinase inhibitor in lung adenocarcinoma. Clin Cancer Res. 2007;13(3):799–805.
74. Taguchi F, Solomon B, Gregorc V, Roder H, Gray R, Kasahara K, et al. Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. J Natl Cancer Inst. 2007;99(11):838–46.
75. Amann JM, Lee JW, Roder H, Brahmer J, Gonzalez A, Schiller JH, et al. Genetic and proteomic features associated with survival after treatment with erlotinib in first-line therapy of non-small cell lung cancer in Eastern Cooperative Oncology Group 3503. J Thorac Oncol. 2010;5(2):169–78.
76. Carbone DP, Ding K, Roder H, Grigorieva J, Roder J, Tsao MS, et al. Prognostic and predictive role of the veristrat plasma test in patients with advanced non-small-cell lung cancer treated with erlotinib or placebo in the NCIC Clinical Trials Group BR.21 Trial. J Thorac Oncol. 2012;7(11):1653–60.
77. Carbone DP, Salmon JS, Billheimer D, Chen H, Sandler A, Roder H, et al. VeriStrat classifier for survival and time to progression in non-small cell lung cancer (NSCLC) patients treated with erlotinib and bevacizumab. Lung Cancer. 2010;69(3):337–40.
78. Chung CH, Seeley EH, Roder H, Grigorieva J, Tsypin M, Roder J, et al. Detection of tumor epidermal growth factor receptor pathway dependence by serum mass spectrometry in cancer patients. Cancer Epidemiol Biomarkers Prev. 2010;19(2):358–65.
79. Gautschi O, Dingemans AM, Crowe S, Peters S, Roder H, Grigorieva J, et al. VeriStrat((R)) has a prognostic value for patients with advanced non-small cell lung cancer treated with erlotinib and bevacizumab in the first line: pooled analysis of SAKK19/05 and NTR528. Lung Cancer. 2013;79(1):59–64.
80. Kuiper JL, Lind JS, Groen HJ, Roder J, Grigorieva J, Roder H, et al. VeriStrat((R)) has prognostic value in advanced stage NSCLC patients treated with erlotinib and sorafenib. Br J Cancer. 2012;107(11):1820–5.
81. Lazzari C, Spreafico A, Bachi A, Roder H, Floriani I, Garavaglia D, et al. Changes in plasma mass-spectral profile in course of treatment of non-small cell lung cancer patients with epidermal growth factor receptor tyrosine kinase inhibitors. J Thorac Oncol. 2012;7(1):40–8.
82. Milan E, Lazzari C, Anand S, Floriani I, Torri V, Sorlini C, et al. SAA1 is overexpressed in plasma of non small cell lung cancer patients with poor outcome after treatment with epidermal growth factor receptor tyrosine-kinase inhibitors. J Proteomics. 2012.

Int. J. Mol. Sci. **2013**, *14*, 3440-3455; doi:10.3390/ijms14023440

OPEN ACCESS
International Journal of
Molecular Sciences
ISSN 1422-0067
www.mdpi.com/journal/ijms

Article

Identification of Oxidative Stress Related Proteins as Biomarkers for Lung Cancer and Chronic Obstructive Pulmonary Disease in Bronchoalveolar Lavage

Maria Dolores Pastor [1,†], Ana Nogal [1,†], Sonia Molina-Pinelo [1], Ricardo Meléndez [1], Beatriz Romero-Romero [2], Maria Dolores Mediano [1], Jose L. López-Campos [1,2], Rocío García-Carbonero [1,3], Amparo Sanchez-Gastaldo [1,3], Amancio Carnero [1,4] and Luis Paz-Ares [1,3,*]

[1] Biomedicine Institute of Seville, Seville 41013, Spain;
E-Mails: marilopastor@gmail.com (M.D.P.); anabsnogal@gmail.com (A.N.); pinelo_sonia@hotmail.com (S.M.-P.); ricardomelendezcadenas@gmail.com (R.M.); mmrambla@gmail.com (M.D.M.); josel.lopezcampos.sspa@juntadeandalucia.es (J.L.L.-C.); rgcarbonero@gmail.com (R.G.-C.); klimta77@hotmail.com (A.S.-G.); acarnero@us.es (A.C.)

[2] Respiratory Diseases Medical and Surgical Unit of the Virgen del Rocío University Hospital, Seville 41013, Spain; E-Mail: beatrizromero@terra.es

[3] Oncology Unit of the Virgen del Rocío University Hospital, Seville 41013, Spain

[4] Superior Counsil of Scientific Research, Seville 41013, Spain

[†] The authors contributed equally to this work.

[*] Author to whom correspondence should be addressed;
E-Mail: luis.pazares.sspa@juntadeandalucia.es; Tel.: +34-955-013-414; Fax: +34-955-013-292.

Received: 24 December 2012; in revised form: 23 January 2013 / Accepted: 31 January 2013 / Published: 6 February 2013

Abstract: Lung cancer (LC) and chronic obstructive pulmonary disease (COPD) commonly coexist in smokers, and the presence of COPD increases the risk of developing LC. Cigarette smoke causes oxidative stress and an inflammatory response in lung cells, which in turn may be involved in COPD and lung cancer development. The aim of this study was to identify differential proteomic profiles related to oxidative stress response that were potentially involved in these two pathological entities. Protein content was assessed in the bronchoalveolar lavage (BAL) of 60 patients classified in four groups: COPD, COPD and LC, LC, and control (neither COPD nor LC). Proteins were separated into spots by two dimensional polyacrylamide gel electrophoresis (2D-PAGE) and examined by

matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF/TOF). A total of 16 oxidative stress regulatory proteins were differentially expressed in BAL samples from LC and/or COPD patients as compared with the control group. A distinct proteomic reactive oxygen species (ROS) protein signature emerged that characterized lung cancer and COPD. In conclusion, our findings highlight the role of the oxidative stress response proteins in the pathogenic pathways of both diseases, and provide new candidate biomarkers and predictive tools for LC and COPD diagnosis.

Keywords: bronchoalveolar lavage; lung cancer; screening; biomarker; inflammation; proteomics; ROS; oxidative stress

1. Introduction

Cigarette smoking has been recognized as the most important causative factor of COPD and it is associated with more than 90% of lung cancer cases [1]. Lung cancer accounts for 12% of all cancer diagnoses worldwide, making it the largest cause of cancer-associated death worldwide, accounting for more than one million casualties per year worldwide. COPD is also a major independent risk factor for lung carcinoma, among long-term smokers. In fact, the presence of COPD increases the risk of lung cancer up to 4.5-fold. Indeed, 50%–70% of patients diagnosed with lung cancer have spirometric evidence of COPD [2]. Cigarette smoke (CS) contains over 10^{14} free radicals per puff that include reactive oxygen species (ROS) [3]. Inhaled oxidants from smoke generate cellular damage by directly targeting proteins, lipids, and nucleic acids, and deplete the level of antioxidants in the lung, thereby overwhelming the oxidant/antioxidant balance of the lung, leading to increased oxidative stress [4]. ROS can lead to the activation of various cell signaling components. Examples include the extracellular signal regulated kinases (ERKs), c-jun *N*-terminal kinases (JNKs), p38 MAPKs, PKC, PI3K/Akt, and growth factor tyrosine kinases receptors pathways, all of which lead to increased inflammatory gene transcription. Indeed, oxidative stress in the lungs has been implicated in COPD severity and lung carcinogenesis [5]. This process is one of the mechanisms proposed in the common pathogenesis of lung cancer and COPD, along with inflammation, epithelial-mesenchymal transition (EMT), altered DNA repair, and cellular proliferation [6]. Proteins are important molecular signposts of oxidative damage. Different proteomic approaches have been developed and used for the detection and identification of ROS related proteins [7]. In the last few years, combined proteomics, mass spectrometry (MS), and affinity chemistry-based methodologies have contributed in a significant way to provide a better understanding of protein oxidative modifications occurring in various biological specimens under different physiological and pathological conditions. Bronchoalveolar lavage (BAL) is the clinical biofluid sampling of the soluble proteins contents of the airway lumen. A comparison between serum and BAL proteomes reveals that a certain number of proteins are present at a higher level in BAL than in plasma, suggesting that they are specifically produced in the airways. These proteins are, therefore, potential candidates for becoming lung-specific biomarkers [8]. 2D-PAGE is considered one of the best techniques for separation of complex mixtures of soluble proteins [9]. Several studies of BAL protein profiles obtained by 2D-PAGE analysis aimed at revealing the

differences between smokers and never smokers [10,11] as well as studies directed to determine the risk of developing COPD [12–14]. However, to the best of our knowledge, there are no 2D-PAGE studies in lung cancer using BAL. The 2D-PAGE studies of LC have been performed mainly in plasma and tissue. These analyses have focused on a better understanding of the molecular basis of cancer pathogenesis [15–17], as well as on the identification of new diagnostic, prognostic, and predictive markers for lung cancer [18,19]. In this regard, analyzing the protein composition of BAL mediated by a high-throughput technology, given its vicinity to tumor cells and enrichment in tumor-derived proteins, would be insightful. In this study, we have investigated the changes occurring in the proteome of BAL samples from lung cancer and/or COPD patients, and found a set of redox regulative proteins differentially expressed in each disease.

2. Results

2.1. Proteome Profiles of Comparison of LC and/or COPD

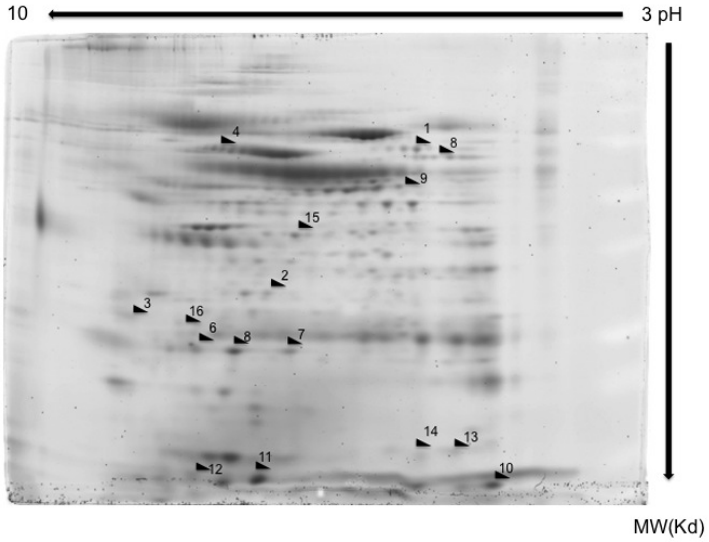
The experiments were performed in BAL samples extracted from a cohort of 60 patients divided into four groups (control group and LC and/or COPD groups) whose characteristics are described in Table 1. The spots that showed significant increment or reduction of their expression compared to control group (neither COPD nor LC) were identified by MALDI-TOF/TOF-MS. The MS/MS data were acquired and compared to the Swiss-Prot database using MASCOT software. Candidate proteins were selected from each spot, taking into consideration several variables such as isoelectric point, molecular mass, matched peptides, and sequence coverage (Figure 1). A total of 123 protein spots were successfully identified. Of these, 40 proteins spots had consistent significant differences (>2-fold, $p < 0.05$) between lung cancer and/or COPD groups and the control group. Among them, a major group of 16 proteins were oxidative stress regulatory proteins (Table 2). The spots corresponding to this group of ROS regulatory proteins are marked on the representative gel 2D image in Figure 1. The rest of identified proteins were distributed in other variable groups such as inflammation, glycolysis and gluconeogenesis (Data not show).

Table 1. Patients characteristics.

	Controls *n* = 15	**COPD *n* = 15**	**LC *n* = 15**	**LC&COPD *n* = 15**
Gender				
Male	100.0% (15)	100.0% (15)	100.0% (15)	100.0% (15)
Female	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
Average age (range)	61.3 (41–80)	61.5 (45–78)	60.7 (46–69)	60.7 (49–68)
Smoking status				
Smokers	73.3% (11)	53.3% (8)	53.3% (8)	80.0% (12)
Ex-smokers	26.7% (4)	46.7% (7)	46.7% (7)	20.0% (3)
Packs-year	21.82	32.20	35.21	30.78
COPD				
Mild	-	20.0% (3)	-	53.3% (8)
Moderate	-	33.3% (5)	-	26.7% (4)
Severe	-	26.7% (4)	-	-
Very severe	-	20.0% (3)	-	20.0% (3)
Histology				
Adenocarcinoma	-	-	73.3% (11)	66.7% (10)
Squamous cell carcinoma	-	-	26.7% (4)	33.3% (5)

Abbreviations: COPD: chronic obstructive pulmonary disease; LC: lung cancer.

Figure 1. 2D-PAGE from a representative patient sample. A representative 7 cm Sypro stained gel of proteins in the non-lineal pH range 3–11.



Venn diagrams were used to display the differentially expressed proteins that were up- or down-regulated in each pathological group. Among these, two proteins from the LC patients group were found to be up-regulated (CTSD, ERZ,) in comparison with the LC/COPD and COPD groups (Figure 2). At the same time, the LC and LC/COPD groups shared one up-regulated protein (PPIA) in comparison with the COPD group (Figure 2). Similarly, the COPD and LC/COPD groups shared four up-regulated proteins (CAT, PRDX1, PRDX2, and PRDX5) in contrast with the LC group (Figure 2). Finally, the three pathological groups shared nine proteins, six of them up-regulated (TXN, CRP, GSR, IDH1, SERPINB1, and ARHGDIB), three down-regulated proteins (GSTA1, GSTA2, and GSTP) (Figure 2).

The 15 identified proteins associated with ROS were subsequently analyzed with IPA, a software tool capable of mapping proteins and existing networks. Interestingly, the transcriptional factor NF-κB was found as a link between the proteins network involving ROS. NF-κB may reflect a functional role of this pathway in lung carcinogenesis (Figure 3).

2.2. Western Blotting

From the results obtained in the previous section, and taking into account the characteristics described in Table 2 for each protein, four proteins (TXN, GSR, GSTA1, and CAT) were selected for validation. The western blot of TXN, GSR, GSTA1, and CAT and the corresponding β-actin are shown in Figure 4. Validation was performed in three random samples from each study group. The results of the western blot experiments indicate that TXN and GSR present similar increment of expression between LC and/or COPD groups in comparison with the control group. The GSTA1 protein showed a decrease of expression in the three pathological groups in comparison with the control group. These data confirm the results obtained from initial proteomic analysis.

Table 2. Protein spots searched by MASCOT software in database.

Spot no	Protein name	Protein symbol	Accession no	Protein MW	Protein PI	Peptide count	Protein score	Score C.I. %	Total ion score	Ion C.I. %	COPD		LC		LC/ COPD	
1	Catalase	CAT	gi|4557014	59946.8	6.90	13	310	100	234	100	Up	2.5	-		Up	2.8
2	Cathepsin D preprotein	CTSD	gi|4503143	45036.8	6.10	8	135	100	95	100	-		Up	3.0	.	
3	Ezrin	EZR	gi|46249758	69312.7	5.94	18	250	100	175	100	-		Up	3.0	-	
4	Glutathione reductase	GSR	gi|119583848	61464.6	8.71	7	131	100	108	100	Up	2.8	Up	3.1	Up	2.0
5	Glutathione *S*-transferase A1 subunit	GSTA1	gi|163310943	25628.7	8.72	15	384	100	268	100	Down	3.0	Down	3.2	Down	3.8
6	Glutathione *S*-transferase A2 subunit	GSTA2	gi|257476	25589.6	8.81	7	105	100	70	100	Down	2.5	Down	3.0	Down	2.6
7	Glutathione *S*-transferase P	GSTP1	gi|4504183	23569.1	5.43	10	633	100	541	100	Down	2.2	Down	2.4	Down	2.5
8	Isocitrate dehydrogenase 1	IDH1	gi|89573979	42091.0	6.19	8	62	100	29	100	Up	2.2	Up	2.1	Up	2.5
9	Leukocyte elastase inhibitor	SERPINB1	gi|13489087	42828.7	5.90	18	403	100	274	100	Up	2.7	Up	2.6	Up	2.5
10	Peptidylprolyl isomerase A (Cyclophilin A)	PPIA	gi|1633054	18097.9	7.82	10	260	100	159	100	-		Up	2.2	Up	2.6
11	Peroxiredoxin 1	PRDX1	gi|55959887	19134.7	6.41	8	170	100	98	100	Up	5.0	-		Up	4.2
12	Peroxiredoxin 5	PRDX5	gi|6166493	22261.6	8.85	11	638	100	537	100	Up	2.3	-		Up	2.4
13	Peroxiredoxin-2 isoform a	PRDX2	gi|32189392	22049.3	5.66	12	451	100	325	100	Up	3.0	-		Up	2.9
14	Rho GDP-dissociation inhibitor 2	ARHGDIB	gi|56676393	23030.6	5.10	7	215	100	170	100	Up	2.9	Up	2.6	Up	2.3
15	Thioredoxin	TXN	gi|135772	12345.0	7.93	10	241	100	203	100	Up	2.4	Up	2.3	Up	2.1

Figure 2. Venn diagram showing the overlap of up-regulated and down-regulated proteins in each pathological group. The up-regulated proteins are represented in black and the down-regulated proteins are represented in grey.

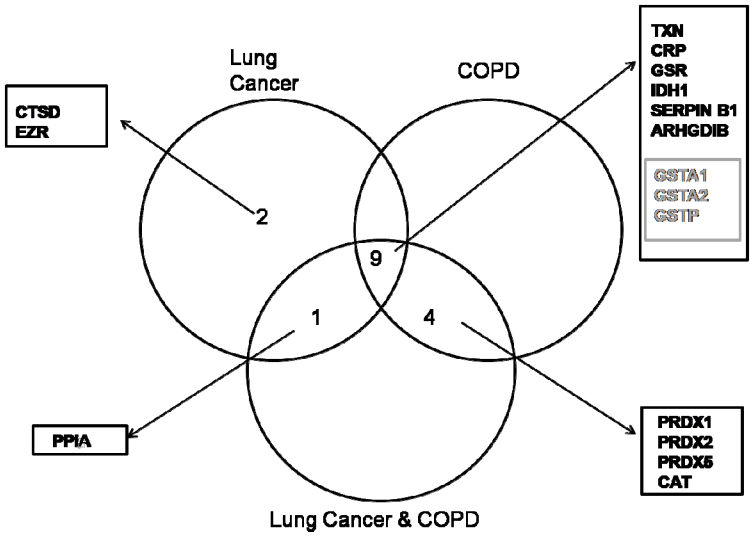


Figure 3. Ingenuity Pathway analysis of lung cancer and/or COPD *versus* controls revealed NF-κB as a major foundation.

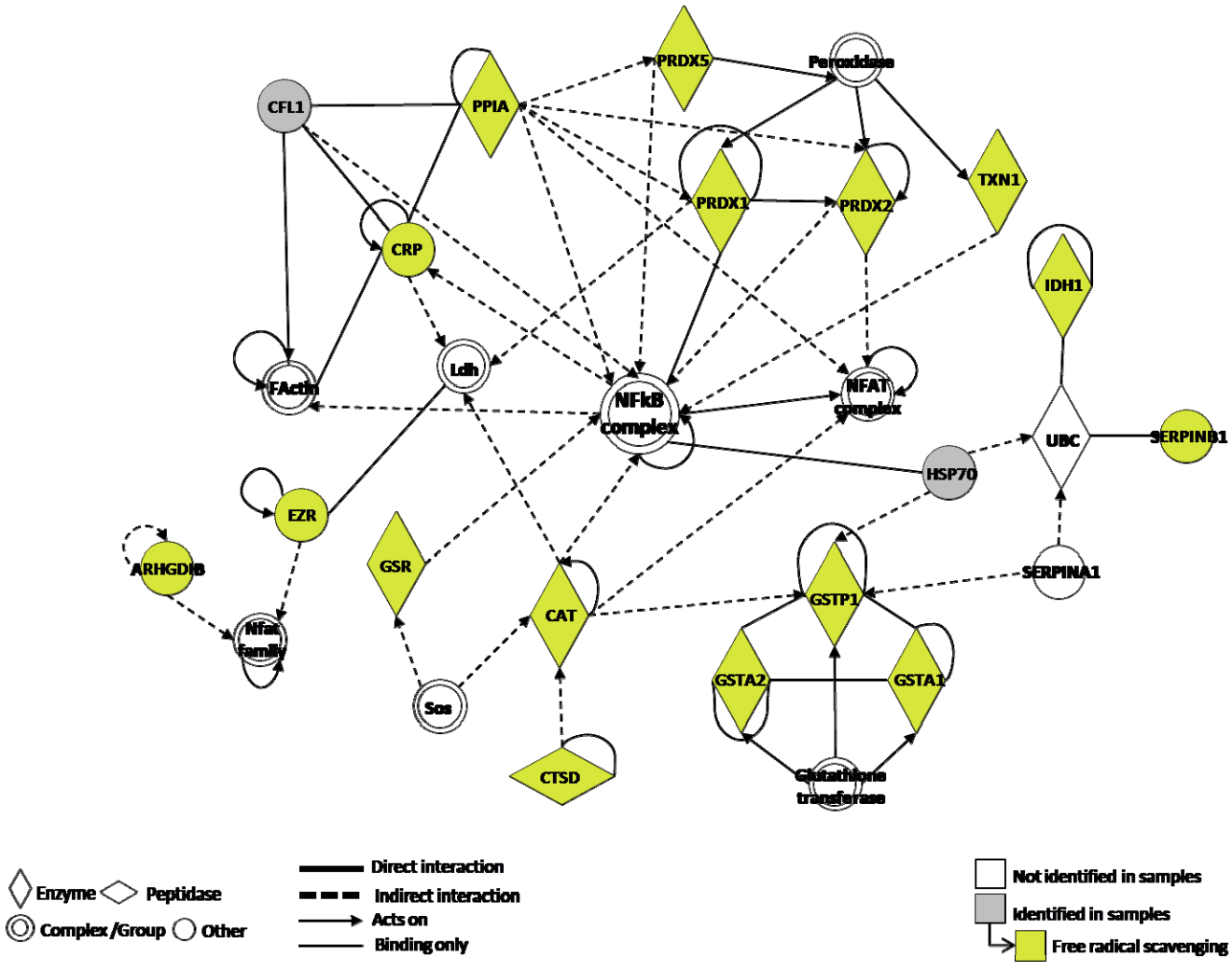
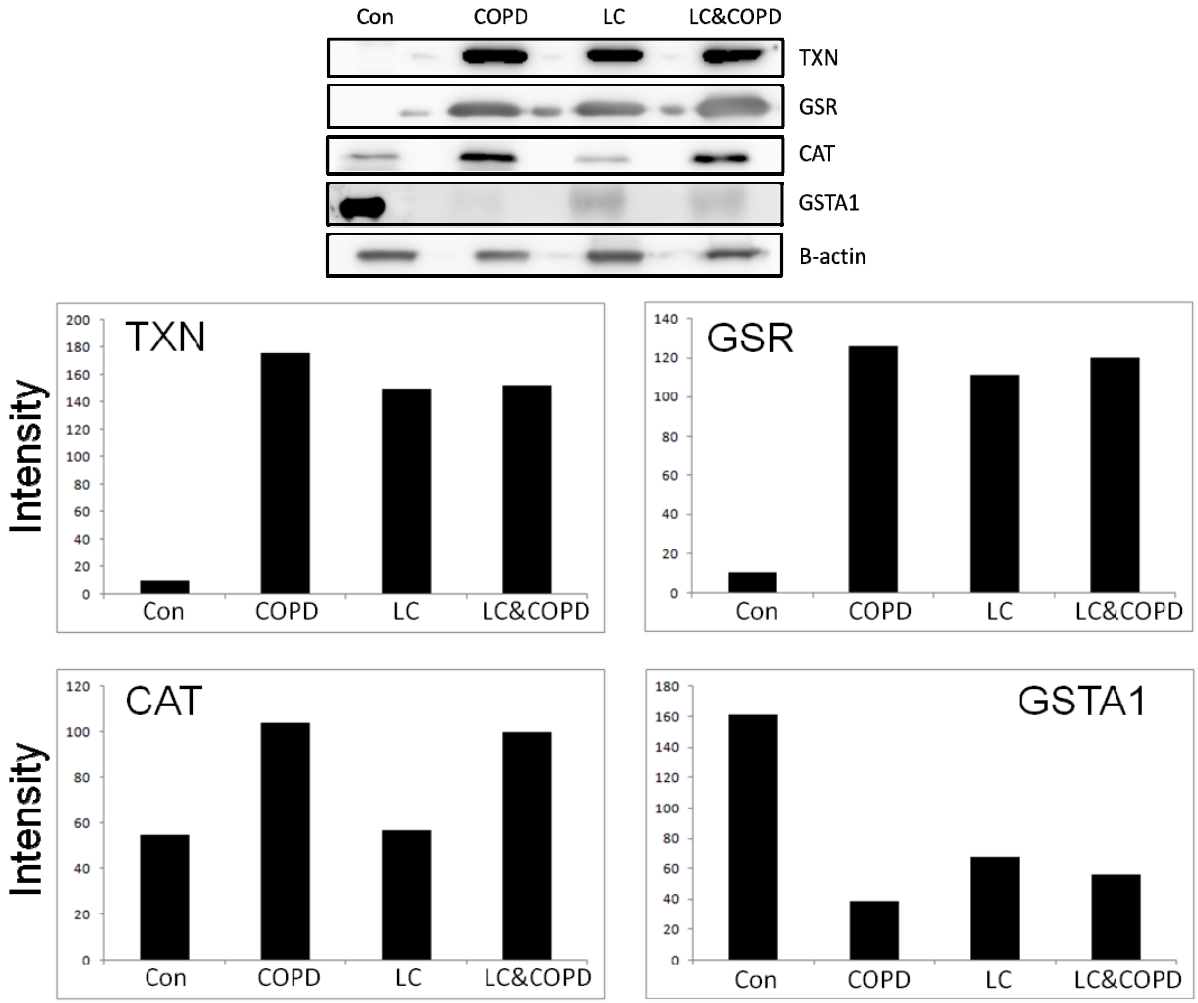


Figure 4. Western blotting for TXN, GSR, GSTA1, and CAT. The different expressions are seen in each of the groups and for each of the proteins. TXN and GSR present a similar increment of expression in all pathological groups. GSTA1 show a decrease of expression in the three pathological groups in comparison with the control group. These differences are also illustrated below with a bar chart.



3. Discussion

In this study, we presented a 2D-PAGE proteomic evaluation of BAL fluid in patients with the two most relevant smoking related diseases, lung cancer and COPD. Our results indicate that the protein composition of BAL showed relevant expression differences among the groups, especially between the control group and the disease groups. In concrete, we have observed 15 differentially expressed proteins involved in ROS metabolism.

ROS are involved in a large number of diseases, degenerative changes, leading to tissue degradation, a hallmark in carcinogenesis [20]. In a metabolically active cell, these redox system pathways maintain the balance between oxidant and antioxidant factors, by regulating the activation of specific transcription factors and the production of substances that neutralize oxidants [21,22]. However, in cancer settings, alterations in these redox pathways occur and the cell is no longer able to produce antioxidant substances to adjust the balance between oxidant and antioxidant factors, being

therefore unable to respond appropriately to the body's needs. This is usually the reason why many anticancer agents, including radiation, are ineffective, because the cytotoxicity induced by them to make the cancer regress affects the antioxidant activity of the redox system pathways [23]. Alterations in the physiological pathways involved in the regulation of the redox system have been identified in tumors [24].

The proteins TXN and PPIA (Cyclophilin A) were up-regulated in lung cancer BALs as compared to the other groups. TXN is a potent growth and cell survival factor, which activates specific transcription factors such as NF-κB, p53, HIFa, and AP-1. TXN regulates gene decoding for the production of substances that protect cells from oxidative stress induced by oxygen free radicals [25,26]. Its expression rises in several types of tumors [27,28] and it is generally related to tumor aggressiveness and inhibition of the immune system. TXN has also been evaluated as a biomarker and therapeutic target for cancer [29,30], and it is known that TXN levels can be used to indicate potential chemotherapy resistance [31]. Indeed, an increase in TXN1 levels has been associated with decreased survival in patients with tumors [32]. PPIA is secreted in response to ROS from vascular smooth muscle cells (VSMC) [33]. This protein is a chaperone protein that has several functions including protein trafficking, such as the nuclear translocation of ERK1/2 [34] and apoptosis-inducing factor (AIF) [35]. In addition, there is evidence that PPIA might be a valuable biomarker. Several studies evaluating the proteomic profile of different types of cancer as gastric, colorectal [36], and prostate have associated PPIA with a favorable outcome.

The proteins GSTA1, GSTA2, and GSTP1 were down-regulated in pathological groups compared with control group. These proteins belong to GST family of proteins, which are Phase II detoxification enzymes that catalyze detoxifying endogenous reactions with reduced GSH and protect cellular macromolecules from damage caused by cytotoxic and carcinogenic agents [37].

A sub-network of interacting peroxiredoxins (PRDX1, PRDX2, and PRDX5) and catalase enzymes were up-regulated in the COPD groups. The overexpression of these proteins is protective to cells given that they increase life span and decreases injuries that arise from ROS generation [38,39]. In addition, the increased levels of PRDX1 in BAL were observed in patients with acute lung injury compared with normal subject [40].

Finally, we have observed several deregulated proteins involved with the second line of defense against oxidative stress such as cathepsin D and ezrin. These proteins act when the first defense mechanism by non-enzymatic molecules and enzymatic scavengers, such as superoxide dismutases, catalase, and glutathione peroxidase, does not work properly against oxidative stress. Cathepsin D has been involved in the oxidative stress-induced apoptotic pathways. Furthermore, cathepsin D and ezrin are secreted aberrantly and excessively in various types of cancers [41,42], and are associated with increased cancer growth, invasion, and metastasis [41–45].

The re-establishment of homeostasis within the physiological pathways of the redox and immunological system is an important therapeutic goal in oncology. For this achievement, the identification of adequate biomarkers and molecular targets is essential. One advantage of integrating our proteomic approach with network analysis is its potential ability to provide an insight into existing molecular mechanisms. The analysis of the proteins found in our study by Ingenuity System Pathway Analysis software, in order to identify any common links beyond oxidative stress, revealed NF-κB, an important transcription factor, which is a redox-sensitive transcription factor [46,47].

NF-κB regulated genes include cytokines, adhesion molecules, angiogenic factors, anti-apoptotic factors, and matrix metalloproteinases (MMPs), which are involved in different steps of carcinogenesis. It has been suggested that NF-κB promotes lung cancer mainly through mediating inflammatory cytokines secretion to establish a cancer-prone inflammatory microenvironment [48]. Similarly, NF-κB pathways play a crucial role in the pathogenesis/development of COPD by increasing the release of pro-inflammatory mediators leading to chronic inflammation in the lung. In bronchial biopsies of airway mucosa from patients with COPD, protein expression of the p65 subunit of NF-κB was increased compared with its expression in non-smokers, and correlated with airflow obstruction [49]. Our results suggest that oxidative stress induced in lung airways might alter redox detoxifying enzymes, which end up activating NF-κB node. This activation might contribute to cancer development and therapy resistance. Both chemo and radiation therapies induce NF-κB activation in cancer cells, which contributes to resistance to those same therapies [50]. Indeed, inhibition of NF-κB signaling by various approaches has been shown to augment the efficacy of chemotherapeutics and radiation in killing cancer cells *in vitro* and *in vivo* [51,52]. Some of the NF-κB inhibitors that enhanced lung cancer cell death induced by chemotherapeutics are genistein with cisplatin or docetaxel [53,54], embelin with paclitaxel [55], expression of IκBα mutant with cisplatin, gemcitabline, adriamicin and etoposide [56,57]. Increasing evidences shows that NF-κB plays a critical role in lung cancer development and suggests NF-κB as a target for lung cancer chemoprevention.

In summary, this study in BAL suggests that cigarette smoking produces a free radical scavenging and oxidative stress response shared by the pathogenic pathways of lung cancer and COPD. Furthermore, the pivotal networking signaling is NF-κB. The proteins included in each specific disease signature may provide new biomarkers and predictive tools for LC and COPD. Additional validation of the identified proteins in independent patient cohorts is warranted.

4. Materials and Methods

4.1. Patients and Samples

Samples were obtained from four groups of patients: control group (without COPD or LC), COPD group, LC group, and LC with COPD. A description of all included patients can be found on Table 1. From 2009 to 2011, a total of 60 patients who had required flexible bronchoscopy for diagnostic purposes, were chosen for the study. All samples were collected from patients of the Virgen del Rocío Hospital (Seville, Spain). The selection criteria to be included in the study were (1) patients had to have been evaluated by pneumology services by haemoptysis and/or a pulmonary nodule, (2) patients were smokers or ex-smokers of >20 pack year, (3) over 40 years of age. The exclusion criteria for this study were (1) Prior diagnosis of malignancy, (2) active pulmonary tuberculosis, (3) previous lung resection, (4) history of drug abuse, and (5) presence of other acute or chronic inflammatory disease. The present study was approved by the Hospital's Ethical Committee and a written informed consent was obtained from all patients prior to their inclusion in the study.

Subjects were prepared with a combination of topical anaesthesia (20% benzocaine spray to the pharynx plus 2% topical lidocaine as needed) and conscious sedation using midazolam and meperidine

according to institutional guidelines. Bronchioalveolar lavage (BAL) samples were obtained by installation and aspiration of 4 to 6 mL aliquots of 0.9% sterile saline in the bronchopulmonary segment. Recovered fluid was immediately passed through a 100 μm sterile nylon filter (Becton Dickinson, San Jose, CA, USA) to remove mucus and transported on ice to the laboratory. The total volume was then centrifuged for 10 min at 1800*g* and 4 °C. The supernatant was aliquoted into 2 mL tubes and frozen at −80 °C until further use.

4.2. Sample Treatment

Approximately 4–8 mL of sample was available to our experiments. Due to its low protein content, BAL samples needed to be concentrated before use. BAL samples were thawed on ice with a Protease Inhibitor Cocktail kit (Thermo Scientific, Franklin, MA, USA). Samples were then aliquoted into new tubes and placed on the vacuum concentrator (Concentrator plus-Eppendorf, Hamburg, Germany). The next step of the sample treatment protocol was the depletion of two of the most abundant proteins present in BAL (albumin and immunoglobulin G) that may obscure the presence of low abundant ones. This was accomplished by the use of SpinTrap columns (GE Healthcare) following the manufacturer's instructions. The depleted BAL samples were then cleaned to remove contaminants, such as salts, thiols, denaturants, that would interfere with the two dimensional gel electrophoresis (2D-PAGE) protocol. A 2-D Clean-up kit (GE Healthcare) was used according to the manufacturer's instructions with a rehydration solution containing urea (7 M), thiourea (2 M) and CHAPS (2%). Protein quantitation was assessed by the RCDC method (Bio-Rad, Hercules, CA, USA).

4.3. 2D-PAGE

BAL samples were used from the 60 patients included in the four groups (all samples were analyzed independently), by taking equal amounts of protein from each individual sample. A mixture of 75 μg of protein from each sample, DeStreak rehydration solution (GE Healthcare) and 0.5% of 3–10 nL pH IPG Buffer (GE Healthcare), in a final volume of 125 μL, were submitted to isoelectric focusing (IEF) in 7 cm IPG DryStrips (GE Healthcare) with a 3–10 nL pH range. Subsequent to IEF, strips were then placed on 12.5% acrylamide gels and proteins separated by electrophoresis. After this second dimension, gels were SYPRO® stained according to the manufacturer's instructions (Bio-Rad, Hercules, CA, USA). Gel images were scanned by using Typhoon TRIO (Amersham Biosciences, Piscataway, NJ, USA).

4.4. Image Analysis and Mass Spectrometry

Two-dimensional gel image analysis protein spot detection, spot matching, and semi-quantitative statistical analysis were performed using the Progenesis SameSpots (Nonlinear Dynamics, Durham, NC, USA). For each study group, four different gel images were analyzed, and a corresponding synthetic reference image was obtained. After computer matching, detected spots and spot matches were manually edited for greater accuracy. The detection of differentially expressed protein spots was performed using the test INCA volume and proteins that were two-fold or higher differentially

expressed were considered significant. Protein spots of interest were excised from the stained gel using a ProXcision robot (PerkinElmer, Boston, MA, USA) and sent for MS analysis.

4.5. Protein Identification by Mass Spectrometry

MS experiments were performed by the Proteomics platform of the University of Cordoba Proteomics Service (Cordoba, Spain). Protein spots of interest were washed with water and their trypsin digestion was performed. Tryptic peptides were mixed with a CCA matrix solution. The mixture was analyzed with a Voyager System DE-STR 7307 MALDI-TOF/TOF Mass Spectrometer (ABI) to obtain a peptide mass fingerprint (PMF). Peptide matching and protein searches against the Swiss-Prot database were done using the Mascot search engine with a mass tolerance of ±50 ppm. Protein scores >60 (threshold) indicated identity or extensive homology ($p < 0.05$) and were considered significant.

4.6. Functional Analysis of the Identified Proteins

From each spot only identified proteins with a probability higher than 95% and with at least two matched unique peptides were considered in the analysis, except for keratins, which were not considered. The experimental molecular weight and isoelectric point of each identified protein were determined based on the location of the original spot on the 2-D gel using the Progenesis software.

4.7. Western Blot

BAL proteins (30 μg) were separated in 12% gels polyacrylamide (SDS-PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes (Bio-Rad, Hercules, CA, USA). After blocking, the blots were incubated overnight at 4 °C with primary antibodies according to the manufacturer's instructions: anti-TXN, anti-GSR, anti-GSTA1, anti-CAT (1:1000, EPITOMICS, Burlingame, CA, USA), Secondary antibodies, peroxidase-conjugated anti-mouse (GE Healthcare, Uppsala, Sweden) and anti-rabbit (Cell Signaling, Beverly, MA, USA), were then applied to the individual membranes (1:2000) for 1 h at room temperature. Protein bands were revealed using enhanced chemiluminescence ECL (GE Healthcare, Uppsala, Sweden) and visualized in an image analyzer (Mini LAS-3000, Fujifilm, Tokyo, Japan). The relative protein levels were calculated by comparison to the amount of β-actin protein (1:1000 Abcam, Cambridge, MA, USA). The analysis of the expression values of the proteins of interest obtained by western blot was performed by densitometry, using Image J software and the results were expressed as fold change relative to the control protein (β-actin). The experiments were repeated three times independently.

4.8. Bioinformatics

The biological functions, in terms of gene ontology and interaction network, were analyzed using Ingenuity Pathways Analysis (IPA, version 7.1). Based on the local networks created by computational algorithms, identified proteins were connected with hub proteins, forming a functional protein cluster.

Acknowledgments

We would like to thank A. Salinas for her continuous technical support. LPA is funded by Fondo de Investigación Sanitaria (PI081156) and Fondo de Investigación Sanitaria (PI1102688). MDP is funded by Fondo de Investigación Sanitaria (CD 09/00148). AN is funded by the Portuguese Ministry of Science, Technology and Superior Education—FCT (Fundação para a Ciência e para a Tecnologia: SFRH/ BD/48341/2008). SMP is funded by Fondo de Investigación Sanitaria (CD 11/00153).

References

1. Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA Cancer J. Clin.* **2011**, *61*, 69–90.
2. Yao, H.; Rahman, I. Current concepts on the role of inflammation in COPD and lung cancer. *Curr. Opin. Pharmacol.* **2009**, *9*, 375–383.
3. Pryor, W.A.; Hales, B.J.; Premovic, P.I.; Church, D.F. The radicals in cigarette tar: Their nature and suggested physiological implications. *Science* **1983**, *220*, 425–427.
4. Yoshida, T.; Tuder, R.M. Pathobiology of cigarette smoke-induced chronic obstructive pulmonary disease. *Physiol. Rev.* **2007**, *87*, 1047–1082.
5. Psarras, S.; Caramori, G.; Contoli, M.; Papadopoulos, N.; Papi, A. Oxidants in asthma and in chronic obstructive pulmonary disease (COPD). *Curr. Pharm. Des.* **2005**, *11*, 2053–2062.
6. Lee, G.; Walser, T.C.; Dubinett, S.M. Chronic inflammation, chronic obstructive pulmonary disease, and lung cancer. *Curr. Opin. Pulm. Med.* **2009**, *15*, 303–307.
7. Ghezzi, P.; Bonetto, V. Redox proteomics: Identification of oxidatively modified proteins. *Proteomics* **2003**, *3*, 1145–1153.
8. Hermans, C.; Dong, P.; Robin, M.; Jadoul, M.; Bernard, A.; Bersten, A.D.; Doyle, I.R. Determinants of serum levels of surfactant proteins A and B and Clara cell protein CC16. *Biomarkers* **2003**, *8*, 461–471.
9. Baraibar, M.A.; Hyzewicz, J.; Rogowska-Wrzesinska, A.; Ladouce, R.; Roepstorff, P.; Mouly, V.; Friguet, B. Oxidative stress-induced proteome alterations target different cellular pathways in human myoblasts. *Free Radic. Biol. Med.* **2011**, *51*, 1522–1532.
10. Plymoth, A.; Yang, Z.; Lofdahl, C.G.; Ekberg-Jansson, A.; Dahlback, M.; Fehniger, T.E.; Marko-Varga, G.; Hancock, W.S. Rapid proteome analysis of bronchoalveolar lavage samples of lifelong smokers and never-smokers by micro-scale liquid chromatography and mass spectrometry. *Clin. Chem.* **2006**, *52*, 671–679.
11. Kelsen, S.G.; Duan, X.; Ji, R.; Perez, O.; Liu, C.; Merali, S. Cigarette smoke induces an unfolded protein response in the human lung: A proteomic approach. *Am. J. Respir. Cell Mol. Biol.* **2008**, *38*, 541–550.
12. Pinto-Plata, V.; Toso, J.; Lee, K.; Bilello, J.; Mullerova, H.; De Souza, M.; Vessey, R.; Celli, B. Use of proteomic patterns of serum biomarkers in patients with chronic obstructive pulmonary disease: Correlation with clinical parameters. *Proc. Am. Thorac. Soc.* **2006**, *3*, 465–466.

13. Merkel, D.; Rist, W.; Seither, P.; Weith, A.; Lenter, M.C. Proteomic study of human bronchoalveolar lavage fluids from smokers with chronic obstructive pulmonary disease by combining surface-enhanced laser desorption/ionization-mass spectrometry profiling with mass spectrometric protein identification. *Proteomics* **2005**, *5*, 2972–2980.
14. Casado, B.; Iadarola, P.; Pannell, L.K.; Luisetti, M.; Corsico, A.; Ansaldo, E.; Ferrarotti, I.; Boschetto, P.; Baraniuk, J.N. Protein expression in sputum of smokers and chronic obstructive pulmonary disease patients: A pilot study by CapLC-ESI-Q-TOF. *J. Proteome Res.* **2007**, *6*, 4615–4623.
15. Hoagland, L.F.M., IV; Campa, M.J.; Gottlin, E.B.; Herndon, J.E., II; Patz, E.F., Jr. Haptoglobin and posttranslational glycan-modified derivatives as serum biomarkers for the diagnosis of nonsmall cell lung cancer. *Cancer* **2007**, *110*, 2260–2268.
16. Jessie, K.; Pang, W.W.; Haji, Z.; Rahim, A.; Hashim, O.H. Proteomic analysis of whole human saliva detects enhanced expression of interleukin-1 receptor antagonist, thioredoxin and lipocalin-1 in cigarette smokers compared to non-smokers. *Int. J. Mol. Sci.* **2010**, *11*, 4488–4505.
17. Brower, V. Biomarker studies abound for early detection of lung cancer. *J. Natl. Cancer Inst.* **2009**, *101*, 11–13.
18. Kang, S.M.; Sung, H.J.; Ahn, J.M.; Park, J.Y.; Lee, S.Y.; Park, C.S.; Cho, J.Y. The Haptoglobin beta chain as a supportive biomarker for human lung cancers. *Mol. Biosyst.* **2011**, *7*, 1167–1175.
19. Tian, T.; Hao, J.; Xu, A.; Luo, C.; Liu, C.; Huang, L.; Xiao, X.; He, D. Determination of metastasis-associated proteins in non-small cell lung cancer by comparative proteomic analysis. *Cancer Sci.* **2007**, *98*, 1265–1274.
20. Luo, C.; Urgard, E.; Vooder, T.; Metspalu, A. The role of COX-2 and Nrf2/ARE in anti-inflammation and antioxidative stress: Aging and anti-aging. *Med. Hypotheses* **2011**, *77*, 174–178.
21. Kakolyris, S.; Giatromanolaki, A.; Koukourakis, M.; Powis, G.; Souglakos, J.; Sivridis, E.; Georgoulias, V.; Gatter, K.C.; Harris, A.L. Thioredoxin expression is associated with lymph node status and prognosis in early operable non-small cell lung cancer. *Clin. Cancer Res.* **2001**, *7*, 3087–3091.
22. Holmgren, A.; Lu, J. Thioredoxin and thioredoxin reductase: Current research with special reference to human disease. *Biochem. Biophys. Res. Commun.* **2010**, *396*, 120–124.
23. Malmberg, K.J. Effective immunotherapy against cancer: A question of overcoming immune suppression and immune escape? *Cancer Immunol. Immunother.* **2004**, *53*, 879–892.
24. Kim, S.J.; Miyoshi, Y.; Taguchi, T.; Tamaki, Y.; Nakamura, H.; Yodoi, J.; Kato, K.; Noguchi, S. High thioredoxin expression is associated with resistance to docetaxel in primary breast cancer. *Clin. Cancer Res.* **2005**, *11*, 8425–8430.
25. Lillig, C.H.; Holmgren, A. Thioredoxin and related molecules—From biology to health and disease. *Antioxid. Redox. Signal.* **2007**, *9*, 25–47.
26. Jones, D.T.; Pugh, C.W.; Wigfield, S.; Stevens, M.F.; Harris, A.L. Novel thioredoxin inhibitors paradoxically increase hypoxia-inducible factor-alpha expression but decrease functional transcriptional activity, DNA binding, and degradation. *Clin. Cancer Res.* **2006**, *12*, 5384–5394.

27. Ceccarelli, J.; Delfino, L.; Zappia, E.; Castellani, P.; Borghi, M.; Ferrini, S.; Tosetti, F.; Rubartelli, A. The redox state of the lung cancer microenvironment depends on the levels of thioredoxin expressed by tumor cells and affects tumor progression and response to prooxidants. *Int. J. Cancer* **2008**, *123*, 1770–1778.
28. Chaiswing, L.; Bourdeau-Heller, J.M.; Zhong, W.; Oberley, T.D. Characterization of redox state of two human prostate carcinoma cell lines with different degrees of aggressiveness. *Free Radic. Biol. Med.* **2007**, *43*, 202–215.
29. Iwasawa, S.; Yamano, Y.; Takiguchi, Y.; Tanzawa, H.; Tatsumi, K.; Uzawa, K. Upregulation of thioredoxin reductase 1 in human oral squamous cell carcinoma. *Oncol. Rep.* **2011**, *25*, 637–644.
30. Lincoln, D.T.; Al-Yatama, F.; Mohammed, F.M.; Al-Banaw, A.G.; Al-Bader, M.; Burge, M.; Sinowatz, F.; Singal, P.K. Thioredoxin and thioredoxin reductase expression in thyroid cancer depends on tumour aggressiveness. *Anticancer Res.* **2010**, *30*, 767–775.
31. Nguyen, P.; Awwad, R.T.; Smart, D.D.; Spitz, D.R.; Gius, D. Thioredoxin reductase as a novel molecular target for cancer therapy. *Cancer Lett.* **2006**, *236*, 164–174.
32. Welsh, S.J.; Bellamy, W.T.; Briehl, M.M.; Powis, G. The redox protein thioredoxin-1 (Trx-1) increases hypoxia-inducible factor 1alpha protein expression: Trx-1 overexpression results in increased vascular endothelial growth factor production and enhanced tumor angiogenesis. *Cancer Res.* **2002**, *62*, 5089–5095.
33. Jin, Z.G.; Melaragno, M.G.; Liao, D.F.; Yan, C.; Haendeler, J.; Suh, Y.A.; Lambeth, J.D.; Berk, B.C. Cyclophilin A is a secreted growth factor induced by oxidative stress. *Circ. Res.* **2000**, *87*, 789–796.
34. Pan, H.; Luo, C.; Li, R.; Qiao, A.; Zhang, L.; Mines, M.; Nyanda, A.M.; Zhang, J.; Fan, G.H. Cyclophilin A is required for CXCR4-mediated nuclear export of heterogeneous nuclear ribonucleoprotein A2, activation and nuclear translocation of ERK1/2, and chemotactic cell migration. *J. Biol. Chem.* **2008**, *283*, 623–637.
35. Zhu, C.; Wang, X.; Deinum, J.; Huang, Z.; Gao, J.; Modjtahedi, N.; Neagu, M.R.; Nilsson, M.; Eriksson, P.S.; Hagberg, H.; *et al*. Cyclophilin A participates in the nuclear translocation of apoptosis-inducing factor in neurons after cerebral hypoxia-ischemia. *J. Exp. Med.* **2007**, *204*, 1741–1748.
36. Peng, Y.; Li, X.; Wu, M.; Yang, J.; Liu, M.; Zhang, W.; Xiang, B.; Wang, X.; Li, G.; Shen, S. New prognosis biomarkers identified by dynamic proteomic analysis of colorectal cancer. *Mol. Biosyst*. **2012**, *8*, 3077–3088.
37. Strange, R.C.; Spiteri, M.A.; Ramachandran, S.; Fryer, A.A. Glutathione-S-transferase family of enzymes. *Mutat. Res.* **2001**, *482*, 21–26.
38. Matsumoto, A.; Okado, A.; Fujii, T.; Fujii, J.; Egashira, M.; Niikawa, N.; Taniguchi, N. Cloning of the peroxiredoxin gene family in rats and characterization of the fourth member. *FEBS Lett*. **1999**, *443*, 246–250.
39. Lim, Y.S.; Cha, M.K.; Kim, H.K.; Uhm, T.B.; Park, J.W.; Kim, K.; Kim, I.H. Removals of hydrogen peroxide and hydroxyl radical by thiol-specific antioxidant protein as a possible role *in vivo*. *Biochem. Biophys. Res. Commun.* 1 **1993**, *92*, 273–280.

40. Gharib, S.A.; Nguyen, E.; Altemeier, W.A.; Shaffer, S.A.; Doneanu, C.E.; Goodlett, D.R.; Schnapp, L.M. Of mice and men: Comparative proteomics of bronchoalveolar fluid. *Eur. Respir. J.* **2010**, *35*, 1388–1395.
41. Garcia, M.; Derocq, D.; Pujol, P.; Rochefort, H. Overexpression of transfected cathepsin D in transformed cells increases their malignant phenotype and metastatic potency. *Oncogene* **1990**, *5*, 1809–1814.
42. Rochefort, H.; Capony, F.; Garcia, M.; Cavailles, V.; Freiss, G.; Chambon, M.; Morisset, M.; Vignon, F. Estrogen-induced lysosomal proteases secreted by breast cancer cells: A role in carcinogenesis? *J. Cell Biochem.* **1987**, *35*, 17–29.
43. Berchem, G.; Glondu, M.; Gleizes, M.; Brouillet, J.P.; Vignon, F.; Garcia, M.; Liaudet-Coopman, E. Cathepsin-D affects multiple tumor progression steps *in vivo*: Proliferation, angiogenesis and apoptosis. *Oncogene* **2002**, *21*, 5951–5955.
44. Kim, H.; Kang, H.J.; You, K.T.; Kim, S.H.; Lee, K.Y.; Kim, T.I.; Kim, C.; Song, S.Y.; Kim, H.J.; Lee, C. Suppression of human selenium-binding protein 1 is a late event in colorectal carcinogenesis and is associated with poor survival. *Proteomics* **2006**, *6*, 3466–3476.
45. Zeng, H.; Xu, L.; Xiao, D.; Zhang, H.; Wu, X.; Zheng, R.; Li, Q.; Niu, Y.; Shen, Z.; Li, E. Altered expression of ezrin in esophageal squamous cell carcinoma. *J. Histochem. Cytochem.* **2006**, *54*, 889–896.
46. Wang, S.J.; Wu, S.T.; Gokemeijer, J.; Fura, A.; Krishna, M.; Morin, P.; Chen, G.; Price, K.; Wang-Iverson, D.; Olah, T.; *et al*. Attribution of the discrepancy between ELISA and LC-MS/MS assay results of a PEGylated scaffold protein in post-dose monkey plasma samples due to the presence of anti-drug antibodies. *Anal. Bioanal. Chem.* **2012**, *402*, 1229–1239.
47. Nicolay, N.; Callaghan, M.A.; Domegan, L.M.; Oza, A.N.; Marsh, B.J.; Flanagan, P.C.; Igoe, D.M.; O'Donnell, J.M.; O'Flanagan, D.M.; O'Hora, A.P. Epidemiology, clinical characteristics and resource implications of pandemic (H1N1) 2009 in intensive care units in Ireland. *Crit. Care. Resusc.* **2010**, *12*, 255–261.
48. Wong, K.K.; Jacks, T.; Dranoff, G. NF-kappaB fans the flames of lung carcinogenesis. *Cancer Prev. Res.* **2010**, *3*, 403–405.
49. Di Stefano, A.; Caramori, G.; Oates, T.; Capelli, A.; Lusuardi, M.; Gnemmi, I.; Ioli, F.; Chung, K.F.; Donner, C.F.; Barnes, P.J.; *et al*. Increased expression of nuclear factor-κB in bronchial biopsies from smokers and patients with COPD. *Eur. Respir. J.* **2002**, *20*, 556–563.
50. Nakanishi, C.; Toi, M. Nuclear factor-κB inhibitors as sensitizers to anticancer drugs. *Nat. Rev. Cancer* **2005**, *5*, 297–309.
51. Denlinger, C.E.; Rundall, B.K.; Jones, D.R. Modulation of antiapoptotic cell signaling pathways in non-small cell lung cancer: The role of NF-kappaB. *Semin. Thorac. Cardiovasc. Surg.* **2004**, *16*, 28–39.
52. Shen, H.M.; Tergaonkar, V. NFκB signaling in carcinogenesis and as a potential molecular target for cancer therapy. *Apoptosis* **2009**, *14*, 348–363.
53. Cheng, Q.; Lee, H.H.; Li, Y.; Parks, T.P.; Cheng, G. Upregulation of Bcl-x and Bfl-1 as a potential mechanism of chemoresistance, which can be overcome by NF-κB inhibition. *Oncogene* **2000**, *19*, 4936–4940.

54. Li, Y.; Ahmed, F.; Ali, S.; Philip, P.A.; Kucuk, O.; Sarkar, F.H. Inactivation of nuclear factor kappaB by soy isoflavone genistein contributes to increased apoptosis induced by chemotherapeutic agents in human cancer cells. *Cancer Res.* **2005**, *65*, 6934–6942.
55. Osaki, S.; Nakanishi, Y.; Takayama, K.; Pei, X.H.; Ueno, H.; Hara, N. Transfer of IkappaBalpha gene increase the sensitivity of paclitaxel mediated with caspase 3 activation in human lung cancer cell. *J. Exp. Clin. Cancer Res.* **2003**, *22*, 69–75.
56. Chen, W.; Wang, X.; Bai, L.; Liang, X.; Zhuang, J.; Lin, Y. Blockage of NF-kappaB by IKKbeta- or RelA-siRNA rather than the NF-κB super-suppressor IkappaBalpha mutant potentiates adriamycin-induced cytotoxicity in lung cancer cells. *J. Cell Biochem.* **2008**, *105*, 554–561.
57. Jones, D.R.; Broad, R.M.; Madrid, L.V.; Baldwin, A.S., Jr.; Mayo, M.W. Inhibition of NF-κB sensitizes non-small cell lung cancer cells to chemotherapy-induced apoptosis. *Ann. Thorac. Surg.* **2000**, *70*, 930–936.

JOURNAL OF PROTEOMICS 89 (2013) 227–237

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Identification of proteomic signatures associated with lung cancer and COPD

M.D. Pastor[a,1], A. Nogal[a,d,1], S. Molina-Pinelo[a], R. Meléndez[a], A. Salinas[a], M. González De la Peña[a,b], J. Martín-Juan[c], J. Corral[a,b], R. García-Carbonero[a,b], A. Carnero[a], L. Paz-Ares[a,b,*]

[a]Instituto de Biomedicina de Sevilla (IBIS), (HUVR, CSIC, Universidad de Sevilla), Sevilla, Spain
[b]Medical Oncology Department, Hospital Universitario Virgen del Rocio, Sevilla, Spain
[c]Pneumology Department, Hospital Universitario Virgen del Rocio, Sevilla, Spain
[d]Instituto de Ciências Biomédicas Abel Salazar (ICBAS, Universidade do Porto), Porto, Portugal

ARTICLE INFO

Article history:
Received 7 November 2012
Accepted 18 April 2013
Available online 9 May 2013

Keywords:
Lung cancer
COPD
Proteomic
Biomarker
Inflammation
Bronchoalveolar lavage

ABSTRACT

Lung cancer (LC) and chronic obstructive pulmonary disease (COPD) commonly coexist in smokers, and the presence of COPD increases the risk of developing LC. The aim of this study was to identify distinct proteomic profiles able to discriminate these two pathological entities. Protein content was assessed in the bronchoalveolar lavage (BAL) of 60 patients classified in four groups: COPD, COPD and LC, LC without COPD, and control with neither COPD nor LC. Proteins were separated into spots by bidimensional polyacrylamide gel electrophoresis (2D-PAGE) and examined by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF/TOF). A total of 40 proteins were differentially expressed in the LC and/or COPD groups as compared with the control group. Distinct protein profiles were identified and validated for each pathological entity (LC and COPD). The main networks involved were related to inflammatory signalling, free radical scavenging and oxidative stress response, and glycolysis and gluconeogenesis pathways. The most relevant signalling link between LC and COPD was through the NF-κB pathway.

In conclusion, the protein profiles identified contribute to elucidate the underlying pathogenic pathways of both diseases, and provide new tools of potential use as biomarkers for the early diagnosis of LC.

Abbreviations: 2D-PAGE, Two dimensional electrophoresis; AKR1B10, Aldo-keto reductase family 1, member B10; AKR1C3, Aldo-keto reductase family 1, member C3; ALDH3A1, Aldehyde dehydrogenase 3 family, member A1; ALDOA, Aldolase A; AMY1A, Alpha 1 amylase; AMY2A, Alpha 2 amylase; ANXA1, Annexin A1; ANXA2, Annexin A2; ANXA5, Annexin A5; ARHGDIB, Rho GDP dissociation inhibitor beta; BAL, Bronchoalveolar lavage fluid; C3A, Complement C3; CA1, Carbonic anhydrase 1; CAPS, Calcyphosine; CAT, Catalase; CFL1, Cofilin 1; COPD, Chronic obstructive pulmonary disease; CRP, C-reactive protein; CTSD, Cathepsin D; ENO1, Alpha enolase; EZR, Ezrin; FBP1, Fructose-1,6-bisphosphatase 1; GSR, Glutathione reductase; GSTA1, Glutathione S-transferase alpha 1; GSTA2, Glutathione S-transferase alpha 2; GSTP, Glutathione S-transferase pi 1; HSP70, Heat shock protein 70; IDH1, Isocitrate dehydrogenase 1; IPA, Ingenuity Pathways Analysis; IκB, Inhibitor of κB; LC, Lung cancer; LCN2, Lipocalin 2; MALDI-TOF/TOF, Matrix-assisted laser desorption/ionization time of flight-time of flight mass spectrometry; MMP, Matrix metalloproteinase; NF-κB, Nuclear factor κB; NSCLC, Non-small cell lung cancer; PEBP4, Phosphatidylethanolamine-binding protein 4; PKM2, Pyruvate kinase 2; PPIA, Peptidylprolyl isomerase A (cyclophilin A); PRDX1, Peroxiredoxin 1; PRDX2, Peroxiredoxin 2; PRDX5, Peroxiredoxin 5; PYGM, Glycogen phosphorylase; SCC, Squamous cell carcinoma; SELENBP1, Selenium binding protein 1; SERPINB1, Serpin peptidase inhibitor, clade B, member 1 (leukocyte elastase inhibitor); TKT, Transketolase; TPPP3, Tubulin polymerization-promoting protein family member 3 (CGI-38); TXN, Thioredoxin

* *Corresponding author at:* Instituto de Biomedicina de Sevilla (IBIS), 2nd Floor, Laboratory # 215, Avenida Manuel Siurot, s/n, 41013 Seville, Spain. Tel.: +34 955 013414; fax: +34 954 232992.
E-mail address: lpazares@hotmail.com (L. Paz-Ares).

[1] The first two authors contributed equally to this work.

1874-3919/$ – see front matter
http://dx.doi.org/10.1016/j.jprot.2013.04.037

Biological significance
Sequence coverage. The protein sequence coverage (95%) was estimated for specific proteins by the percentage of matching amino acids from the identified peptides having confidence greater than or equal to 95% divided by the total number of amino acids in the sequence.
Ingenuity Pathways Analysis. Mapping of our proteins onto biological pathways and disease networks demonstrated that 22 proteins were linked to inflammatory signalling (p-value: $1.35 * 10^{-08}$–$1.42 * 10^{-02}$), 15 proteins were associated with free radical scavenging and oxidative stress response (p-value: $4.93 * 10^{-11}$–$1.27 * 10^{-02}$), and 9 proteins were related with glycolysis and gluconeogenesis pathways (p-value: $7.39 * 10^{-09}$–$1.58 * 10^{-02}$).

1. Introduction

Lung cancer (LC) is the leading cause of cancer death worldwide, with a 5-year survival rate of 15%, which is among the lowest of all cancers [1]. The most important risk factor for the development of lung cancer is the smoking of tobacco, with smokers presenting a 10-fold higher risk to develop LC, on average, than non-smokers. Cigarette smoking is also the most important risk factor for chronic obstructive pulmonary disease (COPD). Interestingly, COPD, which is characterized by a chronic inflammation of the lower respiratory tract, is also a major independent risk factor for the development of lung carcinoma among long-term smokers. In fact, the presence of COPD increases the risk of LC up to 4–5 fold [2]. Furthermore, 50–70% of patients diagnosed with LC have spirometric evidence of COPD [3]. Lung cancer and COPD share risk factors and may therefore also likely share similar pathogenic pathways. Numerous mechanisms have been proposed for the common pathogenesis of LC and COPD, which involve chronic inflammation, epithelial-to-mesenchymal transition, oxidative stress, altered DNA repair and cellular proliferation [3–5]. Understanding the common mechanisms of these lung diseases is critical for the development of new methods of prevention, diagnosis and treatment.

Bronchoalveolar lavage (BAL) is a unique biological fluid that allows for appropriate sampling of the soluble protein content of the airway lumen. Comparative analyses have documented that certain proteins are present at higher levels in BAL than in plasma, suggesting that they are specifically produced in the airway tract [6]. However the plasma could be used as a source of biomarkers for early lung cancer diagnosis, blood has the great advantage of being readily accessible and that its collection is minimally invasive [7].

These proteins are therefore good candidates for becoming lung-specific biomarkers [8]. Proteomic technologies such as 2D-PAGE, 2D-DIGE, SELDI-TOF-MS technology, ICAT and iTRAQ, are valuable tools to expand the understanding of these processes. These technologies have been used to analyse different biological fluids including cell lysates, cell secretome, serum, plasma, tumour tissue, BAL, sputum and saliva [9,10]. Several studies have assessed differential BAL protein profiles obtained by 2D-PAGE analysis in smokers versus never smokers [11,12] and have also tried to identify profiles associated with a higher risk of developing COPD [13–15]. Regarding LC patients, a number of 2D-PAGE studies have been performed to explore the protein content of plasma or tumour tissue, focusing on understanding the molecular basis of cancer pathogenesis [16–18], as well as on the identification of new diagnostic, prognostic or predictive markers in lung cancer [19,20]. However, to the best of our knowledge, no study has evaluated to date the protein profile by 2D-PAGE in BAL of lung cancer patients. We believe that analysing the protein composition of BAL by a high-throughput technology, given its vicinity to tumour cells and potential enrichment in tumour-derived proteins may provide more reliable information and may therefore prove to be a more accurate tool for clinical use.

In the present study, the soluble protein content of the airway tract was assessed by 2D-PAGE and MALDI-TOF/TOF in the BAL of 60 patients classified in four groups: COPD, COPD and LC, LC without COPD, and control (neither COPD nor LC). As COPD and LC may share some common pathogenic pathways, the aim of this study was to identify distinct proteomic profiles that were able to discriminate these two pathological entities. Analysis of the main protein networks involved could also contribute to improve our understanding of the underlying mechanisms implicated in the development of each disease, and lead to the discovery of new biomarkers of potential diagnostic or therapeutic utility.

2. Materials and methods

2.1. *Patients and samples*

From 2009 to 2011, a total of 60 patients from the Hospital Universitario Virgen del Rocío were included in the study. Samples were obtained from four groups of patients: patients with COPD (n = 15), patients with LC (n = 15), patients with LC and COPD (n = 15), and patients with neither LC nor COPD (control group; n = 15). Characteristics of the study population are summarized in Table 1. Patients were required to meet all the following eligibility criteria: 1) patients were under pneumologist consultation due to haemoptysis and/or a pulmonary nodule; 2) patients were requested by their treating physician a flexible bronchoscopy for diagnostic purposes; 3) patients were smokers or ex-smokers of >20 pack-year; and 4) patients had to be older than 40 years of age. Exclusion criteria included: 1) diagnosis of a neoplastic disease other than non-small cell lung cancer (NSCLC); 2) active pulmonary tuberculosis; 3) previous lung resection; 4) history of drug abuse; and 5) the presence of other acute or chronic inflammatory disease. The study protocol was approved by the Hospital's Ethical Committee and a written informed consent was obtained from all patients prior to study entry.

Subjects were prepared with a combination of topical anaesthesia (20% benzocaine spray to the pharynx plus 2%

Table 1 – Characteristics of the patients.

	Controls n = 15	COPD n = 15	LC n = 15	LC + COPD n = 15
Gender				
Male	100.0% (15)	100.0% (15)	100.0% (15)	100.0% (15)
Female	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
Average age [range]	61.3 [41–80]	61.5 [45–78]	60.7 [46–69]	60.7 [49–68]
Smoking status				
Smokers	73.3% (11)	53.3% (8)	53.3% (8)	80.0% (12)
Ex-smokers	26.7% (4)	46.7% (7)	46.7% (7)	20.0% (3)
Packs-year	21.82	32.20	35.21	30.78
COPD				
Mild	–	20.0% (3)	–	53.3% (8)
Moderate	–	33.3% (5)	–	26.7% (4)
Severe	–	26.7% (4)	–	–
Very severe	–	20.0% (3)	–	20.0% (3)
LC histology				
ADC	–	–	73.3% (11)	66.7% (10)
SCC	–	–	26.7% (4)	33.3% (5)

topical lidocaine as needed) and conscious sedation using midazolam and meperidine according to institutional guidelines. All BAL samples were obtained by instillation and aspiration of 4 to 6 ml aliquots of 0.9% sterile saline in the appropriate bronchopulmonary segment. Recovered fluid was immediately passed through a 100 μm sterile nylon filter (Becton Dickinson, San Jose, CA) to remove mucus and transported on ice to the laboratory. The total volume was then centrifuged for 10 min at 1800 ×*g* and 4 °C. The supernatant was aliquoted into 2 ml tubes and frozen at –80 °C until further use.

2.2. BAL sample processing

A sample of approximately 4–8 ml of BAL was obtained from each patient for study purposes. Due to its low protein content, BAL samples had to be concentrated before use. BAL samples were thawed on ice with a Protease Inhibitor Cocktail kit (Thermo Scientific, Franklin, MA, USA), aliquoted into new tubes and placed on the vacuum concentrator (Concentrator plus — Eppendorf, Hamburg, Germany). Samples were then depleted of the two most abundant BAL proteins (albumin and immunoglobulin G) to allow an adequate identification of the less abundant ones. This was accomplished by the use of SpinTrap columns (GE Healthcare) following the manufacturer's instructions. The depleted BAL samples were then cleaned to remove contaminants, such as salts, thiols or denaturants that would interfere with the 2D-PAGE protocol. A 2-D Clean-up kit (GE Healthcare) was used according to the manufacturer's instructions with a rehydration solution containing urea (7 M), thiourea (2 M) and CHAPS (2%). Protein quantitation was assessed by the RC–DC method (Bio-Rad, Hercules, CA, USA).

2.3. 2D-PAGE

A mixture of 75 μg of protein from each individual BAL sample, DeStreak rehydration solution (GE Healthcare) and 0.5% of 3–10 NL pH IPG Buffer (GE Healthcare), in a final volume of 125 μl, were submitted to isoelectric focusing (IEF) in 7 cm IPG DryStrips (GE Healthcare) with a 3–10 NL pH range. Subsequent to IEF, strips were equilibrated and placed on 12.5% acrylamide gels to allow protein separation by electrophoresis. After this second dimension, gels were SYPRO® stained according to the manufacturer's instructions (Bio-Rad, Hercules, CA, USA). Typhoon TRIO (Amersham Biosciences, USA) was used to scan gel images.

2.4. Gel image analysis

Two-dimensional gel image analysis, protein spot detection, spot matching and semi-quantitative statistical analysis were performed using the Progenesis SameSpots software (Nonlinear Dynamics, Durham, NC). For each studied group, four different gel images were analysed and a corresponding reference synthetic image was obtained. After computer matching, detected spots and spot matches were manually edited for improved accuracy. The detection of differentially expressed protein spots was performed using the test INCA volume and proteins that had >2-fold differential expression were considered significant. Protein spots of interest were excised from the stained gel using a ProXcision robot (PerkinElmer) and sent for MS analysis.

2.5. Protein identification by mass spectrometry

MS experiments were performed by the Proteomics platform of the Cordoba University Proteomics Service (Cordoba, Spain). Protein spots of interest were washed with water and then trypsin digestion was performed. Tryptic peptides were mixed with a CCA matrix solution. The mixture was analysed with a Voyager System DE-STR 7307 MALDI-TOF/TOF Mass Spectrometer (ABI) to obtain a peptide mass fingerprint (PMF). Peptide matching and protein searches against the Swiss-Prot database were done using the Mascot search engine (http://www.matrixscience.com/) with a mass tolerance of ±50 ppm. Candidate proteins were selected from each spot, taking into consideration several variables such as isoelectric point, molecular mass, matched peptides, and sequence coverage.

2.6. Functional analysis of identified proteins

The protein sequence coverage (95%) was estimated for specific proteins by the percentage of matching amino acids from the identified peptides having confidence greater than or equal to 95% divided by the total number of amino acids in the sequence. The experimental molecular weight and isoelectric point of each identified protein were determined based on the location of the original spot on the 2D-PAGE gel using the Progenesis software.

2.7. Validation by western blotting

BAL proteins (30 μg) were separated in 12% gels polyacrylamide (SDS-PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes (Bio-Rad, Hercules, CA, USA). After blocking, the blots were incubated overnight at 4 °C with the following selected primary antibodies according to the manufacturer's

instructions: anti-HSP 70, AKR1B10, PRDX1 (1:1000, EPITOMICS, CA, USA), anti-PKM, anti-NF-κBp$_{65}$, and anti-pNF-κBp$_{65}$ (1:1000 Cell Signaling, Beverly, MA, USA). Secondary antibodies, peroxidase-conjugated anti-mouse (GE Healthcare, Uppsala, Sweden) and anti-rabbit (Cell Signaling, Beverly, MA, USA), were then applied to the individual membranes (1:2000) for 1 h at room temperature. Protein bands were revealed using enhanced chemiluminescence ECL (GE Healthcare, Uppsala, Sweden) and visualized in an image analyser (Mini LAS-3000, Fujifilm, Tokyo, Japan). The relative protein levels were calculated by comparison to the amount of the control protein β-actin (1:1000 Abcam, Cambridge, MA, USA). Analysis of the expression values obtained by western blot of the proteins of interest was performed by densitometry, using Image J software (http://rsbweb.nih.gov/ij/), and the results were expressed as fold change relative to the control protein (β-actin). Each experiment was independently repeated three times.

2.8. Bioinformatics

The biological functions, in terms of gene ontology and interaction network, were analysed using Ingenuity Pathways Analysis (IPA, version 7.1; http://www.ingenuity.com). Based on the local networks created by computational algorithms, identified proteins were connected with hub proteins, forming a functional protein cluster.

3. Results

3.1. Distinct proteomic profiles of LC and COPD

A total of 123 protein spots were successfully identified. Of these, the expression of 40 proteins had significant differences (≥2-fold, $p < 0.05$) between LC and/or COPD groups and the control group. All spots are marked on a representative gel 2D image in (Fig. 1). Detailed information of these identified proteins is summarized in Table 2.

Venn diagrams were used to display the differentially expressed proteins in each pathological subgroup studied. Among patients belonging to the LC group, 6 proteins were found to be up-regulated (CTSD, ALDO A, FBP1, ERZ, AKR1B10, TKT) and 1 down-regulated (SELENBP1) as compared to those in the control group, that were not found to be differentially expressed in patients belonging to the COPD or COPD/LC groups (Fig. 2). On the other hand, the LC and LC/COPD groups shared 5 proteins (ALDH3A1, AKR1C3, PKM2, PYGM and PPIA) that were significantly up-regulated compared to the control group, that were not deregulated in the COPD group (Fig. 2). Similarly, the COPD and COPD/LC groups shared 4 up-regulated proteins (CAT, PRDX1, PRDX2 and PRDX5) in contrast with the LC group (Fig. 2). Finally, the three pathological groups shared 24 deregulated proteins, 17 of them up-regulated (HSP70, TXN, ENO1, AMY1A, AMY2A, ANXA1, ANXA2, ANXA5, CA1, CRP, TPPP3, C3A, GSR, IDH1, SERPINB1, PEBP4, and ARHGDIB), 5 down-regulated (GSTA1, GSTA2, GSTP, LCN2, and UCHL1), and 2 proteins (CAPS 2 and CFL1) down-regulated in the COPD group and up-regulated in the LC and LC/COPD groups (Fig. 2) as compared to the control group.

3.2. Validation of selected proteins

Four proteins with high fold changes (PKM2, HSP70, AKR1B10 and PRDX1) representative of each study subgroup were selected for validation by western blot. These proteins showed a high search score, which is a probability based on the observed frequency of peptides in a given molecular weight range and is used to determine the quality of a given match, with higher scores indicating greater confidence. A close-up of the region of the gels showing differentially expressed proteins between LC and/or COPD and control groups is depicted in Fig. 3A.

Western blotting of PKM2, HSP70, AKR1B10, PRDX1 and the corresponding β-actin are shown in Fig. 3B. Validation was performed in three random samples from each study group. Results indicate that HSP70 expression is increased in all 3 pathological study groups (COPD, LC, LC/COPD) as compared with the control group. The PKM2 protein expression is clearly increased only in patients with LC, with or without COPD, but not in the COPD (without LC) or control groups. On the other hand, PDX1 shows an increment of expression in COPD groups (±LC), as compared to the LC (without COPD) and control groups. Finally, ARK1B10 shows greater expression only in the LC group (without COPD) (Fig. 3B). These results confirmed the data obtained in the 2D-PAGE analysis, and illustrate that there are protein profiles that are common to COPD and LC, whereas others are able to discriminate among both pathological entities.

3.3. Bioinformatic analysis

Human homologues of the 40 unique proteins identified in this study were subsequently analysed with IPA, software capable of mapping proteins onto known existing networks and pathways. Cellular compartments as designated by IPA (gene ontology based) indicated that the majority of identified proteins were of cytoplasmatic origin. Mapping of our proteins onto biological pathways and disease networks demonstrated that 22 proteins were linked to inflammatory signalling (p-value: $1.35 * 10^{-08}$–$1.42 * 10^{-02}$), 15 proteins were associated with free radical scavenging and oxidative stress response (p-value: $4.93 * 10^{-11}$–$1.27 * 10^{-02}$), and 9 proteins were related with glycolysis and gluconeogenesis pathways (p-value: $7.39 * 10^{-09}$–$1.58 * 10^{-02}$). The complex NF-κB appeared as a hallmark in the identified IPA networks, although it was not detected in the 2D-PAGE study (Fig. 4A).

In order to confirm the presence of the NF-κB complex in the BAL of our study subjects, NFkBp65 and pNFkBp65 expressions were assessed by western blot in the study samples as described before (Fig. 4B). The results show that NFkBp65 is expressed in all study groups including controls, with slightly higher expression in the COPD group. However, the phosphorylated NFkBp65 showed increased expression only in pathological groups (COPD, LC, COPD/LC), particularly in the LC group. The result suggests that the activation of NFkBp65 is involved in the pathogenesis of both diseases (COPD and LC).

4. Discussion

We present the first shotgun 2D-PAGE proteomic study of lung cancer in BAL fluid where we have established specific protein

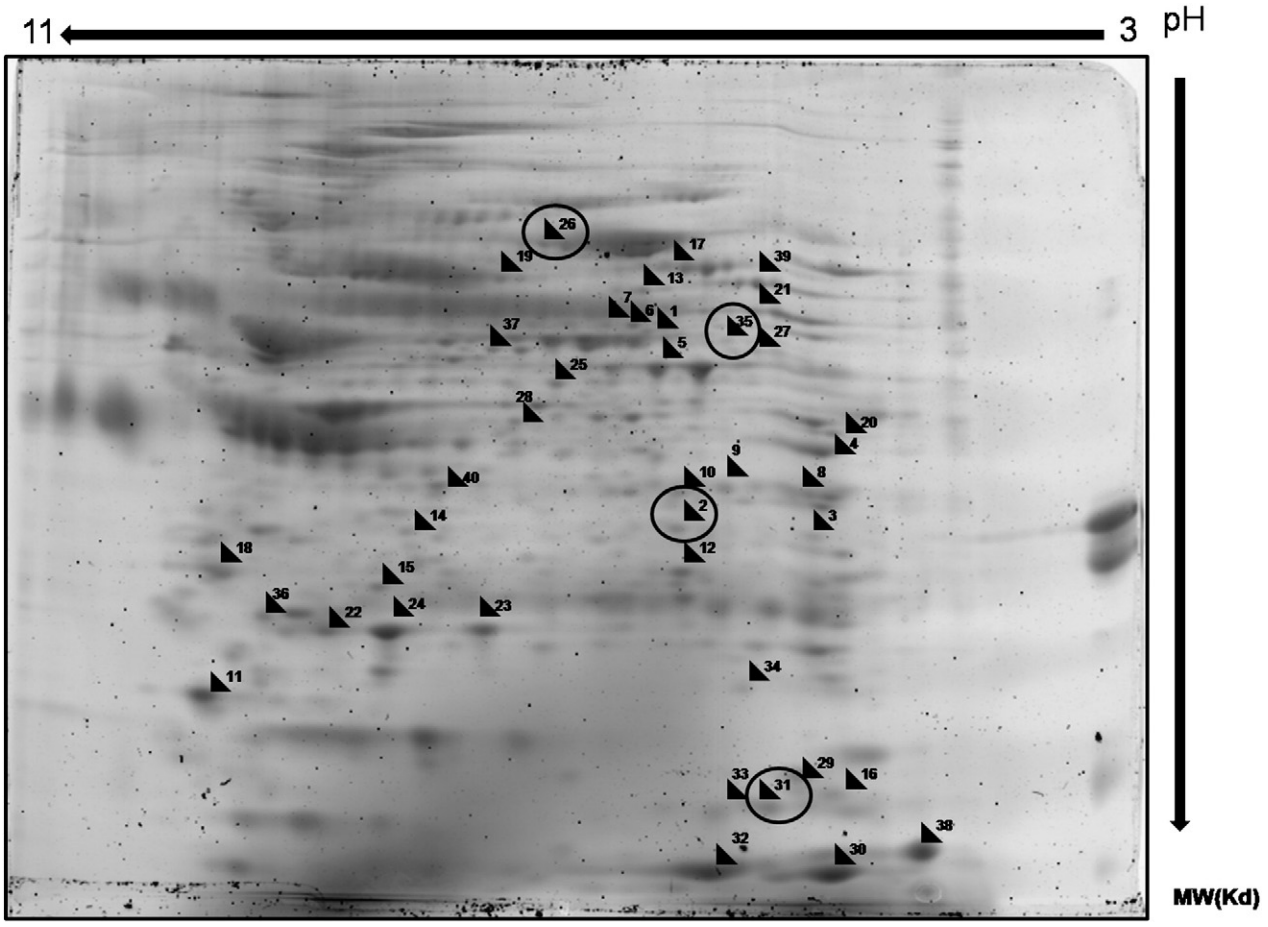


Fig. 1 – 2D-PAGE of a representative patient sample. (A) A representative 7 cm Sypro® stained gel of proteins in the pH range 3–11 non-lineal. The spots represented with a circle correspond to the selected proteins for validation.

profile characteristics of LC and COPD. Our results indicate that the protein composition of BAL shows significant differences among the four groups of studied individuals (LC, COPD, LC & COPD and control), especially when comparisons are made between the control group and the disease groups. These alterations were observed in 40 differentially expressed proteins, which showed changes in expression ≥2 fold ($p < 0.05$). Most proteins were up-regulated in LC and COPD, reflecting a common link. We have validated the bidimensional proteomic results for some proteins by western blotting. Further validation of the described proteomic signature independent cohorts is warranted.

Some of the identified proteins with differential expression appeared only in the LC and LC/COPD groups (Fig. 2). For example, CTSD, Ezrin, and SELENBP1 have been associated with increased cancer growth, invasion, and metastasis in various types of cancers [21–23]. Indeed, different studies have highlighted the potential role of these proteins as prognostic markers of progression in lung cancer [24–26]. ALDH3A1 belongs to the phase II group of drug-metabolizing enzymes and is highly expressed in the stomach, lung, keratinocytes, and cornea. Cytosolic ALDH3A1, which is induced by polycyclic aromatic hydrocarbons or chlorinated compounds, plays an important role in alveolar pneumocyte physiology, and increases during carcinogenesis [27]. Over-expression of AKR1B10, is associated, in most cases with smoking [28,29], suggesting a possible involvement of this enzyme in tobacco-induced lung cancer. Indeed, AKR1B10 has a high catalytic efficiency for the reduction of retinoids, and retinoic acid deficiency has been linked to airway epithelial squamous metaplasia and epithelial-to-mesenchymal transition [30]. Thus, increased airway epithelial expression of AKR1B10 may play an important role in the pathogenesis of lung cancer [31]. In addition, AKR1B10 has also been involved in the resistance to different chemotherapeutic drugs, such as cisplatin [32].

Another group of identified proteins up-regulated in the LC and/or LC/COPD groups in our comparative proteomic study included TKT, FBP1, ALDOA, PYGM and PKM2. All these proteins are involved in the regulation of glycolysis, probably contributing to the Warburg effect, that leads to a state that has been termed "aerobic glycolysis". Previous studies showed the role of PKM2 in the production of lactic acid through glycolysis rather than producing energy through mitochondrial oxidative phosphorylation. This may help tumour cells to survive in low glucose and low oxygen environments, and facilitate tumour invasion [33,34].

The largest number of proteins found in our profiles appears commonly up-regulated in three pathological groups (LC, COPD and LC/COPD) compared with control group. The expression of one of these proteins, HSP70, in NSCLC has been linked to the number of daily smoked cigarettes [35]. HSP70 has also been

Table 2 – Protein spots searched by MASCOT software in database.

Spots no.	Protein name	Symbol	Accession no.	Protein MW	Protein PI	Pep. count	Protein score	Score C.I. %	Total ion score	Ion C.I. %	COPD		LC		LC/COPD	
1	Aldehyde dehydrogenase	ALDH3A1	gi|178375	50702.9	5.99	15	445	100	339	100	–		Up	2.9	Up	2.9
2	Aldo-keto reductase family 1 member B10	AKR1B10	gi|4503285	37111.1	7.13	7	108	100	70	100	–		Up	4	–	
3	Aldo-keto reductase family 1 member C3	AKR1C3	gi|24497583	37229	8.06	10	248	100	186	100	–		Up	2.5	Up	2.1
4	Aldolase A	ALDOA	gi|28614	39706.4	8.34	11	229	100	156	100	–		Up	2	–	
5	Alpha enolase	ENO	gi|2661039	36627.7	6.53	9	128	100	70	100	Up	2	Up	2.5	Up	2.3
6	Alpha-amylase 1 precursor	AMY1A	gi|40254482	58415.2	6.47	18	469	100	326	100	Up	2.1	Up	2.3	Up	2.2
7	Amylase alpha	AMY2A	gi|224980	58443.2	6.47	15	323	100	221	100	Up	2.5	Up	2.2	Up	2.9
8	Annexin A1	ANX1	gi|4502101	38918.1	6.57	17	495	100	354	100	Up	2.3	Up	2.2	Up	2.3
9	Annexin A2	ANX2	gi|18645167	38779.9	7.57	20	326	100	159	100	Up	2.5	Up	2.3	Up	2.7
10	Annexin A5	ANX5	gi|4502107	35971.4	4.94	19	472	100	301	100	Up	2.4	Up	2.2	Up	2.1
11	Calcyphosin 2	CAPS 2	gi|4757908	21068.3	4.74	11	543	100	460	100	Down	2.2	Up	2.8	Up	2.6
12	Carbonic anhydrase 1	CA1	gi|4502517	28909.4	6.59	14	770	100	627	100	Up	2.2	Up	2.5	Up	2.7
13	Catalase	CAT	gi|4557014	59946.8	6.9	13	310	100	234	100	Up	2.5	–		Up	2.8
14	Cathepsin D preproprotein	CTSD	gi|4503143	45036.8	6.1	8	135	100	95	100	–		Up	3	–	
15	CGI-38 protein	TPPP3	gi|4680715	19144.7	9.1	8	79	100	30	88	Up	2.4	Up	2	Up	2.1
16	Cofilin-1	CFL1	gi|5031635	18718.7	8.22	6	98	100	57	100	Down	3	Up	2.2	Up	2.8
17	Complement C3	C3A	gi|119589476	144417.1	8.24	25	472	100	332	100	Up	2.4	Up	3	Up	2.6
18	C-reactive protein	CRP	gi|1942435	23146.7	5.28	8	341	100	281	100	Up	3	Up	3.2	Up	2.7
19	Ezrin	EZR	gi|46249758	69312.7	5.94	18	250	100	175	100	–		Up	3	–	
20	Fructose-1, 6-bisphosphatase	FBP1	gi|3293553	37058.9	6.6	6	124	100	101	100	–		Up	2.5	–	
21	Glutathione reductase	GSR	gi|119583848	61464.6	8.71	7	131	100	108	100	Up	2.8	Up	3.1	Up	2
22	Glutathione S-transferase A1 subunit	GSTA1	gi|163310943	25628.7	8.72	15	384	100	268	100	Down	3	Down	3.2	Down	3.8
23	Glutathione S-transferase A2 subunit	GSTA2	gi|257476	25589.6	8.81	7	105	100	70	100	Down	2.5	Down	3	Down	2.6
24	Glutathione S-transferase P	GSTP1	gi|4504183	23569.1	5.43	10	633	100	541	100	Down	2.2	Down	2.4	Down	2.5
25	Glycogen phosphorylase	PYGM	gi|225897	97386.8	6.57	17	515	100	436	100	–		Up	3	Up	2.1
26	HSP70	HSP 70	gi|4529893	70280.1	5.48	12	325	100	269	100	Up	4	Up	3.1	Up	3.8
27	Isocitrate dehydrogenase 1	IDH1	gi|89573979	42091	6.19	8	62	100	29	85	Up	2.2	Up	2.1	Up	2.5
28	Leukocyte elastase inhibitor	SERPIN B1	gi|13489087	42828.7	5.9	18	403	100	274	100	Up	2.7	Up	2.6	Up	2.5
29	Lipocalin 2	LCN2	gi|119608155	28089.4	9.53	9	513	100	426	100	Down	3	Down	3.5	Down	2.9
30	Peptidylprolyl isomerase A (cyclophilin A)	PPIA	gi|1633054	18097.9	7.82	10	260	100	159	100	–		Up	2.2	Up	2.6
31	Peroxiredoxin 1	PRDX 1	gi|55959887	19134.7	6.41	8	170	100	98	100	Up	5	–		Up	4.2
32	Peroxiredoxin 5	PRDX 5	gi|6166493	22261.6	8.85	11	638	100	537	100	Up	2.3			Up	2.4
33	Peroxiredoxin-2 isoform a	PRDX 2	gi|32189392	22049.3	5.66	12	451	100	325	100	Up	3	–		Up	2.9
34	Phosphatidylethanolamine-binding protein 4	PEBP4	gi|4505621	21157.7	7.01	11	311	100	209	100	Up	2.9	Up	2.3	Up	2.6
35	Pyruvate kinase 2	PKM2	gi|35505	58411.2	7.58	22	416	100	249	100	–		Up	3	Up	3.2
36	Rho GDP-dissociation inhibitor 2	ARHGDIB	gi|56676393	23030.6	5.1	7	215	100	170	100	Up	2.9	Up	2.6	Up	2.3
37	Selenium-binding protein 1	SELENBP1	gi|197097476	52987.9	6.13	12	278	100	207	100	–		Down	2.2	–	
38	Thioredoxin	TXN		12345	7.93	10	241	100	203	100	Up	2.4	Up	2.3	Up	2.1
39	Transketolase	TKT	gi|388891	68528	7.89	17	405	100	246	100	–		Up	2.5	–	
40	Ubiquitin carboxyl-terminal hydrolase isozyme L1	UCHL1	gi|21361091	25150.6	5.33	7	126	100	76	100	Down	2.7	Down	2.2	Down	2.8

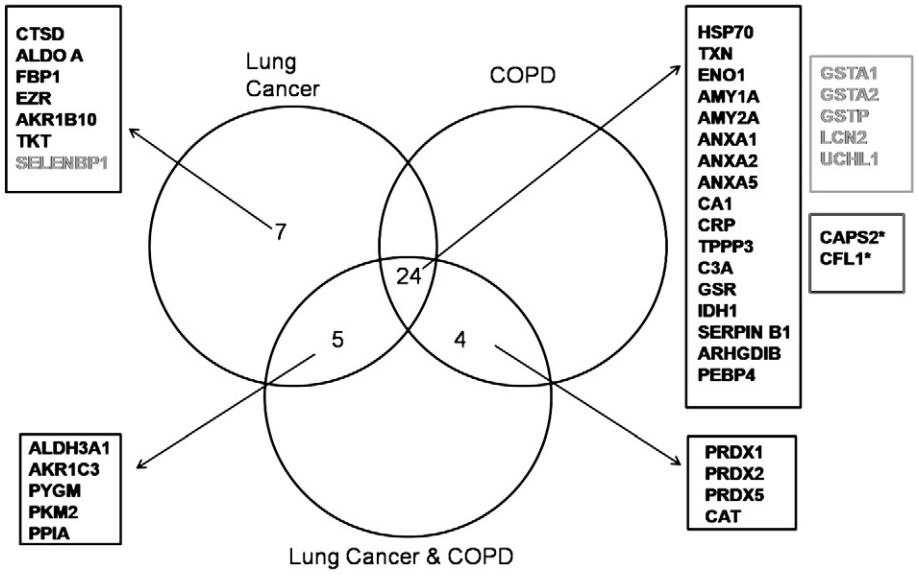


Fig. 2 – A Venn diagram showing the overlaps of up- or down-regulated proteins in each pathological group. Up-regulated proteins are represented in black. Down-regulated proteins are represented in grey. In black and with an asterisk are the proteins up-regulated in LC and down-regulated COPD.

studied in various inflammatory diseases as COPD, where it was significantly elevated in serum samples of patients suffering from COPD compared to controls [36]. Different proteomic approaches have confirmed the correlation between HSP70 over-expression and the differentiation level and/or aggressiveness of various types of cancer, such as gastric adenocarcinomas [37], hepatocarcinoma [38] and oesophageal cancer [39]. In addition, various oncoproteomics studies correlate the elevated levels of HSP70 with therapeutic resistance [40–43]. Other interesting protein is carbonic anhydrase 1 (CA1). This protein constitutes an endogenous marker of cellular hypoxia, a natural phenotype of solid tumours, and its predictive and prognostic potential has been demonstrated in various clinical studies [44]. Up-regulated expression of CA1 was reported in carcinoma cells derived from various organs including breast, cervix, bladder, oesophagus, lung, kidney, colon and rectum, head and neck [45]. IDH1 is an enzyme essential for cell metabolism and energy production. Several studies suggest the potential oncogenicity of IDH1 in acute myeloid leukaemia [46]. A recent report on IDH1 showed that it is up-regulated in more than 70% of NSCLC tumours and that it is correlated with lower 5-year survival rates. Furthermore, IDH1 was proved to be an independent unfavourable prognostic factor for overall survival of NSCLC patients by multivariate analysis [47].

Recent studies have described the role of the protein PEBP4 in the inhibition of the MAPK signalling pathway and its involvement in the inhibition of the JNK pathway that promotes the activation of AKT [48]. It has also been shown that the over-expression of PEBP4 was related to the development, invasion, and metastasis of a variety of tumours [49,50].

Finally, four proteins appeared up-regulated in the COPD and LC/COPD groups. These proteins are enzymatic antioxidants, and represent the first line of defence against oxidative stress. Studies evaluating the involvement of these genes on COPD patients and healthy smokers have revealed conflicting results; some studies have claimed for a down-regulation of enzymatic antioxidants such as peroxiredoxins and catalase in COPD patients [51–53], while others stated their up-regulation in this population [54]. Studies in LC have showed that up-regulation of different members of the peroxiredoxin family correlates with poor prognosis [55–57].

The proteomic signatures found in our study were analysed by Ingenuity System Pathway Analysis software, in order to identify any common pathways or relationships between proteins within the data. The three major pathways associated with the diseases involved within the data were inflammation, free radical scavenging and oxidative stress response, and glycolysis and gluconeogenesis. The results indicated NF-κB, an important transcription factor involved in inflammation, as the main link between the three pathways. This fact contributes to demonstrate the important role of inflammation as a link between the development of lung cancer and COPD. NF-κB regulated genes include cytokines, adhesion molecules, angiogenic factors, anti-apoptotic factors, and matrix metalloproteinases (MMPs) which are involved in the different steps of carcinogenesis. It has been suggested that NF-κB promotes lung cancer mainly through mediating inflammatory cytokine secretion to establish a cancer-prone inflammatory microenvironment [58]. Similarly, NF-κB pathways play a crucial role in the pathogenesis/development of COPD by increasing the release of pro-inflammatory mediators leading to chronic inflammation in the lung. In bronchial biopsies of airway mucosa in patients with COPD, protein expression of the p65 subunit of NF-κB was increased compared with expression in non-smokers, and correlated with airflow obstruction [59]. In addition, both chemotherapy and radiation treatments induce NF-κB activation in cancer cells, which contributes to resistance to those same therapies [60,61]. Indeed, inhibition of NF-κB signalling by various approaches has been shown to augment the efficacy of chemotherapeutics and radiation in killing cancer cells in vitro and in

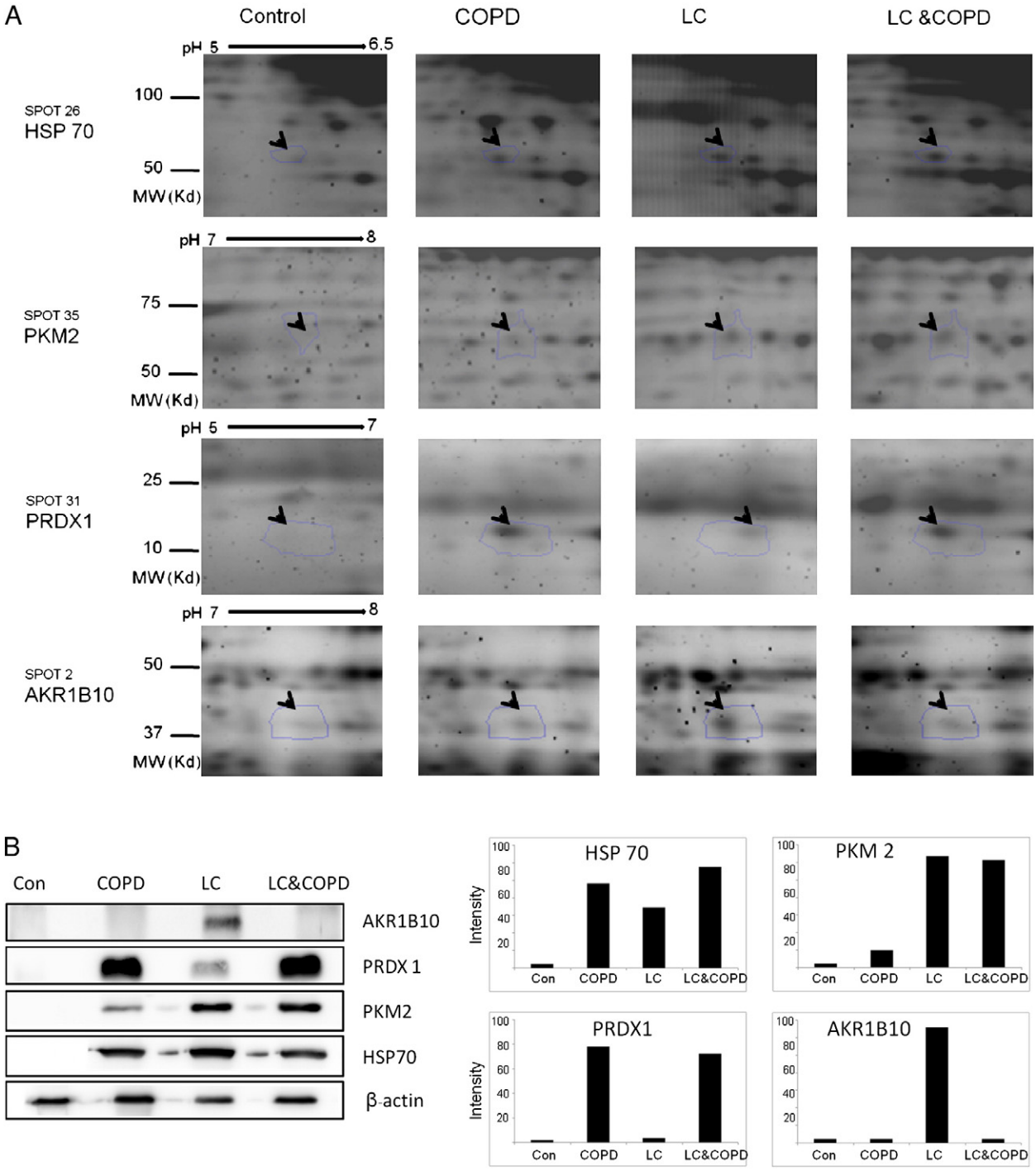


Fig. 3 – (A). The panel shows the different densities of spots 2, 35, 26 and 31 (in Fig. 1) in each study group, which correspond to proteins ARK1B10, PKM2, HSP70 and PDRX1, respectively. (B) Western blotting for ARK1B10, PDX1, PKM2 and HSP70. The different expressions are seen in each of the groups for each of the selected proteins. ARK1B10 is only expressed in the LC group, PDX1 presents an increment of expression in COPD and LC & COPD groups, PKM2 expression is increased in LC and LC & COPD groups, and HSP70 presents a similar increment of expression in all pathological groups. These differences are also illustrated on the right panel with a bar chart.

vivo [62,63]. Some of the NF-κB inhibitors that enhanced lung cancer cell death induced by chemotherapeutics are genistein [64] with cisplatin or docetaxel, embelin with paclitaxel [65], expression of IkBα mutant with cisplatin, gemcitabline, adriamicin [66] and etoposide [67].

5. Conclusions

Our proteomic study in BAL suggests that lung cancer and COPD share some pathogenic pathways such as inflammation,

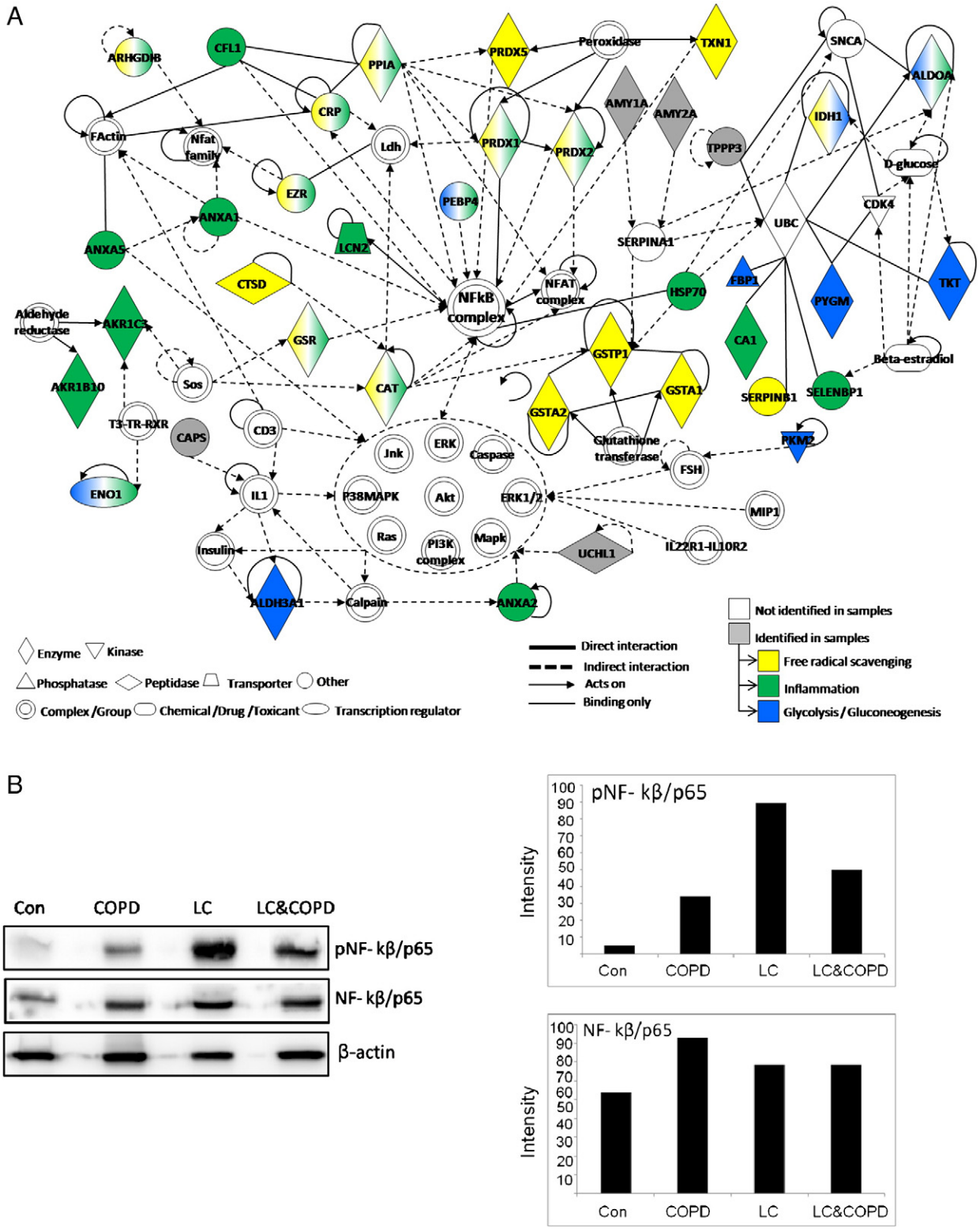


Fig. 4 – (A) Ingenuity Pathways Analysis of lung cancer and/or COPD versus controls revealed three networks that were statistically significant based on the number of differentially expressed proteins. Network 1 revealed processes associated with inflammatory signalling and family members are represented in green colour. Network 2 revealed processes associated with free radical scavenging and oxidative stress response and family members are represented in yellow colour. And Network 3 revealed processes associated with glycolysis and gluconeogenesis pathways and family members are represented in blue colour. (B) Western blotting for the NF-κB activation. Increased expression was seen in all pathological groups, particularly in the LC group. These differences are also illustrated on the right panel with a bar chart.

free radical scavenging and oxidative stress response, and glycolysis and gluconeogenesis. In fact, the pivotal networking signalling is through NF-κB. In addition, we have identified distinct proteomic profile characteristic of LC, COPD and LC & COPD. The proteins included in this signature deserve validation and warrant further study as potential biomarkers for early diagnosis and prognosis, and as therapeutic targets.

Acknowledgements

The present study was partially financed by the Fondo de Investigación Sanitaria (PI081156, PI1102688). MDP is funded by the Fondo de Investigación Sanitaria (CD 09/00148). AN is funded by the Portuguese Ministry of Science, Technology and Superior Education — FCT (Fundação para a Ciência e para a Tecnologia: SFRH/BD/48341/2008). SMP is funded by the Fondo de Investigación Sanitaria (CD 11/00153).

REFERENCES

[1] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin 2012;62:10–29.
[2] Punturieri A, Szabo E, Croxton TL, Shapiro SD, Dubinett SM. Lung cancer and chronic obstructive pulmonary disease: needs and opportunities for integrated research. J Natl Cancer Inst 2009;101:554–9.
[3] Yao H, Rahman I. Current concepts on the role of inflammation in COPD and lung cancer. Curr Opin Pharmacol 2009;9:375–83.
[4] Brody JS, Spira A. State of the art. Chronic obstructive pulmonary disease, inflammation, and lung cancer. Proc Am Thorac Soc 2006;3:535–7.
[5] Lee G, Walser TC, Dubinett SM. Chronic inflammation, chronic obstructive pulmonary disease, and lung cancer. Curr Opin Pulm Med 2009;15:303–7.
[6] Hermans C, Dong P, Robin M, Jadoul M, Bernard A, Bersten AD, et al. Determinants of serum levels of surfactant proteins A and B and Clara cell protein CC16. Biomarkers 2003;8:461–71.
[7] Indovina P, Marcelli E, Pentimalli F, Tanganelli P, Tarro G, Giordano A. Mass spectrometry-based proteomics: the road to lung cancer biomarker discovery. Mass Spectrom Rev 2013;32(2):129–42.
[8] Molina-Pinelo S, Suárez R, Pastor MD, Nogal A, Márquez-Martín E, Martín-Juan J, et al. Association between the miRNA signatures in plasma and bronchoalveolar fluid in respiratory pathologies. Dis Markers 2012;32:221–30.
[9] Yee J, Sadar MD, Sin DD, Kuzyk M, Xing L, Kondra J, et al. Connective tissue-activating peptide III: a novel blood biomarker for early lung cancer detection. J Clin Oncol 2009;27:2787–92.
[10] Ilie MI, Hofman V, Ortholan C, Ammadi RE, Bonnetaud C, Havet K, et al. High levels of carbonic anhydrase IX in tumour tissue and plasma are biomarkers of poor prognostic in patients with non-small cell lung cancer. Br J Cancer 2010;102:1627–35.
[11] Plymoth A, Yang Z, Löfdahl CG, Ekberg-Jansson A, Dahlbäck M, Fehniger TE, et al. Rapid proteome analysis of bronchoalveolar lavage samples of lifelong smokers and never-smokers by micro-scale liquid chromatography and mass spectrometry. Clin Chem 2006;52:671–9.
[12] Kelsen SG, Duan X, Ji R, Perez O, Liu C, Merali S. Cigarette smoke induces an unfolded protein response in the human lung: a proteomic approach. Am J Respir Cell Mol Biol 2008;38:541–50.
[13] Pinto-Plata V, Toso J, Lee K, Bilello J, Mullerova H, De Souza M, et al. Use of proteomic patterns of serum biomarkers in patients with chronic obstructive pulmonary disease: correlation with clinical parameters. Proc Am Thorac Soc 2006;3:465–6.
[14] Merkel D, Rist W, Seither P, Weith A, Lenter MC. Proteomic study of human bronchoalveolar lavage fluids from smokers with chronic obstructive pulmonary disease by combining surface-enhanced laser desorption/ionization-mass spectrometry profiling with mass spectrometric protein identification. Proteomics 2005;5:2972–80.
[15] Casado B, Iadarola P, Pannell LK, Luisetti M, Corsico A, Ansaldo E, et al. Protein expression in sputum of smokers and chronic obstructive pulmonary disease patients: a pilot study by CapLC-ESI-Q-TOF. J Proteome Res 2007;6:4615–23.
[16] Hoagland 4th LF, Campa MJ, Gottlin EB, Herndon 2nd JE, Patz Jr EF. Haptoglobin and posttranslational glycan-modified derivatives as serum biomarkers for the diagnosis of nonsmall cell lung cancer. Cancer 2007;110:2260–8.
[17] Jessie K, Pang WW, Haji Z, Rahim A, Hashim OH. Proteomic analysis of whole human saliva detects enhanced expression of interleukin-1 receptor antagonist, thioredoxin and lipocalin-1 in cigarette smokers compared to non-smokers. Int J Mol Sci 2010;11:4488–505.
[18] Brower V. Biomarker studies abound for early detection of lung cancer. J Natl Cancer Inst 2009;101:11–3.
[19] Kang SM, Sung HJ, Ahn JM, Park JY, Lee SY, Park CS, et al. The haptoglobin beta chain as a supportive biomarker for human lung cancers. Mol Biosyst 2011;7:1167–75.
[20] Tian T, Hao J, Xu A, Hao J, Luo C, Liu C, et al. Determination of metastasis-associated proteins in non-small cell lung cancer by comparative proteomic analysis. Cancer Sci 2007;98:1265–74.
[21] Garcia M, Derocq D, Pujol P, Rochefort H. Overexpression of transfected cathepsin D in transformed cells increases their malignant phenotype and metastatic potency. Oncogene 1990;5:1809–14.
[22] Kim H, Kang HJ, You KT, Kim SH, Lee KY, Kim TI, et al. Suppression of human selenium-binding protein 1 is a late event in colorectal carcinogenesis and is associated with poor survival. Proteomics 2006;6:3466–76.
[23] Zeng H, Xu L, Xiao D, Zhang H, Wu X, Zheng R, et al. Altered expression of ezrin in esophageal squamous cell carcinoma. J Histochem Cytochem 2006;54:889–96.
[24] Zhang XQ, Chen GP, Wu T, Yan JP, Zhou JY. Expression and clinical significance of ezrin in non-small-cell lung cancer. Clin Lung Cancer 2012;13:196–204.
[25] Chen G, Wang H, Miller CT, Thomas DG, Gharib TG, Misek DE, et al. Reduced selenium-binding protein 1 expression is associated with poor outcome in lung adenocarcinomas. J Pathol 2004;202:321–9.
[26] Wang Z, Zhao X. Expression and prognostic relation of cathepsin D in non-small cell lung cancer tissues and lymph nodes. Zhonghua Jie He He Hu Xi Za Zhi 1998;21:164–6.
[27] Muzio G, Maggiora M, Paiuzzi E, Oraldi M, Canuto RA. Aldehyde dehydrogenases and cell proliferation. Free Radic Biol Med 2012;52:735–46.
[28] Kang MW, Lee ES, Yoon SY, Jo J, Lee J, Kim HK, et al. AKR1B10 is associated with smoking and smoking-related non-small-cell lung cancer. J Int Med Res 2011;39:78–85.
[29] Li CP, Goto A, Watanabe A, Murata K, Ota S, Niki T, et al. AKR1B10 in usual interstitial pneumonia: expression in squamous metaplasia in association with smoking and lung cancer. Pathol Res Pract 2008;204:295–304.
[30] Tang XH, Gudas LJ. Retinoids, retinoic acid receptors, and cancer. Annu Rev Pathol 2011;6:345–64.
[31] Ruiz FX, Gallego O, Ardèvol A, Moro A, Domínguez M, Alvarez S, et al. Aldo-keto reductases from the AKR1B subfamily: retinoid specificity and control of cellular retinoic acid levels. Chem Biol Interact 2009;178:171–7.
[32] Matsunaga T, Wada Y, Endo S, Soda M, El-Kabbani O, Hara A. Aldo-keto reductase 1B10 and its role in proliferation capacity of drug-resistant cancers. Front Pharmacol 2012;3:5.

[33] Eigenbrodt E, Kallinowski F, Ott M, Mazurek S, Vaupel P. Pyruvate kinase and the interaction of amino acid and carbohydrate metabolism in solid tumors. Anticancer Res 1998;18:3267–74.
[34] Christofk HR, Vander Heiden MG, Harris MH, Ramanathan A, Gerszten RE, Wei R, et al. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. Nature 2008;452:230–3.
[35] Volm M, Mattern J, Stammler G. Up-regulation of heat shock protein 70 in adenocarcinomas of the lung in smokers. Anticancer Res 1995;15:2607–9.
[36] Hacker S, Lambers C, Hoetzenecker K, Pollreisz A, Aigner C, Lichtenauer M, et al. Elevated HSP27, HSP70 and HSP90 alpha in chronic obstructive pulmonary disease: markers for immune activation and tissue destruction. Clin Lab 2009;55:31–40.
[37] Yoshihara T, Kadota Y, Yoshimura Y, Tatano Y, Takeuchi N, Okitsu H, et al. Proteomic alteration in gastic adenocarcinomas from Japanese patients. Mol Cancer 2006;5:75.
[38] Lee IN, Chen CH, Sheu JC, Lee HS, Huang GT, Yu CY, et al. Identification of human hepatocellular carcinoma-related biomarkers by two-dimensional difference gel electrophoresis and mass spectrometry. J Proteome Res 2005;4:2062–9.
[39] Jazii FR, Najafi Z, Malekzadeh R, Conrads TP, Ziaee AA, Abnet C, et al. Identification of squamous cell carcinoma associated proteins by proteomics and loss of beta tropomyosin expression in esophageal cancer. World J Gastroenterol 2006;12:7104–12.
[40] Allal AS, Kähne T, Reverdin AK, Lippert H, Schlegel W, Reymond MA. Radioresistance-related proteins in rectal cancer. Proteomics 2004;4:2261–9.
[41] Smith L, Welham KJ, Watson MB, Drew PJ, Lind MJ, Cawkwell L. The proteomic analysis of cisplatin resistance in breast cancer cells. Oncol Res 2007;16:497–506.
[42] Castagna A, Antonioli P, Astner H, et al. A proteomic approach to cisplatin resistance in the cervix squamous cell carcinoma cell line A431. Proteomics 2004;4:3246–67.
[43] Bottoni P, Giardina B, Scatena R. Proteomic profiling of heat shock proteins: an emerging molecular approach with direct pathophysiological and clinical implications. Proteomics Clin Appl 2009;3:636–53.
[44] Potter CP, Harris AL. Diagnostic, prognostic and therapeutic implications of carbonic anhydrases in cancer. Br J Cancer 2003;89:2–7.
[45] Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nat Rev Cancer 2011;11:85–95.
[46] Chotirat S, Thongnoppakhun W, Promsuwicha O, Boonthimat C, Auewarakul CU. Molecular alterations of isocitrate dehydrogenase 1 and 2 (IDH1 and IDH2) metabolic genes and additional genetic mutations in newly diagnosed acute myeloid leukemia patients. J Hematol Oncol 2012;5:5.
[47] Tan F, Jiang Y, Sun N, Chen Z, Lv Y, Shao K, et al. Identification of isocitrate dehydrogenase 1 as a potential diagnostic and prognostic biomarker for non-small cell lung cancer by proteomic analysis.Mol Cell Proteomics 2012;11 [M111 008821].
[48] Yu GP, Chen GQ, Wu S, Shen K, Ji Y. The expression of PEBP4 protein in lung squamous cell carcinoma. Tumour Biol 2011;32:1257–63.
[49] Yu GP, Huang B, Chen GQ, Wu S, Ji Y, Shen ZY. PEBP4 gene expression and its significance in invasion and metastasis of non-small cell lung cancer. Tumour Biol 2012;33:223–8.
[50] Liu H, Kong Q, Li B, He Y, Li P, Jia B. Expression of PEBP4 protein correlates with the invasion and metastasis of colorectal cancer. Tumour Biol 2012;33:267–73.
[51] Comandini A, Marzano V, Curradi G, Federici G, Urbani A, Saltini C. Markers of anti-oxidant response in tobacco smoke exposed subjects: a data-mining review. Pulm Pharmacol Ther 2010;23:482–92.
[52] Alexandre BM, Charro N, Blonder J, Lopes C, Azevedo P, Bugalho de Almeida A, et al. Profiling the erythrocyte membrane proteome isolated from patients diagnosed with chronic obstructive pulmonary disease. J Proteomics 2012;76 Spec No:259–69. http://dx.doi.org/10.1016/j.jprot.2012.04.008 (Epub 2012 Apr 17).
[53] Bandow JE, Baker JD, Berth M, Painter C, Sepulveda OJ, Clark KA, et al. Improved image analysis workflow for 2-D gels enables large-scale 2-D gel-based proteomics studies—COPD biomarker discovery study. Proteomics 2008;8:3030–41.
[54] Pierrou S, Broberg P, O'Donnell RA, Pawłowski K, Virtala R, Lindqvist E, et al. Expression of genes involved in oxidative stress responses in airway epithelial cells of smokers with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2007;175:577–86.
[55] Kim JH, Bogner PN, Ramnath N, Park Y, Yu J, Park YM. Elevated peroxiredoxin 1, but not NF-E2-related factor 2, is an independent prognostic factor for disease recurrence and reduced survival in stage I non-small cell lung cancer. Clin Cancer Res 2007;13:3875–82.
[56] Lehtonen ST, Svensk AM, Soini Y, Pääkkö P, Hirvikoski P, Kang SW, et al. Peroxiredoxins, a novel protein family in lung cancer. Int J Cancer 2004;111:514–21.
[57] Alfonso P, Catalá M, Rico-Morales ML, Durante-Rodríguez G, Moro-Rodríguez E, Fernández-García H, et al. Proteomic analysis of lung biopsies: differential protein expression profile between peritumoral and tumoral tissue. Proteomics 2004;4:442–7.
[58] Wong KK, Jacks T, Dranoff G. NF-kappaB fans the flames of lung carcinogenesis. Cancer Prev Res (Phila) 2010;3:403–5.
[59] Di Stefano A, Caramori G, Oates T, Capelli A, Lusuardi M, Gnemmi I, et al. Increased expression of nuclear factor-kappaB in bronchial biopsies from smokers and patients with COPD. Eur Respir J 2002;20:556–63.
[60] Nakanishi C, Toi M. Nuclear factor-kappaB inhibitors as sensitizers to anticancer drugs. Nat Rev Cancer 2005;5:297–309.
[61] Cheng Q, Lee HH, Li Y, Parks TP, Cheng G. Upregulation of Bcl-x and Bfl-1 as a potential mechanism of chemoresistance, which can be overcome by NF-kappaB inhibition. Oncogene 2000;19:4936–40.
[62] Denlinger CE, Rundall BK, Jones DR. Modulation of antiapoptotic cell signaling pathways in non-small cell lung cancer: the role of NF-kappaB. Semin Thorac Cardiovasc Surg 2004;16:28–39.
[63] Shen HM, Tergaonkar V. NFkappaB signaling in carcinogenesis and as a potential molecular target for cancer therapy. Apoptosis 2009;14:348–63.
[64] Li Y, Ahmed F, Ali S, Philip PA, Kucuk O, Sarkar FH. Inactivation of nuclear factor kappaB by soy isoflavone genistein contributes to increased apoptosis induced by chemotherapeutic agents in human cancer cells. Cancer Res 2005;65:6934–42.
[65] Osaki S, Nakanishi Y, Takayama K, Pei XH, Ueno H, Hara N. Transfer of IkappaBalpha gene increase the sensitivity of paclitaxel mediated with caspase 3 activation in human lung cancer cell. J Exp Clin Cancer Res 2003;22:69–75.
[66] Chen W, Wang X, Bai L, et al. Blockage of NF-kappaB by IKKbeta- or RelA-siRNA rather than the NF-kappaB super-suppressor IkappaBalpha mutant potentiates adriamycin-induced cytotoxicity in lung cancer cells. J Cell Biochem 2008;105:554–61.
[67] Jones DR, Broad RM, Madrid LV, et al. Inhibition of NF-kappaB sensitizes non-small cell lung cancer cells to chemotherapy-induced apoptosis.Ann Thorac Surg 2000;70:930–6 [discussion 936–937].

Chapter 7

Oncoproteomic Approaches in Lung Cancer Research

Mª Dolores Pastor, Ana Nogal, Sonia Molina-Pinelo, Luis Paz-Ares and Amancio Carnero

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/53873

1. Introduction

With more than 1 million annual deaths, among both females and males, lung cancer is the world leading cause of cancer-related death (1). The most important risk factor for lung cancer is smoking, with smokers presenting a 10 fold risk increase compared to non-smokers. Lung cancers are usually divided into two categories: small-cell lung cancer (SCLC), representing approximately 15% of cases, and non-small cell lung cancer (NSCLC). This sub-division represents around 85% of all lung cancer cases and includes the histological sub-types adenocarcinoma, large-cell carcinoma and squamous cell carcinoma (2). The lung cancer 5-year survival rate is one of the lowest at 10-15% and treatment depends on the extent of the disease at the time of diagnosis (3). Approximately 30% of patients have early stage lung cancer when diagnosed and those tumours can be surgically removed, 20% have local and/or regionally advanced tumours and are treated with chemo and radiotherapy, and almost half of the patients have advanced metastatic disease when only palliative treatments are available (4). Consequently there is a pressing need for new screening and early diagnostic techniques that are specific and non-invasive, and also for tools that can predict prognosis, optimize treatments and identify new therapeutic targets. Genomic approaches have been used to that end in the last years. Nonetheless, given the importance of proteins to a cells' phenotype, post-translational modifications, and the poor correlation between mRNA and protein expression levels (5, 6), proteomic analyses may enlighten the pathogenesis of lung cancer. A variety of techniques such as two dimensional gel electrophoresis (2D-PAGE, 2D-DIGE), protein arrays, protein labelling and tagging (ICAT, iTRAQ, SILAC), are being used in cancer research (7, 8) and have the potential to aid clinical practice as a complement to histopathology, as a selection method for individualized therapy, and in the assessment of drug efficacy, resistance, and toxicity (9).

2. Lung cancer

In the beginning of the 20th century, lung cancer was a rare disease. Nowadays it has the highest incidence and mortality rates in the world with lifestyle and environmental factors thought to be the major contributors to the development of this disease (10). Epidemiological evidence has shown that two to three decades after a peak in smoking prevalence in a given population, there is a peak in lung cancer deaths, making tobacco smoking the main cause of lung cancer development. This relationship was established in the 1950's and 60's (10-12). Other causes include environmental tobacco smoking, air pollution, indoor radon, occupational exposure to respiratory carcinogens, asbestos, and fumes from cooking stoves and fires (10). Even though smoking is undeniably the major cause of lung cancer, making it the leading cause of preventable death in the world, it is important to recognize that the majority of smokers will not develop this neoplasia over time and that this is probably due to individual variation in the susceptibility to respiratory carcinogens and the existence of a previous lung disease (13, 14). Tobacco components can induce DNA damage through several mechanisms including gene point mutations, deletions, insertions, recombinations, rearrangements, and chromosomal alterations, which drive the development of the disease (15). Nonetheless, the current classification of lung cancer does not emphasize the important of specific molecular and genetic alterations that can differentiate between SCLC and NSCLC. This is also true for the NSCLC subtypes adenocarcinoma, large cell carcinoma, and squamous cell carcinoma, that were until recently, treated similarly, regardless of their biological heterogeneity (16). Lung cancer is characterized by genetic instability of the chromosomes, nucleotides, and the transcriptome. These abnormalities are usually targeted to proto-oncogenes, tumour suppressor genes, DNA repair genes, among others. The silencing of telomerase is present in normal cells, but in almost all SCLC and over 80% of NSCLC, telomerase is activated, promoting cell immortalization (17). The epidermal growth factor receptor (EGFR) is overexpressed or abnormally activated by mutation in 50-90% of all NSCLC, especially in squamous cell carcinomas, leading to increased cell proliferation and survival through the RAS/RAF/MEK/MAPK and PI3K/AKT pathways (18). Activating mutations of the *KRAS* gene from the RAS proto-oncogene family are present in 20% of all NSCLS and between 30-50% of lung adenocarcinomas (19). The fusion of the echinoderm microtubule-associated protein-like 4 (*EML4*) and the anaplastic lymphoma kinase (*ALK*) genes occurs in approximately 7% of NSCLC and is associated with a persistent mitogenic signal. The *EML4-ALK*, *EGFR*, and *KRAS* mutations are almost always mutually exclusive (19). Tumour suppressor genes are also affected in lung cancer. Mutations in *TP53* are the most common genetic alterations found in human cancers and occur in approximately 75% of SCLC and in 50% of NSCLC (17). Alterations in the PI3K/AKT pathway, the CDKN2A/RB1 pathway, *VEGF*, and epigenetic changes are also present in lung cancer (19). Several drugs have been developed to target these alterations and improve survival of lung cancer patients, such as tyrosine kinase inhibitors and monoclonal antibodies, revealing the importance of the molecular characterization of tumours in order to improve detection, diagnosis, treatment and prognosis of lung cancer.

Proteins are crucial operators in the majority of biological systems and a comprehensive knowledge of their expression, modifications, and function in the lung cancer setting, may be more informative than DNA and RNA studies alone. New technologies are being developed that allow the analysis of thousands of cancer cell proteins, possibly generating new therapeutic targets and biomarkers that will have an impact on early detection, therapy and prognostic evaluation of lung cancer patients.

3. Proteomic techniques in lung cancer research

The proteomic technologies which are being implemented in lung cancer research are mainly based on two dimensional gel electrophoresis, as seen on Figure 1 where the 2D-PAGE and 2D-DIGE workflows are represented, or proteomics based on isotope labelling methods as ICAT, iTRAQ, SILAC, followed by mass spectrometry (MS) analysis.

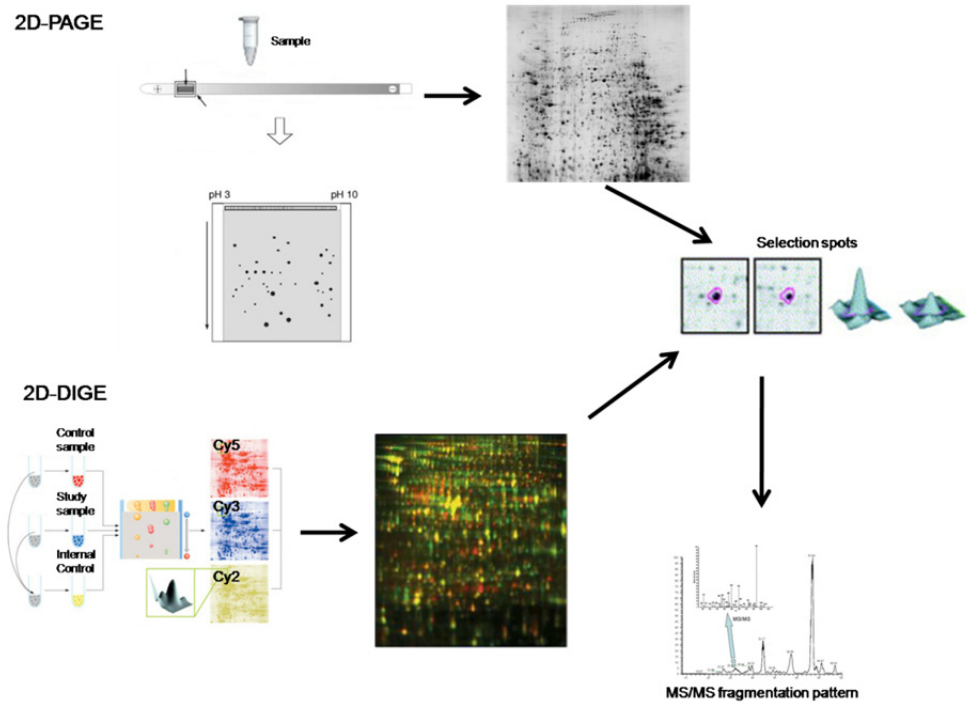


Figure 1. Basic workflow of gel-based proteomic approaches. In 2D-PAGE, protein samples are separated according to their isoelectric point in a process termed isoelectric focusing, using gel strips with a fixed pH range. Then, the focused strip is placed on top of a polyacrylamide gel to allow proteins to separate according to their molecular weight during electrophoresis, generating a gel with protein spots. In 2D-DIGE, proteins from up to three samples are labelled with fluorescent dyes prior to their isoelectric focusing and subsequent gel electrophoresis. Gels are scanned with different wavelengths revealing spots and differences in expression between analysed samples. Protein spots of interest in both techniques are then excised, digested, and identified by MS.

4. Two dimensional gel electrophoresis 2D-PAGE

2D-PAGE is the most used proteomic technique for studying the proteome as well as to search for cancer biomarkers (20, 21). In this methodology intact proteins are firstly separated by their isoelectric point (pI) and then according to their molecular weight. This procedure generates protein spots that are separated from the gel and digested into peptides for MS identification. Multidimensional separation of peptides may also be required given that, although the digestion step facilitates the identification process, it increases sample complexity, decreasing the sensitivity and coverage of the technique. Disadvantages of 2D-PAGE include the separation of low abundant proteins and of membrane proteins. The use of fractioning methods or higher protein concentrations for less detectable proteins and the use of mild detergents to increase the solubility of membrane proteins may be a solution for the aforementioned issues (22, 23). Other problems include co-migration of different proteins, the separation of a protein with different post-translational modifications, proteins with pI values below 4 or above 9, or the separation of very small or very large proteins. Differential gel electrophoresis (2D-DIGE), a modification of 2D-PAGE with fluorescent dyes (Cy3, Cy5 and Cy2), is able to increase reproducibility and throughput and also allows the accurate quantitation of protein expression difference (24). Differential analysis software can recognize the differentially expressed proteins and these can later be trypsin digested into peptides generating peptide mass fingerprints (PMF). The absolute masses of these peptides can be measured by matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS), a technique that is both relatively easy to use and reasonably sensitive for identifying proteins. Additionally other MS techniques, such as electrospray ionization (ESI-MS/MS), are capable of providing amino acid sequence information on peptide fragments of the initial protein (25). Liquid chromatography coupled to tandem mass spectrometry workflow (LC-MS/MS) has become a standard method to identify proteins from complex biological samples. Also, direct MS analysis of tissue, known as MALDI imaging, is a method that has been used to elucidate proteome features characterizing histological differences in lung cancer between adenocarcinoma and squamous- cell carcinoma (26). Another example of a novel way to generate proteomic data is presented in the study of dynamic proteome changes on lung cancer cells (H1299) treated with the cytotoxic drug camptothecin using single-protein labelling on large scale (27).

5. Isotope-labelled mass spectrometry

Isotope-labelling methods, as seen on Figure 2, are gel-free procedures that introduce stable isotope tags to proteins through chemical reactions using isotope-coded affinity tags (ICAT) (28) and isobaric tag for relative and absolute quantitation (iTRAQ) (29), or through metabolic labelling with isotope labelled amino acids in cell culture (SILAC) (30).

ICAT is used to analyse pairs of protein samples, such as a treated sample and its control. Extracted proteins from both samples are labelled with a light or heavy ICAT reagent by reacting with a specific amino acid (cysteine). Samples are then mixed, trypsin digested, fractioned, and analysed by LC-MS/MS (31). Isotope peak ratios for each peptide determine

the differential protein expression. The drawback of this technique is that it can only analyse cysteine containing proteins, two samples, and it can only identify 300-400 peptides.

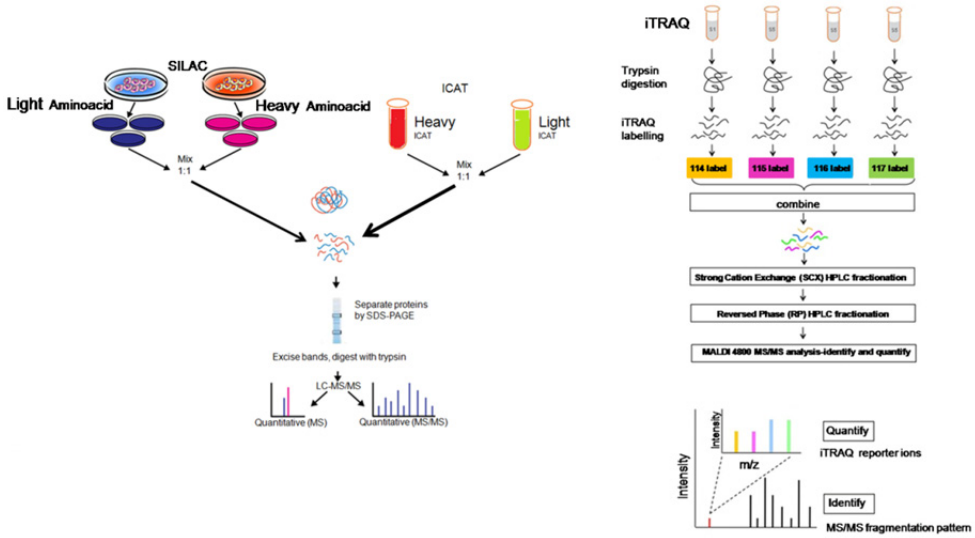


Figure 2. Basic workflow of gel-free quantitative approaches in proteomics. In SILAC, one cellular culture is grown in normal medium and the other with a growth medium with heavy labelled amino-acids. In ICAT, one protein extract is labelled with a light ICAT reagent and the other with a heavy ICAT reagent. In both techniques, samples are mixed, digested, separated and analysed by MS to determine protein identity and differential expression. In iTRAQ, special isobaric tags are applied in 4 to 8 samples up for comparison. They are then pooled together, fractionated and analysed by MS, allowing protein identification and quantitation among studied samples.

iTRAQ is another labelling technique first developed by Ross and co-workers (32) which uses isobaric tags to label and compare proteins extracted from samples. iTRAQ contains a set of four or eight isobaric reagents and therefore can analyse up to four or eight protein samples at one time. After trypsin digestion samples are labelled with four or eight (4-plex or 8-plex) independent iTRAQ reagents. The reporter groups of the iTRAQ reagents separate from the peptides and generate small fragments for each sample with mass-to-charge (m/z) of 114, 115, 116, and 117 for 4-plex, plus 113, 118, 119, and 121 for 8-plex. The intensity of each peak correlates with the quantity of each reporter group and thus with the quantity of the peptide. This method allows the analysis of various samples at a time and also, given that most peptides are suitable to be labelled by iTRAQ, it minimizes information loss and allows the identification of proteins with different post-translational modifications. Disadvantages of iTRAQ include a separate lengthy sample processing, that increases the chances of experimental errors, and the generation of chemical side products during the labelling process that can reduce the sensitivity of the method (33).

SILAC, first developed by Mann and co-workers, is based on the metabolic incorporation of “heavy” and “light” forms of amino acids into the proteins of living cultured cells (34) .

Typically, heavy (13C or 15N) arginine or lysine are used in the culture medium of a cell culture while the other cell culture is supplied with regular amino acids. After several division rounds, these amino acids are incorporated into the newly synthesized proteins. Following trypsin digestion, peptides are analysed by MS and the light and heavy peptides appear in two distinct peaks and, by comparing the signal intensities differences, relative quantitation can be performed. This technique has been widely used for cancer biomarker discovery (35), and cell signalling dynamics (36).

6. Label-free mass spectrometry

Multidimensional Protein Identification Technology (MudPIT) is a generic label-free LC-MS shotgun screening method (36). It separates peptides according to two independent physicochemical properties using liquid chromatography (LC/LC) online with the ion source of a mass spectrometer, allowing the separation and identification of peptides without labelling. The success of this technique depends on the experimental workflow, from protein extraction to sample stability, given that the reproducibility of technical replicates is better than that of experimental replicates. Drawbacks of this method include the fact that not all peptides are equally detectable given the competition between ions, dynamic range limitations and MS sensitivity (37). With time and improvements, label-free MS could be widely used for biomarker discovery and validation.

7. Detection of post-translational modifications (PTMs)

PTMs are the chemical alterations that occur to a protein after translation. They include proteolytic cleavage, glycosylation, phosphorylation, acetylation, ubiquitination, farnesylation, methylation, sialylation, oxidation, prolyl isomerization and hydroxylation (38). Glycosylation and phosphorylation are two of the most biologically relevant PTMs and appear to be key processes in tumour progression in many types of cancers including lung cancer (39, 40)

Glycosylation, the process of adding saccharides to proteins, plays a fundamental role in protein stabilization, molecular and cellular recognition, growth and cellular communication, and can also be a part of immune responses and cancer progression (41). The comparative study of the carbohydrate chains of glycoproteins may provide useful information for the diagnosis, prognosis, and immunotherapy of tumours (42). The proteomic analysis of glycoproteins starts with the enrichment of these molecules from a complex protein sample by the use lectins. This step is followed by a separation of glycoproteins by procedures such as 2D-PAGE and 2D-DIGE coupled with glycoprotein staining methods, for example Pro-Q Emerald 488 glycoprotein stain (43), lectin fluorescence stain (44), and isotope labelling (45). Identification of separated glycoproteins and their glycan structures can be accomplished by chromatographic methods (nano-LC with hydrophilic columns, nano-LC with graphitized carbon packing, anion-exchange chromatography), electromigration approaches (capillary electrophoresis, capillary electrochromatography), capillary LC/MALDI-TOF/TOF MS & tandem MS (MS/MS), and

chip-based approaches (46). Although there are some difficulties when analysing lung tumours, one study has identified 34 glycoproteins with significant differences between lung adenocarcinomas and healthy controls. The α1,6-fucosylation levels were incremented in the lung cancer group in comparison with healthy group (47).

Phosphorylation is the addition of a phosphate group to a protein and is a key regulatory mechanism of cellular signalling processes. Phosphoproteomics and the characterization of phosphorylation sites, which less than 2% are currently known, are some of the most challenging tasks in current proteomic research (48). To isolate and identify phosphorylated proteins one must use immunoaffinity or immunoprecipitation with a specific antibody, chromatofocusing, ion exchange chromatography and affinity chromatography, such as immobilized metal ion affinity chromatography (IMAC) (49). Separation methods include electrophoresis, 2D-PAGE or 2D-DIGE coupled with phosphoprotein staining (Pro-Q Diamond phosphoprotein gel stain) or isotope labelling (ICAT, SILAC) (50, 51). Analysis and identification methods of phosphoproteins and phosphopeptides are mass spectrometry-based approaches, such MALDI-TOF MS, LC-ESI-MS and MS/MS (52). Given that the key regulators of signalling cascades are kinases and phosphatases, lung cancer phosphoproteomics might reveal the correlation between phosphorylation and cancer mechanisms.

8. Samples in lung cancer proteomics

The lung is a heterogeneous organ composed by several highly differentiated cells (bronchial, alveolar, inflammatory) and vascular structures. Its main function is to perform gas exchanges between the atmosphere and the bloodstream. When studying lung cancer with proteomic tools, several different samples can be used: tumour tissue, blood, pleural effusions, among others (53). The accessibility of blood makes for a great sample for oncoproteomic studies. Moreover, it contains many circulating molecules secreted by the tumour that can be used as biomarkers. Nonetheless, due to the abundance of plasma proteins, depletion of these proteins is necessary to reveal the presence of less abundant ones. Tumour tissue samples, fresh-frozen or formalin-fixed and paraffin-embedded, are the ideal for any oncoproteomic study. However, adjacent normal tissue, inflammatory cells, stromal components, and others might also be present. This will result in non-tumour derived protein contamination. To compensate tumour heterogeneity careful sample cell content analysis and the increase of sample numbers is required to obtain relevant results. The pleura is a thin double-layered tissue that surrounds the lung and it is filled with pleural fluid. This liquid is constantly produced and reabsorbed, and its main function is to facilitate respiratory movements and reduce attrition between the lungs and the thorax wall. Pleural effusion is the pathological accumulation of fluid that occurs in inflammatory conditions and lung cancer. In the latter case, pleural effusion is often drained to search for cancer cell infiltration. Its protein composition is similar to plasma, but its proximity to tumour cells makes it useful for lung cancer biomarker detection by proteomic techniques.

9. Proteomics in the discovery and validation of lung cancer biomarkers

9.1. Diagnostic biomarkers

To discover a lung cancer diagnostic biomarker, a molecule that is specific and directly correlates with the presence of this disease, the majority of studies perform a comparison between the protein profiles of tumour samples and normal lung tissue. The ideal would be to study the development of the carcinogenic process from normal tissue, to metaplasia, to dysplasia, and finally to invasive cancer, in order to discover early markers of disease before the onset of clinical features.

In response to inflammation, a cancer enabling characteristic, acute-phase reactant proteins (APRPs) are produced. Recent proteomic studies have shown that APRPs haptoglobin (Hp) β chain (54), serum amyloid A (SAA) (55), and apolipoprotein A-1 (Apo A-1) (56) proteins are potential lung cancer diagnostic biomarkers. SAA proteins are involved in the transport of cholesterol to the liver, the recruitment of immune cells, and the induction extracellular matrix degrading enzymes. SAA1 and SAA2, which are synthesised in response to activated monocytes/macrophages, were recently identified, by LC-MS/MS, ELISA and immunohistochemistry analyses, as lung cancer biomarkers given their higher expression levels in blood and tissue from lung cancer patients when compared to healthy subjects and patients with other cancers and respiratory diseases (55). In another related study, serum and pleural effusions from NSCLC patients were compared by 2D-DIGE to those from patients with benign lung diseases. Gelsolin, possibly involved in cancer invasion, metalloproteinase inhibitor 2 (TIMP2), involved in lung parenchyma disorganization, and pigment epithelium derived factor (PEDF), an angiogenesis inhibitor, were among the candidate biomarkers (57). A study by Patz and co-workers, that aimed to test the diagnostic performance of four lung cancer biomarkers (carcinoembryonic antigen and squamous-cell carcinoma antigen, and 2D-PAGE and MALDI-MS discovered retinol binding protein – RBP - and α-1 antitrypsin), demonstrated that the four markers have inadequate diagnostic power when tested independently but proved useful when used in combination (58). A glycoproteomic study revealed plasma kallikrein (KLKB1), pleural effusion periostin, multimerin-2, CD166 and lysosome-associated membrane glycoprotein-2 (LAMP-2) as potential lung cancer biomarkers (59).

9.2. Prognostic biomarkers

Prognostic biomarkers, those that have expression levels correlating with the natural history of the disease, have the potential to influence survival by identifying high-risk patients and thus improve their management. The study of prognostic biomarkers in lung cancer has been made by correlating the expression of a molecule to the patient survival. An alternative approach is to compare groups of patients with different clinical stages of disease, based on the assumption that a more advanced tumour is more aggressive and may express proteins that drive the metastatic process. Proteomic studies have aimed at discovering altered protein levels and subsequently validating those differences using immunohistochemistry on archive samples. Using 2D-PAGE, Chen and co-workers associated 11 components of the

glycolysis pathway to poor survival in lung adenocarcinoma (39) and also demonstrated their prognostic role in lung cancer at the mRNA level. Nonetheless, glycolysis involved enzyme phosphoglycerate kinase 1 was found to limit tumour growth in mice subcutaneously injected with the Lewis lung carcinoma cell line, by promoting antitumor immunity (60). A study using 2D-DIGE, MS, western blot, and immunohistochemistry correlated the up-regulation of annexin A3, a protein associated with cancer metastasis by angiogenic promotion, with advanced clinical stage, lymph node metastasis, increased relapse time, and overall decreased survival in lung adenocarcinoma, indicating that annexin A3 might be a prognostic lung cancer biomarker (61). The involvement of S100A11, a small calcium-binding protein implicated in the prognosis and metastasis in several tumours, has also been evaluated in lung cancer. Comparative proteomic analysis of two NSCLC cell lines, the non-metastatic CL1-0 and highly metastatic CL1-5, revealed that S100A11 was up-regulated in metastatic CL1-5 cells (62). Moreover, immunohistochemical analyses in NSCLC tissues showed that the up-regulation of S100A11 was significantly associated with a higher TNM stage and a positive lymph node status, indicating its importance in promoting invasion and metastasis of NSCLC. Altered expression of S100A6 was also implicated in NSCLC progression: elevated levels of this protein were associated with longer survival compared to S100A6-negative cases (63). Cytoskeletal reorganization is a central process regulating cell migration and metastasis and cytokeratins (CKs), a family of cytoskeletal intermediate filaments, have been suggested to play a role in carcinogenesis, by promoting cellular architecture reorganization during tumour development and progression. A 2D-PAGE and MS analysis has revealed that isoforms of CK7, 8, 18, and 19 were found in higher levels in adenocarcinoma samples than in adjacent tissues (64). Specific isoforms of the CKs were associated with unfavourable prognosis, CYFRA21-1 was a more accurate diagnostic marker, and CK18 was a stronger prognostic factor (65). Other cytoskeletal proteins found to be correlated with a poor prognosis in lung adenocarcinoma are non-muscle myosin IIA and vimentin proteins, involved in epithelial-mesenchymal transition, a process at the basis of invasive and metastatic behaviour (66). Phosphohistidine phosphatase (PHP14) was proposed to be another lung cancer prognostic biomarker, regulating cell migration and invasion by cytoskeleton rearrangement. Indeed, it has been shown that PHP14 knockdown in highly metastatic lung cancer cells (CL1-5) inhibited migration and invasion, whereas its over-expression in NCI H1299 cells enhanced these processes (67). Calmodulin, a protein implicated in cytoskeletal alterations during cell death, thymosin β4, a regulator of actin polymerization whose over-expression seems to stimulate lung tumour metastasis, thymosin β10 and cofilin proteins, regulators of actin dynamics, were identified and their expression and prognostic role validated on cohort of 188 lung cancer cases (68).

9.3. Predictive biomarkers

The discovery of predictive biomarkers, those on which the efficacy of a specific treatment can be foreseen, has been based on studying clinical samples from responding and non-responding patients and then validating results on selected cohorts. This type of biomarker

aims at individualizing therapies in lung cancer but relies on extremely well characterized samples from cohorts of patients receiving a uniform treatment and closely monitored therapeutic responses. A recent MALDI-TOF-MS study that profiled serum from patients treated with cisplatin-gemcitabine in combination with the proteasome inhibitor bortezomib, revealed a 13-peptide signature that was able to distinguish with high accuracy, sensitivity, and specificity, patients with short and long progression-free survival (69). The epidermal growth factor receptor (EGFR) tyrosine kinase is an important target for treatment of NSCLC, and EGFR-inhibitor-based therapies have showed promising results. The serum MALDI-MS study conducted by Taguchi and co-workers in NSCLC patients

Type of Biomarker	Proteins	Techniques
Diagnostic	Hp β chain (54)	LC-ESI-MS/MS, ELISA
	SAA1 SAA2 (55)	LC-MS/MS, ELISA, IHC
	Apo A1 (56)	2D-PAGE, MALDI-TOF
	Gelsolin TIMP2 PEDF (57)	2D-DIGE
	RBP α-1 antitrypsin (58)	2D-DIGE, MALDI-TOF-MS
	KLKB1 Periostin Multimerin-2 CD166 LAMP-2 (59)	LC-MS/MS
Prognostic	Glycolysis (11 components) (39)	2D-PAGE
	Annexin A3 (61)	2D-DIGE, MS, IHC*
	S100A11 (62)	2D-PAGE, MALDI-TOF-MS/MS, IHC
	S100A6 (63)	SELDI-TOF-MS
	CK 7, 8, 9 and 19 (64)	2D-PAGE, MS
	CYFRA21-1 CK18 (65)	ELISA
	Myosin IIA Vimentin (66)	LC-MS/MS
	PHP14 (67)	2D-PAGE, ESI-TOF-MS/MS
	Calmodulin Thymosin β4 Thymosin β10 (68)	MALDI-MS, IHC
Predictive	13-peptide signature (69)	MALDI-TOF-MS
	8-peak signature (70)	MALDI-MS

* Immunohistochemistry

Table 1. Potential lung cancer biomarkers discovered by the use of proteomic tools

treated with gefitinib and erlotinib revealed an 8-peak profile predictive of outcome (70). This 8-peak signature was commercially launched as a commercial product (Veristrat ®, Biodesix, Broom field, CO, US) and its clinical relevance is being validated in the context of a randomized phase III clinical trial where patients with advanced NSCLC progressing after first-line treatment, stratified according to serum MALDI-MS profiling, are subsequently randomly allocated to receive either erlotinib or chemotherapy as second-line therapy (PROSE, Proteomics Stratified Erlotinib trial). To the best of our knowledge, this is the only clinical trial investigating the predictive role of a proteomics biomarker in lung cancer patients. A summary of all mentioned biomarkers can be found on Table 1.

10. Conclusions

Proteomic approaches are improving rapidly and the development of high-throughput platforms is showing promising results as the list of candidate biomarkers for lung cancer is continuously growing. However, there is a great need for careful interpretation of this intricate data in order to generate biologically relevant hypotheses. The proteome is highly complex and current tools cannot yet provide a definitive solution for its exploration. In addition, cancer is a multifactorial disease so diverse that a great deal of time and effort will be necessary to define its associated proteome modifications and to translate these into practical clinical applications. In fact for many of the identified proteins, their functional role in lung cancer development is not yet known and a solid clinical validation is still lacking. Nonetheless, it is likely that some of these candidate biomarkers will serve to identify new possible therapeutic strategies.

Author details

Mª Dolores Pastor, Ana Nogal, Sonia Molina-Pinelo, Luis Paz-Ares and Amancio Carnero
Institute of Biomedicine of Seville (HUVR/CSIC/University of Seville), Spain

Ana Nogal
Biomedical Sciences Institute of Abel Salazar (University of Porto), Portugal

Amancio Carnero
Consejo Superior de Investigaciones Científicas (CSIC), Spain

11. References

[1] Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. CA A Cancer Journal for Clinicians. 2010 Sep-Oct;60(5):277-300.
[2] Lehtio J, De Petris L. Lung cancer proteomics, clinical and technological considerations. Journal of Proteomics. 2010 Sep 10;73(10):1851-63.
[3] Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA A Cancer Journal for Clinicians. 2009 Jul-Aug;59(4):225-49.

[4] Groome PA, Bolejack V, Crowley JJ, Kennedy C, Krasnik M, Sobin LH, et al. The IASLC Lung Cancer Staging Project: validation of the proposals for revision of the T, N, and M descriptors and consequent stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumours. Journal of Thoracic Oncology. 2007 Aug;2(8):694-705.

[5] Chen G, Gharib TG, Wang H, Huang CC, Kuick R, Thomas DG, et al. Protein profiles associated with survival in lung adenocarcinoma. Proceedings of the National Academy of Sciences of the United States of America. 2003 Nov 11;100(23):13537-42.

[6] Ocak S, Chaurand P, Massion PP. Mass spectrometry-based proteomic profiling of lung cancer. Proceedings of the American Thoracic Society. 2009 Apr 15;6(2):159-70.

[7] Cho WC, Cheng CH. Oncoproteomics: current trends and future perspectives. Expert Review of Proteomics. 2007 Jun;4(3):401-10.

[8] van der Merwe DE, Oikonomopoulou K, Marshall J, Diamandis EP. Mass spectrometry: uncovering the cancer proteome for diagnostics. Advances in Cancer Research. 2007;96:23-50.

[9] Au JS, Cho WC, Yip TT, Law SC. Proteomic approach to biomarker discovery in cancer tissue from lung adenocarcinoma among nonsmoking Chinese women in Hong Kong. Cancer Investigation. 2008 Mar;26(2):128-35.

[10] Alberg AJ, Ford JG, Samet JM. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007 Sep;132(3 Suppl):29S-55S.

[11] Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. British Medical Journal. 1950 Sep 30;2(4682):739-48.

[12] Mills CA, Porter MM. Tobacco smoking habits and cancer of the mouth and respiratory system. Cancer Research. 1950 Sep;10(9):539-42.

[13] Alberg AJ, Samet JM. Epidemiology of lung cancer. Chest. 2003 Jan;123(1 Suppl):21S-49S.

[14] Brody JS, Spira A. State of the art. Chronic obstructive pulmonary disease, inflammation, and lung cancer. Proceedings of the American Thoracic Society. 2006 Aug;3(6):535-8.

[15] Adcock IM, Caramori G, Barnes PJ. Chronic obstructive pulmonary disease and lung cancer: new molecular insights. Respiration. 2011;81(4):265-84.

[16] Borczuk AC, Toonkel RL, Powell CA. Genomics of lung cancer. Proceedings of the American Thoracic Society. 2009 Apr 15;6(2):152-8.

[17] Meyerson M, Franklin WA, Kelley MJ. Molecular classification and molecular genetics of human lung cancers. Seminars in Oncology. 2004 Feb;31(1 Suppl 1):4-19.

[18] Breuer RH, Postmus PE, Smit EF. Molecular pathology of non-small-cell lung cancer. Respiration. 2005 May-Jun;72(3):313-30.

[19] West L, Vidwans SJ, Campbell NP, Shrager J, Simon GR, Bueno R, et al. A novel classification of lung cancer into molecular subtypes. PLoS One. 2012;7(2):e31906.

[20] Lilley KS, Razzaq A, Dupree P. Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. Current Opinion in Chemical Biology. 2002 Feb;6(1):46-50.

[21] Yu KH, Rustgi AK, Blair IA. Characterization of proteins in human pancreatic cancer serum using differential gel electrophoresis and tandem mass spectrometry. Journal of Proteome Research. 2005 Sep-Oct;4(5):1742-51.
[22] Qin S, Ferdinand AS, Richie JP, O'Leary MP, Mok SC, Liu BC. Chromatofocusing fractionation and two-dimensional difference gel electrophoresis for low abundance serum proteins. Proteomics. 2005 Aug;5(12):3183-92.
[23] Luche S, Santoni V, Rabilloud T. Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. Proteomics. 2003 Mar;3(3):249-53.
[24] Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis. 1997 Oct;18(11):2071-7.
[25] Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annual Review of Biochemistry. 2001;70:437-73.
[26] Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proceedings of the National Academy of Sciences of the United States of America. 2000 Aug 15;97(17):9390-5.
[27] Kameshita I, Ishida A, Fujisawa H. Analysis of protein-protein interaction by two-dimensional affinity electrophoresis. Analytical Biochemistry. 1998 Aug 15;262(1):90-2.
[28] Haqqani AS, Nesic M, Preston E, Baumann E, Kelly J, Stanimirovic D. Characterization of vascular protein expression patterns in cerebral ischemia/reperfusion using laser capture microdissection and ICAT-nanoLC-MS/MS. FASEB Journal. 2005 Nov;19(13):1809-21.
[29] Shadforth IP, Dunkley TP, Lilley KS, Bessant C. i-Tracker: for quantitative proteomics using iTRAQ. BMC Genomics. 2005;6:145.
[30] Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular & Cellular Proteomics. 2002 May;1(5):376-86.
[31] Graves PR, Haystead TA. Molecular biologist's guide to proteomics. Microbiology and Molecular Biology Reviews. 2002 Mar;66(1):39-63; table of contents.
[32] Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular & Cellular Proteomics. 2004 Dec;3(12):1154-69.
[33] DeSouza L, Diehl G, Rodrigues MJ, Guo J, Romaschin AD, Colgan TJ, et al. Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. Journal of Proteome Research. 2005 Mar-Apr;4(2):377-86.
[34] Mann M. Functional and quantitative proteomics using SILAC. Nature Reviews Molecular Cell Biology. 2006 Dec;7(12):952-8.
[35] Gronborg M, Kristiansen TZ, Iwahori A, Chang R, Reddy R, Sato N, et al. Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. Molecular & Cellular Proteomics. 2006 Jan;5(1):157-71.

[36] Guerrera IC, Keep NH, Godovac-Zimmermann J. Proteomics study reveals cross-talk between Rho guanidine nucleotide dissociation inhibitor 1 post-translational modifications in epidermal growth factor stimulated fibroblasts. Journal of Proteome Research. 2007 Jul;6(7):2623-30.

[37] Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, Hettich RL, Samatova NF. Detecting differential and correlated protein expression in label-free shotgun proteomics. Journal of Proteome Research. 2006 Nov;5(11):2909-18.

[38] Krueger KE, Srivastava S. Posttranslational protein modifications: current implications for cancer detection, prevention, and therapeutics. Molecular & Cellular Proteomics. 2006 Oct;5(10):1799-810.

[39] Chen G, Gharib TG, Huang CC, Thomas DG, Shedden KA, Taylor JM, et al. Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. Clinical Cancer Research. 2002 Jul;8(7):2298-305.

[40] Okano T, Kondo T, Kakisaka T, Fujii K, Yamada M, Kato H, et al. Plasma proteomics of lung cancer by a linkage of multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis. Proteomics. 2006 Jul;6(13):3938-48.

[41] Dennis JW, Granovsky M, Warren CE. Glycoprotein glycosylation and cancer progression. Biochimica et Biophysica Acta. 1999 Dec 6;1473(1):21-34.

[42] Lowe JB. Glycosylation, immunity, and autoimmunity. Cell. 2001 Mar 23;104(6):809-12.

[43] Hart C, Schulenberg B, Steinberg TH, Leung WY, Patton WF. Detection of glycoproteins in polyacrylamide gels and on electroblots using Pro-Q Emerald 488 dye, a fluorescent periodate Schiff-base stain. Electrophoresis. 2003 Feb;24(4):588-98.

[44] Patwa TH, Zhao J, Anderson MA, Simeone DM, Lubman DM. Screening of glycosylation patterns in serum using natural glycoprotein microarrays and multi-lectin fluorescence detection. Analytical Chemistry. 2006 Sep 15;78(18):6411-21.

[45] Ueda K, Katagiri T, Shimada T, Irie S, Sato TA, Nakamura Y, et al. Comparative profiling of serum glycoproteome by sequential purification of glycoproteins and 2-nitrobenzensulfenyl (NBS) stable isotope labeling: a new approach for the novel biomarker discovery for cancer. Journal of Proteome Research. 2007 Sep;6(9):3475-83.

[46] Mechref Y, Novotny MV. Miniaturized separation techniques in glycomic investigations. Journal of Chromatography B Analytical Technologies in the Biomedical and Life Sciences. 2006 Sep 1;841(1-2):65-78.

[47] Yang Z, Hancock WS, Chew TR, Bonilla L. A study of glycoproteins in human serum and plasma reference standards (HUPO) using multilectin affinity chromatography coupled with RPLC-MS/MS. Proteomics. 2005 Aug;5(13):3353-66.

[48] Mukherji M. Phosphoproteomics in analyzing signaling pathways. Expert Review of Proteomics. 2005 Jan;2(1):117-28.

[49] Schmidt SR, Schweikart F, Andersson ME. Current methods for phosphoprotein isolation and enrichment. Journal of Chromatography B Analytical Technologies in the Biomedical and Life Sciences. 2007 Apr 15;849(1-2):154-62.

[50] Goshe MB, Veenstra TD, Panisko EA, Conrads TP, Angell NH, Smith RD. Phosphoprotein isotope-coded affinity tags: application to the enrichment and

identification of low-abundance phosphoproteins. Analytical Chemistry. 2002 Feb 1;74(3):607-16.

[51] Raggiaschi R, Gotta S, Terstappen GC. Phosphoproteome analysis. Bioscience Reports. 2005 Feb-Apr;25(1-2):33-44.

[52] Carr SA, Annan RS, Huddleston MJ. Mapping posttranslational modifications of proteins by MS-based selective detection: application to phosphoproteomics. Methods in Enzymology. 2005;405:82-115.

[53] De Petris L, Pernemalm M, Elmberger G, Bergman P, Orre L, Lewensohn R, et al. A novel method for sample preparation of fresh lung cancer tissue for proteomics analysis by tumor cell enrichment and removal of blood contaminants. Proteome Science. 2010;8:9.

[54] Kang SM, Sung HJ, Ahn JM, Park JY, Lee SY, Park CS, et al. The Haptoglobin beta chain as a supportive biomarker for human lung cancers. Molecular BioSystems. 2011 Apr;7(4):1167-75.

[55] Sung HJ, Ahn JM, Yoon YH, Rhim TY, Park CS, Park JY, et al. Identification and validation of SAA as a potential lung cancer biomarker and its involvement in metastatic pathogenesis of lung cancer. Journal of Proteome Research. 2011 Mar 4;10(3):1383-95.

[56] Maciel CM, Junqueira M, Paschoal ME, Kawamura MT, Duarte RL, Carvalho Mda G, et al. Differential proteomic serum pattern of low molecular weight proteins expressed by adenocarcinoma lung cancer patients. Journal of Experimental Therapeutics and Oncology. 2005;5(1):31-8.

[57] Rodriguez-Pineiro AM, Blanco-Prieto S, Sanchez-Otero N, Rodriguez-Berrocal FJ, de la Cadena MP. On the identification of biomarkers for non-small cell lung cancer in serum and pleural effusion. Journal of Proteomics. 2010 Jun 16;73(8):1511-22.

[58] Patz EF, Jr., Campa MJ, Gottlin EB, Kusmartseva I, Guan XR, Herndon JE, 2nd. Panel of serum biomarkers for the diagnosis of lung cancer. Journal of Clinical Oncology. 2007 Dec 10;25(35):5578-83.

[59] Soltermann A, Ossola R, Kilgus-Hawelski S, von Eckardstein A, Suter T, Aebersold R, et al. N-glycoprotein profiling of lung adenocarcinoma pleural effusions by shotgun proteomics. Cancer. 2008 Apr 25;114(2):124-33.

[60] Tang SJ, Ho MY, Cho HC, Lin YC, Sun GH, Chi KH, et al. Phosphoglycerate kinase 1-overexpressing lung cancer cells reduce cyclooxygenase 2 expression and promote anti-tumor immunity in vivo. International Journal of Cancer. 2008 Dec 15;123(12):2840-8.

[61] Liu YF, Xiao ZQ, Li MX, Li MY, Zhang PF, Li C, et al. Quantitative proteome analysis reveals annexin A3 as a novel biomarker in lung adenocarcinoma. Journal of Pathology. 2009 Jan;217(1):54-64.

[62] Tian T, Hao J, Xu A, Luo C, Liu C, Huang L, et al. Determination of metastasis-associated proteins in non-small cell lung cancer by comparative proteomic analysis. Cancer Science. 2007 Aug;98(8):1265-74.

[63] De Petris L, Orre LM, Kanter L, Pernemalm M, Koyi H, Lewensohn R, et al. Tumor expression of S100A6 correlates with survival of patients with stage I non-small-cell lung cancer. Lung Cancer. 2009 Mar;63(3):410-7.

[64] Gharib TG, Chen G, Wang H, Huang CC, Prescott MS, Shedden K, et al. Proteomic analysis of cytokeratin isoforms uncovers association with survival in lung adenocarcinoma. Neoplasia. 2002 Sep-Oct;4(5):440-8.

[65] De Petris L, Branden E, Herrmann R, Sanchez BC, Koyi H, Linderholm B, et al. Diagnostic and prognostic role of plasma levels of two forms of cytokeratin 18 in patients with non-small-cell lung cancer. European Journal of Cancer. 2011 Jan;47(1):131-7.

[66] Maeda J, Hirano T, Ogiwara A, Akimoto S, Kawakami T, Fukui Y, et al. Proteomic analysis of stage I primary lung adenocarcinoma aimed at individualisation of postoperative therapy. British Medical Journal. 2008 Feb 12;98(3):596-603.

[67] Xu A, Hao J, Zhang Z, Tian T, Jiang S, Liu C, et al. 14-kDa phosphohistidine phosphatase and its role in human lung cancer cell migration and invasion. Lung Cancer. 2010 Jan;67(1):48-56.

[68] Xu BJ, Gonzalez AL, Kikuchi T, Yanagisawa K, Massion PP, Wu H, et al. MALDI-MS derived prognostic protein markers for resected non-small cell lung cancer. Proteomics - Clinical Applications. 2008 Oct;2(10-11):1508-17.

[69] Voortman J, Pham TV, Knol JC, Giaccone G, Jimenez CR. Prediction of outcome of non-small cell lung cancer patients treated with chemotherapy and bortezomib by time-course MALDI-TOF-MS serum peptide profiling. Proteome Science. 2009;7:34.

[70] Taguchi F, Solomon B, Gregorc V, Roder H, Gray R, Kasahara K, et al. Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. Journal of the National Cancer Institute. 2007 Jun 6;99(11):838-46.