

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.

For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Oncoproteomic Approaches in Lung Cancer Research

---

M<sup>a</sup> 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).

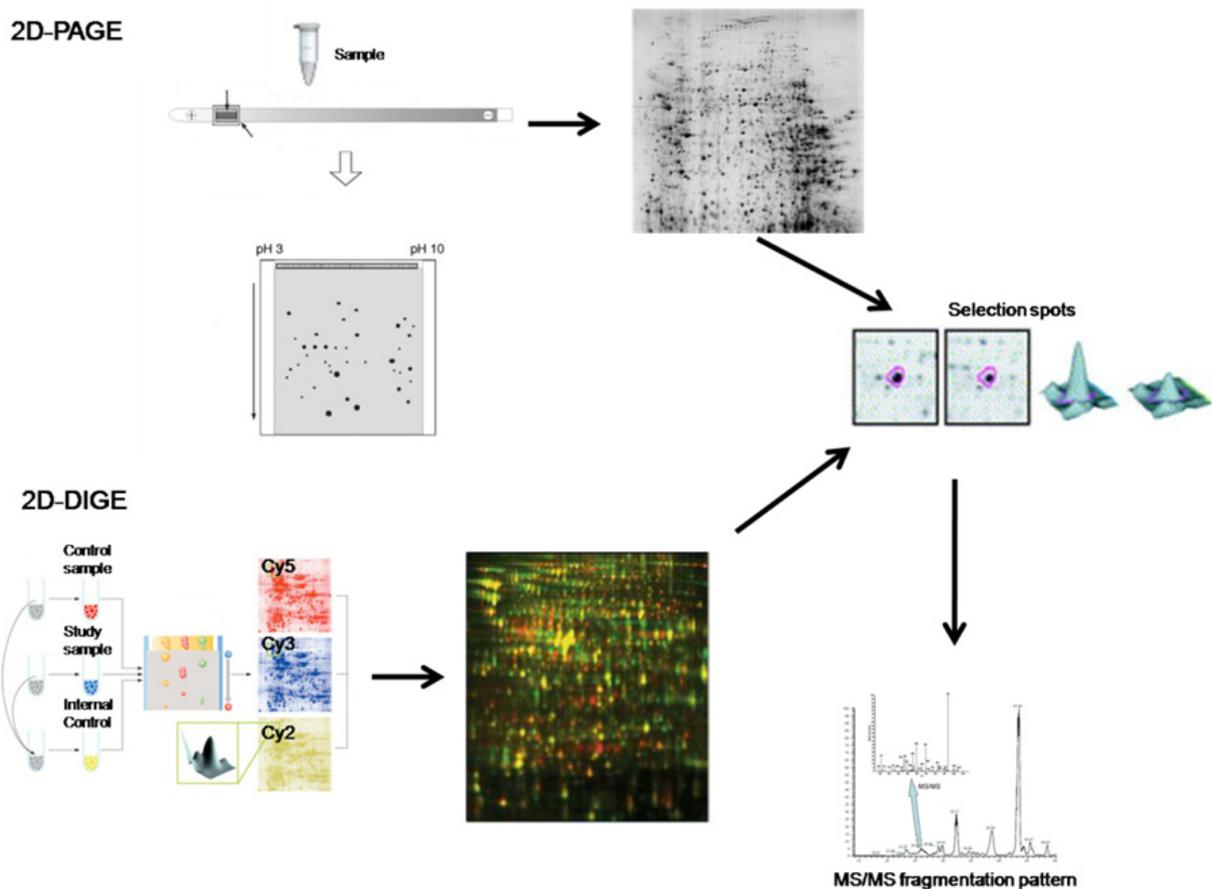
## 2. Lung cancer

In the beginning of the 20<sup>th</sup> 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 importance 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 NSCLC 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**

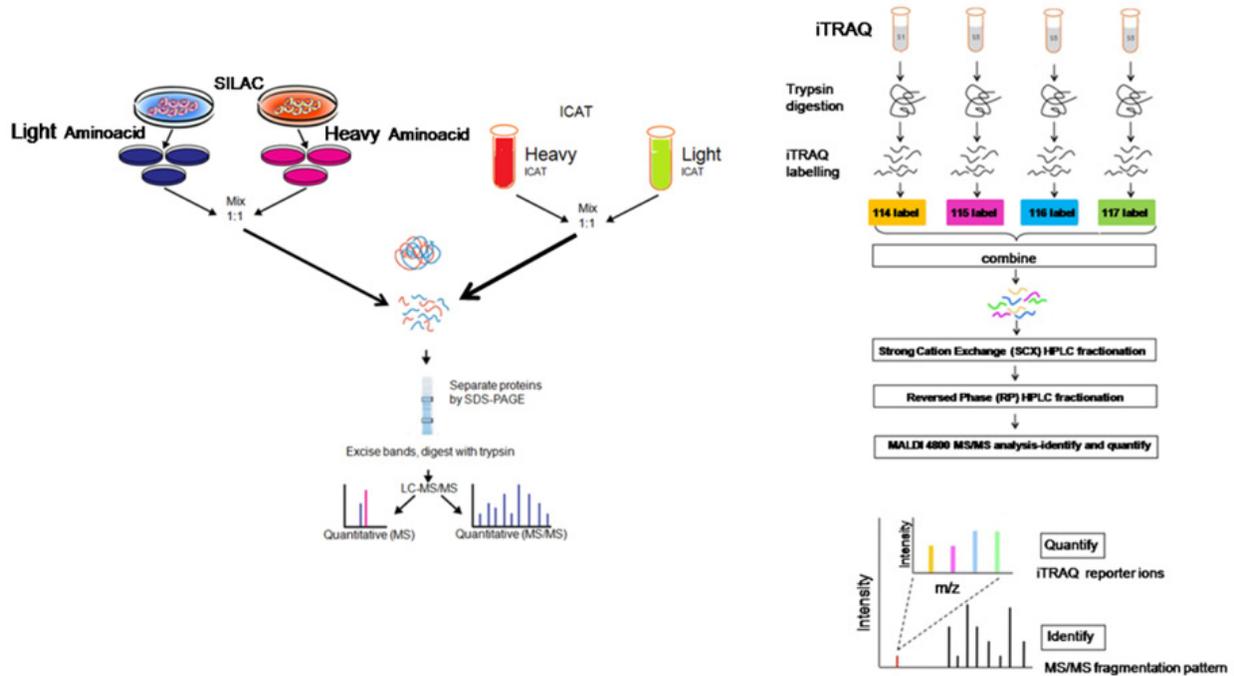
2D-PAGE is the most used proteomic technique for studying the proteome as well as to search for cancer biomarkers (20, 21). In this methodology intact proteins are firstly separated by their isoelectric point (pI) and then according to their molecular weight. This procedure generates protein spots that 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 ( $^{13}\text{C}$  or  $^{15}\text{N}$ ) 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  $\alpha$ 1,6-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 as 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)  $\beta$  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  $\alpha$ -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  $\beta$ 4, a regulator of actin polymerization whose over-expression seems to stimulate lung tumour metastasis, thymosin  $\beta$ 10 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 $\beta$ 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 $\alpha$ -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 $\beta$ 4 Thymosin $\beta$ 10 (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 (Veristat<sup>®</sup>, 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

M<sup>a</sup> 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-nitrobenzensulfonyl (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.