Evaluation of Relationship Between Serum Protein Profiles and Lung Cancer by SELDI-TOF-MS (Surface Enhanced Laser Desorption Ionization Time of Flight Mass Spectrometry) Method

Akciğer Kanseri ile Serum Protein Profilleri Arasındaki İlişkinin SELDI-TOF-MS (Surface Enhanced Laser Desorption Ionization Time of Flight Mass Spectrometry) Yöntemi ile Değerlendirilmesi

ABSTRACT
Objective: Although proteomic profiles detected by using SELDI-TOF-MS method distinguish lung cancer patients from healthy individuals, there are no studies concerning proteomic patterns of tumor subsets in lung cancer. The goal of this study was to establish proteomic patterns of tumor subsets in lung cancer patients with surface enhanced laser desorption- ionization time of flight mass spectrometry (SELDI-TOF-MS) method.

Material and Methods: A total of 169 patients diagnosed histopathologically including 142 non-small cell lung cancer (NSCLC) and 27 small cell lung cancer (SCLC) patients were included in this study. In NSCLC group, there were 60 squamous cell carcinoma, 38 non-squamous cell carcinoma and 44 unclassified type NSCLC patients. Venous blood samples were obtained from all cases. All of the serum samples were analyzed by SELDI-TOF-MS method for proteomics investigation.

Results: Three peaks (9065 m/z, 9175 m/z, 9394 m/z) were found to be discriminatory in serum SELDI profiles between NSCLC and SCLC groups (p < 0.05). All of these three peaks showed higher intensity in patients with SCLC. The analysis between non-squamous and squamous cancer groups of NSCLC revealed eight discriminatory proteomic features. Among these peaks, only two (5815 m/z, 5906 m/z) showed higher intensity in patients in the non-squamous group (p < 0.05).

Conclusion: Proteomic patterns could provide some valuable clues on the carcinogenetic mechanism of different types of lung cancer and may help us to discover some potential subtype-specific biomarkers of lung cancer by SELDI-TOF-MS method.

Key Words: Lung neoplasms; proteomics; spectrometry, mass, matrix-assisted laser desorption-ionization.

ÖZET
Amaç: SELDI-TOF-MS yöntemi ile tespit edilen proteomik pikler akciğer kanserinin sağlığa kişilikli kişilerden ayrıldığı ayırt edebilmesine rağmen akciğer kanserinin alt tiplerinin proteomik paternleri hakkında çalışma yoktur. Çalışmanın amacı; akciğer kanserleri olguların serum örneklerinde SELDI-TOF-MS yöntemi ile serum protein profillerinin analizlerini yaparak akciğer kanserleri olgularında histolojik tip ile serum protein profilleri arasındaki ilişikiye değer lendirmektir.

Gercek ve Yöntemler: Histopatolojik olarak akciğer kanserleri tanısı alan 169 hasta, 60 skuamöz hücreli, 38 non-skuamöz hücreli ve 142 tip ayrımı yapamayan hasta sayıldı. Tüm olguların venöz kan örnekleri alınadyı. Alınan bu örneklerin proteomik analizleri SELDI-TOF-MS yöntemi kullanılarak yapıldı. Bulgular: SELDI-TOF analizi ile KHDAK’lı hastalar ile KHAK’lı hastalar arasında ayırt edici protein pikler 9065 m/z, 9175 m/z ve 9394 m/z olarak bulundu (p < 0.05). Söz konusuz üç proteinin KHAK’lı olgularında yüksek yoğunlukta bulunduğunu gösterdi. Ayrıca KHDAK’lı grup içerisinde yer alan non-skuamöz ve skuamöz gruptlardan arasında sekiz farklı proteomik pik tespit edildi. Bu pikler arasında yer alan 5815 m/z ve 5906 m/z pikler non-skuamöz grupta yüksek yoğunlukta bulundu (p < 0.05). Sonuç: Proteomik paternler, akciğer kanserinin farklı.histolojik tiplerinde görülen kartinogenetik mekanizmalar hakkında bazı değerli bilgiler sağlayabilir. Ayrıca SELDI-TOF-MS analizi ile tespit edilen protein pikleri akciğer kanserinin alt histolojik tiplerine özgü bazı potansiyel biyo-belirteçlerin keşfedilmesini yardımcı olabilir.

Anahtar Kelimeler: Akciğer tümörleri, proteomik; spektrometri, kütle, matrik-assistli lazer salmali-ionozasyon

Lung cancer is the leading cause of cancer deaths in the world.\(^1\) Even patients who present with clinical stage I have at best a 67\% 5-year survival rate, signifying that a large percentage of patients have undetectable metastatic disease at the time of presentation regardless of their stage.\(^2\) The 5-year survival across all subtypes is about 15\% and has not improved over many decades.\(^3\)

Gene expression profiling has improved our understanding of the histological heterogeneity of lung cancer and has identified potential biomarkers and gene signatures for classifying patients with different survival outcomes and treatments.\(^4\) A comprehensive understanding of the mechanism behind carcinogenesis, tumor progression, and metastasis requires an in-depth analysis of not only the genome, but also the proteome.\(^5,6\)

Proteins are responsible for the function of biological systems and phenotypes of cells. Cancer cells express proteins that distinguish them from normal cells.\(^7\) New technologies are being developed to allow the rapid and systematic analysis of thousands of proteins.

Proteomics aims to characterize proteins to obtain a more integrated view of the biology. Proteomics-based early detection strategies for cancer diagnosis include the analysis of complex mixtures such as tissue samples, serum, plasma, sputum and exhaled breath condensate. There are different methods for discovering and assessing proteins. Surface enhanced laser desorption ionization time of flight mass spectrometry (SELDI-TOF-MS), which is an affinity-based mass spectrometry method that uses a protein chip modified with a specific chromatographic surface, is a modified matrix-assisted laser desorption/ionization mass spectrometry system, and it has overcome many of the limitations of 2-dimensional electrophoresis and matrix-assisted laser desorption/ionization time of flight (TOF) mass spectrometry.\(^8,9\) The proteomic analysis studies have been successfully performed for the diagnosis of cancer, prognosis of the disease and efficacy of the treatment.\(^10-15\) These methods are currently used for cancer biomarker screening including ovarian, breast, prostate and liver cancers.\(^16-19\)

Today, lung cancer is classified according to histological criteria. The two main subtypes are non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). In addition, in NSCLC group, there are two subtypes which are non-squamous and squamous cancer groups.\(^20,21\)

Although SELDI-TOF MS has been used to discover potential protein biomarkers to differentiate protein profile of patients with lung cancer from healthy individuals, no studies have been performed to differentiate protein profile of tumor subsets in patients with lung cancer using this technology.

The goal of this study was to discover potential biomarkers in serum samples of patients with different tumor subsets of lung cancer using SELDI-TOF MS.

### MATERIAL AND METHODS

#### PATIENTS AND SERUM SAMPLES

A total of 169 patients diagnosed histopathologically including 142 non-small cell lung cancer and 27 small cell lung cancer patients admitted to university and education and research hospitals between 2009 and 2011 were included in this study. In NSCLC group, there were 60 squamous cell carcinoma, 38 non-squamous cell carcinoma and 44 unclassified type of non-small cell lung cancer patients according to pathological investigation. Venous blood samples were obtained from all cases. All of the serum samples were analyzed using SELDI-TOF-MS method in the university hospital. All serum samples were aliquoted and stored at -80 °C until used. The samples were thawed more than once. All of the patients gave their written informed consents, and informed about the details of the study. This study was conducted in accordance with the Declaration of Helsinki amended the 52nd WMA General Assembly (Edinburgh, 2000), and approved by local ethics committees.

#### SELDI-TOF PROTEIN ANALYSIS

Regarding the number and resolution of the protein peaks, three different ProteinChip surfaces...
(cationic, anionic, and Cu metal binding, Ciphergen Biosystems, Fremont, CA, USA) were tested. The IMAC30 (Cu metal binding surface) protein chip, which displayed the best serum profile, was selected for further analysis. Briefly, 5 µl of each serum sample was denatured by adding U9 solution (9M urea, 2% CHAPS and 150 mM Tris-HCl, pH 9) in a 1:3 ratio and mixed on a platform shaker for 35 min at 4 ºC. The array spots were pre-equilibrated twice with 100 mM CuSO₄ for 5 min at room temperature, followed by three washes with 1 mM Hepes (pH 7) for 2 min. Arrays were then incubated three times with 150 µl of binding buffer (PBS 1X, 0.25 M NaCl, 0.1% Triton X-100) for 5 min. Ten microliters of each diluted serum sample was randomly added on preactivated spots with 120 µl of binding buffer for 1 hr. Each array was then washed two times with binding buffer (5 min for each), then without triton-binding buffer and washed twice with 1 mM HEPES. The air-dried arrays were saturated with sinapinic acid matrix (Ciphergen Biosystems) in 0.5% trifluoroacetic acid and 50% acetonitrile before being read on the instrument SELDI ProteinChip System 4000 Mass spectrometer (Ciphergen Biosystems). The data were analyzed with Ciphergen Express software, version 3.0 (Ciphergen Biosystems). All the raw data with protein peaks ranging from 2000 to 35000 Da were further normalized to total ion current and aligned.

REPRODUCIBILITY ANALYSIS OF SELDI PROTEIN SPECTRA
To evaluate machine reproducibility, serum sample of a healthy volunteer was applied onto different spots on every chip and run as a quality control (QC). In these protein profiles, several independent control peaks were identified to calculate the coefficient of variance (CV) of intensity and CV of m/z in intra-assay and inter-assay. In all spectra, autodetected peak labeling was achieved with 5.0 S/N (signal to noise ratio), 3 valley depth for the first pass; minimal peak threshold: 20% of all spectra and 3.0 S/N, 1 valley depth for the second pass with 0.2% cluster mass window and the estimated peaks were added.

STATISTICAL ANALYSIS
Protein peaks were clustered with the Ciphergen Express software, version 3.0 performing Expression Difference Mapping (EDM). Discriminatory peaks, depending on peak intensity, were identified using the Mann-Whitney non-parametric test. To determine the best discriminative proteomic index, receiver operating characteristics curves (AUROC) were used as a measurement. Statistically significant discriminatory peaks between groups were determined with p<0.05.

RESULTS
When the Mann-Whitney non-parametric test was used, only 3 peaks were found to be discriminatory between serum SELDI profiles of NSCLC and SCLC groups, with p values less than 0.05 (Table 1). All of these 3 peaks showed higher intensity in patients with SCLC group.

When analysis was performed between non-squamous and squamous groups of NSCLC, 8 pro-

<table>
<thead>
<tr>
<th>Proteomic feature (m/z)</th>
<th>**p value</th>
<th>*AUROC</th>
<th>**Average intensity of SCLC cases relative to NSCLC cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>9065</td>
<td>0.003</td>
<td>0.66</td>
<td>1.3</td>
</tr>
<tr>
<td>9175</td>
<td>0.017</td>
<td>0.63</td>
<td>2.3</td>
</tr>
<tr>
<td>9394</td>
<td>0.026</td>
<td>0.63</td>
<td>1.25</td>
</tr>
</tbody>
</table>

*The area under the ROC curve; **Mann-Whitney U test.
Proteomic features were found to be discriminatory, with \( p \) values less than 0.05 (Table 2). Among these peaks, only 2 of them (5815 m/z, 5906 m/z) showed higher intensity in patients with non-squamous group. In addition to non parametric test, a decision tree classification was performed with the help of Biomarker Patern Software (BPS). However, due to the poor \( p \) value results, any classification tree could not be created.

**DISCUSSION**

Lung cancer is the most common cause of cancer deaths in industrialized countries and its incidence is steadily increasing in many countries. Proteins are responsible for the function of biological systems and phenotypes. Cancer cells express proteins that distinguish them from normal cells. New technologies are being developed to allow the rapid and systematic analysis of thousands of proteins.

SELDI-TOF-MS proteinChip technology is a new technique that allows multiple serum samples obtained directly from patients to be analyzed in a relatively short time. It is a high-throughput approach used to generate protein expression profiles, which in combination with bioinformatics tools to extract information for biomarker discovery, has been essential in identifying novel protein biomarkers.

Proteomic patterns by SELDI-TOF methods has been applied in identifying early detection markers in multiple cancers. However, there is no study about proteomic patterns of tumor subsets in lung cancer patients with SELDI-TOF-MS method.

In our study, three peaks were found to be discriminatory between NSCLC and SCLC groups and all of these three peaks showed higher intensity in patients with SCLC. Yang et al. found that the sensitivity of the SELDI marker pattern for NSCLC was significantly higher than for SCLC (91.4%, 64.3%, respectively), indicating that the pattern may be more effective in discriminating NSCLC patients than SCLC patients.

Diagnosis and classification of cancers are critical for the selection of appropriate therapies. Today, lung cancer is treated according to histological criteria. The two main subtypes are NSCLC and SCLC. In addition, NSCLC is divided into two major groups as non-squamous and squamous. The advent of effective targeted therapies for the lung cancer, such as epidermal growth factor receptor inhibitors erlotinib and gefitinib, and the prospect of developing additional targeted therapies, has emphasized the importance of accurate diagnosis.

In NSCLC group, eight proteomic features were found to be discriminatory between squamous and non-squamous groups. Among these peaks only two of them (5815 m/z, 5906 m/z) showed higher intensity in the non-squamous group compared to squamous group. Proteomics patterns obtained directly from serum could be used to accurately classify histological groups in lung cancer.

<table>
<thead>
<tr>
<th>Proteomic feature (m/z)</th>
<th><strong>p value</strong></th>
<th><em>AUROC</em></th>
<th><strong>Average intensity of squamous cases relative to non-squamous cases</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>15879</td>
<td>0.0085</td>
<td>0.64</td>
<td>1.3</td>
</tr>
<tr>
<td>3145</td>
<td>0.0093</td>
<td>0.65</td>
<td>1.5</td>
</tr>
<tr>
<td>16051</td>
<td>0.0098</td>
<td>0.66</td>
<td>1.4</td>
</tr>
<tr>
<td>2855</td>
<td>0.0115</td>
<td>0.64</td>
<td>2.3</td>
</tr>
<tr>
<td>5683</td>
<td>0.0267</td>
<td>0.63</td>
<td>1.4</td>
</tr>
<tr>
<td>15172</td>
<td>0.0321</td>
<td>0.63</td>
<td>1.8</td>
</tr>
<tr>
<td>5815</td>
<td>0.0392</td>
<td>0.64</td>
<td>0.7</td>
</tr>
<tr>
<td>5906</td>
<td>0.0408</td>
<td>0.63</td>
<td>0.81</td>
</tr>
</tbody>
</table>

*The area under the ROC curve; ** Mann-Whitney U test.
Petricoin et al. used SELDI-TOF method for early detection of ovarian cancer through the identification of unique malignant proteomic patterns in serum. SELDI-based proteomics has also been shown by the identification of new biomarkers for many tumors. In addition, a proteomics-based approach was used for the preliminary search of tumor proteins that elicit a humoral response in breast carcinoma and may occur as circulating antigens. However, there are no studies about the subsets of NSCLC proteomic profiles in lung cancer. In previous studies, it has been shown that cells intensively express proteins at low mass range, such as 7522, 7649, and 7948, in premalignant lesions, atypical adenomatous hyperplasia (AAH). High mass range profile peaks (47102, 66576) were found at significantly greater intensities in normal lung epithelial cells when compared with malignant or atypical samples. This may represent decreasing expression, or possible breakdown of these proteins to shorter peptides during malignant changes in lung epithelium. In our study, we also found moderate and low mass range profiles, such as 9394, 15879, 5906, 8683.

In this study, we instituted various preventive measures to avoid generation of biased results caused by artifacts related to the nature of the clinical samples. All serum samples were collected and processed within the same clinical and laboratory settings. To avoid variation in the procedure, freshly collected sera were immediately aliquotted, stored at -80°C and thawed only once. Standard protocols must be developed to minimize unwanted fluctuation, and CVs between ProteinChips must be calculated by using common peaks across different spectra. The use of paired serum samples from individual patients in the study excluded most of genetic and environmental variables and made it likely that the changes in the protein profile reflected the disease state more exactly. We used QC serum to allow detection of any unusual features during the process. Such precautions led to very good reproducibility of the protein peak patterns.

SELDI-TOF-MS method, an innovative proteomic technology introduced by Hutchens and Yip, has overcome many of the limitations of the two-dimensional electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) method. This is a high throughput technique for analysis of complex biological specimens such as serum. We used serum materials in our study by using SELDI-TOF-MS method.

Proteomics-based strategies for cancer diagnosis and treatment include the analysis of complex mixtures such as tissue samples, serum, plasma, sputum and exhaled breath condensate. SELDI-TOF-MS technology utilizes patented biochip arrays to capture individual protein or groups of proteins with common biochemical properties such as hydrophobicity or charge from complex mixtures. These retained proteins are subsequently resolved directly by time-of-flight mass spectrometry. It is fast, sensitive, and scalable for high-throughput sample processing. These characteristics make SELDI technology suitable for the studies of cancer. Despite its advantages, this technology is unsuitable for high molecular weight proteins, whereas the proteins below 20 kilodaltons (kDa) is well resolved.

Among limitations of our study is the heterogeneity in our study groups in regard of number of patients. For this reason, this promising findings need to be confirmed with further studies with larger patient populations.

In conclusion, we have shown that using SELDI-TOF MS method could find new potential tumor markers for NSCLC and SCLC and also two major subset groups of NSCLC, non-squamous and squamous groups. More samples should be collected to validate these protein peaks and future studies must aim to establish proteomics as a useful tool in the discovery of novel biomarkers with potential diagnostic relevance.