HHS Public Access

Author manuscript

Proteomics Clin Appl. Author manuscript; available in PMC 2017 February 22.

Published in final edited form as:

Proteomics Clin Appl. 2012 June ; 6(5-6): 244–256. doi:10.1002/prca.201100042.

Application of Glycoproteomics in the Discovery of Biomarkers for Lung Cancer

Qing Kay Li, MD, PhD [Assistant Professor], **Edward Gabrielson, MD [Professor]**, and **Hui Zhang, PhD [Associate Professor]**

Department of Pathology, the Johns Hopkins Medical Institutions, Baltimore, MD 21287

Abstract

Lung cancer is the leading cause of cancer-related deaths in the United States. Approximately 40–60% of lung cancer patients present with locally advanced or metastatic disease at the time of diagnosis. In order to improve the survival rate of lung cancer patients, the discovery of early diagnostic and prognostic biomarkers is urgently needed. Lung cancer development and progression are a multistep process which is characterized by abnormal gene and protein expressions ultimately leading to phenotypic change. In lung cancer, the expression of cellular glycoproteins directly reflects the physiological and/or pathological status of the lung parenchyma. Glycoproteins have long been recognized to play fundamental roles in many physiological and pathological processes, particularly in cancer genesis and progression. Although numerous papers have already acknowledged the importance of the discovery of cancer biomarkers, the systemic study of glycoproteins in lung cancer using glycoproteomic approaches is still suboptimal. Herein, we review the recent technological development of glycoproteomics in highlighting their utility and limitations for the discovery of glycoprotein biomarkers in lung cancer.

Keywords

glycoproteins; lung cancer; non-small cell lung cancer (NSCLC); glycoproteomics; protein biomarkers

1. Introduction

Lung cancer is the leading cause of cancer-related deaths in the United States [1]. Among all lung cancers, non-small cell lung cancer (NSCLC) accounts for approximately 85% of them [1,2]. Clinically, approximately 40–60% of NSCLC patients present with locally advanced or metastatic disease at the time of diagnosis [2,3]. Only a small portion of lung cancer is diagnosed at an early stage (stage I or II) when the disease is amenable by surgical resection [2,3]. In order to improve the early detection and outcome of lung cancer patients, several lung cancer screening clinical trials using highly sensitive image technology such as low-dose computed tomography (LDCT) have been implemented to surveillance high risk

Address Correspondence to: Qing Kay Li MD PhD, Department of Pathology, The Johns Hopkins Medical Institutions, Bayview Medical Center, 4940 Eastern Ave., Baltimore, MD 21224, Tel: (410)550-0671, Fax: (410)550-0075, qli23@jhmi.edu.

Conflict of interest

The authors declare no conflict of interest.

patients in developing lung cancers [4, 5]. These national trials have demonstrated an increased detection proportion of early stage lung cancers. However, even when NSCLC is detected at an early stage and a curative surgery is performed, about 37% of stage I lung cancer patients will still experience a recurrent disease eventually [2, 3]. Furthermore, although targeted therapies of lung caner have progressed rapidly, largely based upon the discovery of new molecular markers such as *EGFR* (epidermal growth factor receptor) and *KRAS* (V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog) mutations and *EML4-ALK* (echinoderm microtubule-associated protein-like 4 (*EML4*) and the anaplastic lymphoma kinase (*ALK*)) gene fusion [6–8], the overall progression-free survival rate of patients is still suboptimal [5, 9]. Therefore, in order to improve clinical outcomes of NSCLC patients, both early detection and identification of patients with a high risk of tumor progression must take place together.

Lung cancer development and progression are a multistep process. It is characterized by aberrant genetic and protein expressions, which subsequently lead to phenotypic transformation of cells and progression of the tumor. This process involves complex intracellular signaling pathways and various cellular proteins [10–13]. Over the past decade, numerous studies have been published to report findings of candidate protein biomarkers in the lung cancers. For example, a keyword search of “protein biomarkers and lung cancer” on PubMed has retrieved more than eleven thousands articles; approximately ten thousands articles are lung cancer studies related to glycoproteins (Table 1). Among them, more than six hundreds articles are related to predictive biomarkers, and more than three hundreds articles are related to prognostic biomarkers. These large numbers of data demonstrate the general efforts and interests in the discovery of potential protein biomarkers for detection and monitoring progression of lung cancer [14–22] (Table 1).

More than 50% of cellular proteins, including most secreted proteins, cell surface and intracellular proteins are glycoproteins. The glycosylation is one of the most critical post-translation modifications [23, 24]. Glycoproteins play critical roles in the regulation of cellular functions, including cell growth, differentiation and migration [25–27]. In lung cancer, the expression of glycoproteins may directly reflect the physiological and/or pathological status of the lung parenchyma. Furthermore, lung cancers have significant molecular heterogeneity and involve a large number of genetic and protein alterations, individual protein biomarker is unlikely to be representative in all NSCLC. Therefore, study of glycoprotein profile is particularly important for understand lung cancer biology and the identification of candidate protein biomarkers [28–31].

Recent advances in high throughput glycoproteomics allow for a way to evaluate thousands of glycoproteins in a single experiment. These state-of-art technologies provide new platforms for systemic study of glycoproteins and characterization of the complex alveolar microenvironment in lung cancers. However, it is worth noticing that there are only a handful of papers published in profiling glycoproteins of lung cancer using glycoproteomics (Table 1).

In this review, we focus on the discussion of recent advances of glycoproteomics in the field of lung cancers. The purpose of this review is to summarize current strategies and methods

used in glycoproteomics, and recent achievements toward clinical application of these approaches in studying lung cancer. This important insight also serves as another resource to understand the pathogenesis of lung cancer, and facilitate the identification of tumor-associated protein biomarkers.

2. Glycoproteomics Techniques

Glycosylation is one of the most common post-translational modifications in proteins. More than 50% of known eukaryotic proteins are glycosylated [23,24]. Glycoproteins play important roles in the regulation of cellular functions, differentiation and migration as well as angiogenesis [25–27]. They involve in cell-to-cell interactions, cancer cell growth, invasion and metastasis [28–31]. In O-glycosylation, glycans are attached to serine or threonine, whereas in N-glycosylation, glycans are attached to asparagines [25,32,33]. In eukaryotic cells, O- and N-linked glycosylations are mediated through different enzymatic and biosynthetic pathways, which regulate different cellular functions [25,32]. For example, in the study of functional role of O- and N-glycosylated glycoprotein hormones, Fares F *et al.* have shown that the O-linked glycosylation is critical in regulating the half-life of glycoprotein hormones whereas the N-linked glycosylation plays an important role in regulating the receptor binding and bioactivity of glycoprotein hormones [33]. In addition, precise glycosylation of proteins is heavily required for cancer-related biological processes [25]. Cancer cells are known to express aberrant glycosylation patterns such as increased branching, sialylation and/or fucosylation of N-linked glycans, and truncation of O-linked glycans [34–36].

To examine disease related changes in glycoproteins, efforts have been focused on the discovery of sensitive and robust analytic technologies, which ultimately lead to development of several high throughput methods [36–43]. These analytical approaches can broadly be divided into two categories: glycoprotein-based and glycopeptides-based analyses. The workflow of two approaches is summarized in Figure 1. In the glycoprotein-based approach, glycoproteins are enriched first by several techniques, such as size exclusion, ion exchange chromatography, affinity chromatography and chemical immobilization. Enriched glycoproteins are then digested into peptides and analyzed by mass spectrometry (MS). This approach can characterize the primary structure of the glycoprotein; however, it may not be able to identify glycosites, where glycans attach to amino acids in the glycoprotein backbone [30]. In the glycopeptide-based approach, on the other hand, the peptide mixture is enriched for glycopeptides after the proteins are digested enzymatically and/or chemically. Then, glycopeptides are deglycosylated and identified by MS analysis. This strategy has been widely used not only for identification of glycoproteins but also for identification of glycosylation sites [30].

No matter what approaches are used in the glycoproteomic analysis, it is essential to enrich glycoproteins or glycopeptides in complex biological samples prior to MS analysis. Two major strategies have been developed to enrich glycoproteins or glycopeptides: chemical capture and/or affinity capture based on the glycan moieties using affinity reagents [30,40–43].

2.1 Enrichment by Chemical Capture or Immobilization

Zhang H *et al.* has developed a method based on the oxidation of glycans to aldehydes and subsequent reactions between aldehyde groups and hydrazide on solid support using solid-phase extraction of N-linked glycopeptides (SPEG) [40]. In the experiment, the glycans on the glycoprotein are oxidized to form aldehyde groups, then, the newly formed aldehyde reacts with hydrazide, which is attached to a solid support and forms covalent hydrazone bonds. Using SPEG, both complex N-linked and O-linked glycoproteins or glycopeptides can be conjugated to solid support via covalent bonds [40,41]. The N-linked glycopeptides can be released by enzyme peptide-N-glycosidase (PNGase) [40,41]; however, O-linked glycopeptides cannot be easily released from the solid support due to the lack of specific enzymes. Chemical approaches for releasing O-linked glycosylated peptides such as alkaline beta-elimination have been used with a limited success [44, 45].

The advantage of chemical immobilization on solid-phase extraction is that this method has more than 90% specificity in the identification of N-linked glycopeptides from complex biological samples [37,40,41]. The method can detects approximately one to two de-glycosylated peptides from each glycoprotein in the peptide mixture, thus it reduces the complexity and provides an accurate analysis for low abundant glycoproteins. In comparison to other enrichment methods including lectin affinity chromatography, the chemical immobilization on solid-phase extraction has a highly specific capture of glycopeptides based upon the covalent conjugation between glycans and solid support. One limitation of the method is if glycopeptides are not detected by MS, the corresponding glycoproteins will not be identified from the isolation step. Currently, the sensitivity of the detection is markedly improved by using multiple enzymatic digestions of glycoproteins, such as using trypsin, pepsin and thermolysin during isolation step prior to glycopeptide capturing step [40,46,47].

2.2 Enrichment by Lectin Affinity Capture

Lectins are carbohydrate binding proteins. They bind to oligosaccharide epitopes on glycoproteins [48,49]. Based on their binding specificities, various lectins are currently used for enrichment of glycoproteins or glycopeptides [50–54]. The carbohydrate specificities of lectins are summarized in Table 2. Lectins with a relatively broad specificity such as wheat germ agglutinin (WGA) and concanavalin A (con A) are preferentially used for enriching proteins containing glycans interacting with specific lectins [51,52]. For example, Con A has specificities for mannose, glucose and galactose residues (Table 2). In lectin affinity chromatography-based methods, enrichment with certain lectin is shown to be particularly useful for identifying glycoproteins or glycopeptides with particular glycan structures [48,49].

The profiling of glycoprotein is a necessary step in the discovery of protein biomarkers. The limitation of selective capturing a subset of glycoproteins with a given lectin column can be overcome by a technique that involves multiple-lectins chromatography [52–54]. Recently, a multi-lectin affinity column has been developed that allows for an almost complete enrichment of glycoproteins from biological fluids [52–53]. In addition, lectin microcolumns have been developed for high-pressure analytical schemes. These

microcolumns are directly coupled on-line to reversed-phased HPLC (high-pressure liquid chromatography) in generating highly sensitive semi-automated profiling of glycoproteins [46].

Lectin affinity approach is relative simple and easy to use. A single, combination or series of lectin affinity columns may be used during the enrichment process. However, it has a limitation of non-specific binding of non-glycoproteins or non-glycopeptides to the lectin column. Pan S. *et al* has conducted a study using both hydrazide chemistry immobilization and lectin affinity column for enrichment of glycoproteins in the cerebrospinal fluid (CSF) [55]. Disease-related glycoproteins in the CSF are usually low-abundance proteins; therefore, in order to comprehensive characterization of CSF proteome, they have compared the capturing specificity and capability of these two methods. In the study, they have found that the hydrazide chemical immobilization method had a higher specificity than that of the lectin affinity method. They have also found that the combination of these two methods can greatly increase the detection ability of glycoproteins in CSF.

Finally, different subsets of glycoproteins can be enriched by the lectin affinity chromatography as opposed to the chemical immobilization method. From the study of rat liver membrane glycoproteins by Lee A *et al.*, glycoproteins enriched by lectins are mostly of high molecular weight proteins which are involved in intracellular signal transduction and cell adhesion. Conversely, glycoproteins enriched by hydrazine chemistry are mostly of low molecular weight proteins which function as enzymes [56].

Taken together, these two approaches, i.e. chemical immobilization and lectin affinity capture, use different mechanism and are complementary to each other for glycoprotein enrichment.

2.3 Boronic Acid Capture

Glycoprotein enrichment can also be achieved by the reaction with boronic acid [57,58]. The principle of this method is based on the formation of stable boronic diesters by the reaction of geminal diols under basic conditions. Recently, Sparbier K *et al.* have successfully detected low-abundance glycoproteins in human blood samples using this method [57]. Xu Y *et al.* have synthesized a novel diboronic acid functionalized mesoporous silica material (FDU-12-GA) and used it for specific glycopeptide enrichment [58]. Their data has demonstrated a markedly improvement of detection of glycopeptides by using this method.

2.4 Other Methods

During the glycoprotein enrichment, non-glycoprotein may interact with glycoprotein. Hagglund P *et al* have used a hydrophilic interaction liquid chromatography (HILIC) to reduce the non-specific interaction of proteins and the complexity of peptide/glycopeptide mixtures through depletion of hydrophobic peptides and retention of hydrophilic glycopeptides [59]. They enable to detect glycoprotein from plasma samples using hydrophilic interaction solid-phase extraction [59]. In addition, Alvarez-Manilla G *et al.* have used the size-exclusion chromatography (SEC) to enrich N-linked glycopeptides. They have found a three-fold increase in the total number of glycopeptides indentified in sera by LC-MS/MS [60].

3. Type of Specimen for Glycoproteomic Analysis in Lung Cancer: Specimen Collection and Limitations

The main function of the lung is to contain and exchange air to the vascular system [61]. Figure 2 represents the anatomical structures in normal lung and the major subtypes of NSCLC such as adenocarcinoma and squamous cell carcinoma. The lung parenchyma is composed of vascular structures, stroma and several types of highly differentiated cells including bronchial epithelium, alveolar type I and type II pneumocytes [61]. In the right panels of Figure 2, well-differentiated adenocarcinoma and squamous-cell carcinoma are shown. These two types of lung cancer are the most common types of NSCLC [2]. Within tumor tissues, various amounts of inflammatory cells and necrosis are commonly present depending on the histological subtype and the differentiation of the tumors. If the tumor is poorly differentiated, the tumor cells may be scant and presented as isolated clusters. When analyzing the tumor tissue with glycoproteomic approaches, stromal and inflammatory cells within tumor tissue must be avoided to reduce the contamination from non-tumor components.

Several types of biological samples can be used in the study of lung cancers. Blood and/or serum are well known specimens for the study of potential biomarkers in cancer patients [30,31,34,35,62,63]. They are relatively easy to obtain from patients. However, the tumor-specific biomarkers are usually present as low-abundance proteins which are difficult to detect by conventional MS methods. Candidate protein biomarkers may be differentially expressed during tumor progression. Certain biomarkers may not be expressed by the early stage of the tumor and cannot be detected in the blood. Therefore, in addition to using serum or plasma, tumor tissues and fluids from lung airway (bronchoalveolar lavage) and/or body cavity (pleural effusion) have been used for study of potential biomarkers. Data have shown successes in identifying glycoproteins from these specimens. Because of the abundance and enrichment of proteins within these samples, they are particularly useful in the identification and validation of potential biomarkers of lung cancers.

3.1 Tumor Tissue

Tumor tissue is a commonly used specimen in general proteomics [10–22], glycoproteins have not been studied in depth for lung cancers. Theoretically, tumor tissues can be obtained from all NSCLC patients by several techniques such as transbronchial fine needle aspiration biopsy (TBNA), transthoracic fine needle biopsy (FNA) and surgical resection of tumor. Among these techniques, the surgical resected tumor tissue is preferred [10–13,19–21] for its ability to provide large volume of tumor samples. Tumor tissues can be stored as fresh-frozen or formalin-fixed [64,65] and/or paraffin-embedded (FFPE) in laboratory [39–41,66]. When extracting glycoproteins or glycopeptides using a direct tissue lysis, contamination from the blood cells, inflammatory cells and plasma is a major concern. Therefore, selective isolation of cancer cells from tumor tissue should be considered to limit non-specific contamination during analytic process. The most effective method for isolating tumor cells is laser capture microdissection (LCM) [15,19,20]. Tumor cells can be isolated using LCM with high purity, however, the technique requires special instrumentation and time-consuming. The major drawback of the technique is that it only yields a limited quantity of

tumor sample. Another alternation is to micro-dissect tumor tissue using microscope and analyses glycoproteins from selective areas of the tumor [37,40,64–66]. The technique of micro-dissection of tumor tissue is more convenient and provides a relatively large quantity of tumor samples. Currently, the micro-dissection of tumor tissues is the technique used by most of proteomic studies.

3.2 Bronchoalveolar lavage fluid (BAL)

In addition to using tumor tissue, the other promising biological material for the study of biomarkers in lung disease and cancer is bronchoalveolar lavage fluid (BAL) [67–70]. Proteins in BAL are derived from secretion or leakage from lung parenchymal cells [67–70]. The protein levels directly reflect the physiological or pathological status of the lung. As aforementioned, most extracellular proteins are glycoproteins, such as secreted proteins, cell surface receptors and intracellular proteins [23–25,28–31]. The analysis of glycoproteins can characterize the complex alveolar microenvironment and provide protein profile in the discovery of potential biomarkers of lung diseases and cancers. BAL specimen is much easier to obtain than tumor tissue, and is collected by brochoscopy [67–70]. The proteomic analysis of BAL specimen has been applied to the study of a variety of benign lung diseases such as asthma and interstitial lung disease [67,68]. Thus, proteomic analyses, particularly glycoproteomic approaches, provide important insights into the pathogenesis of lung diseases and cancers, which can lead to the identification of tumor-associated glycoprotein biomarkers. By using this approach, airway glycoproteins can be recovered from a large area of lung parenchyma. This is especially important in the study of preinvasive and early cancer, since these lesions may not have visible histological changes under bronchoscopy. It is also an important method to study peripheral located lung cancers (particularly adenocarcinomas), since adenocarcinomas arise from lung parenchyma away from main bronchus and may not be reached by bronchoscopic biopsy needles.

3.3 Pleural Effusion

The surface of the lung is covered by a thin layer of pleura, which is frequently involved by lung cancer during tumor progression. The pleura are lined by a single layer of mesothelial cells, which covers the surface of lung (visceral surface) and inner surface of the chest wall (parietal surface). The pleural cavity is formed between these two layers of cells. Normally, the pleural cavity contains only a small amount of fluid to lubricate the visceral and parietal surface when they move against each other during respiration. The protein composition in the pleural effusion is similar to that of the plasma [61]. In the presence of variety of diseases, particularly when lung cancer metastases to the pleura, a larger amount of fluid, known as pleural effusion, can accumulate in the pleural cavity, due to the increased leakage of protein and/or decreased reabsorption of fluid [61]. Pleural effusion has been considered as a biological specimen with an enrichment of tumor-derived proteins in lung cancers [71,72]. Proteomic analysis may also provide an important insight into tumor-related biomarkers during lung cancer progression [71,72]. Other advantages of using pleural effusion are: (a) easily to obtain by the thoracentesis, (b) to have a minimal risk for the patient, (c) to provide a large quantity of samples, and (d) can be performed repeatedly during disease progression for studying progressive biomarkers.

3.4 Exhaled Breath Condensate (EBC)

EBC has been used as a biological sample in the study of benign lung diseases such as asthma and chronic obstructive pulmonary disease (COPD) as well as lung cancer [73]. EBC is easy to collect from mouth and nose of the patient via a special colleting instrument. The procedure is non-invasive and safe to the patient. The EBC contains a range of nonvolatile substances, including macromolecules such as proteins; thus, a change of protein expression in EBC may represent the physiologic and pathologic status of the lung. For example, vascular endothelial growth factor (VEGF), the tumor necrosis factor-α, and 8-isoprostane are found to be elevated in the EBC in NSCLC patients [74]. However, the abundance of proteins is usually low in EBC, which limits its utility in the proteomics. The detection of low-abundance protein is dependent not only on the sensitivity of the technique, but also on the complexity of the proteins. Glycoproteomics provide a highly sensitive platform to study these low abundance proteins. Thus, the analysis of glycoproteins in EBC may aid to identify potential biomarkers that are not necessarily detected by routine analytic methods.

3.5 Lung Cancer Cell Lines

Using lung cancer cell lines in studying protein signatures can help us understand the mechanism and intracellular signal pathways involved in lung cancers, particularly the mechanism of cancer progression. Lung cancer cell lines, non-metastatic CL1-0 and highly metastatic CL1–5, have been used in various studies of lung cancer progression and metastasis [75]. In the study of Tian T *et al*, thirty-three proteins are found to be differentially expressed between CL1–5 and CL1-0; among them 16 proteins are up-regulated and 17 proteins are down-regulated. In the group of elevated proteins, the high level of S100A11 is shown to correlate with the positive nodal status in lung cancer, indicating that S100A11 may be an important regulatory protein in tumor progression and metastasis [75]. Xu A. *et al*. have utilized the same cell lines to study the role of 14 kDa phosphohistidine phosphatase (PHP14) in metastatic lung cancer, and found that PHP14 regulates intracellular cytoskeleton reorganization and cancer cell migration [76]. The advantages of using lung cancer cell lines are several folds. Lung cancer cell lines can provide a relatively pure population of tumor cells and minimize non-tumor components during the assay. They are easy to grow and maintain in the laboratory and without the potential risk of contracting infectious diseases from clinical samples. However, it should be noted that the complexity and branching of the N-glycan in cell line may be different than that in vivo. Although it exhibits high fidelity in recapitulating tumor-specific phenotypes, cancer cell line may express high level of N-glycan than that of tumor cell in vivo [77]. In addition, the glycosylation of cell line may also be affected by cell culture condition.

4. Application of Glycoproteomics in the Discovery of Biomarkers in Lung Cancer

By definition, protein biomarker is an indicator of normal biological or pathological processes, disease progression or pharmacological responses to a therapeutic intervention [78]. From a biochemical point of view, a biomarker may be defined as an objectively measured biomolecule and its level changes significantly in a specific disease and during the

disease progression [79]. Potential biomarkers should be able to predict the biological behavior of the disease or cancer and the probability of the disease in response to the chemotherapy [78,79]. Although thousands of publications have highlighted numerous potential biomarkers in lung cancer research over the years, the systemic analysis of glycoproteins by glycoproteomics is still suboptimal in the field (Table 1). Table 3 summarizes the current data in lung cancers using glycoproteomic approaches. These data demonstrates the importance of glycoprotein discovery in lung cancer. However, glycoproteomics study of lung cancer is just emerging and the results from initial studies (8 studies in table 3) are difficult to compare due to (1) different experimental approaches used by different investigators, (2) different samples from different pathological stages of tumors, and (3) different histological subtypes of lung cancer. Among these studies, there is no overlap of glycoproteins identified among serum, fresh tumor tissue and pleural effusion samples. Nevertheless, several glycoproteins identified from these studies have been previously reported. For example, annexin is detected in the fresh tumor tissue by glycoproteomics [65], and has been reported previously [19.20]. Mucin protein has been detected in the pleural effusion [72], and has also been reported previously [12,13]. Further large scale studies, including a well selected sample set, are necessary to profile glycoproteins in lung cancer.

4.1 Diagnostic Biomarkers

If the expression of a biomarker is specifically and directly correlated with the presence of the disease, it can be considered as a diagnostic biomarker [79]. The most commonly used approach in identifying diagnostic biomarkers or lung cancer specific glycoproteins is comparing the glycoprotein profiles between tumor and normal lung tissue [64,65]. The normal lung samples can be from a patient-matched normal lung tissue adjacent to tumor, or from unrelated benign lung disease controls. In addition, another strategy for studying diagnostic markers is to directly profile glycoproteins using serum or plasma from cancer patients, and compare the result with healthy or non-cancer controls [80–84].

By using hydrazide chemistry on solid-phase extraction, Zeng X. *et al.* have studied pooled sera from NSCLC patients. After hydrazide chemistry enrichment and high resolution LC-MS/MS analysis, they have found that twenty-two proteins are differentially expressed in NSCLC patients, compared to individuals without NSCLC [80]. In the study, they enable to further verify three glycoproteins, alpha-1-antichymotrypsin, insulin-like growth factor-binding protein and lipocalin-type prostaglandin D synthase, using commercially available enzyme-linked immunosorbent assay (ELISA) kits [80]. They have also performed a hierarchical clustering analysis of glycoprotein expression between different cancer types. Their data indicate that the identified glycoprotein biomarkers may be used in the separation of NSCLC from controls.

Hongsachart P. *et al.* have used WGA lectin affinity enrichment followed by co-immunoprecipitation, in-gel electrophoresis and MS analysis to identify serum biomarkers [81]. In this study, they have analyzed ten serum samples from stage II/III lung adenocarcinoma patients, and found that twenty-seven glycoproteins are up-regulated and twelve are down-regulated in comparison to the healthy controls. The three up-regulated

proteins, adiponectin, cerulolasmin and glycosylphosphatidyl-inositol-80, and two down-regulated glycoproteins, cyclin H and proto-oncogene tyrosine-protein kinase Fyn, are verified by Western blot analysis. These data highlight the potential utility of glycoprotein biomarkers in the early detection of lung cancers.

4.2 Prognostic and Predictive Biomarkers

The progression of lung cancer is a multi-step process. It is believed that tumor at a later clinical stage is more aggressive than a tumor at an early stage; therefore, tumor at different stages may express a unique subset of proteins that can be used in monitoring tumor progression. If a biomarker is correlated with the natural history of the disease and reflects the tumor progression such as invasiveness or metastatic potential independent of the therapeutic intervention, it is defined as prognostic biomarker [29]. On the other hand, a biomarker is defined as predictive if the efficacy of a specific therapy on the tumor can be predicted by using the biomarker [29,79]. The most commonly used strategy to identify such potential biomarkers is to compare protein expression from different stage of tumor tissue or in the blood from patients responding or not responding to a certain treatment. An alternative approach is to study protein expression within different stages of tumors and to correlate them with patients' clinical survival rates. Glycoproteomics is particular useful in discovering prognostic and predictive biomarkers, since it allows not only for evaluating the differential expression of glycoproteins, but also for identifying the potential changes of glycosylation sites during tumor progression.

In a study of fresh frozen lung adenocarcinoma and patient-matched normal lung tissue, Rho JH *et al.* have used a comprehensive glycoproteomic enrichment by lectins of ConA, WGA and amylas inhibitor-like protein (AIL), then, analyzed glycoproteins by 2-D PAGE and MS/MS approaches [64,65]. They have analyzed both N-glycoproteins (by ConA- and WGA-capture) and O-glycoproteins (by AIL-capture) in sixteen lung adenocarcinomas and matched controls [65]. They have found that eight glycoproteins are up-regulated, including alpha1-antitrypsin, fructose-bisphosphate aldolase A, annexin A1, calreticulin, alpha-enolase, protein disulfide isomerase A1, proteasome subunit beta type1, and mitochondrial superoxide dismutase. In comparison, seven glycoproteins are down-regulated including annexin A3, carbonic anhydrase 2, fetuin A, hemoglobin subunit beta, peroxiredoxin-2, receptor for advanced glycosylation end products and vimentin. They have also found a high level of mannose glycan structures on fetuin A protein in cancers, but not in normal tissue [65]. In addition, they have also identified that transgelin is overexpressed in stromal compartment whereas transgelin-2 is overexpressed in lung cancer tissue [64].

Soltermann A. *et al.* have applied the hydrazide chemistry capture and LC-MS/MS approach in analyzing five cases of pleural effusions from metastatic lung adenocarcinoma patients [72]. In the study, they have identified three lung-specific proteins, mucin 5B, thyroid transcript factor 1 (TTF1) and surfactant protein A. They have also found six tumor-progression or metastasis-associated glycoproteins, including CA-125, CD44, CD166, lysosomal-associated membrane glycoprotein 2 (LAMP-2), multimerin 2, and periostin [72]. In their study, pleural effusion from clinically cancer free (for a period of one year) patients are used as controls. Although they are able to identify potential progression biomarker in

the pleural effusion, the utility of cancer free patients as controls may have been a setback, since these patients may harbor cancer cells that are not detectable by current clinical tests. Nevertheless, their data indicate the potential utility of pleural effusion in the discovery of progression biomarkers.

Glycan structures of glycoprotein are organ-specific and correlate with disease states. In order to identify N-linked glycosylation sites between glycans and glycoproteins, Ueda K. *et al.* have used an approach called the isotopic glycosidase elution and labeling on lectin-column chromatography (IGEL) [82]. This technique is based on the lectin affinity chromatography and site-directed tagging of N-linked glycosylation sites by ^{18}O during the elution with N-glycosidase [36,43], and the combination of the IGEL method with iTRAQ stable isotope labeling [82]. They are able not only to identify N-glycosylation sites, but also to quantitatively compare glycan structures on each glycosylation site in a single LC-MS/MS analysis. They have studied eight sera samples from lung adenocarcinoma patients and found that six peptides have shown more than two-fold affinity to ConA lectin whereas two peptides have shown less than 0.5-fold affinity to ConA lectin. Among these glycoproteins, TIMP1 (metalloproteinase inhibitor 1) has shown an up-regulated modification of high-mannose motif in stage IV lung cancer in comparison to stage I/II lung cancer and controls [82]. In addition to profile serum glycoproteomic using lectin chromatography described above, Ueda K *et al.* have also performed SELDI-TOF MS (surface-enhanced laser desorption/ionization-time-of-flight mass spectrometry) analysis using lectin-coupled ProteinChip (Jacalin or SNA lectins), and found a higher frequency of loss of SNA binding in serum apolipoprotein C-III, indicating a cancer-associated aberrant glycosylation in NSCLC patients [83].

Recently, Tsai HY *et al.* have studied the alteration of glycans in sera of lung cancer patients using 2-DE, Western blot analysis and lectin staining, and MALDI (matrix-assisted laser desorption/ionization) MS and MS/MS analysis [84]. They have found that the fucosylated haptoglobin significantly increased in serum of lung cancer patients, compared to the normal controls. Further analysis has shown that the alteration of glycan in the cancer serum samples is primarily on the trisialylated triantennary N-glycan of haptoglobin [84]. Their data have suggested that specific glycans may be associated with certain subtype of NSCLC. In addition of glycoproteins, aberrant glycan may be used as a potential biomarker to monitor lung cancer progression. In conclusion, signature studies of glycoproteins provide tremendous potential to discover prognostic and predictive biomarkers in lung cancers.

4.3 Potential Limitations

In glycoproteomic biomarker discovery, it is important to use clinical materials in both discovery process and subsequent validation phase. Despite the rapid progress in glycoproteomics, the workflow in analysis of clinical samples hinders the necessary throughput in a large scale studies. Most current glycoproteomic studies are performed using a limited number of clinical samples (Table 3). Therefore, the improvement of analytic throughput ability is needed for study of a large scale of patient cohort. In addition, technologies still need to be further improved in terms of accuracy and sensitivity in measurements of clinical material. This is evident from the serum glycoproteomic studies

where the analysis of dynamic range and glycoproteome coverage need to be more sufficient, particularly in large scale of patient cohorts. Finally, the potential success of glycoproteomic biomarker study largely depends on the quality and availability of patient cohorts. This requires large number of carefully selected patient cohorts to determine the potential utility of the biomarker. Clinical validation of potential glycoprotein biomarkers must be conducted in a fashion to avoid the occurrence of false positive results. For each candidate biomarker, robust and reproducible assay need to be developed and used in the validation phase.

5. Future Role of Glycoproteomics in Lung Cancer

In the United States, ninety-four million current or former smokers are at a high risk for developing lung cancers [1]. The estimated new cases of lung cancer have already reached 222,520 in the US in 2010, in addition to the statistics on lung cancer-related death being the number one cause of cancer-related death in both man and woman [1]. In order to increase survival rate and decrease mortality in lung cancer patients, potential biomarkers for early diagnosing and monitoring lung cancer progression are urgently needed. The recent advances of glycoproteomics technology clearly facilitate the discovery of novel biomarkers in lung cancer. The advances of glycoproteome have significantly improved our knowledge in the field of lung cancer biology. It also promotes the potential clinical utility of them in the targeted therapy of lung cancer.

5.1 Contribution to Personalized Lung Cancer Therapy and Potential Drug Development

In the past decade, the genetic profile of cancer cells has changed the anti-cancer treatment toward a targeted and personalized therapy [6,7]. To facilitate the new era of personalized medicine, the profile of glycoprotein by glycoproteomics in lung cancer may play a major role. Glycoproteomic analysis can further define glycoprotein signature related to early diagnosis (cancer versus benign disease), prognosis (likelihood of cure or risk of progression and metastasis), and prediction (probability of response to therapy) in lung cancers. Therefore, glycoprotein profiling provides a strong candidacy to the next step in discovering potential biomarkers for early diagnosis and targeted therapy.

5.2 Monitoring the Chemotherapy Response

In lung cancers, the *EGFR* mutation is associated with a 70–80% response rate to tyrosine-kinase inhibitors (TKIs) therapy and a longer progression free survival rate in patients [8,9]. The growth factor receptors are *N*-glycosylated transmembrane proteins, their biological functions are in part regulated by intracellular endogenous lectins, such as galectins [85,86]. Galecin-3 has been shown to bind to *N*-glycans of EGFR and limits its distribution on the plasma membrane [86]. Furthermore, deletional mutation of *N*-glycan sites and variation of *N*-glycan induce receptor dimerization and signaling [87,88]. Recent data also suggests that the function of EGFR may be regulated by sialylation and fucosylation [89]. These data indicate that the interaction of *N*-glycan and receptor protein may regulate the distribution and residency of the growth receptor on the cell membrane. The study may help us understanding the mechanism of drug resistance in lung cancer, particularly among patients who are treated with *EGFR* inhibitors.

5.3 Chemoprevention of lung cancer

Recent data have shown that lung cancer is the result of accumulation of genotypic and phenotypic abnormalities, and only a minority of preinvasive lung lesions progress to invasive cancer [78,90]. These preinvasive lesions can be subtyped into the mild, moderate and server dysplasia and carcinoma in situ. Studies using serial bronchoscopic biopsies have suggested that 3.5% of mild or moderate dysplasias progressed to severe dysplasia, 37% of severe dysplasias remains or progress, and 50% of carcinoma in situ progress to invasive carcinoma within a two- to three-year period [78, 9]. Currently, several clinical trials to treat these patients with bronchial epithelium dysplasia (chemoprevention) have shown the regression of the lesion [78]. It is also known that differentiation and proliferation of cells are regulated by glycosylation [88]. Thus, the analysis of glycoprotein expression during the process may identify potential tumor-associated biomarkers. The development a quantitative measurement of probability of having lung cancer based on the glycoprotein analysis of the bronchial epithelium may have an important clinical implication. The identification of preinvasive lesion with a high risk of progression can also improve the early detection of lung cancer.

In summary, glycoproteomics presents itself as a prominent technology in the field of lung cancer research. Selected candidate biomarkers have been identified and studied in the small size of clinical samples. Further improvement of the workflow and validations are both needed in the discovery of lung cancer glycoprotein biomarkers.

Acknowledgments

This work is partially supported by Drs. Ji and Li Family Foundation (QKL), and the federal fund from the National Institutes of Health/the National Cancer Institute/Early Detection Research Network grant (NIH/NCI/EDRN) U01CA152813 (HZ). Authors thank Ms. Caitilin Choi and Makeda Heard for proofreading.

References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. CA Cancer J Clin. 2010; 60:227–300.
2. Travis WD. IASLC Staging Committee. Reporting lung cancer pathology specimens. Impact of the anticipated 7th Edition TNM classification based on recommendations of the IASLC Staging Committee. Histopathology. 2009; 54:3–11. Review. [PubMed: 19187176]
3. Harpole DH, Herndon JE, Young WG, Wolfe WG, Sabiston DC. Stage I non-small cell lung cancer. Cancer. 1995; 76:787–796. [PubMed: 8625181]
4. Swensen SJ, Jett JR, Hartman TE, Mandrekar SJ, Hillman SL, Sykes AM, Aughenbaugh GL, Bungum AO, Allen KL. CT screening for lung cancer: Five-year prospective experience. Radiology. 2005; 235:259–265. [PubMed: 15695622]
5. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD. National Lung Screening Trial Research Team. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. N Engl J Med. 2011; 365:395–409. Epub 2011 Jun 29. [PubMed: 21714641]
6. Mok TS, Wu YL, Thongprasent S, Yang CH, Chu DT, Saijo N, Sunpaweravong P, Han B, Margono B, Ichinose Y, Nishiwaki Y, Ohe Y, Yang JJ, Chewaskulyong B, Jiang H, Duffield EL, Watkins CL, Armour AA, Fukuoka M. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009; 361:947–945. [PubMed: 19692680]
7. Chen HY, Yu SL, Chen CH, Chang GC, Chen CY, Yuan A, Cheng CL, Wang CH, Terng HJ, Kao SF, Chan WK, Li HN, Liu CC, Singh S, Chen WJ, Chen JJW, Yang PC. A five-gene signature and

clinical outcome in non-small-cell lung cancer. N Engl J Med. 2007; 356:11–20. [PubMed: 17202451]

8. Coate LE, John T, Tsao MS, Shepherd FA. Molecular predictive and prognostic markers in non-small-cell lung cancer. Lancet Oncol. 2009; 10:1001–1010. [PubMed: 19796752]
9. Klabunde CN, Marcus PM, Silvestri GA, Han PK, Richards TB, Yuan G, Marcus SE, Vernon SWUS. primary care physicians' lung cancer screening beliefs and recommendations. Am J Prev Med. 2010; 39:411–420. [PubMed: 20965378]
10. Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, Roberts JR, Edgerton M, Gonzalez A, Nadaf S, Moore JH, Caprioli RM, Carbone DP. Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet. 2003; 362:433–439. [PubMed: 12927430]
11. Buhrens RI, Amelung JT, Reymond MA, Beshay M. Protein expression in human non-small cell lung cancer: a systematic database. Pathobiology. 2009; 76:277–285. [PubMed: 19955839]
12. Lehtio J, De Petris L. Lung cancer proteomics, clinical and technological considerations. J Proteome. 2010; 73:1851–1863.
13. Vegvari A, Marko-Varga G. Clinical protein science and bioanalytical mass spectrometry with an emphasis on lung cancer. Chem Rev. 2010; 110:3278–3298. [PubMed: 20415473]
14. Tian T, Hao J, Xu A, Hao J, 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. [PubMed: 17537172]
15. Yao H, Zhang Z, Xiao Z, Chen Y, Li C, Zhang P, Li M, Liu Y, Guan Y, Yu Y, Chen Z. Identification of metastasis associated proteins in human lung squamous carcinoma using two-dimensional difference gel electrophoresis and laser capture microdissection. Lung Cancer. 2009; 65:41–48. [PubMed: 19058872]
16. Rahman SM, Shyr Y, Yildiz PB, Gonzalez AL, Li H, Zhang X, Chaurand P, Yanagisawa K, Slovis BS, Miller RF, Ninan M, Miller YE, Franklin WA, Caprioli RM, Carbone DP, Massion PP. Proteomic patterns of preinvasive bronchial lesions. Am J Respir Crit Care Med. 2005; 172:1556–1562. [PubMed: 16179643]
17. 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 Investig. 2008; 26:128–135. [PubMed: 18259942]
18. 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; 63:1652–1656. [PubMed: 12670919]
19. Kawamura T, Nomura M, Tojo H, Fujii K, Hamasaki H, Mikami S, Bando Y, Kato H, Nishimura T. Proteomic analysis of laser-microdissected paraffin-embedded tissues: (1) stage-related protein candidates upon non-metastatic lung adenocarcinoma. J Proteome. 2010; 73:1089–1099.
20. Nishimura T, Nomura M, Tojo H, Hamasaki H, Fukuda T, Fujii K, Mikami S, Bando Y, Kato H. Proteomic analysis of laser-microdissected paraffin-embedded tissues: (2) MRM assay for stage-related proteins upon non-metastatic lung adenocarcinoma. J Proteome. 2010; 73:1100–1110.
21. Oshita F, Morita A, Ito H, Kameda Y, Tsuchiya E, Asai S, Miyagi Y. Proteomic screening of completely resected tumors in relation to survival in patients with stage I non-small cell lung cancer. Onco Rep. 2010; 24:637–645.
22. Hirano T, Kato H. Present status of clinical proteomic analysis for the early detection and determination of therapeutic strategy in lung cancer. Ann Thorac Cardiovasc Surg. 2006; 12:4–9. [PubMed: 16572067]
23. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim. Biophys. Acta. 1999; 1473:4–8. [PubMed: 10580125]
24. Wong CH. Protein glycosylation: new challenges and opportunities. J. Org. Chem. 2005; 70:4219–4225. [PubMed: 15903293]
25. Helenius A, Aebi M. Intracellular functions of N-linked glycans. Science. 2001; 291:2364–2369. [PubMed: 11269317]
26. Isaji T, Gu J, Nishiuchi R, Zhao Y, Takahashi M, Miyoshi E, Honke K, Sekiguchi K, Taniguchi N. Introduction of bisecting GlcNAc into integrin alpha5beta1 reduces ligand binding and down-

regulates cell adhesion and cell migration. J. Biol. Chem. 2004; 279:19747–19754. [PubMed: 14998999]

27. Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA. Glycosylation and the immune system. Science. 2001; 291:2370–2376. [PubMed: 11269318]
28. Lowe JB, Marth JD. A genetic approach to Mammalian glycan function. Annu. Rev. Biochem. 2003; 72:643–691. [PubMed: 12676797]
29. Drake PM, Cho W, Li B, Prakobphol A, Johansen E, Anderson NL, Regnier FE, Gibson BW, Fisher SJ. Sweetening the pot: adding glycosylation to the biomarker discovery equation. Clini. Chem. 2010; 56:223–236.
30. Tian Y, Zhang H. Glycoproteomics and clinical applications. Proteomics Clin. Appl. 2010; 4(124): 132.
31. Kam RKT, Poon TCW. The potentials of glycomics in biomarker discovery. Clin. Proteom. 2008; 4:67–79.
32. Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology. 1993; 3:97–130. [PubMed: 8490246]
33. Fares F. The role of O-linked and N-linked oligosaccharides on the structure-function of glycoprotein hormones: development of agonists and antagonists. Biochim. Biophys. Acta. 2006; 1760:560–567. [PubMed: 16527410]
34. Tong L, Baskaran G, Jones MB, Rhee JK, Yarema KJ. Glycosylation changes as markers for the diagnosis and treatment of human disease. Biotechnol. Genet. Eng. Rev. 2003; 20:199–244. [PubMed: 14997853]
35. Powlesland AS, Hitchen PG, Parry S, Graham SA, Barrio MM, Elola MT, Mordoh J, Dell A, Drickamer K, Taylor ME. Targeted glycoproteomic identification of cancer cell glycosylation. Glycobiology. 2009; 19:899–909. Epub 2009 May 11. [PubMed: 19433864]
36. Ueda K, Katagiri T, Shimada T, Irie S, Sato TA, Nakamura Y, Daigo Y. 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. J Proteome Res. 2007; 6:3475–3483. Epub 2007 Aug 18. [PubMed: 17705522]
37. Tian Y, Kelly-Spratt KS, Kemp CJ, Zhang H. Identification of glycoproteins from mouse skin tumors and plasma. Clin. Proteom. 2008; 4:117–136.
38. Kaji H, Isobe T. Liquid chromatography/mass spectrometry (LC/MS)-based glycoproteomics technologies for cancer biomarker discovery. Clin. Proteom. 2008; 4:14–24.
39. Zhang H, Cotter RJ. Glycoproteomics: New technology developments and applications provide renewed interest in glycoproteins. Clin. Ptoteom. 2008; 4:1–4.
40. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat. Biotechnol. 2003; 21:660–666. [PubMed: 12754519]
41. Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H. Solid phase extraction of N-linked glycopeptides. Nat. Protoc. 2007; 2:334–339. [PubMed: 17406594]
42. Sun B, Ranish JA, Utleg AG, White JT, Yan X, Lin B, Hood L. Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. Mol. Cell. Proteomics. 2007; 6:141–149. [PubMed: 17074749]
43. Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, Kasai K, Takahashi N, Isobe T. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. Nat. Biotechnol. 2003; 21:667–672. [PubMed: 12754521]
44. Zauner G, Koeleman CA, Deelder AM, Wuhrer M. Mass spectrometric O-glycan analysis after combined O-glycan release by beta-elimination and 1-phenyl-3-methyl-5-pyrazolone labeling. Biochim Biophys Acta. 2011 Jul.:20. [Epub ahead of print].
45. Zheng Y, Guo Z, Cai Z. Combination of beta-elimination and liquid chromatography/quadrupole time-of-flight mass spectrometry for the determination of O-glycosylation sites. Talanta. 2009; 78:358–363. [PubMed: 19203595]
46. Madera M, Mechref Y, Klouckova I, Novotny MV. Semiautomated high-sensitivity profiling of human blood serum glycoproteins through lectin preconcentration and multidimensional

chromatography/tandem mass spectrometry. J. Proteome Res. 2006; 5:2348–2363. [PubMed: 16944947]

47. Chen R, Jiang X, Sun D, Han G, Wang F, Ye M, Wang L, Zou H. Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. J. Proteome Res. 2009; 8:651–661. [PubMed: 19159218]
48. Hirabayashi J, Kuno A, Tateno H. Lectin-based structural glycomics: a practical approach to complex glycans. Electrophoresis. 2011; 32:1118–1128. [PubMed: 21544837]
49. Kuno A, Matsuda A, Ikehara Y, Narimatsu H, Hirabayashi J. Differential glycan profiling by lectin microarray targeting tissue specimens. Methods Enzymol. 2010; 478:165–179. [PubMed: 20816479]
50. Cummings RD, Kornfeld S. Fractionation of asparaginelinked oligosaccharides by serial lectin-Agarose affinity chromatography. A rapid, sensitive, and specific technique. J. Biol. Chem. 1982; 257:11235–11240. [PubMed: 7118881]
51. Wiener MC, van Hoek AN. A lectin screening method for membrane glycoproteins: application to the human CHIP28 water channel (AQP-1). Anal. Biochem. 1996; 241:267–268. [PubMed: 8921199]
52. Bunkenborg J, Pilch BJ, Podtelejnikov AV, Wisniewski JR. Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. Proteomics. 2004; 4:454–465. [PubMed: 14760718]
53. Yamamoto K, Tsuji T, Osawa T. Analysis of asparaginelinked oligosaccharides by sequential lectin-affinity chromatography. Methods Mol. Biol. 1998; 76:35–51. [PubMed: 9664345]
54. Wang Y, Wu SL, Hancock WS. Approaches to the study of N-linked glycoproteins in human plasma using lectin affinity chromatography and nano-HPLC coupled to electrospray linear ion trap--Fourier transform mass spectrometry. Glycobiology. 2006; 16:514–523. [PubMed: 16497783]
55. Pan S, Wang Y, Quinn JF, Peskind ER, Waichunas D, Wimberger JT, Jin J, Li JG, Zhu D, Pan C, Zhang J. Identification of glycoproteins in human cerebrospinal fluid with a complementary proteomic approach. J. Proteome Res. 2006; 5:2769–2779. [PubMed: 17022648]
56. Lee A, Kolarich D, Haynes PA, Jensen PH, Baker MS, Packer NH. Rat liver membrane glycoproteome: enrichment by phase partitioning and glycoprotein capture. J. Proteome Res. 2009; 8:770–781. [PubMed: 19125615]
57. Sparbier K, Wenzel T, Kostrzewa M. Exploring the binding profiles of ConA, boronic acid and WGA by MALDITOF/TOF MS and magnetic particles. J. Chromatogr. B Analyt Technol. Biomed. Life Sci. 2006; 840:29–36.
58. Xu Y, Wu Z, Zhang L, Lu H, Yang P, Webley PA, Zhao D. Highly specific enrichment of glycopeptides using boronic acid-functionalized mesoporous silica. Anal. Chem. 2009; 81:503–508. [PubMed: 19117470]
59. Hagglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P. A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. J. Proteome Res. 2004; 3:556–566. [PubMed: 15253437]
60. Alvarez-Manilla G, Atwood J, Guo Y, Warren NL, Orlando R, Pierce M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. J. Proteome Res. 2006; 5:701–708. [PubMed: 16512686]
61. Kobzik, L. Lung. In: Cotran, RS.Kumer, V., Collins, T., editors. Robbins pathologic basis of disease. 7th. Philadelphia: Saunders; 2004. p. 697-755.
62. Zhang H, Chan DW. Cancer biomarker discovery in plasma using a tussue-targeted proteomic approach. Cancer Epidemiol. Biomarkers Prev. 2007; 16:1915–1917. [PubMed: 17932335]
63. Li Y, Sokoll LJ, Barker PE, Zhang H, Chan DW. Mass spectrometric identification of proteotypic peptides from clinically used tumor markers. Clin. Proteom. 2008; 4:58–66.
64. 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; 8:5610–5618. [PubMed: 19848416]
65. Rho JH, Roehrl MH, Wang JY. Glycoproteomic analysis of human lung adenocarcinomas using glycoarrays and tandem mass spectrometry: differential expression and glycosylation patterns of vimentin and fetuin A isoforms. Protein J. 2009; 28:148–160. [PubMed: 19412661]

66. Tian Y, Gurley K, Meany DL, Kemp CJ, Zhang H. N-linked glycoproteomic analysis of formalin-fixed and paraffin-embedded tissues. J. Proteome. Res. 2009; 8:1657–1662. [PubMed: 19714870]

67. Magi B, Bargagli E, Bini L, Rottoli P. Proteome analysis of bronchoalveolar lavage in lung diseases. Proteomics. 2006; 6:6354–6369. [PubMed: 17133372]

68. Rottoli P, Bargagli E, Landi C, Magi B. Proteomic analysis in interstitial lung diseases: a review. Curr Opin Pulm Med. 2009; 15:470–478. [PubMed: 19593140]

69. Wattiez R, Falmagne P. Proteomics of BAL fluid. J Chromatogr B Analyt Technol Biomed Life Sci. 2005; 815:169–178.

70. Plymoth A, Löfdahl CG, Ekberg-Jansson A, Dahlbäck M, Lindberg H, Fehniger TE, Marko-Varga G. Human bronchoalveolar lavage: biofluid analysis with special emphasis on sample preparation. Proteomics. 2003; 3:962–972. [PubMed: 12833521]

71. Tyan YC, Wu HY, Lai WW, Su WC, Liao PC. Proteomic profiling of human pleural effusion using two-dimensional nano liquid chromatography tandem mass spectrometry. J Proteome Res. 2005; 4:1274–1286. [PubMed: 16083277]

72. Soltermann A, Ossola R, Kilgus-Hawelski S, von Eckardstein A, Suter T, Aebersold R, Moch H. N-glycoprotein profiling of lung adenocarcinoma pleural effusions by shotgun proteomics. Cancer Cytopathol. 2008; 114:124–133.

73. Chan HP, Lewis C, Thomas PS. Exhaled breath analysis: Novel approach for early detection of lung cancer. Lung Cancer. 2009; 63:164–168. [PubMed: 18599152]

74. Dalaveris E, Kerenidi T, Katsabeki-Katsafli A, Kiropoulos T, Tanou K, Gourgoulianis KI, Kostikas K. VEGF,TNF-alpha and 8-isoprostane levels in exhaled breath condensate and serum of patients with lung cancer. Lung Cancer. 2009; 64:219–225. [PubMed: 18845357]

75. 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 Sci. 2007; 98:1265–1274. [PubMed: 17537172]

76. 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; 67:48–56. [PubMed: 19344975]

77. West MB, Hanigan MH. γ-Glutamyl Transpeptidase is a heavily *N*-glycosylated heterodimer in HepG2 cells. Arch Biochem Biophys. 2010; 504:177–181. [PubMed: 20831856]

78. Kelloff GJ, Lippman SM, Dannenberg AJ, Sigman CC, Pearce HL, Reid BJ, Szabo E, Jordan VC, Spitz MR, Mills GB, Papadimitrakopoulou VA, Lotan R, Aggarwal BB, Bresalier RS, Kim J, Arun B, Lu KH, Thomas ME, Rhodes HE, Brewer MA, Follen M, Shin DM, Parnes HL, Siegfried JM, Evans AA, Blot WJ, Chow WH, Blount PL, Maley CC, Wang KK, Lam S, Lee JJ, Dubinett SM, Engstrom PF, Meyskens FL Jr, O'Shaughnessy J, Hawk ET, Levin B, Nelson WG, Hong WK. AACRTask Force on Cancer Prevention. Progress in chemoprevention drug development: the promise of molecular biomarkers for prevention of intraepithelial neoplasia and cnacer—a plan to move forward. Clin Cancer Res. 2006; 12:3661–3697. [PubMed: 16778094]

79. Boja E, Hiltke T, Rivers R, Kinsinger C, Bahbar A, Mesri M, Rodriguez H. Evolution of clinical proteomics and its role in medicine. J. Proteom. Res. 2011; 10:66–84.

80. Zeng X, Hood BL, Sun M, Conrads TP, Day RS, Weissfeld JL, Siegfried JM, Bigbee WL. Lung cancer serum biomarker discovery using glycoprotein capture and liquid chromatography mass spectrometry. J Proteome Res. 2010; 9:6440–6449. [PubMed: 20931982]

81. Hongsachart P, Huang-Liu R, Sinchaikul S, Pan FM, Phutrakul S, Chuang YM, Chen ST. Glycoproteomic analysis of WGA-bound glycoprotein biomarkers in sera from patients with ling adenocarcinoma. Electrophoresis. 2009; 30:1206–1220. [PubMed: 19294700]

82. Ueda K, Takami S, Saichi N, Daigo Y, Ishikawa N, Kohno N, Katsumata M, Yamane A, Ota M, Sato TA, Nakamura Y, Nakagawa H. Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. Mol Cell Proteomics. 2010; 9:1819–1828. [PubMed: 20811073]

83. Ueda K, Fukase Y, Katagiri T, Ishikawa N, Irie S, Sato TA, Ito H, Nakayama H, Miyagi Y, Tsuchiya E, Kohno N, Shiwa M, Nakamura Y, Daigo Y. Targeted serum glycoproteomics for the discovery of lung cancer-associated glycosylation disorders using lectin-coupled ProteinChip arrays. Proteomics. 2009; 9:2182–2192. [PubMed: 19322776]

84. Tsai HY, Boonyapranai K, Sriyam S, Yu CJ, Wu SW, Khoo KH, Phutrakul S, Chen ST. Glycoproteomics analysis to identify a glycoform on haptoglobin associated with lung cancer. Proteomics. 2011; 11:2162–2170. [PubMed: 21538882]

85. Lee RT, Lee YC. Affinity enhancement by multivalent lectin-carbohydrate interaction. Glycoconj J. 2000; 17:543–551. [PubMed: 11421347]

86. Lajoie P, Partrideg EA, Guay G, Goetz JG, Pawling J, Lagana A, Dennis JW, Nabi IR. Plasma membrane domain organization regulates EGFR signaling in tumor cells. J Cell Bio. 2007; 179:341–356. [PubMed: 17938246]

87. Fernandes H, Cohen S, Bishayee S. Glycosylation-induced conformational modification positively regulates receptor-receptor association: A study with an aberrant epidermal growth factor receptor (EGFRvIII/deltaEGFR) expressed in cancer cells. J Boil Chem. 2001; 276:5375–5383.

88. Lau K, Partridge EA, Silvescu CI, Grigorian A, Pawling J, Reinhold VN, Demetriou M, Dennis JW. Complex *N*-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. Cell. 2007; 129:123–124. [PubMed: 17418791]

89. Liu YC, Yen HY, Chen CY, Chen CH, Cheng PF, Juan YH, Chen CH, Khoo KH, Yu CJ, Yang PC, Hsu TL, Wong CH. Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. Proc Natl Acad Sci USA. 2011; 108:11332–11337. Epub 2011 Jun 27. [PubMed: 21709263]

90. Venmans BJ, van Boxem TJ, Smith EF, Postmus PE, Sutedja TG. Outcome of bronchial carcinoma in situ. Chest. 2000; 117:1572–1576. [PubMed: 10858385]

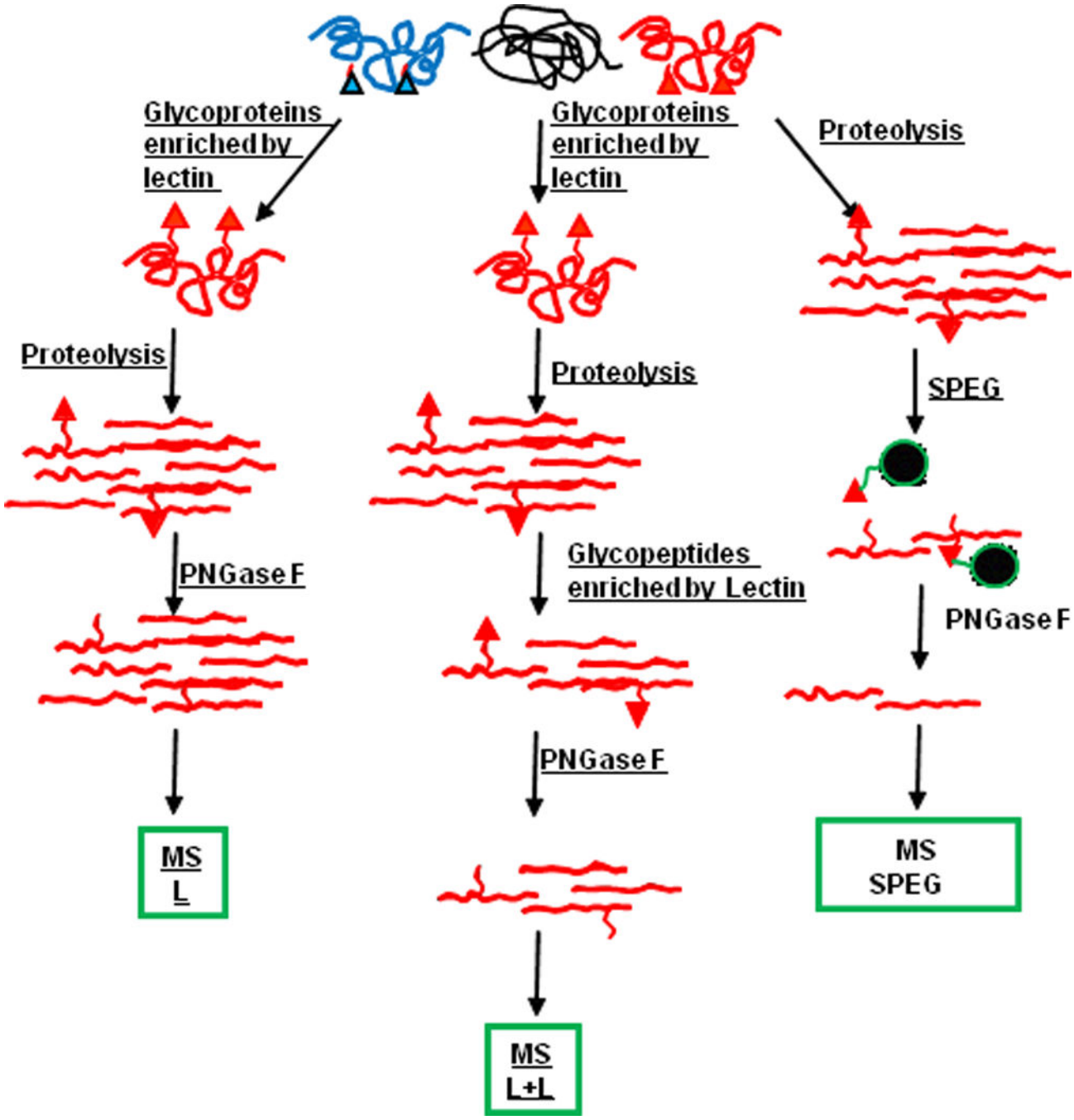

Figure 1.

The workflow of glycoprotein-based and glycopeptides-based analyses. Briefly, in the glycoprotein-based approach, glycoproteins are enriched first, then digested into peptides and analyzed by mass spectrometry (MS). In the glycopeptide-based approach, proteins are digested enzymatically and/or chemically; then, the peptide mixture is enriched for glycopeptides and identified by MS analysis. SPEG: solid-phase extraction of N-linked glycopeptides).

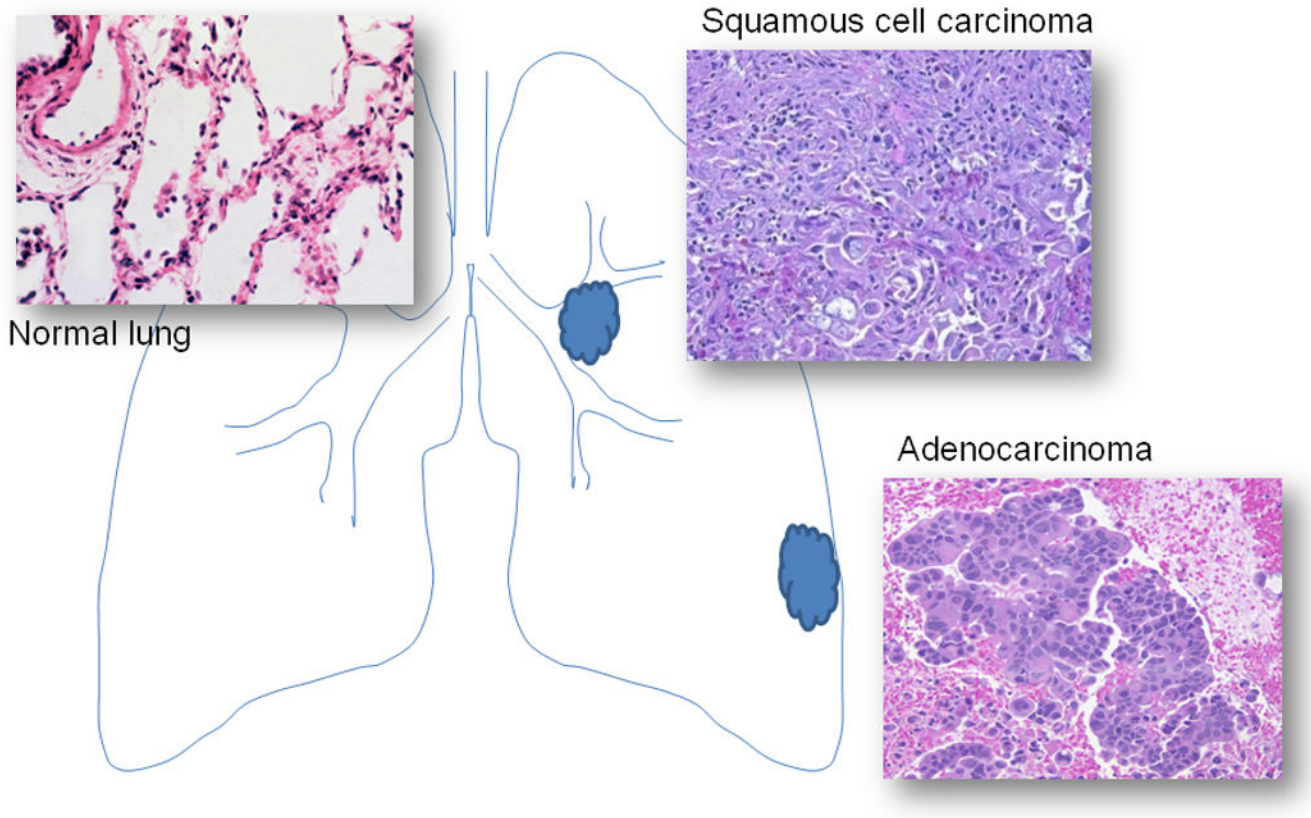

Figure 2.
Histomorphology of normal lung parenchyma, squamous cell carcinoma and adenocarcinoma.

Table 1

Keywords search of the Pub med in lung cancer research.

Keywords search	Number of Publications
Protein biomarkers and Lung cancer	11659
Glycoprotein and lung cancer	9435
Prognostic glycoproteins and lung cancer	670
Predictive glycoproteins and lung cancer	343
Lung cancer and Glycoproteomics	8

Table 2

Lectins and their carbohydrate specificities.

Lectin	Specificity
Concanavalin A (Con A)	mannose, glucose, galactose
Wheat germ agglutinin (WGA)	N-acetylglucosamine, sialic acid
Peanet agglutinin (PNA)	Gal-Gal-Nac
M. amurensis agglutinin (MAL1)	Sialic acid
E. cristagalli agglutinin (ECL)	Galβ1–4GlcNAc
A aurantia lectin (AAL)	Fucα1–2Galβ1
U. europaeus agglutinin (UEA)	Fucose

Table 3

Glycoproteomics in lung cancers.

Investigators	Samples	Methods	Major Findings
Soltermann A, et al, 2008 [72]	Pleural effusion of 5 adenoca (advance stage) lung cancer cell line A549	Hydrazide enrichment and LC-MS/MS	• 6 Glycoproteins were identified and correlated with tumor tissue by IHC. • Periostin was identified in pleural effusion.
Hongsachart P, et al, 2009 [81]	Sera of 10 adenoca (stage II/III)	Con A, WGA and others lectin-affinity 2-D PAGE, 2-D DIGE MALDI-TOF MS, MS/MS	• 27 glycoproteins were up-regulated and 12 were down-regulated. • Among them, 3 up- and 2 down-regulated glycoprotein were validated by Western blotting.
Rho JH, et al, 2009 [65]	Fresh tumor tissue of 16 adenoca (stage II/III)	Glycoarray 2-D PAGE LC-MS/MS	• 8 glycoproteins were up- and 7 were down-regulated. • Differential glycosolation pattern was identified between cancer and matched normal tissues.
Rho JH, et al, 2009 [64]	Fresh tumor tissue of 5 adenoca (stage II/III)	2-D DIGE LC-MS/MS	• 14 spots were differentially expressed. • 3 proteins were overexpressed by LC-MS/MS. • Differential protein expression pattern was identified between cancer cells and stroma.
Ueda K, et al, 2009 [83]	Sera of 10 adenoca (stage IV)	SELDI-TOF MS Lectin-coupled ProteinChip array	• 41 protein peaks showed significant differences. • Cancer-associated aberrant glycosylations were identified
Ueda K, et al, 2010 [82]	Sera of 8 adenoca *	Con A lectin-chromatography Combination of IGEL with 8-plex iTRAQ	• 107 glycoproteins were identified. • 8 glycoproteins showed significant differences.
Zeng X, et al, 2010 [80]	Pooled sera of 31 adenoca * and 23 SqCC *	Hydrazide enrichment and LC-MS/MS	• 22 glycoproteins showed significantly differential expressions • Three of them were verified by ELISAs.
Tsai HY, et al, 2011 [84]	Sera of 19 adenoca *, 8 SqCC *, 11 SCLC *, and 7 unknown types *	2-D DIEG MALDI MS/MS Nanospray-MS^n	• Fucosylated haptoglobin was elevated in all cancers.

* tumor stage is not specified. Adenoca: adenocarcinoma. SqCC: squamous cell carcinoma. SCLC: small cell lung cancer.