Current Serum Lung Cancer Biomarkers

Jung-Mo Ahn and Je-Yoel Cho*
Department of Biochemistry, BK21 and Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul, Korea

Abstract

The development of proteomics technology over the past decade has advanced the identification and validation of new biomarkers. Clinical applications of serum biomarkers will benefit us by being cost-effective and non-invasive, as well as by allowing the simple detection of diseases. Numerous serum biomarkers for lung cancer have been identified and reported in publications. Despite extensive studies, cancer detection via a single marker remains difficult due to the low sensitivity, specificity and reproducibility of the identified lung cancer serum biomarkers. An optimal combination of multiple biomarkers, rather than a single biomarker, should be considered for the development of a lung cancer detection biomarker panel and should be validated with these combinatory markers. This review summarizes the lung cancer biomarkers that have been published within the last five years.

Keywords: Serum biomarkers; Lungs cancer; Adenocarcinoma

Introduction

Lung cancer is the most common cause of cancer mortality worldwide, accounting for 25% of all cancer deaths with an incidence rate of 1.2 million people per year [1]. The main risk factor of lung cancer is smoking; the risk of lung cancer is tenfold higher in smokers than in non-smokers [2]. Lung cancer can typically be grouped into two large categories: small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC), which account for 15% and 85% of lung cancers, respectively. NSCLC consists of 3 major histological subtypes: Adenocarcinoma (ADC), Squamous Cell Lung Carcinoma (SQC), and Large Cell Carcinoma (LCC). The five-year survival rate for lung cancer remains much lower than that of other cancers, at approximately 15% [3]. Survival rates remain low despite advances of current lung cancer diagnostic methods such as X-rays and computed tomography (CT) scans. Furthermore, high cost, a high risk of radiation exposure, and poor sensitivity and specificity are problems with current screening techniques [4]. A better diagnostic measure for lung cancer that facilitates the early detection of the disease, therefore allowing for effective intervention, is necessary to lower lung cancer mortality rates. The discovery of cancer biomarkers, specific molecules that help to distinguish between normal and cancerous conditions, may potentially be used to develop a more effective diagnostic tool for lung cancer. Cancer biomarkers consist of either of genetic materials or proteins because cancer is a heterogeneous disease that reflects gene and protein changes within a cancer cell. However, proteins are the main functional units of biological processes. Almost all of the Food and Drug Administration (FDA)-approved cancer biomarkers are protein markers [5].

The recent advances in proteomics technologies, including mass spectrometry, protein labeling, protein array-based approaches, imaging, and protein bioinformatics, have enabled researchers to rapidly discover more biomarkers and to better understand their roles. Various putative biomarkers have been identified from blood samples, and these markers could be used to develop a non-invasive, cost-effective method to identify individuals at high risk for lung cancer. Proteomics strategies provide powerful information toward blood-based biomarker discovery (Figure 1). Cytokeratin-19 fragments (CYFRA 21-1), carcinoembryonic antigens (CEA), cancer antigen-125 (CA-125), and neuron-specific enolase (NSE) are among some of the lung cancer protein biomarkers that have been discovered to date. A multiple biomarker approach, rather than a single biomarker approach, is proposed for cancer detection due to the low sensitivity and specificity of individual biomarkers [2,6]. This review mainly discusses lung cancer serum biomarkers that have been published recently and the benefits of a multiple biomarker approach for lung cancer detection.

Currently Available Biomarkers for Lung Cancer Detection via Blood Samples

The following markers for use with blood samples are clinically available in some areas. They are currently used in a limited capacity as accessory blood biomarkers for lung cancers. Carcinoembryonic antigen (CEA) is an oncofetal protein not typically expressed in adult tissues. Functioning of the cell surface glycoprotein CEA involves cell adhesion and intracellular signaling. Some lung cancer cells are overexpressed by the de-repression of CEA-encoding genes. In lung cancers, the CEA levels in blood are elevated and are inversely correlated with the response to cancer therapy; therefore, this marker is used for the detection of cancer recurrence and the prediction of a poor survival rate [7]. CYFRA 21-1 is a fragment of cytokeratin 19 that is typically associated with epithelial cell cancers, including NSCLC, and is typically of the SLCQ type. CYFRA 21-1 is correlated with disease response and the prognosis of cancer but cannot be used to identify cancer patients among patients with respiratory diseases. The sensitivity of CYFRA 21-1 for NSCLC ranges between 23 and 70% [8,9]. Squamous cell carcinoma antigen is a structural cytoplasmic protein that is present at elevated levels in the blood of NSLC patients. It is also involved in terms of having metastatic potential. The sensitivity for NSCLC ranges from 15 to 55% [8,10]. Neuron-specific enolase (NSE) is a glycolysis enzyme produced in neuronal cells and cells with neuroendocrine differentiation. In lung cancer, SCLC is of a neuroendocrine origin and...
with LC-MS/MS analysis. The serum levels of its fragmentation form, cancer patients using multilectin affinity chromatography combined coagulation, fibrinolysis, hemostasis, and inflammatory response [13].

In this section, we will summarize recently reported lung cancer blood potential lung cancer biomarkers, including plasma kallikrein (KLKB1) that can be tested in the blood. Previously, our group reported several


is therefore found to be elevated in patients’ blood. The sensitivity ofNSE for detection of SCLC may be as high as 74% [11]. Tumor M2-pyruvate kinase (PKM2) is a dimeric form of the pyruvate kinase isoenzyme type M2 that is increased in various cancers. The sensitivity for lung cancer ranges from 50 to 71% and is especially high in ADC [8]. C-reactive protein (CRP) is an acute-phase protein, the levels of which rise in response to inflammatory conditions such as lung cancer. The blood CRP levels were increased in various cancers. However, the association between CRP and the risk of incident lung cancer was not stable in many published results. Recently, a CRP case-control study specifically conducted with respect to lung cancer concluded that high levels of CRP could be used as a prognosis biomarker of lung cancer [12].

Other Markers

Proteomics is a powerful tool for identifying lung cancer biomarkers that can be tested in the blood. Previously, our group reported several potential lung cancer biomarkers, including plasma kallikrein (KLKB1) protein fragments, serum amyloid A (SAA), haptoglobin β chain [Hp β], and complement component 9 (C9), using proteomics approaches. In this section, we will summarize recently reported lung cancer blood biomarkers, including those from our findings, which were discovered by proteomics technologies and were then validated in the patients’ blood (Table 1).

Plasma kallikrein (KLKB1)

The KLKB1 enzyme cleaves Lys-Arg and Arg-Ser bonds in kininogen to release bradykinin and has functions related to blood coagulation, fibrinolysis, hemostasis, and inflammatory response [13]. We identified this protein in the glycoproteome-enriched sera of lung cancer patients using multilectin affinity chromatography combined with LC-MS/MS analysis. The serum levels of its fragmentation form, which is supposed to be glycosylated H4 domain (aa 292-375), were increased in lung cancer samples compared with normal control sera in a Western blot analysis [14].

Serum Amyloid A (SAA)

SAA proteins are a family of apolipoproteins associated with the high-density lipoprotein (HDL) complex that are secreted during the acute phase of inflammation. Three isoforms of SAA have been reported. The acute SAA (SAA1 and SAA2) levels are dominantly expressed in the liver during inflammation, whereas SAA3 is induced in various distinct tissues [15,16]. Our study identified SAA1/2 in the lung adenocarcinoma patients’ sera by comparing the proteomes in healthy donor sera using LC-MS/MS analysis. High levels of SAA in the sera of lung cancer patients were detected by mass spectrometry, verified by Western blot, and quantified by ELISA. Expression of the SAA protein was also detected in lung cancer tissues by immunohistochemistry. The SAA levels were highly expressed in sera and tissues derived from lung cancers of various histological types (ADC, DLC, and SCCLC) in cancer patients compared with those of healthy controls. In using the 50 pg/ml cutoff criteria, the SAA ELISA showed between 50-70% sensitivity with 95% specificity. An in vitro experiment also demonstrated that lung cancer cells secreted SAA1 and SAA2, which stimulated infiltrating macrophages to induce matrix metalloproteinase 9 (MMP9), which has been suggested to help drive the metastasis of cancer cells [17]. A large-scale, isotype-specific proteomics quantification method, multiple reaction monitoring (MRM), was also developed for SAA serum level validation. The MRM results also demonstrated significantly higher levels of SAA in the lung cancer patients’ sera, which agrees with the ELISA results [18]. Our group also manufactured a chromatography-based rapid kit using developed dual monoclonal antibodies.

Haptoglobin β Chain (Hp β)

Haptoglobin (Hp) is a free hemoglobin-binding glycoprotein that inhibits the oxidative stress of hemoglobin and assists in hemoglobin uptake by the hemoglobin scavenger receptor CD163. The molecular structures of Hp consist in tetramer of two α and two β chains that are connected by a disulfide bridge [19]. High levels of blood Hp have been reported in various cancer types, such as breast cancer [20], ovarian cancer [21], pancreatic cancer [22], and bladder cancer [23]. In our study, the Hp β chain was indicated as a diagnostic marker for lung cancer of both NSCLC and SCCLC types [24]. According to the proteomics analysis, Hp β chain region peptide levels were 3 fold higher in lung cancer patients’ sera than in control subjects. The reason that the Hp β chain is a better diagnostic indicator may be due to its increased stability due to N-linked glycosylation at a site that is not found in the Hp α chain. The Western blot and ELISA results revealed significantly higher levels of the Hp β chain in lung cancer patients’ sera and plasma compared with samples from other tumors that were associated with breast cancer, hepatocellular carcinoma and respiratory diseases of tuberculosis, idiopathic pulmonary fibrosis and bronchial asthma. In the ELISA analysis for lung adenocarcinoma and healthy control samples, when the cut-off was set at 2 µg/ml, the AUC value was 0.822 and the specificity and sensitivity for the Hp β chain as a diagnostic marker were estimated at 63.2% and 82.6%, respectively.

Complement Component 9 (C9)

Complement component 9 (C9) protein, a terminal constituent of the membrane attack complex (MAC-C5b, C6, C7, C8 and C9), plays a role in the immune response to cell death by forming pores in the plasma membrane of target cells [25]. C9 is a glycoprotein that has two

**Figure 1:** Proteomic platforms for blood biomarker discovery studies.

Blood is the most commonly used test sample in biomarker discovery studies. There are abundant proteins of all types in the blood; therefore, the depletion of abundant proteins and the enrichment of specific protein groups, such as glycoproteins, are needed before the identification steps. Protein separation is generally performed by gel electrophoresis and liquid chromatography in proteomic studies. Mass spectrometry is predominantly used for protein identification. LC-ESI-MS/MS and MALDI-TOF/MS are most commonly used for high throughput protein identification platforms. Verification steps for MS-based proteomics data involve immunoassays such as immunoblot, antibody arrays, and ELISAs. Recent developments incorporate a liquid chromatography–coupled mass spectrometry (LC-MS)-based verification system called multiple reaction monitoring (MRM).
Autoantibody combined biomarkers provided a sensitivity of 85.7% to IGFBP-2 alone [34]. The serum IGFBP-2 and anti-IGFBP-2 diagnostic sensitivity and specificity for lung cancer compared to IGFBP-2 alone showed increased specificity of the disease. A recent study regarding circulating anti-IGFBP-2 levels in patients with several cancers, fewer cases were detected in the early stages [33]. Although increased levels of IGFBP-2 were found in the advanced stages of several cancers, fewer cases were detected in the early stages of the disease. A recent study regarding circulating anti-IGFBP-2 autoantibodies and IGFBP-2 combined markers showed increased diagnostic sensitivity and specificity for lung cancer compared to IGFBP-2 alone [34]. The serum IGFBP-2 and anti-IGFBP-2 autoantibody combined biomarkers provided a sensitivity of 85.7% and a specificity of 57.5% in the detection of lung cancer.

**Insulin-like Growth Factor-Binding Protein-2 (IGFBP-2)**

Insulin-like growth factor-binding protein-2 [IGFBP-2] is a member of the insulin-like growth factor-binding protein family [IGFBPs]. The main function of IGFBP-2 is to inhibit IGF-mediated growth and development rates [28]. Increased levels of IGFBP-2 have been found in solid tumors and in blood from patients with glomma [29], colorectal [30], prostate [31], ovarian [32] and breast cancers [33]. Although increased levels of IGFBP-2 were found in the advanced stages of several cancers, fewer cases were detected in the early stages of the disease. A recent study regarding circulating anti-IGFBP-2 autoantibodies and IGFBP-2 combined markers showed increased diagnostic sensitivity and specificity for lung cancer compared to IGFBP-2 alone [34]. The serum IGFBP-2 and anti-IGFBP-2 autoantibody combined biomarkers provided a sensitivity of 85.7% and a specificity of 57.5% in the detection of lung cancer.

**Peroxiredoxin 1 (PRX1)**

Peroxiredoxin 1 (PRX1) and 2 belong to a family of ubiquitous multifunctional antioxidant proteins. The main function of PRX1 is to eliminate peroxides such as hydrogen peroxide (H₂O₂) generated during metabolism [35]. PRX1 is also involved in the inhibition of oncogenes such as Ras and c-Myc [36,37]. PRX1 protein levels were found to be higher in human cancer cells and tissue samples for mesothelioma [38], breast [39,40], and head-and-neck cancers [41]. More recently, PRX1 was also identified in lung cancer patients’ plasma by LC-MS/MS-based screening technology. Plasma PRX1 levels were increased in the lung cancer group and the asbestos exposure group [42].

**Endoglin [CD105]**

Endoglin [CD105] is a major cell membrane glycoprotein of the vascular endothelium. It forms a complex with the signaling receptors for transforming growth factor-beta [TGF β R1 or TGF β R2]. It efficiently binds TGF β 1 and TGF β 3 [36,43,44]. The main function of endoglin is to help the binding of endothelial cells to integrins and other receptors [45]. Endoglin promotes angiogenesis by the activation of endothelial cells [46]. Endoglin-overexpression has also been found in the endothelium of vessels in human solid tumors and is closely associated with a poor prognosis and the presence of metastases [47]. Levels of soluble endoglin [s-endoglin], formed by the cleavage of ectodomain of membrane receptors, were higher in patients with various types of solid tumors, including those from breast [48,49], colorectal [50] and liver cancers [51] compared to normal subjects, and its levels were also correlated with the presence of metastases. A recent study comparing s-endoglin serum levels in patients with NSCLC pre- and post-operation found significantly decreased levels of s-endoglin post-operation [52]. Thus, s-endoglin has been confirmed as a potential monitoring and prognostic biomarker for lung cancer.

**Progesterone Receptor Membrane Component 1/ sigma-2 Receptor (Pgrmc1)**

Progesterone receptor membrane component 1 (Pgrmc1) is a...
cytochrome b- related protein [53] that is induced by carcinogens [54,55]. Pgrmc1 levels are elevated in spontaneous ovarian, breast, and lung cancers [56]. Pgrmc1 is known to localize at the endoplasmic reticulum or punctate cytoplasmic sites with epidermal growth factor receptors (EGFR) [57]. Pgrmc1 was identified as a sigma-2 receptor, which is induced in cancers [58]. Recently, Pgrmc1 showed the potential to be a serum biomarker for lung cancer [59]. Pgrmc1 levels in SCLC cell lines and tissues were high compared with controls by using the immune assay and siRNA knockdown of Pgrmc1 -inhibited SCLC cell proliferation and cancer cell survival. Pgrmc1 is reported to be secreted protein neudesin, which binds to unknown neuron receptors and elevates AKT signaling. These findings also indicate that Pgrmc1 is secreted by lung cancer cells as well as demonstrating the localization of Pgrmc1 to secretory vesicles in cancer cells. Furthermore, Pgrmc1 levels in the plasma and in the exosomes fractions of plasma were significantly increased in lung cancer groups (ADC, SCLC, LCC).

Pro-gastrin-releasing Peptide (proGRP)

Pro-gastrin-releasing peptide (proGRP, residue 31-98) is a more stable biochemical precursor of gastrin-releasing peptide [GRP] that is specifically produced by the neuroendocrine origin of SCLC cells [60]. Circulating proGRP is suggested to serve as a potential diagnostic marker for SCLC [61-63]. In a recent report, plasma proGRP levels were increased in SCLC patients relative to the levels in patients with other types of lung cancer [specificity of 95%]. Its levels also gradually increased with the progression of the disease [64].

p21(Cip1)-interacting zinc finger protein (Ciz1)

Ciz1 is a nuclear matrix protein, and its main function is to promote the initiation of mammalian DNA replication [65]. Nuclear matrix proteins offer several advantages as biomarkers. They are localized in the nucleus, which is the heart of dysregulated conditions during carcinogenesis formation. Therefore, nuclear matrix proteins such as NMP22, BCLA4, PC1, and NM179 were suggested as tumor markers for bladder [66], cervix [67], and prostate cancer [68]. Recently, B-variant Ciz1 (24 nucleotides from the 3’ end of exon 14 are excluded, leading to in-frame deletion of eight amino acids ‘VEEELCKQ’) protein levels were significantly increased in the plasma of early stage lung cancer patients compared with the plasma from normal controls and patients with other respiratory diseases. The growth of tumor cells was significantly reduced when B-variant Ciz1 was depleted using RNAi in vitro and in vivo [69]. B-variant Ciz1 is a potential diagnostic lung cancer-specific marker. The sensitivity and specificity of Ciz1 for stage I NSCLC were 95% and 74%, respectively.

Matrix Metalloproteinase-1 [MMP-1]

MMP-1 is a collagenase that cleaves collagen types I, II, III, VII and X at one site in the helical structure [70]. It is well known to be overexpressed in various cancer cells during the invasion and metastasis phases [71]. The up-regulated expression of MMP-1 has been detected in colorectal [72], esophageal [73], pancreatic [74], gastric [75], breast cancer [76] and malignant melanoma [77]. A recent study that measured MMP-1 levels in plasma and tissue samples from lung cancer patients showed that higher plasma levels of MMP-1 and tumor progression (i.e., tumor size, staging and lymphatic invasion) were associated with a lower patient survival rate [78].

Urokinase Plasminogen Activator Receptor (uPAR)

uPAR is a glycosylphosphatidylinositol (GPI)-anchored glycoprotein and cell surface receptor specific to the urokinase Plasminogen Activator (uPA) and is a promoter of plasmin activation. uPAR consists of three homologous domains, denoted by I, II and III [78]. The uPA-catalyzed cleavage of uPAR is a negative feedback loop in which uPA cleaves uPAR between domains I and II, releasing I (uPAR(I)) and leaving the cleaved form uPAR(II-III) attached to the cell surface [79,80]. Increased levels of cleaved uPAR were found in a blood sample from an NSCLC patient [81,82]. A recent analysis of serum and plasma levels in pre-operative NSCLC patients correlated with higher levels of cleaved uPAR and lower survival rates [83].

Conclusions

The recent advances developed in proteomics technology specifically improved techniques for the identification, detection, and verification of biomarkers and have thus improved our understanding of lung cancer. Earlier diagnosis of lung cancer is urgently needed to decrease the lung cancer mortality rate. A noninvasive technique that integrates an understanding of lung cancer biomarkers and identifies lung cancer patients from hemo-analysis would be a better alternative to the current diagnostic methods. From our experience studying lung cancer biomarkers, we have found that many protein/peptide biomarkers discovered in the serum and/or plasma seem to overlap with other diseases, especially other cancers and inflammatory diseases. Thus, the discovery of more lung tissue-specific biomarkers is needed, starting with tissue-specific genes/proteins and the related proteins in lung tissues and lung cancers.

The best method to verify a lung cancer diagnosis would be one that involves a multiple biomarker approach rather than a single biomarker approach. The sensitivity and specificity of anti-IGFBP-2 autoantibodies in lung cancer patients’ sera were 73.2% and 60.6%, respectively. However, combination of IGFBP-2 and anti-IGFBP-2 autoantibody increased the sensitivity and specificity (85.7% and 57.5%) for diagnosing lung cancer [34]. Also, lung cancer patients with lymph node metastasis had higher level of serum vascular endothelial growth factor (VEGF)-C, MMP-9 and VEGF than those without metastasis [84]. Serum VEGF-C reached a sensitivity of 85% and specificity of 68%, while MMP-9 was 63% sensitivity and 75% specificity. VEGF reached a sensitivity of 80% and specificity of 99%. However, combination of three markers had higher sensitivity and specificity (83% and 76%) than single biomarker approach. Although some biomarkers show high sensitivity and specificity, it should be noted that small-scale verification in retrospectively selected samples from one institution may not represent a general test with the samples prospectively collected from multiple institutions. With plenty of lung cancer biomarkers mentioned above and many more that are published, now is the time to extensively validate the markers in large-scale clinical samples and to assess their effectiveness in various combinations to develop an optimized multi-marker approach. Therefore, the combination of multiple biomarkers that best distinguishes cancer samples from controls should be utilized in the development of cancer detection technology for clinical applications.

Acknowledgements

This work was supported by the Post 21C Frontier Project (No. M3CSA105334) from the Korean Ministry of Education, Science and Technology (MEST), the National Research Foundation of Korea (NRF) grant (2012036669).

References


