Serum Protein Profile Analysis in Patients With Head and Neck Squamous Cell Carcinoma

Christine G. Gourin, MD; Zhong-Sheng Xia, MD; Yan Han; Adam M. French, BS; Ashli K. O’Rourke, MD; David J. Terris, MD; Bao-Ling Adam, PhD

Objectives: To analyze surface-enhanced laser desorption and ionization time-of-flight mass spectrometry (SELDI-TOF-MS) protein profiles of patients with head and neck squamous cell carcinoma (HNSCC) and healthy controls and to determine the sensitivity and specificity of SELDI assay for HNSCC detection before and after treatment.

Design: Proteomic analysis and comparison of serum samples.

Setting: Tertiary care academic medical center.

Subjects: Seventy-eight patients with HNSCC and 68 healthy controls.

Main Outcome Measures: Serum samples were prospectively collected from 78 patients with HNSCC and 68 healthy control volunteers. SELDI-TOF-MS was performed on serum samples to identify protein peaks in the range of 0 to 100 kDa. Classification analysis of the spectral data was performed and used to classify the disease status of the patients.

Results: The SELDI-TOF-MS assay generated serum protein profiles ranging from 0 to 100 kDa. After background subtraction, mass calibration, and normalization, 545 protein peaks were identified. Classification tree analysis based on peak expression correctly classified patients with HNSCC with 82% sensitivity and 76% specificity. Subgroup analysis correctly classified 83% of oral cavity tumors, 81% of oropharyngeal tumors, and 88% of laryngeal tumors. Pretreatment and posttreatment samples were available from 12 patients, and the posttreatment samples were correctly classified in 86% of the patients at 3 months and 75% of the patients at 6 months.

Conclusions: Proteomic SELDI-TOF-MS analysis of serum protein profiles distinguishes patients with HNSCC from controls with a high degree of sensitivity and specificity. Further investigation into the clinical utility of this technology in HNSCC detection and surveillance is warranted.


Overall survival rates for patients with head and neck squamous cell carcinoma (HNSCC) have remained unchanged over the past 30 years, despite advances in surgical and nonoperative treatment. Prognosis and treatment are influenced by the stage of disease at diagnosis. Because early-stage disease is associated with few signs and symptoms, the majority of patients present with advanced-stage disease, which is associated with increased treatment-related morbidity and poorer survival rates. These findings have driven the search for new methods to allow early detection of disease. At present, however, no effective screening test exists for HNSCC beyond physical examination.

Research efforts directed at improving early detection of malignancy have focused on the identification of genetic alterations and tumor biomarkers. Genetic alterations in cancer cells result in differences in the response of the cell to its environment, which are manifested by changes in cellular signaling pathways through alterations in protein expression. Molecular techniques directed at identification of specific genetic alterations in HNSCC have revealed the presence of multiple and varied molecular defects, and this heterogeneity limits the efficacy of molecular techniques as a screening test for early detection. In addition, molecular techniques require significant time, labor, and expertise for adequate specimen processing. Similarly, efforts at detecting biomarkers associated with HNSCC have revealed considerable heterogeneity in biomarker expression, and no biomarker, alone or in
combination, has been shown to be effective as a diagnostic or screening tool.

Proteomic analysis has received increasing attention in recent years as a method of distinguishing alterations in protein expression associated with malignancy. The proteome consists of the library of proteins expressed that define the phenotype of the cell, and because proteomic diversity appears to exist among different specimens with similar histologic features owing to genomic diversity, the pattern of protein expression may be more useful as a disease marker than the detection of individual biomarkers. A recent advance in proteomic analysis is the development of ProteinChip surface-enhanced laser desorption and ionization time-of-flight mass spectrometry (SELDI-TOF-MS), which can detect proteins that are affinity bound to a protein chip array. This system is rapid, has high-throughput (multiple samples) capability and high sensitivity, and allows resolution of proteins with low mass. Most importantly, this technology has potential for use as a clinical assay.

Previous studies have suggested that SELDI-TOF-MS may allow detection of HNSCC-associated serum tumor biomarkers. We sought to determine the sensitivity and specificity of serum protein profiling using SELDI-TOF-MS in distinguishing patients with HNSCC from controls, as well as to determine whether protein profiles differed between tumors of differing stages and sites and whether these profiles changed after treatment.

SERUM SAMPLES

Serum samples were prospectively collected from patients who were evaluated by the Head and Neck Tumor Board of the Medical College of Georgia, Augusta, and healthy control volunteers over an 8-month period (April 2004-November 2004) through an institutional review board–approved protocol. Clinical data were collected in compliance with the Health Insurance Portability and Accountability Act. Healthy control volunteers with a history of cancer were excluded from analysis. All patients gave informed consent.

SELDI PROTEIN PROFILING

Serum samples were processed robotically using an automated workstation (Biomek 1000; Beckman Coulter, Inc, Fullerton, Calif) to increase the degree of reproducibility. A copper-treated chip array (IMAC-3 ProteinChip System; Ciphergen Biosystems, Inc, Fremont, Calif) was used for SELDI analysis, as previously described, with modification. Triplicate runs were performed for each serum sample, with the random placement of each sample in a 96-well bioprocessor format. Briefly, serum samples were prepared for SELDI analysis by vortexing 20 µL of serum with 30 µL of 8M urea with 1% CHAPS [3-{(3-cholamidopropyl)dimethylammonio}-1-propanesulfonic acid] in phosphate-buffered saline at 4°C for 10 minutes. Then, 100 µL of 1M urea with 0.125% CHAPS was added to the serum mixture and vortexed briefly. Next, phosphate-buffered saline was added to make a 1:5 dilution of the serum mixture, which was then added to the protein chip array. After 30 minutes of incubation at room temperature, the protein chips were washed with phosphate-buffered saline and air-dried. Afterward, 1 µL of saturated sinapinic acid solution in 0.5% trifluoroacetic acid and 50% acetonitrile was applied to each array twice, allowing the array to dry between each application. The SELDI instrument (Ciphergen Protein Biology System IIc; Ciphergen Biosystems, Inc) was used with an autoloader, which increases the high throughput tremendously. The protein chips were assayed with a laser intensity of 180 and a sensitivity of 8. A total of 192 shots were collected and averaged for each sample. The all-in-1 peptide molecular mass standard (Ciphergen Biosystems, Inc) was used to generate a peptide standard spectrum for mass accuracy calibration.

SELDI-TOF-MS ANALYSIS

All triplicate spectra were compiled after the completion of the SELDI assay. Mass calibration was performed using the all-in-1 peptide standard spectrum. The default background subtraction was applied, and the peak intensities were normalized using the total ion current from a mass-charge of 1000 to 100 000 Da. A biomarker detection software package (Ciphergen Biomarker Wizard; Ciphergen Biosystems, Inc) was used to autodetect protein peaks. Protein peaks were selected based on a first pass of a signal-noise ratio of 3 and a minimum peak threshold of 20% of all spectra. This process was completed with a second pass of peak selection at 0.2% of the mass window, and the estimated peaks were added. These selected protein peaks were averaged as clusters and exported to a commercially available software package (Biomarker Patterns; Ciphergen Biosystems, Inc) for further classification analysis.

CLASSIFICATION AND REGRESSION TREE ANALYSIS

Classification and regression tree analysis (CART) was performed using the common protein peaks identified by SELDI, as previously described. A decision classification tree algorithm was generated based on the identification of protein peaks differentially expressed between HNSCC and control samples. Classification and regression tree analysis splits the data into 2 groups or nodes by separating samples by rules based on the presence or absence of a peak sequentially until terminal nodes are reached. An optimal classification tree is then built, and 10-fold cross-validation is used to estimate the error rate of this tree. Tenfold cross-validation randomly splits the data into 10 partitions, 9 of which are used for training, and the remaining data set is used as a pseudo test sample to validate and test the classification tree. The validation process is performed 10 times. This technique allows use of the full complement of patients so that a classification tree with predictive value can be built when the sample size is small.

STATISTICAL ANALYSIS

The P value of each cluster was calculated using the nonparametric analysis to indicate the discriminate power of each cluster between groups. Analysis of clinical parameters was performed using the Fisher exact test, with significance defined as P<.05.
Portance of each cluster in comparison of HNSCC and
solved 348 clusters in the range of 2 to 100 KDa. The
ards) was used to analyze a total of 438 spectra and re-
Peak detection software (Ciphergen Biomarker Wiz-
equally distributed with respect to primary tumor size.
patients with HNSCC were
ccluded sinonasal and unknown primary sites (15%). The
HNSCC, the most common primary site was the oral cav-
mors, and 88% of those with laryngeal primary tumors
years after treatment compared with pretreatment
files resulted in correct classification of 86% of patients
went nonoperative treatment. Analysis of protein pro-
Posttreatment samples were available for 12 patients, 6
of whom were treated surgically and 6 of whom under-
went nonoperative treatment. Analysis of protein pro-
files resulted in correct classification of 86% of patients
months after treatment compared with pretreatment
classification analysis, which was performed with a commercial
software package (Ciphergen Biomarker Pattern Software). This software
age was developed specifically for tree-structured data
analysis, which is derived from a nonparametric regression
method based on the recursive partitioning method.10
The classification analysis constructed a decision tree (Figure 1) that correctly classified patients with HNSCC with 96% sensitivity and 93% specificity. The classification tree used 11 masses (8134.78, 3526.49, 2661.08, 2812.08, 2522.58, 5881.14, 2863.83, 2635.54, 6960.85, 3453.94, and 2886.12 Da) to construct a tree with 12 terminal nodes (Figure 2).

The classification rule is simple and straightforward. If the sample has a peak at 8134.78 Da, with an intensity of 3.74 or less and a peak at an intensity higher than 0.15, then it is placed in terminal node 5 and classified as HNSCC. Therefore, if the sample is placed in terminal node 8, it will be designated as normal and has a protein peak at 5881.14 Da with an intensity lower than 3.74, a protein peak at 8134.78 Da with an intensity higher than 1.21, and a protein peak at 2863.83 Da with an intensity higher than 0.63. The P value of each mass used in this classification tree is displayed in Table 2. Five of 11 peaks used in the classification tree of HNSCC and controls had significantly different intensity levels between these 2 groups. The remaining 6 peaks did not demonstrate significantly different intensity levels in the whole study cohort, but they served to give significant discriminate power in the subset of samples.

The small sample size precluded separation of training and blinded test sets. The whole sample set was used as a training set, and then a 10-fold validation was performed on samples as a test set. This 10-fold validation test correctly classified patients with HNSCC with 82% specificity, 76% sensitivity, and a positive predictive value of 83% (Table 3).

Analysis of patients with HNSCC by primary tumor stage correctly classified 59% of patients whose disease was stage T1/T2 and 46% of patients whose disease was stage T3/T4 (Table 3). Analysis by TNM stage correctly classified 80% of the patients with stage I/II disease and 45% of the patients with stage III/IV disease. Classification analysis correctly classified 83% of the patients with oral cavity tumors, 81% of those with oropharyngeal tumors, and 88% of those with laryngeal primary tumors (Table 3).

RESPONSE TO TREATMENT

Posttreatment samples were available for 12 patients, 6 of whom were treated surgically and 6 of whom underwent nonoperative treatment. Analysis of protein profiles resulted in correct classification of 86% of patients at 3 months after treatment and 75% of patients at 6 months after treatment compared with pretreatment samples from the same patients (Figure 3).

REPRODUCIBILITY OF SELDI ASSAY

The assay reproducibility was examined by running a pooled serum quality control sample in each chip array.
A total of 63 chip arrays were used for our study, which contained 63 quality control spectra. A total of 271 clusters were detected, and the coefficient of variance was calculated for mass and intensity. Average coefficients of variance of 0.06% and 27% were observed for mass and intensity, respectively.

Early detection strategies have been the focus of increased investigation in HNSCC to optimize outcomes and to minimize treatment-related morbidity and mortality. Patients with HNSCC often present with advanced-stage disease, which is associated with a poorer prognosis and requires more aggressive therapy, which in turn results in increased functional disability. Conventional diagnostic techniques, including direct inspection or increasingly sophisticated imaging technology, such as positron emission tomography–computed tomography, are limited in their ability to detect HNSCC at its earliest stages and are ineffective for use as a screening tool in high-risk populations. These limitations have fueled interest in the development of novel approaches for early detection through the identification of tumor-associated biomarkers and genetic alterations that are associated with HNSCC.

The molecular and cellular heterogeneity of HNSCC results in the expression of a variety of tumor cell products, with no single marker able to reliably distinguish patients with HNSCC from normal subjects. Analysis of the resultant protein profile may have greater utility in early diagnosis by selecting a combination of protein alterations (pattern recognition) rather than by focusing on specific tumor markers, which may vary between individual patients. Evaluation of conventional serologic markers, such as carcinoembryonic antigen, basic fibroblast growth factor, vascular endothelial growth factor, matrix metalloproteinase 2, squamous cell carcinoma antigen, cytokeratin 19, and antibodies to p53, have yielded mixed results and poor sensitivity and specificity for use.

Figure 1. Diagram of biomarker pattern analysis in the classification of patients with head and neck squamous cell carcinoma (HNSCC) and healthy normal controls. Each node represents a splitting rule where the samples will be split into 2 daughter nodes. Each node also displays the peak mass (M), the cutoff intensity level (I), the number of samples, and the composition of the samples. Each terminal node is classified as either normal or HNSCC based on the majority population in that terminal node.
in early detection.\textsuperscript{11-13} Investigations of nucleic acid–based microsatellite analysis for specific markers, tumor-specific aberrant promotor hypermethylation, and mitochondrial DNA mutations in blood and saliva suffer the same problem of tumor heterogeneity.\textsuperscript{14-16} These techniques are also technologically challenging and require significant expertise and time for processing, limiting the widespread use and application of these techniques for early detection. By identifying a pattern of protein characteristics, protein profiling takes into consideration not only the gain but also the loss of proteins in the disease state. Because cancer induces changes in cellular signaling pathways, it has been hypothesized that the most important tumor-associated biomarkers may be altered versions or reduced numbers of normal host molecules in the disease state and not necessarily increased numbers of new biomarkers.\textsuperscript{4}

Two-dimensional gel electrophoresis has traditionally been used to identify differences in protein expression in serum, saliva, or tissue specimens, with identi-
However, SELDI-TOF-MS analysis was able to correctly classify 83% of the patients with oral cavity tumors, 81% of those with oropharyngeal tumors, and 88% of those with laryngeal primary tumors, suggesting that proteomic differences exist between primary tumor sites in the head and neck. This finding suggests that there is true phenotypic variability between tumors arising in different primary sites, which may account for clinically observed differences in tumor behavior based on primary site.

Analysis of the protein profiles of a subset of 12 patients for whom posttreatment samples were available resulted in correct classification of 86% of patients at 3 months after treatment and 75% of patients at 6 months after treatment compared with their pretreatment samples. These data suggest that treatment of HNSCC is associated with alterations in the serum protein profile that can be detectable as early as 3 months after treatment and 75% of patients at 6 months after treatment compared with their pretreatment samples. The biomarker pattern software is simple and straightforward for use in data analysis; however, overfitting of the data can potentially occur with small sample sizes. Analysis of differences between primary tumor sites and response to treatment is limited by small numbers and length of follow-up. The control group and the group size limited us to validate our classification using 10-fold validation without a blinded independent test set. The biomarker pattern software is simple and straightforward for use in data analysis; however, overfitting of the data can potentially occur with small sample sizes. Analysis of differences between primary tumor sites and response to treatment is limited by small numbers and length of follow-up. The control group and the group

### Table 2. Statistical Comparison of the Intensity Levels of the Peak Used in the Decision Tree of Patients With HNSCC and Healthy Normal Controls

<table>
<thead>
<tr>
<th>Node No.</th>
<th>m-z, Da</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8134.78</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>2</td>
<td>3526.49</td>
<td>.21</td>
</tr>
<tr>
<td>3</td>
<td>2661.08</td>
<td>.11</td>
</tr>
<tr>
<td>4</td>
<td>2812.08</td>
<td>.45</td>
</tr>
<tr>
<td>5</td>
<td>2522.58</td>
<td>.02</td>
</tr>
<tr>
<td>6</td>
<td>5881.14</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>7</td>
<td>2863.83</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>8</td>
<td>2635.54</td>
<td>.11</td>
</tr>
<tr>
<td>9</td>
<td>6960.85</td>
<td>.002</td>
</tr>
<tr>
<td>10</td>
<td>3453.94</td>
<td>.14</td>
</tr>
<tr>
<td>11</td>
<td>2866.32</td>
<td>.045</td>
</tr>
</tbody>
</table>

Abbreviations: HNSCC, head and neck squamous cell carcinoma; m-z, mass-charge ratio.

In the present study, we found that SELDI-TOF-MS detected HNSCC with 82% sensitivity, 76% specificity, and a positive predictive value of 83%. These results are slightly lower than those of Wadsworth et al, who demonstrated a sensitivity of 83.3% and specificity of 90.0% for HNSCC detection with SELDI-TOF-MS. These differences may be related to intrinsic differences in the populations studied, with a younger mean age and greater ethnic diversity in the present study. Also, in our study, there was no attempt to separate controls with a history of smoking for analysis because of the small number of control smokers, whereas Wadsworth and colleagues analyzed control smokers separately and detected similarities in the protein profiles between control smokers and patients with HNSCC. If similarities exist, the inclusion of control smokers in the control group would be expected to decrease the specificity of SELDI-TOF-MS analysis by potentially increasing the number of false-positive classifications. Analysis of similarities between the protein profiles of control smokers and patients with HNSCC will be necessary to determine whether the effect of tobacco exposure on protein expression is independent of the effect of HNSCC or part of a spectrum of changes induced by tobacco exposure, with HNSCC representing one end of the spectrum.

We found that SELDI-TOF-MS analysis of patients with HNSCC by primary tumor stage and TNM stage showed a low sensitivity and specificity for subgroup detection. However, SELDI-TOF-MS analysis was able to correctly classify 83% of the patients with oral cavity tumors, 81% of those with oropharyngeal tumors, and 88% of those with laryngeal primary tumors, suggesting that proteomic differences exist between primary tumor sites in the head and neck. This finding suggests that there is true phenotypic variability between tumors arising in different primary sites, which may account for clinically observed differences in tumor behavior based on primary site.

There are several limitations to the present study. The small sample size limited us to validate our classification using 10-fold validation without a blinded independent test set. The biomarker pattern software is simple and straightforward for use in data analysis; however, overfitting of the data can potentially occur with small sample sizes. Analysis of differences between primary tumor sites and response to treatment is limited by small numbers and length of follow-up. The control group and the group
of patients with HNSCC differed significantly by age, sex, and smoking history, with the controls being significantly younger, more often female, and unlikely to be smokers, while the reverse was true for the HNSCC population. Thus, the control group is not truly representative of the population of interest for HNSCC screening and early detection.

Exposure to tobacco has been suggested to alter the profile of serum proteins; therefore, control smokers would appear to represent a target population for screening of high-risk individuals. Ideally, the control group for HNSCC detection analysis would consist primarily of control smokers to increase the validity of serum protein profile analysis. The inclusion of healthy smokers in control groups for SELDI-TOF-MS analysis may in fact increase the number of false-positive results and confound classification algorithms.\textsuperscript{7,20} It is unclear whether proteomic differences in otherwise healthy smokers represent (1) changes related to tobacco exposure alone, such as release of inflammatory mediators or cytokines; (2) premalignant changes; or (3) occult malignancy. If the latter is correct, then the use of control smokers would have the undesired effect of eliminating those proteomic profiles associated with early disease. Analysis of larger numbers of control smokers and elucidation of the underlying biomarkers associated with alterations in the protein expression profile in this high-risk group will be required to answer these questions.

**CONCLUSIONS**

Serum protein profiling can distinguish patients with HNSCC from controls with a high degree of sensitivity and specificity and may distinguish between tumors arising from different primary head and neck sites. Further-

Figure 3. Representative spectra from matched pretreatment and posttreatment samples from 6 patients with head and neck squamous cell carcinoma (HNSCC). The most significant difference consistently identified by biomarker pattern analysis between posttreatment and posttreatment samples was a peak at 47 205 Da, which was expressed in pretreatment samples but absent in posttreatment samples.
more, protein profiling appears to be able to distinguish between pretreatment and posttreatment samples based on alterations in the serum protein profile that can be detected as early as 3 months after treatment. The identification of differentially expressed protein peaks is useful for group classification and has the potential to identify protein biomarkers previously unsuspected to have a causal association because of small size and number. Further investigation into the clinical utility of this technology in HNSCC detection and surveillance is warranted.

Submitted for Publication: May 2, 2005; final revision received October 25, 2005; accepted November 9, 2005.

Correspondence: Bao-Ling Adam, PhD, Center for Biotechnology and Genomic Medicine, Medical College of Georgia, 1120 15th St, CA 4124, Augusta, GA 30912-4060 (badam@mcg.edu).

Financial Disclosure: None.

Previous Presentation: This study was presented in part at the annual meeting of the American Head and Neck Society; May 15, 2005; Boca Raton, Fla.

REFERENCES