Published in final edited form as:

Arch Pathol Lab Med. 2015 April ; 139(4): 494–507. doi:10.5858/arpa.2014-0131-OA.

Proteomic Analysis of Oral Cavity Squamous Cell Carcinoma Specimens Identifies Patient Outcome–Associated Proteins

Dr. Thomas M. Harris, PhD, Dr. Peicheng Du, PhD, Dr. Nicole Kawachi, MS, Dr. Thomas J. Belbin, PhD, Dr. Yanhua Wang, MD, PhD, Dr. Nicolas F. Schlecht, PhD, Dr. Thomas J. Ow, MD, Dr. Christian E. Keller, MD, Dr. Geoffrey J. Childs, PhD, Dr. Richard V. Smith, MD, Dr. Ruth Hogue Angeletti, PhD, Dr. Michael B. Prystowsky, MD, PhD, and Dr. Jihyeon Lim, PhD

Departments of Pathology (Drs Harris, Belbin, Wang, Ow, Childs, Prystowsky, and Lim and Ms Kawachi), Epidemiology & Population Health (Dr Schlecht), Developmental and Molecular Biology (Dr Angeletti), and Biochemistry (Dr Angeletti) and the Laboratory for Macromolecular Analysis & Proteomics (Drs Angeletti and Lim), Albert Einstein College of Medicine, Bronx, New York; High Performance and Research Computing, Department of Information Systems and Technology, UMDNJ, Newark, New Jersey (Dr Du); the Department of Otorhinolaryngology–Head & Neck Surgery, Montefiore Medical Center, Bronx, New York (Drs Ow and Smith); and the Department of Pathology, Henry Ford Hospital and Medical Group, Detroit, Michigan (Dr Keller)

Abstract

Context—Global proteomic analysis of oral cavity squamous cell carcinoma was performed to identify changes that reflect patient outcomes.

Objectives—To identify differentially expressed proteins associated with patient outcomes and to explore the use of imaging mass spectrometry as a clinical tool to identify clinically relevant proteins.

Design—Two-dimensional separation of digested peptides generated from 43 specimens with high-resolution mass spectrometry identified proteins associated with disease-specific death, distant metastasis, and loco-regional recurrence. RNA expressions had been correlated to protein levels to test transcriptional regulation of clinically relevant proteins. Imaging mass spectrometry explored an alternative platform for assessing clinically relevant proteins that would complement surgical pathologic diagnosis.

Results—Seventy-two peptide features were found to be associated with 3 patient outcomes: disease-specific death (9), distant metastasis (16), and loco-regional recurrence (39); 8 of them were associated with multiple outcomes. Functional ontology revealed major changes in cell adhesion and calcium binding. Thirteen RNAs showed strong correlation with their encoded proteins, implying transcriptional control. Reduction of DSP, PKP1, and TRIM29 was associated with significantly shorter time to onset of distant metastasis. Reduction of PKP1 and TRIM29 correlated with poorer disease-specific survival. Additionally, S100A8 and S100A9 reductions

Reprints: Jihyeon Lim, PhD, Laboratory for Macromolecular Analysis & Proteomics, Albert Einstein College of Medicine, 1300 Morris Park Ave, Bronx, NY 10461 (jihyeon.lim@einstein.yu.edu).

The authors have no relevant financial interest in the products or companies described in this article.

were verified for their association with poor prognosis using imaging mass spectrometry, a platform more adaptable for use with surgical pathology.

Conclusions—Using global proteomic analysis, we have identified proteins associated with clinical outcomes. The list of clinically relevant proteins observed will provide a means to develop clinical assays for prognosis and optimizing treatment selection.

Head and neck squamous cell carcinoma (HNSCC) is the seventh most prevalent cancer worldwide, representing a major international health problem.¹ Although treatment paradigms have evolved significantly over time, there has been relatively little change in 5-year survival since the 1970s, with the exception of survival among patients with human papillomavirus-driven oropharynx cancer.² To improve patient survival, significant effort has been made to guide treatment decisions based on using histologic criteria to prognosticate, but these efforts have been to no avail to date.^{3,4} As a result, some patients may receive additional treatment to eliminate the possibility of metastasis, and consequently develop unnecessary morbidity. On the other hand, less aggressive therapy runs the risk of the development of local or regional persistent or recurrence disease, which carries significant mortality.⁵ Hence, there is an urgent need of innovative technology and biomarkers to better evaluate tumor aggressiveness and potential for metastasis, as well as to predict response to treatments. Although effort has been made to classify HNSCCs using genomic patterns with the goal of guiding treatment decisions,^{6–8} proteomic analysis reflects the functional phenotype more directly than profiling DNA or RNA aberrations alone. For example, changes such as DNA copy number amplification could be canceled by transcriptional silencing via DNA methylation, or increased mRNA expression could be reversed by posttranscriptional silencing as with microRNAs.⁹ Therefore, efforts to use proteomics as an alternative classification method have increased.^{10–13}

Many previous attempts have been made to capture these differential protein expression patterns using patient tissue or plasma, or even closely related cancer cell lines.^{14–18} Many proteins have been detected that show strong potential to be developed further as biomarkers for clinical applications; these previous studies have had some limitations. Generally these studies have had small cohort sizes with limited clinical information. They have been focused on diagnostic markers that simply distinguish tumor from normal tissue. They have been restricted to relatively high-abundance proteins because of the limited nature of the separation method used, such as 2-dimensional gel electrophoresis.¹⁹ Our study differs in that we have used (1) an extensive tissue repository, with some cases having more than 10 years of detailed follow-up data; (2) a label-free quantitative proteomic approach to compare tumor tissues, and (3) 2-dimensional separation prior to mass spectrometry (MS) detection to provide wider proteome coverage. The MS-based proteomic analysis platform used in this study has been validated previously with 2 different cell lines.²⁰ This platform enables the study of several thousand proteins per experiment, yielding a robust data set from which to assess differential expression of proteins associated with different physiological events during tumorigenesis.

Here, we present protein expression that we have found to be associated with patient outcomes in oral cavity squamous cell carcinoma (OCSCC) identified from global

proteomic analysis. Previously, we have found that HNSCC produces site-specific molecular signatures by anatomic site for the oral cavity, oropharynx, and larynx.²¹ Hence, we focused this study solely on OCSCC. We also characterized the regulation of proteome changes at the transcriptional level as a part of a systems biology approach by integrating our global proteomic results with RNA expression data.

MATERIALS AND METHODS

Characteristics of Patient Population

Patients recruited for this study were undergoing treatment for histologically confirmed OCSCC at Montefiore Medical Center (Bronx, New York). All patients consented to participate in this study under protocols approved by the institutional review board. Only patients undergoing treatment with curative intent and no prior treatment for OCSCC were included in the study. From each patient, we obtained invasive primary tumor tissue at the time of surgical treatment. Patients were assessed, as appropriate, following their initial surgery. Medical records were kept noting the disease status at follow-up, the presence (or absence) of disease, the location and extent of failure, and death. The preliminary patient characteristics are described in Table 1. De-identified clinical and research data were used for the study. Primary tumor taken at surgery underwent multiple analyses including routine histopathology, genomic, transcriptional, and proteomic analyses. All OCSCC tissue samples were processed for total RNA, genomic DNA, and protein, and these banked samples were the source of the material to be used in this study.

Quality Control for Tumor Tissues

Study samples derived from resection specimens were taken fresh. Approximately 0.5 cm³ of tissue from invasive tumor was harvested in each case. This was further trisected and the middle one-third served as tumor control. Percentage of tumor and extent of necrosis were estimated by determining the number of microscopic fields containing tumor or necrotic tissue in relation to the total number of microscopic fields of the control sample on the contiguous slide. Normal mucosa was taken at a site at a significant distance away from the tumor identified by the surgeon (excisional biopsy) intraoperatively, or by the pathologist (resection specimen) when processing the specimen for frozen section or final diagnosis.

RNA Extraction and Detection

The RNA expression profiles used in the current study are a selected subset from our larger head and neck profiling project. All of the tissues were snap frozen within 30 minutes of surgical resection and kept at -80°C until analysis. Tissues were homogenized using a Brinkmann model PT 10/35 tissue homogenizer (KINEMATICA Inc, Bohemia, New York). RNA was isolated from tumor specimens in 1 mL volume of TRIzol (15596; Invitrogen, Carlsbad, California) following the manufacturer's suggested protocol and quantified on a NanoDrop ND-1000 spectrophotometer (NanoDrop, Wilmington, Delaware). Five hundred nanograms of total RNA was subjected to a linear amplification and biotin labeling protocol (Illumina TotalPrepRNA Amplification Kit; Ambion, Grand Island, New York) and used on Illumina HumanHT-12 v3 Expression BeadChips (Illumina, San Diego, California). These chips contained 48 000 probes targeting genes and known alternative splice variants from

the RefSeq database release 17 and UniGene build 188, as well as more than 1000 control beads to assay sample RNA quality, labeling reaction success, hybridization stringency, and signal generation. All expression data were quantile normalized and background subtracted before analysis using BeadStudio software (Illumina).

Protein Extraction and Digestion

Proteins were extracted and digested as previously reported.²⁰ Briefly, the protein pellets left after RNA isolation were reconstituted, reduced with Tris(2-carboxyethyl)phosphine, and digested successively with cyanogen bromide, endoproteinase Lys-C, and trypsin. Using the same tissue pieces for extractions facilitated the accurate comparison of RNA and protein expression levels. The concentration of the digested peptides was measured using a NanoDrop spectrophotometer at A₂₈₀, and a 100-μg aliquot was used for strong cation-exchange separation.

Two-Dimensional Liquid Chromatography–MS and Peptide Feature Array Generation

Digested peptide mixtures obtained were separated into 6 fractions using strong cation-exchange spin columns (polysulfoethyl A resin, 12 μm, 300 Å; PolyLC, Columbia, Maryland). Peptides were eluted stepwise with aqueous ammonium acetate solutions at salt concentrations of 60, 70, 80, 120, 200, and 300 mM in 25% acetonitrile at pH adjusted to 3 with formic acid. A mixture of 4 standard peptides (with different retention times and masses)²⁰ was added to each salt cut fraction as an internal control and analyzed using liquid chromatography–MS (LC-MS). The peptides served as markers for retention time alignment of the LC-MS spectra and used for correcting mass measurement drift. Each fraction was separated online by reversed-phase chromatography using a microflow UltiMate 3000 HPLC System (Dionex Corporation, Sunnyvale, California) interfaced with a QStar Pulsar i mass spectrometer (AB/Sciex, Foster City, California). Once acquired, the data were analyzed as previously described, as is shown in Figure 1.²⁰ Peak picking was performed with a program built in house,^{22,23} which reduced the raw spectra to peptide feature lists containing monoisotopic mass, retention times, and intensities. Next, to alleviate any between-run retention time variation, retention times in different runs were aligned by first correcting with internal standards added as described previously,²⁰ then by robust spline fitting. Subsequently, peptide features from different samples were matched using accurate mass and time tag approach with 0.3 Da and 1 minute tolerance each. Peptide feature arrays were constructed where each row contained the intensity value of each peptide obtained from different samples and statistical analysis performed.

Identification of Significant Peptides With Liquid Chromatography–Tandem MS

The generated peptide feature arrays contain peptide masses only. The sequence of each peptide in the array is unknown. Once associated with clinical outcomes, target lists of masses and retention times are generated for peptide sequence identification. Patient samples containing these targeted peptides and having the highest intensities were reanalyzed by liquid chromatography–tandem MS as described in Figure 2. Nanospray liquid chromatography–tandem MS analysis was performed on an Agilent 1100 HPLC system (Agilent, Palo Alto, California) and a QStar XL mass spectrometer (AB/Sciex). Acquired tandem MS spectra were manually curated to ensure high spectral quality and

peptide identification. Tandem MS spectra were searched against the human International Protein Index database (ver. 3.79) for sequence identification using the Mascot algorithm (Matrix Science, Boston, Massachusetts) and adjusted for a false discovery rate of 1%.

Imaging In Situ MS for Additional Verification

Samples were prepared as 8- μ m sections from 5 frozen OCSCC specimens and thaw mounted directly onto indium tin oxide-coated conductive glass slides (Delta Technologies, Loveland, Colorado). Sections were dried under vacuum for 30 minutes followed by fixation in graded ethanol (70%, 95%, and 98%; 30 seconds each) baths. Samples were dried again under vacuum for 10 minutes and spray coated with 10 mg/mL sinapinic acid (Fluka, St Louis, Missouri) in 60% acetonitrile and 0.2% trifluoroacetic acid aqueous solution using ImagePrep (Bruker Daltonics, Fremont, California). Matrix-assisted laser desorption/ionization (MALDI) mass spectra were acquired directly from the tissue sections with an ultrafleXtreme MALDI-TOF/TOF instrument (Bruker Daltonics). Data acquisition was performed in positive linear mode at constant laser power, optimized for detection of peaks in the 4- to 25-kDa range. A total of 800 laser shots were accumulated from each pixel of 200 $\times$ 200 μ m. The data were baseline subtracted and rendered into an ordered array of spectra with xy coordinates. Images were visualized with selected ion images using FlexImaging (Bruker Daltonics). Once the MS spectra were acquired, the slides were washed with ethanol to remove sinapinic acid, stained with hematoxylin and eosin, and photomicrographed at high resolution (Mirax; 3DHISTECH Kft, Budapest, Hungary). Photomicrographs were examined and then marked with regions of interest and overlaid with the images extracted from MS spectra for evaluation.

Identification of Proteins Discovered on Imaging In Situ MS

Imaging in situ MS (IMS) generates an array of proteins with molecular weight versus intensities. Hence, an additional step is required to find the identity of the peaks with the certain molecular weight measured in IMS. To achieve this, remaining tissues (~50 mg) after making sections were mechanically homogenized in 1 mL of 60% acetonitrile and 0.2% trifluoroacetic acid aqueous solution on ice and ultracentrifuged at 40 000g and 4°C for 20 minutes. Soluble protein concentration was determined using a 2xD Quant Kit (GE Healthcare Life Sciences, Pittsburgh, Pennsylvania). About 30 μ g proteins from each sample was separated on 16.5% Mini-PROTEAN Tris/Tricine Urea Precast Gel (Biorad, Hercules, California) with peptide molecular weight marker (Sigma, St Louis, Missouri) to target small protein ranges better. Entire lanes were excised using disposable grid cutters (The Gel Company, San Francisco, California) to produce 40 bands each, and band slices were placed in 96-well plates. Slices in the molecular weight range of 3 to 25 kDa were digested with trypsin. Tryptic peptides were separated on an online nano-HPLC interfaced to a QStar XL mass spectrometer (AB/Sciex). MS and alternating MS/MS data were acquired in data-dependent acquisition mode to acquire m/z and peptide sequences to be searched against SwissProt_AC 2013_04 using Mascot (Matrix Science) version 2.4.1. Protein hits were filtered to less than 1% false discovery rate and included those that were identified with at least 2 unique peptides.

Statistical Analysis of Protein and RNA Expression

As part of our initial screen for candidate proteins with possible clinical relevance, we restricted our analyses to a subset ($n = 28$) of OCSCC patients for whom we had at least 2 years of clinical follow-up, and assessed differences in peptide levels isolated from the 6 salt cuts between patients with and without evidence of disease progression within 2 years of diagnosis. Clinical outcome measures included disease-related mortality, defined as death with active HNSCC disease (either persistence, resistance, or metastasis) present; presence or absence of distant metastasis (DM); and presence or absence of a local/regional recurrence. Student t tests were performed using the aforementioned clinical parameters comparing mean peptide levels measured in OCSCC tumor samples. After the identities of the significantly associated peptides were established, proteins for further analysis were selected based upon the presence of multiple peptides with clinical associations corresponding to the same protein or protein family if a specific isotype could not be clearly determined. RNA expression data for each of these proteins were also examined using a gene expression array database generated on the same patient population. Pearson correlation analysis was done to identify RNAs that have strong correlation with protein expression (Figure 3, A through C). The protein and RNA expression data were transformed to remove the scalar differences using the function $[X - M]/[(Q_3 - Q_1)/1.35]$ where X is the intensity, M is the median and $(Q_3 - Q_1)$ represents the interquartile range. RNAs that showed significant correlations were then analyzed to determine whether significant associations were also observed with clinical outcomes that were in line with those observed with the matching proteins. Finally, univariate Kaplan-Meier analysis was performed to reassess the clinical associations on the full cohort of 43 OCSCC samples, taking into account time to outcome for cancer mortality, DM, and local/ regional recurrence. Patients were classified into 2 groups based on their tumor protein/RNA expression levels at diagnosis using quartile cut points.

RESULTS

Patient Characteristics

Table 1 shows the population demographics and clinicopathologic characteristics for the patient cohort used in the current study. All of the tumors in this study were squamous cell carcinomas of the oral cavity. Most of the cases presented with a TNM stage of III or IV (81%; $n = 35$) with 67% ($n = 29$) being positive for lymph node metastasis. The gender ratio was 3:1 male to female with a mean age at diagnosis of 62.6 years (± 14.9 SD).

Seventy-two Peptides Derived From 7269 Features Are Found to Be Associated With Patient Outcomes

Comparative LC-MS approaches rely on the observation that the peak intensity is proportional to the concentration of the peptide in the sample in most cases.²⁴ Therefore, measuring peak intensities and comparing these among multiple LC-MS data from different samples leads to a relative quantitation of peptide concentrations among samples. Figure 1 shows this procedure used to generate, separate, analyze, and identify peptide features enabling detection of differential expression of proteins among samples. After removing bins with confounding values due to peak overlap, this process generated an array with 7269

peptide features. To focus on features with potential clinical relevance, the lists from each of the strong cation-exchange fractions were examined for clinical association by taking each peptide and separating it into 2 populations based upon the binary value of the clinical parameter. Twenty-eight OCSCC specimens for which we had at least 24 months of clinical follow-up data were tested for 3 outcomes; disease-specific death (DSD), DM, and loco-regional recurrence. The intensity values from the MS run for each patient were compared using a Student *t* test and those with a significant difference between the 2 populations were kept. Overall, 72 unidentified features were selected that discriminated clinical outcomes with a *P* value of at least .01. Nine of the features associated only with DSD, 16 with DM, and 39 with loco-regional recurrence, as is shown in Table 2. Eight peptides among the 72 were associated with more than one clinical outcome, which is to be expected because many of the same pathways could lead to the same clinical outcome. Selected patient samples with the highest levels of the clinically significant features were reanalyzed for sequence identification of the peptides as shown in Figure 2. All 72 peptides were down-regulated in patients with poor outcomes by 2- to 200-fold from the mean of each group.

Functional Classification of Peptides Associated With Patient Outcomes

Proteins containing the digested peptide features associated with the different clinical parameters were independently separated into functional groups using the PANTHER classification system project²⁵ based upon their gene ontologies as well as Qiagen's (Valencia, California) Ingenuity Pathway Analysis software. Most of the proteins identified as having clinical associations in the current study related to structural components of the cytoskeleton and calcium ion binding (Figure 4, A through C) from the PANTHER analysis, whereas Ingenuity Pathway Analysis's functional classification system showed significant functional clusters in inflammatory disease, cell death and survival, cancer, and cellular growth and proliferation (Table 3). Down-regulation of members of the structural molecule group would result in major changes in the structural architecture of the cell affecting cell-cell and cell-matrix interactions, intracellular trafficking, and cell signaling. The largest related group of proteins consisted of members of the keratin (KRT) family, specifically KRT 5, 8, 14, and 16 and one or more members of the KRT 6 family. Most of these have known functions in the differentiation state of keratinocytes or have been seen to be coregulated in genetic syndromes or cancer.²⁶⁻³¹ Another group of proteins that had a significant association with DM are associated with cell-cell contact. Desmoplakin (DSP) and plakophilins (PKPs) are proteins associated with desmosome formation although they also have been shown to have other functions, for example transcriptional regulation.³² The remaining proteins, specifically TPM1, ACTB, ACTN4, LMNA, and TUBB, are important elements in the formation of the cytoskeleton itself. Reorganization of the cytoskeleton using proteins in this class is a vital process in the modulation of cell shape and migration during oncogenesis.

The binding class was comprised of 2 major functional subclasses: cytoskeletal structural element binding and calcium binding. As expected many proteins possessed both functions and are part of the cytoskeletal components mentioned above. Examples of these would be DSP, filamin-A, α -actinin-4, and spectrin α chain. Other calcium ion binding proteins include calmodulin or S100A8 and S100A9, which complex to form calprotectin, which

explains almost identical localization in Figure 5, a through e and f through j. Ca^{2+} homeostasis has been demonstrated to be a key element in mechanisms such as migration and differentiation and its disruption is a common theme in events involving tumorigenesis.³³ Some of the proteins in these 2 categories are implicated in several signaling pathways, for example HMGB (p53 pathway) and HSPB1 (p38 MAPK and VEGF signaling pathway).

Protein and RNA Expressions of TRIM29, DSP, and PKP1 Are Correlated and Their Reductions Are Associated With Poor Prognosis

To gain insights into possible regulatory mechanisms, we analyzed our RNA expression data for the peptides that exhibited clinical associations. For the current study we limited the analysis to a single clinical outcome, DM, to focus on studying OCSCC's predilection to early DM. Thirty-three of the 43 patients in the current study had corresponding RNA expression data on the identical tumor samples. There is no obligatory correlation between RNA expression levels and protein levels because posttranscriptional modulation and epigenetic influence, such as suppression via microRNA targeting, can reduce protein translation in the presence of high RNA levels. A correlation analysis between clinically significant peptide features from the array and RNA expression levels from the same tumor samples yielded 17 peptides, which directly correlated with RNA expression, suggesting regulation of these proteins at the transcription level. Because outliers can distort the size of the correlation coefficient, we examined scatter plots of protein expression versus RNA expression to see if any outliers were present, and whether the relationship was linear or curved. Examples of the correlation between RNA and protein expression for 3 peptides are shown in Figure 3: DSP, the tripartite motif-containing protein 29 (Trim29), and PKP1. Correlation of protein and RNA expression was not skewed by outliers and is evenly distributed across the range of expression. Mantel-Haenszel hazard ratios for either RNA or protein demonstrate statistically significant differences in 7 of the 17 peptides for time to onset of DM. These differences would not have been observed by RNA expression analysis alone, as is shown in Table 4, which demonstrates the advantage of using the proteomic strategy. Kaplan-Meier survival analysis (Figure 6, A through F) on time to DM demonstrated that both low RNA and peptide levels of Trim29, PKP1, or DSP are associated with decreased time to DM. Even though the initial *t* tests did not show a significant association between these 3 peptides and DSD, we also performed Kaplan-Meier survival analysis on time to death from OCSCC for these 3 proteins and observed a similar trend in the case of Trim29 and PKP1 but not for DSP (Figure 7, A through F).

Low S100A8 Expression Is Associated With Poor Prognosis and Verified Using Imaging Mass Spectral Assessment of Patient Tissues

Correlating RNA and protein expressions for validation can only be applied to the limited number of proteins under transcriptional control. Both S100A8 and S100A9 mRNA expression levels did not correlate with the corresponding protein levels. However, low S100A8 and S100A9 expressions were clearly associated with a poor prognosis in the global proteomic study, and this has been verified using an IMS assessment of patient tissue samples. Unlike LC-MS, IMS can analyze whole proteins directly from tissue sections as expression maps and provides another platform for validation for proteins. We provide an

example of the proteins S100A8 and S100A9 as IMS expression maps in Figure 5, a through e and f through j. Morphologic details of the tissues are illustrated in Figure 5, k through o, and zoomed-in areas around the black arrows in Figure 5, k through o, are shown in Figure 5, p through t. When comparing tissue sections from 5 patients with different outcomes using IMS, the images from 2 patients with poor prognosis (Figure, 5 a, b, f, and g) had significantly lower expression of S100A8 and S100A9 than the images from 3 patients (Figure 5, c through e and h through j) who survived without loco-regional recurrence or DM, consistent with the global proteomic study (Table 2).

COMMENT

Technological advances have allowed us to begin to see the global changes in gene and protein expressions that commonly occur in cancer. The challenge will be to successfully integrate this knowledge to identify markers that could be used to aid in the treatment of cancer. Correlating markers with clinical outcomes will eventually aid physicians in the treatment of their patients with the goal of a more individualized approach to prognosis and therapy. As technology has advanced, we are now gaining a more thorough understanding of how transcriptional regulation through mechanisms like epigenetic DNA modifications or translational regulation via mechanisms like microRNA silencing can alter protein expression and ultimately cell phenotype. To date the major focus for the identification of biomarkers has been RNA expression. The problem with this approach is that clinical outcomes ultimately arise from changes in protein expression or posttranslational modifications. There can be many factors that can diminish RNAs usefulness as a surrogate marker for protein expression. Small noncoding RNA like microRNA can effectively neutralize increases in mRNA levels by interfering with translation. A single microRNA can regulate the effective concentrations of many different mRNAs, which in turn may each be under independent regulatory feedback loops. The more degrees of separation between a marker and an outcome, the greater the probability the marker will be affected by multiple agents and not perform well in a larger independent population. In the current study we have taken a global proteomic approach that bypasses many of these problems because clinical outcomes have to arise from a change in proteins. Because an exhaustive analysis of all of the mechanisms involved in the regulation of the proteins, that is, digested peptides discovered in this study is well beyond the scope of a single report, we have taken the first steps to integrate these analyses in an ad hoc manner to select potentially transcriptionally regulated targets and demonstrate the ability of global RNA and proteomics to verify and complement the strengths of each platform while minimizing the weaknesses.

Reduced Cell-Cell Adhesion Through Down-Regulated DSP, PKP1, and KRTs Contributes to Poor Prognosis

We observed that patients in our cohort who developed DM had significantly lower DSP at both the RNA and peptide levels. Low levels of DSP were also associated with DSD. Desmosomes are highly organized intercellular junctions that provide mechanical integrity to tissues by anchoring intermediate filaments to sites of strong adhesion.³⁴ DSP and PKPs are major components of desmosomes that couple cytoskeletal elements to the plasma membrane at the cell-cell or cell-substrate adhesions.³⁴ DSP binds KRTs via its C terminus,

facilitating strong intercellular adhesion necessary to maintain tissue morphogenesis and architecture.³² DSP and PKP reduction have also been linked to dramatic loss of desmosomal functions in many diseases including cancers. For instance, DSP has been shown to function as a tumor suppressor in non-small cell lung cancer.³⁵

Yang et al³⁵ reported that DSP levels are reduced via DNA methylation as a primary mechanism in 73% of the lung cancer lines and 61% of the primary tumors they examined. We did not observe a strong relationship between DSP expression and DNA methylation in our cohort (data not shown), which suggests an alternative mechanism of silencing in OCSKC. They went on to demonstrate that overexpression of DSP led to increased plakoglobin (γ -catenin). In turn, it led to a decrease in TCF/LEF transcriptional activity and decreased expression of target genes such as Axin2 and MMP14. Silencing DSP via interfering RNA led to the opposite state for the previously mentioned genes. In the current study, tumors that led to DM also had low levels of DSP as well as low RNA levels of γ -catenin and elevated RNA levels of β -catenin. However, we did not observe any consistent statistically significant changes in the RNA expression levels of TCF/LEF, Axin2, or MMP14.

Plakophilins are proteins also involved in desmosome formation.³⁶ In the current study, lower levels of PKP1 protein and RNA were associated with DM and disease-specific survival. Sobolik-Delmaire et al³⁷ demonstrated that unlike the other members of its family, PKP1 is not a major element in the assembly of desmosomes. Yet it is translocated into the nucleus, where it can bind single-stranded DNA even though it lacks a canonical nuclear localization sequence. South et al³⁸ showed that PKP1 appears to play a role in whether the desmosomes formed are calcium dependent or independent. This in turn affects keratinocyte migration, with cells expressing low levels of PKP1 having increased calcium dependent desmosomes and greater migratory capacity.

Several KRTs were identified in our study as having an association with clinical outcome. Among these, Krt5 and Krt14 are known to be coregulated as part of an obligatory acidic/basic heteropolymer expressed in the basal proliferating layer of stratified epithelia. Wang et al³¹ noted that loss of Krt5 and Krt6 staining could distinguish between breast papillomas and papillary carcinomas. Alam et al³⁹ showed that loss of Krt14 expression can lead to decreased proliferation in the presence of a reciprocal increase in activated Notch1. We did not observe an increase in Notch1; however, Notch1 is deactivated in many squamous cell carcinomas of the head and neck, and lower levels of Krt14 and 5 are associated with earlier time to onset of DM in our cohort.^{40,41} Indeed, some level of quiescence could protect clinically occult micrometastases from destruction during adjuvant therapy.

Interestingly, 2 other KRTs, Krt16 and Krt6A or B, are also coregulated as part of an obligatory acidic/basic heteropolymer. Mutations in these KRTs have been shown to cause pachyonychia congenita; a rare genodermatosis characterized by dystrophic, thickened nails and painful palmoplantar keratoderma. Overall, down-regulation of DSP, PKP1, and several KRTs has been observed to associate with DM and DSD in this study.

These observations suggest that in squamous cell carcinoma of the oral cavity, cells that have lower levels of these proteins may have a loss in desmosomal integrity as well as a shift to mitotic quiescence while exhibiting an increased migratory capacity.

Suppressed TRIM29 Expression May Affect Anchorage-Independent Growth, Contributing to DM

TRIM29 is a member of the TRIM protein family containing both zinc finger and leucine zipper motifs but not a RING finger domain. It is also the ataxia-telangiectasia group D-complementing gene⁴² and its expression confers resistance to ionizing radiation.⁴³ In this study, expression of TRIM29 protein and its encoding RNA are down-regulated in patients with DM, indicating that this decrease in expression is regulated at the transcriptional level. Suppressed expression of TRIM29 has been observed in other cancers such as breast and prostate cancers,⁴⁴ implying its relevance to tumorigenesis. Hosoi et al⁴⁴ have shown that expression of TRIM29 leads to suppression of anchorage-independent growth (AIG) in multiple cancer cell lines. AIG is a crucial step in the acquisition of malignancy for cells to have the potential to migrate through the body, colonize other tissues, and grow metastatically.⁴⁵ This observation with our study suggests that suppressed TRIM29 may promote AIG, contributing to poorer outcome in OCSCC.

Additionally, TRIM 29 has been shown to interact with the p53 pathway, either by a direct interaction with p53 itself⁴⁶ and/or by interacting with the Tat-interacting protein 60 (Tip60), resulting in suppression of p53 activity.⁴⁷ Inactivation of the p53 pathway has widespread effects in HNSCC tumorigenesis.⁴⁸ Our data also suggest involvement of p53 pathway indirectly with 14 proteins (bold in Table 2). P53 were indicated as their top upstream regulator in Ingenuity Pathway Analysis. In the case of TRIM 29, its reduction was associated with poorer prognosis, even though it can function as both an oncogenic protein and a tumor suppressor,⁴⁹ possibly because of the state of p53, that is, posttranslational modification and/or mutation, directly involved in the OCSCC tumorigenesis.

Reduced Calcium-Binding Proteins Associated With Poor Prognosis

The discussion has focused on the 17 peptides showing strong correlation with mRNA expression, implying a transcriptional regulatory mechanism. However, that is not meant to imply that proteins with weaker correlations to their mRNAs lack importance in OCSCC progression. They may be regulated at the translational level by the actions of microRNAs or at the posttranslational level by functions like targeted degradation due to modifications such as ubiquitination, or they could be affected by environmental factors such as Ca²⁺ homeostasis. According to the functional classification described above, one of the major functional subclasses belongs to calcium-binding proteins. They are a large and heterogeneous group of proteins that participate in numerous cellular functions (eg, Ca²⁺ homeostasis and Ca²⁺ signaling pathways).⁵⁰ S100A8, S100A9, CALM, ACTN4, FLNA, SPTAN1, ANAXA2, and DSP are calcium-binding proteins associated with poor prognosis in this study. Three of these, S100A8, S100A9, and CALM, are EF hand proteins. EF hand proteins bind Ca²⁺ with high affinity to helix-loop-helix motifs that are repeated from 2 to 12 times. They not only buffer Ca²⁺ but also transduce calcium signals by changing

conformation after binding. The current study shows that they are markedly down-regulated in the tumors of patients with poor prognosis. This observation from global proteomics data has also been independently verified using IMS, which also demonstrated the loss of S100A8 and S100A9 in the patient with poor outcome (Figure 5, a through e and f through j). The S100A8 and S100A9 complex calprotectin has been suggested to act as a switch between differentiation and proliferation. The switch is believed to be triggered by calcium binding, which releases S100A8 from the complex, allowing it to interact with telomerase,⁵¹ disruption of which appears to be required in essentially all tumors for immortalization.

Patient Outcome–Associated Proteins Generated From Global Proteomic Analysis Can Be Potential Prognostic Indicators

It has been shown in the present study that global proteomic analysis can produce a set of peptides associated with clinical outcomes for OCSCC. Overall, it is clear that our global proteomic analysis has revealed outcome-associated proteins that may directly or indirectly be involved in regulatory mechanisms affecting prognosis. With further validation we propose that these outcome-associated proteins generated from global proteomic analysis may be used as prognostic indicators. Additionally, we provide proof-of-concept data for a simple biomarker assay, demonstrating that S100A8 and S100A9 are reduced as measured using IMS on a tumor section taken from a sample associated with poor outcome.

Future studies will be focused on validating these clinically significant peptides, and we plan to assess their clinical utility using more cost- and time-effective analysis platforms. Multiple-reaction monitoring using a triple-quadrupole mass spectrometer⁵² is a potential option, because it can be developed into an assay that screens multiple peptides at once. This will also permit us to test the combination of proteins as a panel to improve the sensitivity and specificity of the assay as development proceeds. The additional cohort can be tested using IMS as well, because it acquires protein expression directly from frozen section samples. Imaging in situ MS is more adaptable to the clinical environment because of its potential to provide data that rival immunohistochemistry data for reliability and reproducibility, with the added benefit of the ability to screen a few hundred proteins at once with a very rapid processing time. Overall, our study is a step toward the development of a proteomics approach to prognostication in OCSCC.

Acknowledgments

This study was supported by grants from the US National Institute of Health (NIH), R33CA103547 (Dr Prystowsky), R21CA103547 (Dr Prystowsky), S10RR021056 (Dr Angeletti), S10RR025128 (Dr Angeletti), S10RR015859 (Dr Angeletti), and P30CA013330 (Dr Goldman). It was also supported in part by CTSA grants UL1RR025750, KL2RR025749, and TL1RR025748 from the National Center for Research Resources. The authors would like to thank Rani Sellers, DVM, PhD, at the Albert Einstein College of Medicine histotechnology and comparative pathology facility for her assistance. The authors also thank Shannon Cornett, PhD, at Bruker Daltonics for helpful discussions.

References

1. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA Cancer J Clin.* 2011; 61(2):69–90. [PubMed: 21296855]
2. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol.* 2010; 11(8):781–789. [PubMed: 20451455]

3. Brandwein-Gensler M, Smith RV, Wang B, et al. Validation of the histologic risk model in a new cohort of patients with head and neck squamous cell carcinoma. *Am J Surg Pathol*. 2010; 34(5): 676–688. [PubMed: 20414102]
4. Brandwein-Gensler A, Teixeira MS, Lewis CM, et al. Oral squamous cell carcinoma—histologic risk assessment, but not margin status, is strongly predictive of local disease-free and overall survival. *Am J Surg Pathol*. 2005; 29(2):167–178. [PubMed: 15644773]
5. Goodwin WJ. Salvage surgery for patients with recurrent squamous cell carcinoma of the upper aerodigestive tract: when do the ends justify the means? *Laryngoscope*. 2000; 110(3):1–18. [PubMed: 10714711]
6. Adrien LR, Schlecht NF, Kawachi N, et al. Classification of DNA methylation patterns in tumor cell genomes using a CpG island microarray. *Cytogenet Genome Res*. 2006; 114(1):16–23. [PubMed: 16717445]
7. Belbin TJ, Singh B, Barber I, et al. Molecular classification of head and neck squamous cell carcinoma using cDNA microarrays. *Cancer Res*. 2002; 62(4):1184–1190. [PubMed: 11861402]
8. Belbin TJ, Singh B, Smith RV, et al. Molecular profiling of tumor progression in head and neck cancer. *Arch Otolaryngol Head Neck Surg*. 2005; 131(1):10–18. [PubMed: 15655179]
9. Harris T, Jimenez L, Kawachi N, et al. Low-level expression of miR-375 correlates with poor outcome and metastasis while altering the invasive properties of head and neck squamous cell carcinomas. *Am J Pathol*. 2012; 180(3):917–928. [PubMed: 22234174]
10. Belbin TJ, Bergman A, Brandwein-Gensler M, et al. Head and neck cancer: reduce and integrate for optimal outcome. *Cytogenet Genome Res*. 2007; 118(2–4):92–109. [PubMed: 18000360]
11. Patel V, Leethanakul C, Gutkind JS. New approaches to the understanding of the molecular basis of oral cancer. *Crit Rev Oral Biol Med*. 2001; 12(1):55–63. [PubMed: 11349962]
12. Matta A, DeSouza LV, Shukia NK, et al. Prognostic significance of head-and-neck cancer biomarkers previously discovered and identified using iTRAQ-labeling and multidimensional liquid chromatography–tandem mass spectrometry. *J Proteome Res*. 2008; 7(5):2078–2087. [PubMed: 18407684]
13. Balog J, Sasi-Szabo L, Kinross J, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci Transl Med*. 2013; 5(194)
14. Bleijerveld OB, Brakenhoff RH, Schaaij-Visser TBM, et al. Protein signatures associated with tumor cell dissemination in head and neck cancer. *J Proteomics*. 2011; 74(4):558–566. [PubMed: 21262401]
15. Chi LM, Lee CW, Chang KP, et al. Enhanced interferon signaling pathway in oral cancer revealed by quantitative proteome analysis of microdissected specimens using (16)O/(18)O labeling and integrated two-dimensional LC-ESI-MALDI tandem MS. *Mol Cell Proteomics*. 2009; 8(7):1453–1474. [PubMed: 19297561]
16. Baker H, Patel V, Molinolo AA, et al. Proteome-wide analysis of head and neck squamous cell carcinomas using laser-capture microdissection and tandem mass spectrometry. *Oral Oncol*. 2005; 41(2):183–199. [PubMed: 15695121]
17. Koike H, Uzawa K, Nakashima D, et al. Identification of differentially expressed proteins in oral squamous cell carcinoma using a global proteomic approach. *Int J Oncol*. 2005; 27(1):59–67. [PubMed: 15942644]
18. Turhani D, Krapfenbauer K, Thurnher D, Langen H, Fountoulakis M. Identification of differentially expressed, tumor-associated proteins in oral squamous cell carcinoma by proteomic analysis. *Electrophoresis*. 2006; 27(7):1417–1423. [PubMed: 16568407]
19. Wu WG, Tang XM, Hu W, et al. Identification and validation of metastasis-associated proteins in head and neck cancer cell lines by two-dimensional electrophoresis and mass spectrometry. *Clin Exp Metastasis*. 2002; 19(4):319–326. [PubMed: 12090472]
20. Sudha R, Kawachi N, Du P, et al. Global proteomic analysis distinguishes biologic differences in head and neck squamous carcinoma. *Lab Invest*. 2007; 87(8):755–766. [PubMed: 17558418]
21. Belbin TJ, Schlecht NF, Smith RV. Site-specific molecular signatures predict aggressive disease in HNSCC. *Head Neck Pathol*. 2008; 2(4):243–256. [PubMed: 20614290]
22. Du P, Sudha R, Prystowsky MB, Angeletti RH. Data reduction of isotope-resolved LC-MS spectra. *Bioinformatics*. 2007; 23(11):1394–1400. [PubMed: 17496000]

23. Du PC, Stolovitzky G, Horvatovich P, et al. A noise model for mass spectrometry based proteomics. *Bioinformatics*. 2008; 24(8):1070–1077. [PubMed: 18353791]
24. America AHP, Cordewener JHG. Comparative LC-MS: a landscape of peaks and valleys. *Proteomics*. 2008; 8(4):731–749. [PubMed: 18297651]
25. Mi HY, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc*. 2013; 8(8):1551–1566. [PubMed: 23868073]
26. Linskey KR, Gimbel DC, Zukerberg LR, et al. BerEp4, cytokeratin 14, and cytokeratin 17 immunohistochemical staining aid in differentiation of basaloid squamous cell carcinoma from basal cell carcinoma with squamous metaplasia. *Arch Pathol Lab Med*. 2013; 137(11):1591–1598. [PubMed: 24168496]
27. Eliason MJ, Leachman SA, Feng BJ, Schwartz ME, Hansen CD. A review of the clinical phenotype of 254 patients with genetically confirmed pachyonychia congenita. *J Am Acad Dermatol*. 2012; 67(4):680–686. [PubMed: 22264670]
28. Trost A, Desch P, Wally V, et al. Aberrant heterodimerization of keratin 16 with keratin 6A in HaCaT keratinocytes results in diminished cellular migration. *Mech Ageing Dev*. 2010; 131(5):346–353. [PubMed: 20403371]
29. Moll R, Franke WW, Schiller DL, Geiger B, Krepler R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell*. 1982; 31(1):11–24. [PubMed: 6186379]
30. Haber RM, Drummond D. Pachyonychia congenita with laryngeal obstruction. *Pediatr Dermatol*. 2011; 28(4):429–432. [PubMed: 21554383]
31. Wang Y, Zhu JF, Liu YY, Han GP. An analysis of cyclin D1, cytokeratin 5/6 and cytokeratin 8/18 expression in breast papillomas and papillary carcinomas. *Diagn Pathol*. 2013; 8:8. [PubMed: 23327593]
32. Green KJ, Simpson CL. Desmosomes: new perspectives on a classic. *J Invest Dermatol*. 2007; 127(11):2499–2515. [PubMed: 17934502]
33. Roderick HL, Cook SJ. Ca^{2+} signalling checkpoints in cancer: remodelling Ca^{2+} for cancer cell proliferation and survival. *Nat Rev Cancer*. 2008; 8(5):361–375. [PubMed: 18432251]
34. Yin TF, Green KJ. Regulation of desmosome assembly and adhesion. *Semin Cell Dev Biol*. 2004; 15(6):665–677. [PubMed: 15561586]
35. Yang LL, Chen Y, Cui TT, et al. Desmoplakin acts as a tumor suppressor by inhibition of the Wnt/beta-catenin signaling pathway in human lung cancer. *Carcinogenesis*. 2012; 33(10):1863–1870. [PubMed: 22791817]
36. Schmidt A, Jager S. Plakophilins—hard work in the desmosome, recreation in the nucleus? *Eur J Cell Biol*. 2005; 84(2–3):189–204. [PubMed: 15819400]
37. Sobolik-Delmaire T, Reddy R, Pashaj A, Roberts BJ, Wahl JK. Plakophilin-1 localizes to the nucleus and interacts with single-stranded DNA. *J Invest Dermatol*. 2010; 130(11):2638–2646. [PubMed: 20613778]
38. South AP, Wan H, Stone MG, et al. Lack of plakophilin 1 increases keratinocyte migration and reduces desmosome stability. *J Cell Sci*. 2003; 116(16):3303–3314. [PubMed: 12840072]
39. Alam H, Sehgal L, Kundu ST, Dalal SN, Vaidya MM. Novel function of keratins 5 and 14 in proliferation and differentiation of stratified epithelial cells. *Mol Biol Cell*. 2011; 22(21):4068–4078. [PubMed: 21900500]
40. Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science*. 2011; 333(6046):1157–1160. [PubMed: 21798893]
41. Agrawal N, Frederick MJ, Pickering CR, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science*. 2011; 333(6046):1154–1157. [PubMed: 21798897]
42. Leonhardt EA, Kapp LN, Young BR, Murnane JP. Nucleotide-sequence analysis of a candidate gene for ataxia-telangiectasia group D (ATDC). *Genomics*. 1994; 19(1):130–136. [PubMed: 8188213]
43. Murnane JP, Zhu Y, Young BR, Christman MF. Expression of the candidate A-T gene ATDC is not detectable in a human cell-line with a normal response to ionizing-radiation. *Int J Radiat Biol*. 1994; 66(6):S77–S84. [PubMed: 7836856]

44. Hosoi Y, Kapp LN, Murnane JP, et al. Suppression of anchorage-independent growth by expression of the ataxia-telangiectasia group D complementing gene, ATDC. *Biochem Biophys Res Commun.* 2006; 348(2):728–734. [PubMed: 16890201]
45. Gassmann P, Haier J. The tumor cell–host organ interface in the early onset of metastatic organ colonisation. *Clin Exp Metastasis.* 2008; 25(2):171–181. [PubMed: 18058027]
46. Yuan ZG, Villagra A, Peng LR, et al. The ATDC (TRIM29) protein binds p53 and antagonizes p53-mediated functions. *Mol Cell Biol.* 2010; 30(12):3004–3015. [PubMed: 20368352]
47. Sho T, Tsukiyama T, Sato T, et al. TRIM29 negatively regulates p53 via inhibition of Tip60. *Biochim Biophys Acta.* 2011; 1813(6):1245–1253. [PubMed: 21463657]
48. Mao L, Hong WK, Papadimitrakopoulou VA. Focus on head and neck cancer. *Cancer Cell.* 2004; 5(4):311–316. [PubMed: 15093538]
49. Zhang ZY, Gerhard DS, Nguyen L, et al. Fine mapping and evaluation of candidate genes for cervical cancer on 11q23. *Genes Chromosomes Cancer.* 2005; 43(1):95–103. [PubMed: 15672406]
50. Yáñez, M.; Gil-Longo, J.; Campos-Toimil, M. Calcium binding proteins. In: Islam, MS., editor. *Calcium Signaling.* Dordrecht, the Netherlands: Springer Netherlands; 2012. p. 461–482. *Advances in Experimental Medicine and Biology.* p. 740
51. Rosenberger S, Thorey IS, Werner S, Boukamp P. A novel regulator of telomerase: S100A8 mediates differentiation-dependent and calcium-induced inhibition of telomerase activity in the human epidermal keratinocyte line HaCaT. *J Biol Chem.* 2007; 282(9):6126–6135. [PubMed: 17197440]
52. Anderson L, Hunter CL. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics.* 2006; 5(4):573–588. [PubMed: 16332733]

Figure 1. Proteomic workflow to generate peptide feature array (first analysis). Features are peptide peaks detected from liquid chromatography–mass spectrometry. Abbreviations: CNBr, cyanogen bromide; Lys-C, endoproteinase Lys-C; MS, mass spectrometry; OCC, oral cavity cancer; RPLC-MS, reverse phase liquid chromatography–mass spectrometry; SCX, strong cation exchange chromatography.

Figure 2.

Proteomic workflow to sequence target features using tandem mass spectrometry (MS/MS) (second analysis). Abbreviation: RPLC, reverse-phase liquid chromatography.

Figure 3.

Examples of the correlation between RNA and protein expression using 3 peptides: scatter plots for TRIM29 (A), PKP1 (B) and DSP (C) of RNA expression (x-axis) versus protein expression (y-axis) demonstrate positive correlation with an even distribution. Pearson correlation values were $R^2 = 0.43$, $P < .001$ (TRIM29); $R^2 = 0.62$, $P < .001$ (PKP1); and $R^2 = 0.31$, $P < .001$ (DSP).

Figure 4.

Molecular functional classification of proteins associated with disease-specific death (A), distant metastasis (B), and loco-regional recurrence (C). The binding indicates selective, noncovalent, often stoichiometric interaction of proteins with other molecules. Two major functional subclasses, cytoskeletal protein binding and calcium binding, were included in the study.

Figure 5.

S100A8 (a–e) and S100A9 (f–j) expression maps (1.44 $\times$) across 5 oral cavity squamous cell carcinoma sections acquired using in situ mass spectrometry with $200 \times 200\text{-}\mu\text{m}$ pixel sizes. Two patients (HN201 and HN224) on the left with poor prognosis, disease-specific death, distant metastasis (DM), and loco-regional recurrence (LRR) were compared with 3 patients (HN227, HN229, and HN341) who survived without DM and LRR; contiguous sections (k–o); higher-power images (p–t) focused around the region noted with black arrows on sections in k–o (hematoxylineosin, original magnifications $\times 150$ [k–o] and $\times 200$ [p–t]).

Figure 6.

Kaplan-Meier plots for onset to distant metastasis: Kaplan-Meier function curves are plotted for low (gray lines) versus high (black lines) tumor expression based on a 25th-percentile expression value cut point. Trim29 protein (A), Trim29 RNA (B), PKP1 protein (C), PKP1 RNA (D), DSP protein (E), DSP RNA (F). Censored subjects are indicated by vertical hash marks.

Figure 7.

Kaplan-Meier plots for cancer survival. Kaplan-Meier function curves are plotted for low (grey lines) versus high (black lines) tumor expression based on a 25th-percentile expression value cut point. Trim29 protein (A), Trim29 RNA (B), PKP1 protein (C), PKP1 RNA (D), DSP protein (E), DSP RNA (F). Censored subjects are indicated by vertical hash marks.

Table 1

Characteristics of Patient Population

	No.	%
Age at diagnosis, y		
58	15	34.9
59–67	12	27.9
68	16	37.2
Sex		
Female	11	25.6
Male	32	74.4
Race ^a		
White	35	81.4
African American	5	11.6
Asian	2	4.7
ND	1	2.3
Ethnicity ^a		
Non-Hispanic	28	65.1
Hispanic/Latino	15	34.9
Nodal stage		
N0	14	32.6
N+	29	67.4
TNM stage		
I/II	8	18.6
III/IV	35	81.4
Total	43	

Abbreviations: ND, no data; N0, node negative; N+, node positive; TNM, tumor, node, metastasis.

^aNumbers may not sum to column totals because of missing data or exclusion of “unknown/other” categories.

Table 2

Peptide Features Associated With Patient Outcome and Their Sequence Identified From the Global Proteomic Study

Gene Symbol ^a	Peptide Sequence ^b	Associated Outcome ^c
ACTB	LLTEAPLNPK	LRR
ACTN4	ASFNHFDKDHGGALGPEEFK	LRR
ANXA2	ISDTSGDFRK	DSD, LRR
ANXA2	DIISDTSGDFR	LRR
CALM3; CALM2; CALM1	DNGYISAAELR	LRR
CASPI	ALIDSVIPK	LRR
CBR1	VVNVSIMSVR	LRR
CKAP4	QREELGQGLQGVEQK	LRR
CRABP2	IAVAAASKPAVEIK	LRR
CRABP2	PNFSGNWK	LRR
CSTB	GAPSATQPATAETQHIADQVR	LRR
CYB5R3	DILLRPELEELR	DSD
DSP	AELIVQPELK	DSD, DM
DSP	QLQNIIQATSR	DSD
DSP	TGSQYDIQDAIDK	DM
DSP	SMVEDITGLR	DM
DSP	LYEIEDEKR	LRR
EEF1A2; EEF1A1	AAGVGEFEAGISK	LRR
EIF6	HGLLVNNTTDQELQHIR	LRR
ENO1	IGAEVYHNLK	LRR
FKBP4	RGEAHLAVNDFELAR	LRR
FLNA	LVSNSLSHETSSVFVDSLTK	DSD, DM
GAPDH	GALQNIIPASTGAAK	LRR
GSTK1	FLTAVNLEHPEMLEK	DM
HMGB2	IKSEHPGLSIGDTAK	DSD
HSP90AA1; HSP90AA2	MGRGTKVILHLKEDQTEYLEER	DSD
HSP90AB1; HSP90AA1	SLIINTFYSNK	LRR
HSPA8	NQVAMNPTNTVFDAK	LRR
HSPB1	LFDQAFGLPR	LRR
IQGAP1	VDQIQEIVTGNPTVIK	LRR
ITGB4	LVFSALGPTSLR	LRR
KRT14	VMDVHDGKVVSTHEQVLR	DSD, DM, LRR
KRT14	VVSTHEQVLR	DM
KRT14	APSTYGGGLSVSSSR	DM
KRT14; KRT16	STYGGGLSVSSSR	DM
KRT14; KRT16	ASLENSLEETK	DM
KRT14; KRT16	LEQEIATYR	DM
KRT16	LLEGEDAHLSSQQASGQSYSSR	DM

Gene Symbol ^a	Peptide Sequence ^b	Associated Outcome ^c
<i>KRT16</i>	TIENAQPILQIDNAR	LRR
<i>KRT16; KRT16P3</i>	IIAATIENAQPILQIDNAR	DSD, DM
<i>KRT5</i>	GLGVGFGSGGGSSSVK	DM
<i>KRT5</i>	NMQDLVEDFK	LRR
<i>KRT5</i>	VLETKWTLLQEQGTK	LRR
<i>KRT5</i>	QLDSIVGER	LRR
<i>KRT6A; KRT6B; KRT6C</i>	EYQELMNVK	DM
<i>KRT6A; KRT6B; KRT6C</i>	DVDAAVMNKKVELQAK	DM, LRR
<i>KRT6A; KRT6B; KRT6C</i>	TAAENEFVTLKK	LRR
<i>KRT6B; KRT6C</i>	SGFSSISVSR	DM
<i>KRT8</i>	LSELEAALQR	LRR
<i>LMNA</i>	LQEKEDLQELNDR	DSD
<i>LMNA</i>	EGELHDLR	LRR
<i>NDRG1</i>	NPGNLHLFINAYNSR	DM
<i>PDCD6IP</i>	LLDEEEATDNDLR	LRR
<i>PKP1</i>	NMLGTLAGANSLR	DSD
<i>PKP1</i>	SPNQNVQAAAAGALR	DSD
<i>PRDX2</i>	RLSEDYGVLK	LRR
<i>PSMB1</i>	GAVYSFDPVGSYQR	LRR
<i>S100A8</i>	YSLIKGNFHAVYR	LRR
<i>S100A9</i>	NIETIINTFHQY	LRR
<i>S100A9</i>	VIEHIMEDLDTNADK	LRR
<i>S100A9</i>	LGHPTLNLQGEFKELVR	LRR
<i>SERPIN13</i>	LGFDLFK	LRR
<i>SPTAN1</i>	LSDDNTIGKEEIQQR	LRR
<i>TF</i>	EFQLFSSPHGK	LRR
<i>TPM1; TPM3</i>	GTEDELDKYSEALKDAQEK	DSD
<i>TRIM29</i>	VIMDALDER	DM
<i>TUBB; TUBB4; TUBB3</i>	ISEQFTAMFR	DM
<i>TXN</i>	EKLEATINELV	DM
<i>UQCRC1</i>	ADLTEYLSHYK	DSD
<i>UQCRC2</i>	YEDFSNLGTTHLLR	DSD, DM
<i>VCL</i>	NQWIDNVEK	LRR
<i>YWHAZ; SFN</i>	NLLSVAYK	LRR, DSD

Abbreviations: DSD, disease-specific death; DM, distant metastasis; LRR, loco-regional recurrence.

^a All 72 peptide features were down-regulated in the tumors used in this study. When specific isoforms are not clear, multiple gene symbols are listed.

^b Peptide sequence was identified with less than 1% false discovery rate in Mascot search algorithm.

^c Outcome association at $P < .001$ for 70–300 mM fractions and $P < .01$ for 60 mM fraction based on the number of total features using Student t test. Bold proteins were linked to a top upstream regulator p53 using Ingenuity Pathway Analysis.

Table 3Ingenuity Pathway Analysis Functional Classification of Proteins With Clinical Associations^a

Function	Molecules	Molecules, No.	P
Inflammatory disease	CBR1, EEF1A1, ENO1, FKBP4, FLNA, KRT14, KRT16, KRT5, KRT6A, PSMB1, S100A8, S100A9, SERPINB13, SFN, TPM1, TPM3, TUBB, TUBB3, TUBB4A, VCL	20	<.001
Cell death and survival	ANXA2, CALM1, CASP1, CBR1, CRABP2, CSTB, EEF1A1, EIF6, ENO1, FLNA, GAPDH, HMGB2, HSP90AA1, HSP90AB1, HSPA8, HSPB1, ITGB4, KRT14, KRT6A, KRT8, LMNA, NDRG1, PDCD6IP, PRDX2, S100A8, S100A9, SFN, TF, TPM1, TRIM29, TUBB, TUBB3, TXN, VCL, YWHAZ	35	<.001
Cancer	ANXA2, ENO1, HSP90AA1, HSP90AB1, HSPB1, ITGB4, TPM1, TPM3, TUBB, TUBB3, TUBB4A	11	<.001
Cellular growth and proliferation	ACTB, ACTN4, ANXA2, CALM1, CASP1, CRABP2, DSP, EEF1A1, ENO1, FKBP4, FLNA, HMGB2, HSP90AA1, HSP90AB1, HSPA8, IQGAP1, ITGB4, KRT14, KRT16, KRT8, LMNA, NDRG1, PDCD6IP, PKP1, S100A8, S100A9, SFN, SPTAN1, TF, TPM1, TUBB, TUBB3, TXN	33	<.001

^a Multiple peptides were grouped into top scoring proteins using Mascot algorithm. When specific isoforms are not clear, multiple gene symbols are listed.

Table 4Univariate Hazard Ratio for Time to Distant Metastasis (DM)^a

	Protein to RNA Correlation Coefficient	Protein ^b		RNA	
		Time to DM Hazard Ratio ^c	Log-Rank Test <i>p</i> ^d	Time to DM Hazard Ratio ^c	Log-Rank Test <i>p</i> ^d
DSP	0.56	11.6 (1.8–76.8)	.01	4.7 (0.9–24.7)	.07
KRT14	0.69	5.5 (1.3–23.3) ^e	.02	1.5 (0.4–6.4) ^e	.57
KRT16	0.59	8.6 (1.7–44.9)	.01	3.0 (0.7–13.0)	.14
KRT5	0.63	6.0 (1.4–25.1) ^e	.01	1.6 (0.4–7.6)	.57
KRT6B	0.70	7.9 (1.9–33.7) ^e	.005	2.5 (0.6–10.7) ^e	.20
PKP1	0.79	6.5 (1.1–37.2)	.04	37.0 (5.3–257.7)	.003
TRIM29	0.71	6.5 (1.1–37.2)	.04	41.8 (5.9–297.3)	.002

^aTumor expression groups were defined upon a 25th-percentile cutoff except where noted.^bIntensities of digested peptides were used to represent those of proteins.^cMantel-Haenszel.^dMantel-Cox.^eTumor expression groups were defined upon a 50th-percentile cutoff.