Laboratory Investigation (2007) 87, 755–766

Global proteomic analysis distinguishes biologic differences in head and neck squamous carcinoma

Rajagopalan Sudha[1], Nicole Kawachi[1], Peicheng Du[2], Edward Nieves[2,3], Thomas J Belbin[1], Abdissa Negassa[4], Ruth Hogue Angeletti[2,3] and Michael B Prystowsky[1]

The goal of this study was to establish a method for detecting biologically significant differences in protein expression of head and neck squamous cell carcinoma (HNSCC) obtained from the same samples utilized in gene expression analyses. Proteins from two head and neck tumor cell lines, SCC-25 and FaDu, were isolated from the denatured protein solution remaining from the TRIzol extraction procedure used for isolation of total RNA for microarray analysis. Peptides resulting from chemical and enzymatic digestion of the proteins were first separated by strong cation-exchange chromatography, followed by liquid chromatography-mass spectrometry (LC-MS) analysis on a QqTOF mass spectrometer. Stable isotope-labeled synthetic peptides were added to each ion-exchange fraction as internal standards, for reversed-phase HPLC retention time alignment. Protein extraction and digestion were repeated three times for each cell line and each extract was analyzed three times by LC-MS. To discriminate between technical *vs* biological variation, the ion-exchange fraction, retention time, normalized mass and signal intensity of these nine data sets were constructed into numerical arrays for statistical analysis. Of the ~50 000 signals, 90 peptide ions were found to discriminate the two cell lines with high stringency. Of those, six peptides were derived from vimentin and four peptides were derived from annexin II; both expressed more in SCC-25. Follow-up analysis of some of these signals by LC-MS/MS and RNA expression profiling revealed both concordance and discordance of RNA and protein expression. This study demonstrates that this procedure is highly reliable for identifying peptides that distinguish biological variability among samples, indicating that this method can be applied to study clinical samples, to identify potential prognostic biomarkers for HNSCC.
Laboratory Investigation (2007) **87,** 755–766; doi:10.1038/labinvest.3700598; published online 11 June 2007

KEYWORDS: biomarker; cancer; global proteomics; squamous cell carcinoma

Head and neck squamous cell carcinoma (HNSCC) is the fifth most common malignancy worldwide, representing a major international health problem.[1] The 5-year survival rate for HNSCC (ca.50%) has improved only marginally over the past decade; as a result, it is estimated that 45 641 cases and 11 210 deaths will occur in 2007 in the United States from HNSCC.[2] These tumors constitute an anatomically heterogeneous group of neoplasms arising from the oral cavity, oropharynx, hypopharynx, larynx and nasopharynx. The initial anatomic site of HNSCC is highly correlated with early metastasis[3–5] and treatment planning depends largely on anatomical staging of the disease at presentation. Conventional treatment with surgery and/or radiation therapy, and/or chemotherapy, is associated with significant morbidity, affecting speech, swallowing and overall quality of life. Despite these interventions, recurrence of the disease is observed in about 50% of patients, locally, regionally or at a distant site, with high rates of associated mortality.[6] Treatment failures for HNSCC can be attributed to multiple factors, but remain difficult to predict.

Although several biomarkers have been associated with HNSCC, none as yet have proven to be useful clinically; that is, no biomarker is used to guide treatment selection at initial diagnosis. For example, studies of individual biomarkers such as p53, EGFR, Bcl-2, MMPs, cyclins and molecular markers have demonstrated inconsistent, and at times contradictory, results.[7–9] Molecular characterization of this genetically complex disease has provided some insight into individual

[1]Department of Pathology, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA; [2]Department of Developmental and Molecular Biology, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA; [3]Department of Biochemistry, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA and [4]Department of Epidemiology and Population Health, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA

Correspondence: Dr MB Prystowsky, MD, PhD, Department of Pathology, Albert Einstein College of Medicine and Montefiore Medical Center, Belfer 713, 1300 Morris Park Avenue, Bronx, NY 10461, USA. E-mail: prystows@aecom.yu.edu

Received 2 April 2007; revised and accepted 11 May 2007

genetic abnormalities that contribute to tumor progression.[10] Tumors originating from different locations can exhibit varying behavior that is not predictable by histopathology of the primary tumor but is discernable by gene profiling.[11] Our ability to provide effective treatment will depend upon our ability to determine that surgical margins are free of tumor, and on our ability to predict tumor behavior, such as, metastatic potential, potential for local recurrence and response to therapy.[12] The goal of this study was to establish a method for detecting biologically significant differences in protein expression in head and neck tumors, using the same samples utilized in analyses of gene expression. This strategy would minimize sampling error and enhance integration of data sets for future analysis of protein, mRNA, DNA methylation and tissue microarray data.

Therefore, we developed a method to take advantage of the total protein mixture remaining after the extraction of total RNA from the TRIzol extraction procedure. The method employs offline strong cation exchange (SCX) and liquid chromatography-mass spectrometry (LC-MS), to find biologically distinct peptide ions, based on charge, hydrophobicity, mass and signal intensity. These peptide ions can be identified subsequently by targeted analysis, using the inclusion list for tandem mass spectrometry (MS/MS). In order to be able to achieve the goal of linking protein profiles to patient outcome or treatment responsiveness in an HNSCC cell carcinoma study, we tested the method for reproducibility. The data presented below demonstrate that the procedure can discriminate biological differences among complex protein mixtures derived from cultured cells, and should be applicable to analysis of other cancers and disease conditions.

MATERIALS AND METHODS

Cell Lines and Tissue

The cell lines obtained from the American Type Culture Collection (Manassas, VA, USA), representing two tumor sites, tongue (SCC-25, CRL-1628),[13] and pharynx (FaDu, HTB-43),[14] were grown under cell culture conditions as recommended by the vendor. Resection of an oral cavity HNSCC, collected at surgery by an IRB approved protocol to study HNSCC, provided both normal squamous oral mucosa and HNSCC carcinoma.

Gel Electrophoresis and Western Blot Analysis

TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) can be used to extract DNA, RNA and protein from tumor tissue. To determine the quality obtained following TRIzol extraction, we compared 2D gel patterns from normal and HNSCC samples extracted either with a mix of detergents and chaotropic agents[15] or with TRIzol (https://www.invitrogen.com/content/sfs/manuals/15596026.pdf). One-dimensional SDS-PAGE and two-dimensional gel electrophoresis (2D gels) were performed by following standard procedures used in other studies.[15] Protein concentration in cell lysates was determined using Pierce's BCA Protein Assay reagents (Rockford, IL, USA). Western blotting was carried out as described previously, using mouse monoclonal anti-human annexin II antibody (BD Biosciences, San Jose, CA, USA), mouse monoclonal anti-human vimentin antibody (BD Biosciences, San Jose, CA, USA) and goat anti-mouse Ig secondary antibody conjugated with horseradish peroxidase (Southern Biotech, Birmingham, AL, USA). To summarize: Western blots were blocked in 1 × TBS/T (20 mM Tris, pH 7.6, 137 mM NaCl, 0.1% Tween 20 containing 5% non-fat dry milk), then incubated for 1 h at room temperature in blocking solution containing mouse monoclonal anti-annexin II at a dilution of 1:5000. After washing in 1 × TBS/T, blots were incubated for 1 h at room temperature in blocking solution containing goat anti-mouse horseradish peroxidase-conjugated secondary antibody at 1:5000. Bands were visualized after further washes using Pierce's ECL Western Blotting Substrate (Rockford, IL, USA). Western blotting for vimentin followed the same protocol; mouse monoclonal anti-vimentin antibody (BD Biosciences, San Jose, CA, USA) was used at a dilution of 1:2500.

Preparation of Protein Digests

To optimize the protocol before applying to human tumor tissue, proteins from two HNSCC cell lines, SCC-25 and FaDu,[14] were extracted using the TRIzol protocol. The protein pellet derived from the TRIzol procedure was washed three times with 0.3 M guanidine HCl in 95% ethanol, each time for 20 min at room temperature, followed by a final wash with 90% isopropanol. This final wash removed any color remaining from the TRIzol that interfered with the LC-MS measurement, and removed any residual guanidine HCl. The pellet was then dissolved in 70% trifluoroacetic acid (TFA; Pierce, Rockford, IL, USA), Tris(2-carboxyethyl)-phosphine hydrochloride (TCEP; Pierce, Rockford, IL, USA) added to 50 μM to reduce disulfide bonds, and cleaved with 20 mM cyanogen bromide (CNBr; Sigma-Aldrich, St Louis, MO, USA) for 18 h at room temperature under nitrogen in the dark. The reaction mixture was evaporated under nitrogen, and the protein pellet resuspended in 100 mM ammonium bicarbonate, pH 8.5 (Sigma-Aldrich, St Louis, MO, USA), with sonication for 30 min at room temperature. To improve solubility, solid urea (Fluka, Sigma-Aldrich, St Louis, MO, USA) was added to make the final urea concentration 8 M. The mixture was first digested with 300 ng of endoproteinase Lys-C, 1.5 ng/μl (Roche, Indianapolis, IN, USA) at 37°C for 18 h. The solution was then diluted to reduce the urea concentration to 4 M with 100 mM ammonium bicarbonate, pH 8.5, calcium chloride (JT Baker, Phillipsburg, NJ, USA) was added to a concentration of 1 mM, followed by the addition of 15 μl of washed immobilized trypsin bead suspension (Pierce, Rockford, IL, USA), and incubated at 37°C for 18 h. After dilution of the sample to 2 M urea with 100 mM ammonium bicarbonate, pH 8.5, additional trypsin (15 μl) was added and the mixture was incubated for an additional 4–6 h at 37°C. Following

digestion, the immobilized trypsin was removed from the digest mixture by centrifugation (16 000 r.p.m., ~1 min) and the pH of the sample adjusted to 3 using formic acid (Acros Organics USA, Morris Plains, NJ, USA). The polypeptide content of the sample was measured using a Nanodrop spectrophotometer and a 100 μg aliquot was removed, and acetonitrile (Fisher, Fairlawn, NJ, USA) added to 25% v/v.

Offline SCX LC-MS Method Development

Since reproducibility in sample processing and analysis is key to the reliability of results, internal and external controls were used during method development. Pilot studies were carried out with serine acetyltransferase from *Haemophilus influenzae*. The initial ion-exchange separation was best performed offline, because of the large volumes. Thus, peptide mixtures obtained from the digestion were separated into 10 fractions using ion-exchange spin columns, as noted below. In addition, a mixture of four standard peptides (with different retention times and masses) was added to each salt cut fraction as an internal control and analyzed using LC-MS on a QqTOF mass spectrometer equipped with a microionspray source. The peptides served as markers for retention time alignment of the LC-MS spectra and used for correcting mass measurement error drift. The mass spectrometer was calibrated every day and instrument performance (retention time and sensitivity) was determined using our quality control 16-peptide mixture before sample analysis.[16] Peak picking, LC alignment and intensity normalization procedures were verified manually in comparison to the raw data and internal standards.

Approximately 100 μg of the sample was loaded on 50 mg of polysulfoethyl A resin (12 μm, 300 Å, PolyLC, Columbia, MD, USA) in a spin column and eluted with 200 μl of ammonium acetate pH 3, stepwise, at salt concentrations of 0, 10, 20, 30, 40, 60, 100, 200, 300 and 600 mM, in 0.1% formic acid, 25% acetonitrile. To reduce the acetonitrile concentration before reversed-phase HPLC separation, 40 μl of each fraction was diluted to 100 μl with 0.1% formic acid. A mixture of four synthetic peptides (200 fmol each, used as internal standards) with different masses and retention times was added to every sample before injection. Sequences and monoisotopic masses of the internal standard peptides are as follows: VFLQYLKN (1023.6 Da), VFL*QYL*KN (1025.6 Da), KYIPGTK (805.5 Da) and NFL*QYL*KD (1041.5 Da), where L* is the ^{15}N-labeled amino acid. These standard peptide peaks were used for retention time alignment of the HPLC chromatograms in data analysis and mass measurement error drift correction.

After the addition of internal standard peptides, samples were separated by reversed-phase chromatography, using an LC Packings Ultimate HPLC System (Dionex Corporation, Sunnyvale, CA, USA), and analyzed using a QStar Pulsar i mass spectrometer (AB/Sciex, Foster City, CA, USA). HPLC-grade acetonitrile (CH_3CN, Fisher, Fairlawn, NJ, USA), 18 Mohm Milli-Q water (Millipore, Bedford, MA, USA) and high-purity formic acid (Acros Organics USA, Morris Plains, NJ, USA) were used to prepare the mobile phases for LC-MS. A C18 Pepmap100 column (3 μm, 100 Å, 300 μm × 15 cm, Dionex, Sunnyvale, CA, USA) at a flow rate of 4 μl/min, with 2% CH_3CN/0.1% formic acid (solvent A) and 80% CH_3CN/0.1% formic acid (solvent B), was eluted with the following gradient: 0–30 min, 5% B (desalting); 30–80 min, 5–55% B; 80–90 min, 55% B; 90–100 min, 55–95% B; 100–110 min, 95% B; 110–125 min, 95–5% B; 125–135 min, 5% B. The QStar was equipped with a microionspray source using 20 μm I.D. fused silica capillary. TOF-MS was acquired in the mass range of 300 to 1800 *m/z*, with a scan time of 1 s. The QStar was calibrated daily using a CsI-peptide standard at 0.1 and 0.15 μM, respectively; consisting of cesium iodide (Sigma-Aldrich, St Louis, MO, USA) and the iPD1 peptide (amino-acid sequence—ALILTLVS; Bachem Bioscience Inc., King of Prussia, PA, USA) dissolved in 50% methanol (Fisher, Fairlawn, NJ, USA)/water containing 0.1% formic acid. A mixture of 16 peptides, each containing 2.5×10^{-13} M, was used as a quality control for monitoring the daily performance of the HPLC and mass spectrometer, before and after each sample set. Peptide ions identified by statistical analysis (see below) to have 90% biological variability were placed in an inclusion list for tandem mass spectrometry fragmentation. An LTQ (Thermo Scientific, Waltham, MA, USA) linear ion trap instrument was used to carry out tandem mass spectrometry measurements to identify a few peptides discriminated by the arrays. DTA files were created from the raw data, merged into one file and searched against the NCBInr database, using the Mascot search engine.

Data Analysis

The aim of our study was to determine whether the biological variability was much larger than the extraction and analytical variability, thus, providing confidence to proceed to the analysis of human tumor samples. In this experiment, three extractions were prepared from each cell line. As described earlier, the protein fraction from each of these extracts was fragmented into peptides and analyzed in triplicate by LC-MS. Use of internal and external protein and peptide standards provided relative quantitation. Each peptide in the final pattern derived from LC-MS had a characteristic mass, elution time from the reversed-phase column and elution characteristic (salt concentration) from a SCX spin column.

Data analysis was performed using an in-house software,[16,17] to build arrays with monoisotopic mass, salt cut, retention time and signal intensity. There are four steps in the process: (1) generation of peptide lists from individual LC-MS runs; (2) retention time alignment to correct any retention time shifts in the LC runs; (3) intensity normalization and (4) construction of arrays for each salt cut. For each salt cut analyzed by LC-MS, a peptide list with monoisotopic mass, retention time and intensity is first produced by performing denoising, deisotoping and charge state assignments. To remove chemical noise due to solvent peaks, all

singly charged ions below 600 Da are removed from the list. To be conservative, higher charge-state ions ($+4$ or $+5$) that do not have lower charge-state ions of the same mass are also removed. In the second step, the retention time drift is first corrected by using the retention time of the internal standard peptides, followed by a refinement procedure to maximize the number of matches, that is, peptides with similar masses and retention times. The refinement procedure corrects non-linear shift in the chromatograms by using robust spline fitting. Dividing the individual intensities with the average intensity of the LC-MS run normalized the peptide intensities. Peak picking, LC alignment and intensity normalization procedures were verified by manually checking the raw data, including, in several cases, the internal standard peptides. Arrays with normalized mass, retention time and intensities for each peptide in the 18 LC-MS runs (nine repeats for each cell line) were constructed. An array was constructed for each of the 10 salt fractions. Peptides having similar mass (± 0.1 Da) and retention time (± 40 s) were considered to be the same peptide, and clustered with average linkage clustering. Their normalized intensities form a row in the peptide array, while each column in the array represents an individual run.

In order to quantify the reliability of our approach, we employed a random effects ANOVA model. Based on this model, we computed the ratio of the variation due to differences between cell lines for a given peptide ion to the total variation. We refer to this variation as the percentage biologic variation.[18] A peptide ion with $\geq 90\%$ biologic variation has high reliability in differentiating the two cell lines. These data are presented in a second set of arrays for each salt fraction and combined into a single array. The statistical array was used to identify those peptide ions that should be studied further.

RESULTS

Our long-term goal is to identify clinically useful biomarkers using a systems approach that will predict tumor behavior or response to therapy in patients with HNSCC. The TRIzol extraction protocol permits the isolation of DNA, RNA and protein from the same sample, which facilitates the comparison of proteomics results with DNA and RNA analyses. Proteins from normal human squamous oral mucosa and HNSCC carcinoma were extracted using a standard protocol employing detergent and chaotropic agents, and the TRIzol protocol. Proteins derived from both extraction methods were separated by 2D gel electrophoresis and compared, as shown in Figure 1. The protein patterns for each extraction method are nearly the same for both normal mucosa (Figure 1a and c) and tumor (Figure 1b and d), indicating that TRIzol extraction can be used to generate a global proteomic profile.

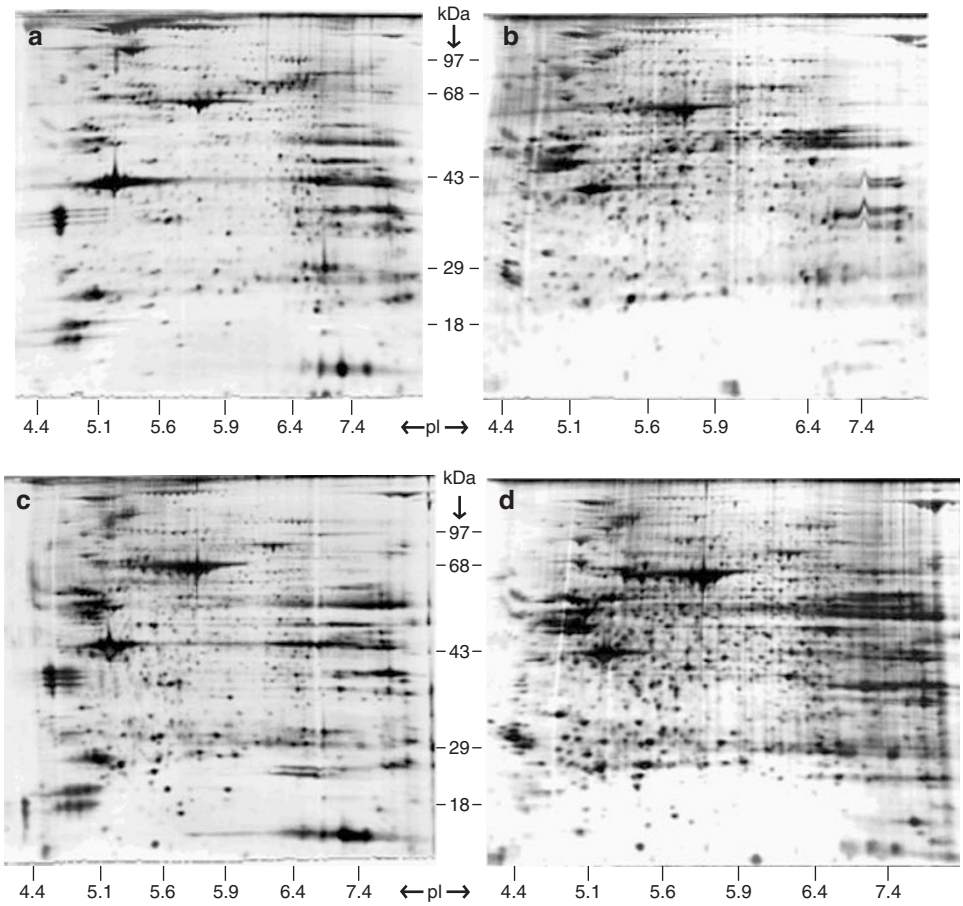

Figure 1 Two-dimensional gel electrophoresis, comparing proteins extracted by direct solubilization of normal human oral mucosa and HNSCC with proteins extracted during the TRIzol procedure. (**a**) Normal tissue, standard extraction (detergent/chaotropic salts); (**b**) HNSCC, standard extraction (detergent/chaotropic salts); (**c**) normal tissue, TRIzol extraction; (**d**) HNSCC, TRIzol extraction.

An overview of the analytical method and data analysis developed using TRIzol extracted proteins is given in the flow chart (Figure 2). Due to the insoluble nature of the proteins, samples were dissolved in aqueous TFA and cleaved with CNBr before digestion with multiple enzymes. Data analysis was performed using in-house, publicly available software, to build arrays of each salt fraction with retention time, normalized mass and signal intensity.[16,17]

Replicate Analysis of Peptide Mixtures

After optimizing the offline SCX and LC-MS procedure, two cell lines (SCC-25 and FaDu) derived from human HNSCC cell carcinoma were compared. To assess variation in sample processing and instrument variation, as compared to biologic variation, three replicates of peptide mixtures were prepared from each cell line and each replicate was analyzed by LC-MS in triplicate, yielding nine data sets for each cell line.

A range of ~4000 to 9000 peptide ions were identified in each LC-MS run as follows: 0 mM salt cut = 9574; 10 mM salt cut = 7021; 20 mM salt cut = 4026; 30 mM salt cut = 4639; 40 mM salt cut = 5337; 60 mM salt cut = 5247; 100 mM salt cut = 5264; 200 mM salt cut = 6685; 300 mM salt cut = 5593; 600 mM salt cut = lack of useful peaks; total = 53 386. The base peak chromatograms and mass spectra produced from analyses of two replicates of the 60 mM salt fraction of the SCC-25 cell line, carried out 2 days apart are, shown in Figure 3. The base peak chromatograms of the two runs shown in Figure 3a and c provide a graphic representation of reproducibility. The accumulated mass spectrum obtained between 65.9 and 66.3 min in Figure 3a is shown in Figure 3b. Nearly 20 peaks are observed in the mass spectrum. The same set of peaks was observed between 66.3 and 66.7 min of Figure 3c, in the replicate LC-MS run (Figure 3d). The spectra are reproducible, but with a 24 s shift between these two runs.

The shift in retention time is corrected in the alignment step of the data analysis program, using internal standard peptides. All the LC-MS runs were aligned using the SCC-25 first extraction data as reference. Dividing the peptide intensities with average intensity of the LC-MS run, normalized the intensities. Average intensities in the LC-MS runs were 80–100 counts. Retention times in different runs were first aligned by using internal marker peptides, then by fitting a smooth curve over a large number of potential matches, that is, peptide pairs from different runs. The alignment curve between two of the replicate runs is shown in the scatter plot in Figure 4. Black circles are matched peptide pairs with similar mass and slightly different retention times. The line is the fitted alignment curve, showing excellent run to run correlation ($r = 0.997$).

Determination of Biological Differences that Exceed Technical Differences

Comprehensive arrays corresponding to each salt cut were constructed including retention time, mass and signal in-

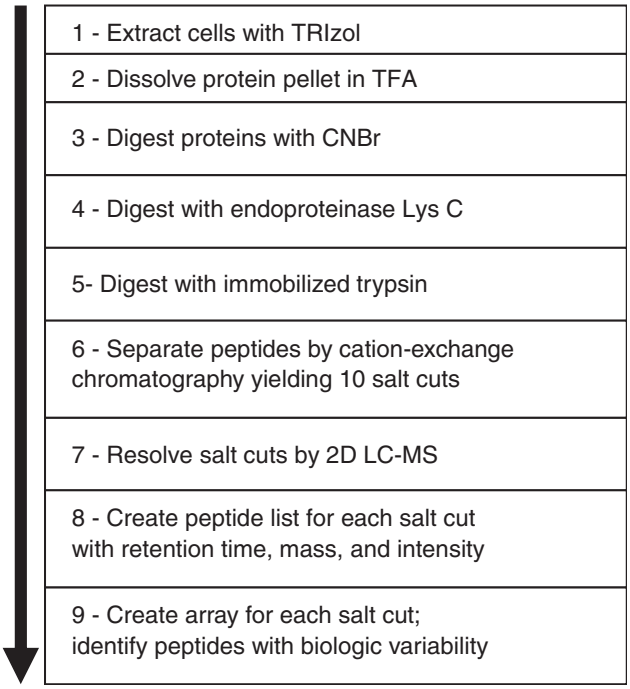

Figure 2 Schema showing overview of analytical method and data analysis. (Step 1) FaDu and SCC-25 cells in culture were lysed and their protein extracted using TRIzol, according to the manufacturer's protocol. (Step 2) The resulting protein pellet was dissolved in 70% TFA. (Step 3) Tris(2-carboxyethyl)-phosphine hydrochloride was added to a final concentration of 50 μM, CNBr to a final concentration of 20 mM, and samples were left overnight at room temperature in the dark. After ~16 h, CNBr and TFA were evaporated using a gentle stream of nitrogen; proteins were redissolved by adding 0.1 M ammonium bicarbonate, sonication for ~30 min and addition of urea to 8 M. pH was checked and adjusted to 8.5, if necessary, using ammonium hydroxide. (Step 4) Endoproteinase Lys-C was added to a concentration of 1.5 ng/μl and sample incubated with shaking, overnight at 37°C. (Step 5) Urea in sample was diluted to 4 M with 0.1 M ammonium bicarbonate, calcium chloride ($CaCl_2$) added to 1 mM and 15 μl washed, immobilized trypsin added. Sample was left overnight at 37°C, with shaking. After ~18 h, 0.1 M ammonium bicarbonate was added to further dilute the urea to 2 M, $CaCl_2$ added to 1 mM, an additional 15 μl of trypsin was added, and the sample incubated for a further 4–6 h. (Step 6) Approximately 100 μg peptide digest was loaded onto 50 mg of polysulfoethyl A resin (12 μm, 300 Å, PolyLC) in a spin column and eluted with 200 μl of ammonium acetate, pH 3, stepwise at salt concentrations of 0, 10, 20, 30, 40, 60, 100, 200, 300 and 600 mM, in 0.1% formic acid, 25% acetonitrile. An aliquot of each fraction was diluted, so that the acetonitrile concentration was 10%. A mixture of four synthetic peptides (200 fmol each) with different masses and retention times was added to every fraction before injection. (Step 7) Each fraction was separated by reversed-phase chromatography using an LC Packings HPLC and analyzed using a QStar Pulsar i mass spectrometer (AB/Sciex, Foster City, CA, USA). A C18 Pepmap100 column (3 μm, 100 Å, 300μm × 15 cm, Dionex, Sunnyvale, CA, USA) at a flow rate of 4 μl/min, with 2% acetonitrile/0.1% formic acid (solvent A) and 80% acetonitrile/0.1% formic acid (solvent B). After desalting for 30 min, the peptides were separated using a 5–55% B gradient over 50 min. (Steps 8–9) Arrays for each salt cut containing HPLC retention time, monoisotopic mass and signal intensity were built after noise removal, deisotoping, retention time alignment and intensity normalization of the raw data.

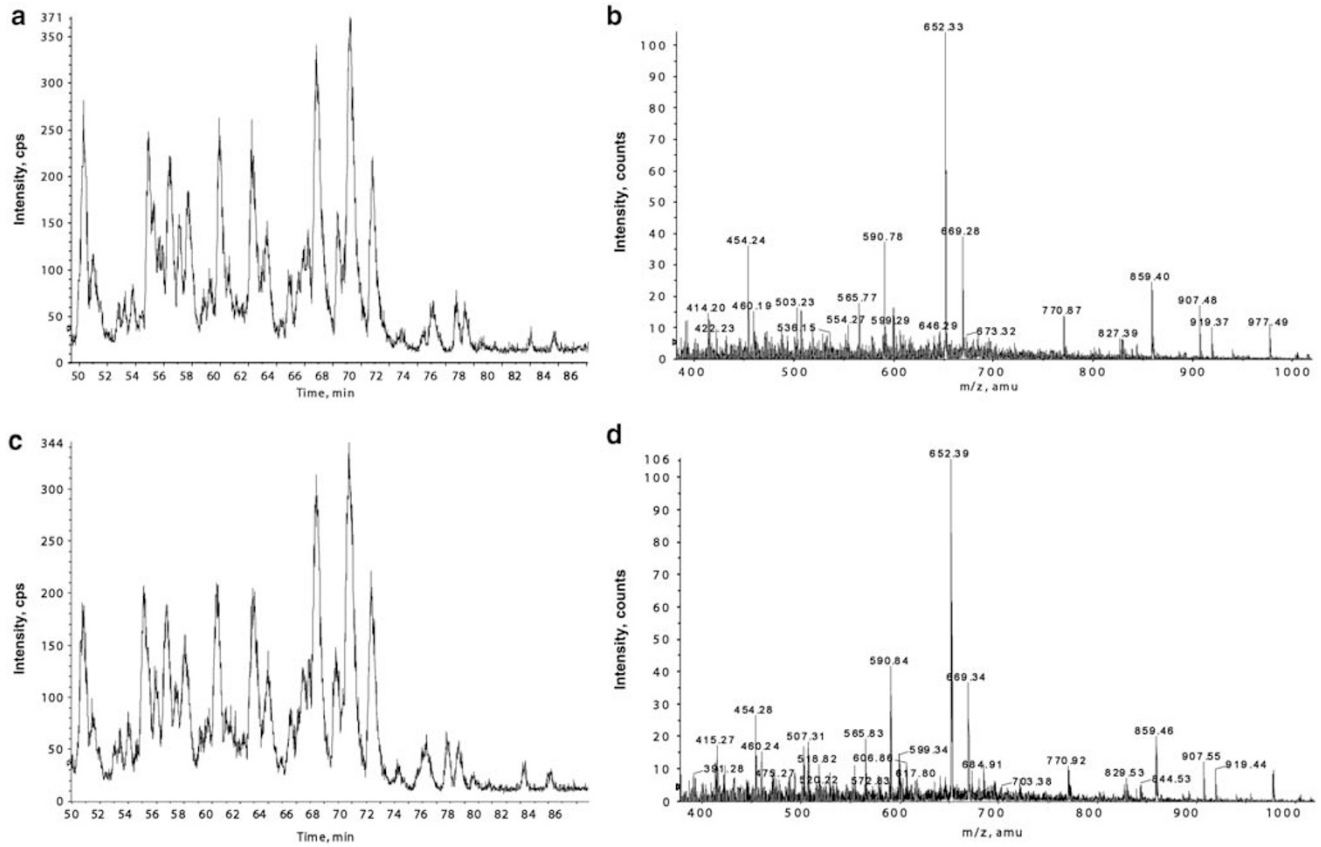

Figure 3 Reproducibility of LC-MS analysis. (**a**) Base peak chromatogram (50–80 min) of the 60 mM salt fraction from the second extract of the cell line SCC-25. (**b**) Mass spectrum between 65.9–66.3 min. Nearly 20 peaks are observed in the mass spectrum. (**c**) Base peak chromatogram of the same sample 2 days later. (**d**) Mass spectrum between 66.3–66.7 min. The same set of peaks is observed in the mass spectrum, showing that the LC-MS analysis is reproducible.

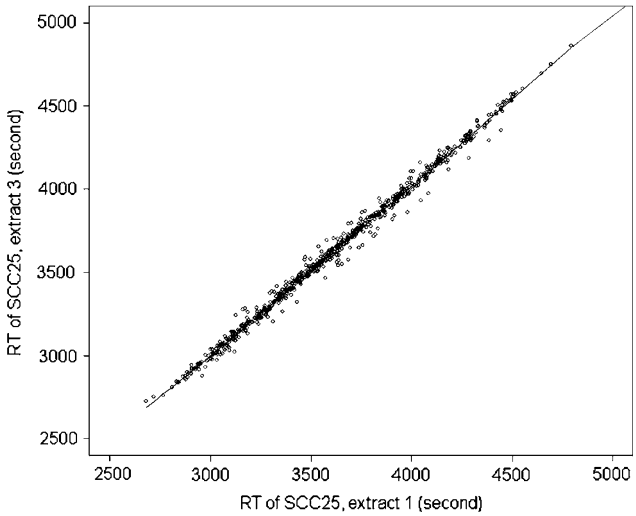

Figure 4 Retention time correlation across different LC-MS runs. The data from SCC-25, extract 1, replicate 1 is compared to the data from SCC-25, extract 3, replicate 2. The line is the fitted RT alignment curve with R statistical package ($r = 0.997$).

tensity of all peptide ions from 180 LC-MS runs. Each salt fraction was analyzed individually for the purpose of comparing the two cell lines. These arrays were then subjected to statistical analysis, generating a new set of arrays for each salt cut that included only those peptide signals with $\geq 90\%$ biological variation. The 300 and 600 mM fractions did not yield biologically significant peptide ions. From the remaining eight ion-exchange fractions, 90 peptide ions were identified that differentiated the two cell lines; an array for the peptide ions differentiating the two cell lines is given in Table 1. Figure 5 shows a mass spectrum of the signal for one of the differentiated ions, *m/z* 714.82, which is set in italic in Table 1 (peptide mass 1427.77 Da). This biologically significant peptide ion is present in SCC-25 cells (Figure 5a, arrow) and absent at the appropriate *m/z* in the FaDu cell line (Figure 5b).

Identification of Two Biologically Differentiated Peptide Ions

To confirm that differences revealed in the statistical arrays represent differential protein expression in the two cell lines, preliminary tandem mass spectrometry experiments were performed. The MS/MS spectrum of one differentially expressed peptide from the above array (Figure 5; Table 1) is represented graphically in Figure 6a. This peptide, SLYASSPGGVYATR from the protein vimentin, is expressed in the SCC-25 cell line but undetectable in the FaDu cell line

Table 1 Statistically generated array of normalized intensity values obtained from the salt fractions showing ≥90% biological differences between the cell lines

Mass	RT0	RT1	SCE11	SCE12	SCE13	SCE21	SCE22	SCE23	SCE31	SCE32	SCE33	FAE11	FAE12	FAE13	FAE21	FAE22	FAE23	FAE31	FAE32	FAE33
644.346	3720.58	3773.26	4.3	3.8	4.5	4.5	4.5	3.7	3.8	4.1	1.5	0	0	0	0	0	0	0	0.6	0
644.366	3115.33	3174.13	2.5	1.7	2.6	3.7	2.8	3	2.2	3.3	1.9	0	0	0	0	0	0	0	0	0
717.459	3150.13	3258.94	0.8	0	1.2	0	0	0	0	0	0	7.2	7.3	4.2	3.8	4.5	4	7.6	5	6.9
733.448	3104.68	3191.99	0	0	0	0	0	0	0	0	0	4.1	4.7	3	1.4	1.7	1.7	3	2.4	3.8
736.431	3226.14	3292.54	2.7	2.1	2.8	3.3	2.8	2.4	1.3	2.1	1.9	0.5	0	0	0.2	0.1	0.2	0.5	0.3	0.5
745.411	3725.32	3829.22	11.7	11.5	14.1	13	13	11.4	11.2	14.3	18	0	0	0	0	0	0.2	0.3	0	0
745.43	3036	3213.68	5.9	4.6	5.3	9.8	9.1	7.5	7.2	8.6	10.4	1	0.4	0	0	0	0.1	0	0	0
757.462	3293.93	3362.37	1	0	0.8	1	0.5	0.2	0.7	1	2.2	8.5	8.2	5.3	4.7	7.3	6.3	8.2	5.8	8.1
770.41	3752.47	3863.84	4.9	4.4	4.7	4.8	4.9	3.9	4.5	3.9	5.1	2.3	1	0.9	2.2	1.6	1.9	1.8	1.7	2
832.465	3223.62	3276.77	7.1	5.6	7.6	7.7	7.1	5.8	6.8	8.2	8.3	2.8	3.4	2.3	1.1	1.3	1.2	2.3	1.4	2.6
836.495	3356.02	3397.31	1.1	0.8	1.3	0.6	0.8	0.9	0.8	1.6	1.8	7.8	8.6	4.7	4.4	6.3	5.8	7.1	5.4	7.2
836.525	3258.67	3445.9	0.7	0.6	1	1.3	1.3	0.9	1.2	1.5	1.9	9	8.3	4	4.8	5.6	5	6.4	5.3	6
869.524	2989.54	3093.88	0.5	0	0.9	0.8	0.4	0	0.4	0.4	0	10.7	11.8	7.7	4.2	5.3	5	6	5.2	7.3
876.454	4535.75	4671.2	7.4	6.8	2.7	4.2	3.8	3.6	2.9	2.5	3.2	0.3	0.4	0.6	0	0.1	0	0.4	0.2	0
876.536	3991.51	4012.22	4.3	4.5	3.5	2.6	2.5	2.2	1.7	2.6	2.5	0	0	0	0	0	0	0	0	0.1
877.415	3763.71	3857.81	8.3	7.6	6.8	5.1	4.9	3.6	2.9	4	5.2	0	0	0	0.2	0.1	0.3	0	0	0
888.452	3185.12	3301.59	0	0	0	0.6	0	0	0	1	1.3	5.7	6.1	2.9	3.3	4.5	4.1	5.9	4.2	5.7
916.016	3914.96	4046.48	3.6	3.4	3.8	3.1	2.9	2.8	2.4	2.3	3.3	0	0	0.3	0.7	0	0.4	0.5	0.3	0.7
967.465	4289.8	4374.28	3.7	3.5	3.2	2.8	2.7	2.4	3.3	3.9	4.5	0	0.4	0	0	0	0	0	0.1	0
967.514	3670.77	3701.29	3	2.7	3	2.3	2.4	1.9	2.5	3.1	3.1	0	0	0	0	0	0	0	0	0.1
971.519	3772.59	3835.48	3	2.2	1.9	1.8	1.6	1.1	2.5	2.1	3.1	0	0	0	0.1	0.3	0.4	0	0	0.4
998.043	4772	4860.32	4.3	5.3	3.9	3	2.9	2.4	2.3	3	3.9	0	0	0	0	0	0.2	0	0	0.3
999.568	3577.61	3809.81	0.3	0.5	1.1	0.7	0.8	0.9	0.7	1.1	0.8	4.1	8.1	5.7	8.4	9	7.2	12.3	8.4	13.5
999.586	3552	3657.45	1.2	0.9	1.2	1.2	1.1	1	0.9	1.6	1.6	9.7	9.3	9	5	5.9	5.2	7.7	6.2	7.7
1011.529	3329.07	3481.75	3.3	2.7	4.1	3.4	3	2.6	2.7	3.6	3.8	1.4	1.4	0.8	0.6	0.9	0.3	0.9	0.6	1.5
1029.534	3846.05	3924.18	0.2	0	0.4	0.2	0.1	0.7	0	0.3	0	3	3.8	2.3	4.2	4.3	3.2	5.1	3.3	4.8
1029.556	3711	3884.96	0.7	0.5	0.5	0.3	0	0.1	0	0.3	0.4	5.6	4	4	3.1	3.7	3.1	4.3	3.7	4.2
1040.455	3617.16	3792.53	0	0	1.4	0.9	0.8	1.8	1.7	2.1	1.5	6.3	11.5	7.5	14.1	17.2	12.6	19.8	12.8	19.6
1040.617	3550.99	3659.87	2.8	2.9	2.3	2.2	2.1	1.7	2.7	2.4	2.8	11.6	11.1	10.9	8.5	9.9	8.2	13.2	10.2	12.1
1064.563	3506.2	3541.13	0.7	0.4	0.9	0.3	0.2	0	0.2	0.4	0.7	2.6	2.6	1.5	2.4	3.3	2.7	3.3	2.3	3.2
1072.411	3380.17	3498.27	0	0	0	0	0	0.2	0	0	0.3	1.8	2.7	1.7	4	4.5	3.6	5.3	4	5.2
1072.46	3338.11	3446.08	0	0	0	0.4	0.6	0.9	0	0.3	0.6	3.3	1.9	1.7	2.6	3	2.4	3.2	2.2	3.1
1072.451	3881	4085.83	0.2	0	0	0	1	0	0.5	1	0	2.8	2.6	3.3	3.2	3.2	2.8	3.2	1.8	1.7
1075.533	4724.06	4783.07	3.3	2.9	2.2	1.8	1.8	1.6	2.1	2.2	1.7	0	0	0	0	0	0	0	0	0
1078.605	3225.13	3271.81	0	0	0	0	0	0.2	0	0	0	2.3	4.4	2.4	3.9	3.9	3.6	4.8	3.7	5.5
1078.605	3130.68	3231.59	0.5	0.2	0.7	0.4	0	0.2	0	0	0	7.8	8.4	5.3	2.4	3.2	2.6	3.8	2.9	4.4
1086.57	***3699.6***	***3807.42***	***2.4***	***2.4***	***2.5***	***2.3***	***2.4***	***1.9***	***1.9***	***2.4***	***2.5***	***0***	***0***	***0***	***0.2***	***0.1***	***0.3***	***0.4***	***0.2***	***0.5***
1088.551	4438.53	4486.55	3	2.9	2.3	2.3	2	1.7	2.3	2	3.2	0	0	0.3	0	0.5	0.5	0.6	0.7	0.5
1089.529	3721.64	3920.5	6.8	6.8	8.4	6.5	6.2	5.5	4.5	6	6.9	0	0	0.3	0.1	0	0	0	0	0
1089.566	3226.48	3328.2	3.8	2.7	4.6	3.3	3.2	3.4	2.5	3.3	4.8	0	0	0	0.1	0	0	0	0	
1089.566	3184.75	3229.34	3.4	2.7	3.2	4.5	4.5	3.4	3.1	3.6	4.2	0.2	0	0	0.1	0	0	0	0	0
1092.522	4013.94	4114.68	5.3	5.5	6.1	4.3	4.3	3.9	4.2	6.1	6.2	0	0	0	0.1	0	0	0	0	0.1
1092.563	3436.14	3475.21	3.8	2.7	3.5	3	2.7	2.3	2.9	3.5	3.9	0	0	0	0	0	0	0	0	0
1128.603	3784.85	4066.3	1.1	1.1	1	0.6	0.6	0.7	1.7	2.1	1.7	6.5	9.9	6.7	8.4	9.8	6.6	11.4	7.8	12.4
1128.632	3739.57	3947.89	1.8	1.8	1.6	1.1	0.9	0.8	0.3	1.1	0.8	8.8	7.9	6.1	5.8	7.7	5.8	8	6.3	8
1128.642	3795.6	3970.92	1.3	1.2	0.5	0.4	0.7	0.5	0.7	0.6	0.7	5.9	4.7	3.1	5.8	5.6	4.9	4.5	4	5.2

Table 1 Continued

Mass	RT0	RT1	SCE11	SCE12	SCE13	SCE21	SCE22	SCE23	SCE31	SCE32	SCE33	FAE11	FAE12	FAE13	FAE21	FAE22	FAE23	FAE31	FAE32	FAE33
1162.658	4290.97	4438.48	4.8	4.6	3.3	2.6	2.3	2.4	3.5	4.8	2	0.2	0.6	0	0	0.6	0.2	0.8	0.4	0.9
1162.649	4276.33	4308.1	4	3	2.8	3.7	3.1	2.8	4.2	4.4	3.2	0	0	0	0.3	0.3	0	0	0	0.2
1164.561	4016.12	4079.95	4.5	4.6	4.6	3.7	3.6	3	3.7	4.9	5.1	0	0	0	0.2	0.1	0	0	0	0.1
1164.62	3419.78	3602.67	2.5	1.9	2.9	1.8	1.5	1.5	1.8	2.7	2.7	0.3	0.3	0.2	0	0	0	0	0	0.1
1164.616	3378.14	3461.21	1.9	1.4	1.8	2.5	2.3	1.7	2.2	2.7	2.8	0.1	0	0	0	0	0	0	0	0
1171.503	5035.6	5105.36	3.6	3.1	2.2	4	3.5	3	3.1	1.7	2.2	0	0	0	0.1	0	0	0	0	0
1171.573	4423.05	4482.99	3.3	2.6	2.3	2	1.8	1.4	2.1	2.7	1.7	0	0	0	0	0	0	0	0	0
1172.637	3399.91	3496.05	1	0.8	1.2	0.8	0.8	0.8	0.6	1.4	1.8	9	9.5	6.8	5.6	7.9	6.1	7.1	4.5	6.4
1209.649	4050.68	4181.21	3.8	4	3.3	2.4	2.2	1.8	2.7	3.8	4.7	0	0	0	0	0	0	0	0	0
1219.586	4059.45	4221.08	9.4	8.8	8.2	6.4	6.2	5.7	5.8	6.8	7.8	0	0	0	0.2	0	0	0	0	0
1219.612	3468.16	3655	4.3	3.8	3.2	3.3	3.5	2.7	3.1	3.1	3.4	0.3	0	0	0	0	0	0	0	0
1224.569	4471.95	4667.45	2.7	3	2.1	2	2	1.6	2.1	2.3	2.5	0	0	0	0	0.1	0.2	0.4	0.7	0.6
1253.571	4316.32	4427.73	4.5	4.3	3.6	3.1	2.9	2.5	3.3	3.8	4.5	0	0.5	0.3	0.2	0.2	0.2	0	0	0
1253.621	3718.78	3811.45	3.2	2.8	2.9	2.1	2.1	1.4	2.4	3.1	2.3	0.4	0.2	0.3	0	0	0	0.1	0	0
1291.666	3380.35	3500.82	0	0.4	0.5	0.3	0.1	0.8	0	0	0.6	3.4	4.3	3.4	3.8	4.1	3.6	7.9	5.7	8.9
1291.692	3412.9	3582.38	1.1	1	1.2	0.6	0.5	0.9	0.2	1.2	1.3	4.9	5.6	4.5	2.6	3.6	3	5.4	3.7	5.4
1300.703	3484.27	3540.02	12.5	9.2	9.6	9.5	8.6	9.1	9.3	13.2	13.9	3.3	3.1	2.2	2.5	3.4	2.7	2.7	2	2.7
1303.677	3453.48	3507.21	3.2	2.4	4	1.9	2	2.2	1.7	2.7	2.5	0	0	0	0	0	0	0	0	0
1303.703	3357.79	3465.22	4.1	3.1	3.7	2.5	2.3	1.9	2.5	3.1	3.3	0	0	0	0	0	0	0	0	0
1308.62	4461.81	4497.65	3.3	3	2.2	2.5	2.2	1.6	2.6	1.8	2.5	0	0	0	0	0	0	0	0	0
1308.665	3837.89	3880.2	3	2.7	2.2	2.1	1.7	1.5	2.4	2.3	1.7	0	0	0	0	0	0	0	0	0
1318.708	3598.36	3744.57	0.5	0	0	0.7	0	1	0	0.9	1	5.2	6.6	4.8	6.9	7.3	5.7	7.2	5.4	8.2
1318.729	3457.26	3669.05	0.8	0.7	1	0	0.5	0	0	0.5	0.3	7.8	8.2	7.8	3.9	4.4	3.8	4.5	3.9	5.2
1319.678	5082.47	5194.34	0.3	0	0.2	0	0	0.2	0.1	0	0	4.4	3.7	4.5	5.4	3.7	3.2	5.4	4.1	3.7
1319.713	4453.05	4565.78	0.6	0.2	0	0	0	0	0	0	0	1.7	1.8	1.1	3.1	3.1	2.4	2.3	2.1	2.6
1322.611	4230.7	4260.11	4.9	5.2	4.7	3.3	3.2	2.5	3.2	3.5	4.2	0	0	0	0	0	0	0	0	0
1322.67	3602.94	3647.53	3.6	2.8	3.5	2.4	2.2	1.7	2	2.7	2.4	0	0	0	0	0	0	0	0	0
1335.68	3905.07	4042.29	7.6	7.2	5.9	5	5.1	4.6	4.7	5.8	6.2	0.8	0.7	0.9	0.9	1.6	1.1	2.2	1.8	2.3
1336.665	3765.84	3813.95	6.3	5	8.7	4.9	3.8	6.3	5	6.4	4.2	13.1	16.2	13.8	13.7	16.2	12.2	20.1	13.7	19.8
1341.751	4065.08	4084	2.8	3.6	1.5	1.6	1.4	1.4	2.7	1.8	1.8	0	0	0	0	0	0	0	0	0
1343.727	4018.01	4063.44	1.7	1.6	0.6	0	0.5	0.4	0.6	0	0	4.5	4.6	3.1	5.8	6.2	5.3	5.2	4.2	5.6
1343.727	3856.76	3954.23	0.8	0.6	1.5	0.4	0.6	1.4	0	0.8	0	3.3	4.2	3.3	3.9	4.2	3.4	6.1	4	5.9
1351.862	4237.05	4292.68	0	0	0.4	0	0	0	0	0.4	0	1.9	3.1	2.6	2.7	3.2	2	4.3	3.2	4.7
1361.708	4299.08	4433.16	0	0	0	0	0	0	0	0	0	3.1	2.9	2.2	2.2	3.5	2.1	2	1.2	1.9
1414.659	3757.85	3853.9	3.7	3.4	3.5	3.2	3	2.7	2.5	3	3	0.3	0	0	0.4	0.4	0.4	0.4	0.3	0.7
1418.68	5091.95	5342.64	1.1	1.1	0.4	1.5	1.2	0.8	0.7	0.4	0.2	6.5	6.2	7.1	8.9	5.1	4.3	6.7	4.9	4.3
1424.699	3979.18	4004.94	14.3	14.5	12.9	9.9	9.6	7.6	5.4	7.3	8.4	0	0	0	0	0	0	0.1	0.1	0
1424.775	3358.74	3386.32	10.1	6.9	7.8	6.6	6.2	4.9	3	4.3	4	0	0	0.2	0	0	0	0	0	0
1427.712	4165.02	4184.98	7.1	6.8	6.6	4.9	4.8	4	3.5	5	5.9	0	0	0	0	0	0	0	0	0
1427.77	***3541.44***	***3572.86***	***4.1***	***3.6***	***3.6***	***3.1***	***2.7***	***2.3***	***2.3***	***3***	***2.6***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
1446.754	4137.89	4255.48	9.2	9.5	6.9	6.5	6.2	5.3	5	6.4	6.9	0	0	0	0	0.1	0	0	0	0
1446.803	3525.3	3566.95	4.4	3.4	3.6	2.3	2.1	1.8	1.9	2.4	2.5	0	0.1	0	0	0	0	0	0	0
1505.743	4892.04	5026.54	0.4	0.5	0.1	0.6	0.3	0.3	0.1	0.2	0	2.6	2.6	3.6	4.9	3.4	3.6	5.3	4.5	3.8
1693.913	3681.11	3767.48	0.6	0.7	0.4	0.2	0	0.8	0	0.4	0.8	1.8	2.5	2.5	2.7	3.6	2.4	4.6	2.7	4.9

RT, retention time; SCE1, first extract of SCC-25; SCE2, second extract of SCC-25; SCE3, third extract of SCC-25; FAE1, first extract of FaDu; FAE2, second extract of FaDu; FAE3, third extract of FaDu.

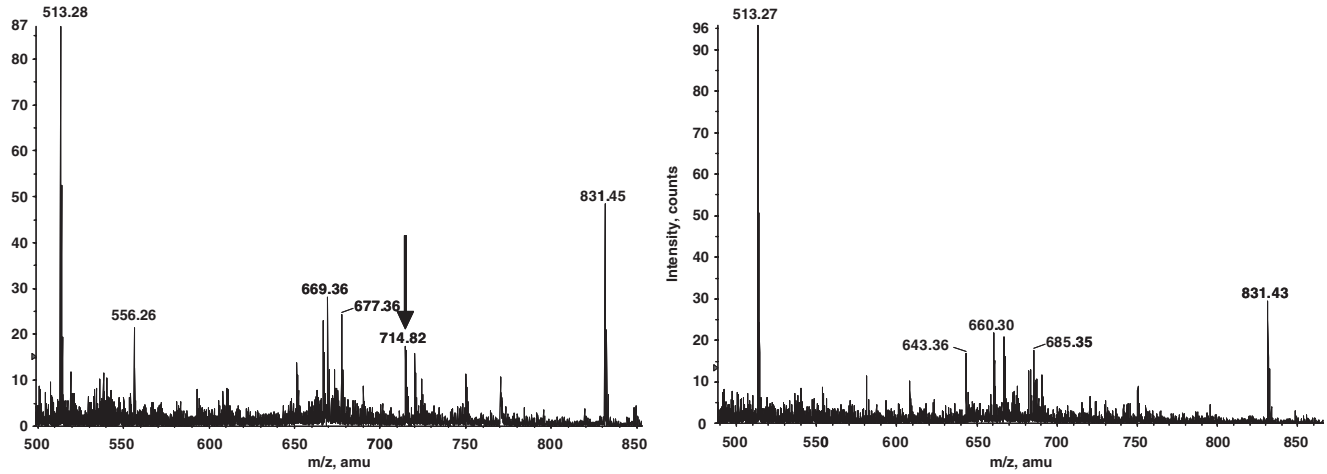

Figure 5 Biological differences between cell lines. Spectrum shown on the left is from one of the nine replicates of cell line SCC-25, and spectrum on the right is from one of the nine replicates of cell line FaDu. Both spectra are from LC-MS runs of the 60 mM salt fraction. The doubly charged ion at *m/z* 513.2 and the singly charged ion at *m/z* 831.4 arise from such proteins, which are present in both cell lines. The doubly charged ion at *m/z* 714.8 is present in all the nine replicates of SCC-25, while it is absent in all nine replicates of FaDu.

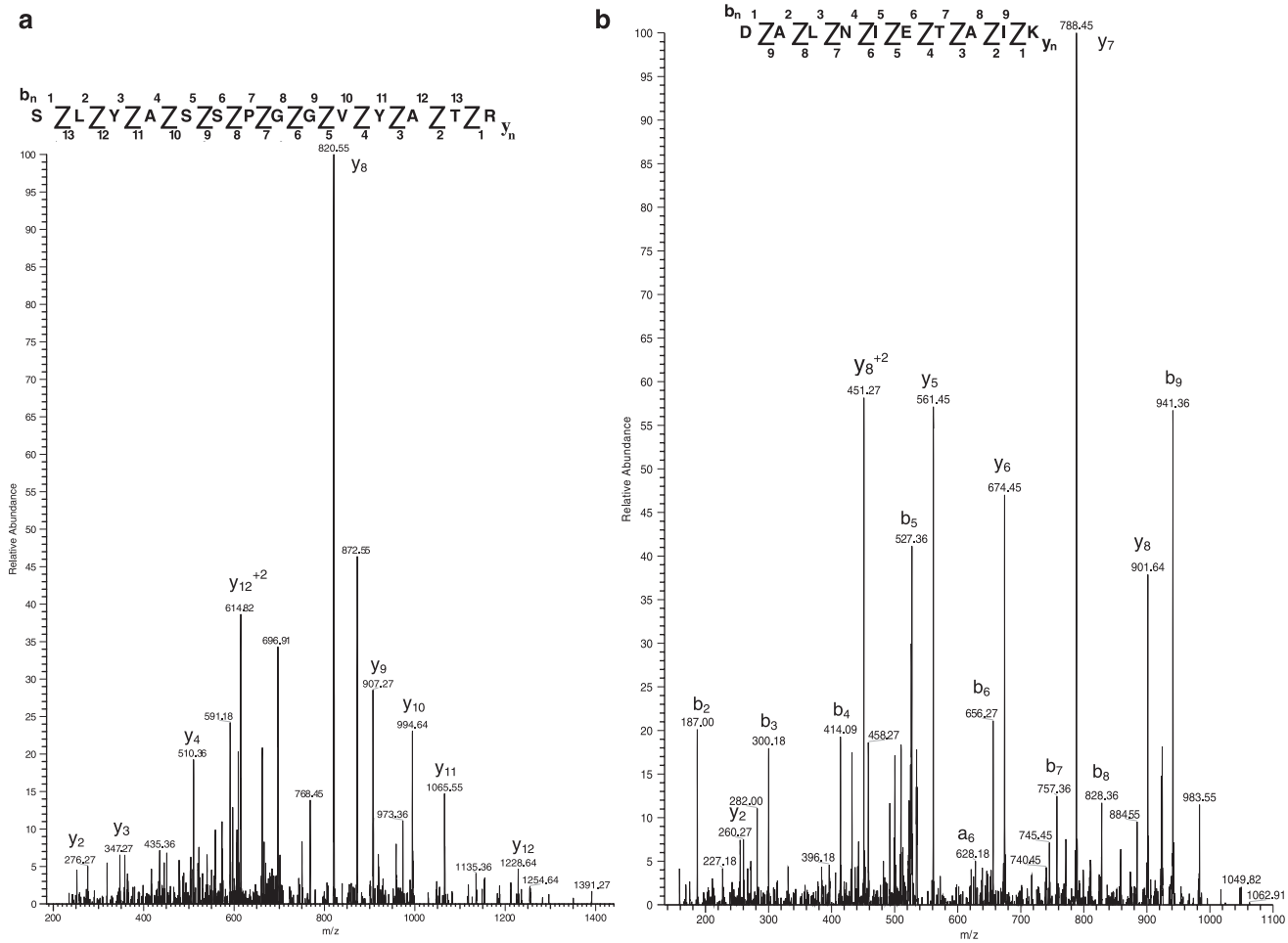

Figure 6 Product ion (MS/MS) spectra of the doubly charged precursor ions *m/z* 714.85 (**a**; shown in Figure 4 and Table 1) and *m/z* 544.30 (**b**) obtained from the 60 and 0 mM salt fractions of SCC-25, respectively. The MS/MS spectrum of *m/z* 714.85 (**a**) matched the peptide SLYASSPGGVYATR of human vimentin (residues 51–6; accession number CAA79613; peptide mass 1427.7 Da). The MS/MS spectrum of *m/z* 544.3 (**b**) matched the peptide DALNIETAIK of human annexin A2 (residues 38–47; accession number AAAH09561, one of many isoforms; peptide mass 1086.57 Da). The matched b and y ions are shown in both MS/MS spectra. MS/MS was obtained on a linear ion trap mass spectrometer (LTQ, Thermo Scientific).

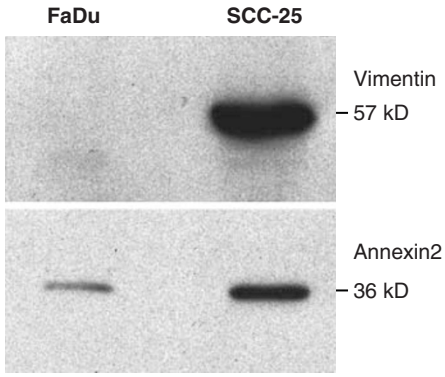

Figure 7 Western blots of annexin/vimentin. Equal amounts of FaDu and SCC-25 cell lysates were separated on 12% SDS polyacrylamide mini-gels and transferred to PVDF membranes. Western blots were blocked in 1 × TBS/T (20 mM Tris, pH 7.6, 137 mM NaCl, 0.1%Tween 20, containing 5% non-fat dry milk), then incubated for 1 h at room temperature in blocking solution containing mouse monoclonal anti-annexin II at a dilution of 1:5000. After washing in 1 × TBS/T, blots were incubated for 1 h at room temperature in blocking solution containing goat anti-mouse horseradish peroxidase-conjugated secondary antibody at 1:5000. Bands were visualized after further washes using Pierce's ECL Western blotting substrate. Protocol for vimentin followed the same outline; mouse monoclonal anti-vimentin antibody (BD Biosciences) was used at a dilution of 1:2500.

(Table 1). The monoisotopic mass of this peptide is 1427.77 Da, with an intensity of 3.0 ± 0.6 counts (normalized average intensity $\pm$ s.d., $n = 9$) and a retention time of 59.7 min. Five additional peptides were identified from this protein. Another protein that found to be more highly expressed in the SCC-25 is annexin II, whose MS/MS spectrum is shown in Figure 6b. The monoisotopic mass of this peptide, DALNIETAIK, is 1086.57 Da (Table 1), retention time 62.6 min, with an intensity of 2.3 ± 0.2 counts for SCC-25 and 0.2 ± 0.2 counts for FaDu (normalized average intensity $\pm$ s.d., $n = 9$). Three additional peptides were identified from this protein. Orthogonal validation of these expression differences was further provided by Western blot analysis with antibodies specific to these proteins (Figure 7). The relative difference in protein expression was found for both peptides, when the results from LC-MS and Western blot analysis were compared. Vimentin was expressed in SCC-25 but undetectable in FaDu, using either method, while annexin II showed a 12-fold difference by LC-MS and a six-fold difference by Western blot analysis.

Discordant RNA and Protein Expression

Our long-term goal is to integrate data from multiple data sets (genomic and proteomic), to identify the most reliable prognostic indicators for HNSCC. RNA expression profiling of SCC-25 and FaDu was performed using our custom cDNA arrays carrying ~28 000 cDNAs.[19,20] Using high stringency (five-fold differential gene expression) we found 290 SCC-25-specific RNAs and 113 FaDu-specific RNAs. LC-MS/MS analysis identified one of the SCC-25-specific peptides as keratin (KRT6E), which was also identified by RNA expression profiling (data not shown). In contrast, vimentin RNA expression was increased slightly in FaDu cells (1.3 to 1.6-fold), but vimentin protein was expressed only in SCC-25 cells, exemplifying discordant expression of RNA and protein. There are many possible mechanisms for the discordance and some may be responsible for pathology, indicating the advantage of measuring both RNA and protein expression in trying to predict tumor behavior. Thus, the ability to integrate genomic and proteomic data will strengthen our ability to dissect pathologic processes and to identify prognostic biomarkers.

DISCUSSION

Treatment selection for patients with squamous cell carcinoma of the head and neck is guided by clinical examination, imaging studies and histopathology.[21] There are no biomarkers in routine use that provide additional information on tumor behavior or predict response to therapy.[7,8] Our long-term goal is to identify biomarkers that can be used to guide initial treatment selection. Our approach is multifaceted, analyzing DNA methylation,[22] RNA expression[11] and *in situ* protein expression[23] in HNSCC, in relation to clinical outcome. We predict that integration of multiple data sets will be more successful than a single analytic approach. This study was undertaken to develop a method to analyze the proteome from the same tissue sample used for high-throughput DNA methylation and RNA expression, which would permit integration of data sets with clinical outcome.

Many methods are available and have been used to study the proteome of HNSCC.[24] Usage of precious human tissue for proteomics demands the usage of reliable methods. There has been significant commentary on biomarker development, emphasizing the need for validated, reproducible protocols.[25–27] Some methods provide a selective analysis of the proteome, while others provide a more global approach. While no single method can assess every protein in a tissue, each experimental approach must be tested for reproducibility, sensitivity, dynamic range of detection and quantitative reliability.

Proteomic Approaches to HNSCC Biomarker Discovery

Surface-enhanced laser desorption ionization-time of flight mass spectrometry (SELDI-TOF-MS) is a profiling method that has been extensively used in cancer biomarker studies.[28,29] It is becoming increasingly recognized that reproducibility and validation of these biomarkers should be addressed carefully.[30,31] In an effort to correlate RNA expression with protein expression, Bosch and co-workers analyzed HNSCC, as compared to normal esophageal mucosa, using cDNA microarrays, quantitative RT-PCR and a variety of MS methods, including, primarily, SELDI-TOF-MS.[32] Because of the multiple approaches used, they were able to identify calgranulins A and B and annexin I and II as proteins that were downregulated in HNSCC tumor tissue. It

is interesting that the calgranulins are expressed in stratum spinosum of normal epithelium; if HNSCC represents an undifferentiated cell, perhaps the calgranulins are not down-regulated but simply not expressed because the differentiation program has not been initiated. In a similar study comparing normal oral mucosa to HNSCC in five patients, Baker *et al*,[33] using laser-capture microdissection, followed by LC-MS/MS, identified several proteins that were differentially expressed but not correlated with clinical outcome. Roesch-Ely *et al*,[34] using SELDI-TOF-MS, showed that features of the proteomic profile of HNSCC can exist in histologically normal tumor-adjacent and tumor-distant mucosa. Expression of the HNSCC protein profile in tumor-distant mucosa correlated with local regional recurrence, suggesting a prognostic role for this profile in predicting recurrent disease.

LC-MS for Detecting Biologically Significant Peptide Ions

LC-MS is an alternative method used for quantitative profiling.[35–39] In quantitative profiling, signal intensity is compared across many LC-MS spectra. Reproducibility in sample processing and analysis is the key factor in determining the reliability of results in these studies. Few reports have tested the reproducibility in LC-MS-based analysis.[40,41] Sample preparation and separation, the resolution capability of the mass spectrometer and the data analysis highly influence the outcome of studies. Stewart and co-workers have systematically studied the technical variation in LC-MS-based analysis, using two breast cancer cell line samples. They could not identify proteins with significant differential expression in the two cell lines. In their study, the differences between the cell lines were less than the technical variation in the method.[41]

Gaspari *et al*[40] have studied the reproducibility using secreted proteins from culture media of human U937 macrophages, under three different culture conditions. Macrophages were exposed to lipopolysaccharide, in order to generate an inflammatory reaction; 2D LC-MS was then used to identify differential protein expression in the absence or presence of β2-adrenergic receptor inhibitors. They differentiated the following two sets of samples, namely the inflammation state and inhibition of inflammation, based on principle component analysis. Although they could differentiate the two states, they have reported that the poor resolution of mass spectrometer used in their study and the limitations in their data analysis have increased the technical variation. These studies show that there is still a need for validating label-free LC-MS-based methods and data analysis programs used for biomarker studies.

In the present study, quantitative profiling was used to study the differential expression of proteins in two human head and neck cancer cell line samples. The method employs offline ion-exchange chromatography and LC-MS to identify differential candidates, based on their normalized salt cut fraction, retention time, mass, and signal intensity. Differential candidates can be identified subsequently by targeted analysis, using MS/MS, thus eliminating the variation due to precursor ion selection for MS/MS. Since signal intensities are compared across mass spectra acquired on several days, quality of the data highly influences the outcome of the analysis. We enhanced reproducibility in two ways. A reference standard peptide mixture was analyzed by LC-MS between samples to ensure optimal instrument performance. Internal standards spiked into each ion-exchange fraction provided points of reference for retention time.

Because our purpose was to develop a rugged and reproducible method, identification of peptide ions with a high reliability coefficient for distinguishing biologically different cells or tissue is more important initially than identification of the peptide itself. Once biologically significant peptide ions have been identified, their identity can easily be obtained, as shown for vimentin and annexin II in Figure 6.

The goal of this study was to develop a reliable mass spectrometry method with high resolution and sensitivity, for cancer biomarker discovery in HNSCC. To be able to compare RNA expression, DNA methylation and protein expression in the same tumor sample, an experimental protocol was developed to isolate and analyze the protein fraction after extraction of RNA. Before embarking on analysis of patient tumor samples, testing and validation of the method were carried out with two HNSCC cell lines to provide a statistical basis for the reproducibility of the combined steps of protein preparation and analysis, and to demonstrate that biologic variation can be discriminated from technical variation. Our results show that the method presented is highly reproducible (ie, with a reliability coefficient $>90\%$), such that combined technical variation due to extraction, digestion, separation and LC-MS analysis is less than the biological variation between the two cell lines. Future clinical studies will integrate the findings from our proteomic analysis with genomic analyses from the same patient tissue, to develop diagnostic biomarkers that will guide treatment selection at initial diagnosis for patients with HNSCC.

ACKNOWLEDGEMENT

We thank Ms Annie D'Alauro for editorial assistance, and acknowledge NIH for financial support: CA101150 (RHA) and CA103547 (MBP).

DISCLOSURE/DUALITY OF INTEREST

The authors have no duality of interest to declare.

1. Ferlay J, Bray F, Pissani P, *et al.* GLOBOCAN 2000: Cancer Incidence, Mortality and Prevalence Worldwide, Version 1.0. IARC CancerBase No. 5. 2001. IARC Press: Lyon. Ref Type: Serial (Book, Monograph).
2. Jemal A, Siegel R, Ward E, *et al.* Cancer statistics. CA Cancer J Clin 2007;57:43–66.
3. Barnes L, Verbin R, Guggenheimer J. Cancer of the oral cavity and oropharynx. In: Barnes L (ed). Surgical Pathology of the Head and Neck, 2nd edn. Marcel Dekker: New York, 2001, pp 370–438.
4. Lindberg R. Distribution of cervical lymph node metastases from squamous cell carcinoma of the upper respiratory and digestive tracts. Cancer 1972;29:1446–1449.
5. Schwartz GJ, Mehta RH, Wenig BL, *et al.* Salvage treatment for recurrent squamous cell carcinoma of the oral cavity. Head Neck 2000;22:34–41.

6. Takes RP, Baatenburg de Jong RJ, Schuuring E, *et al.* Markers for assessment of nodal metastasis in laryngeal carcinoma. Arch Otolaryngol Head Neck Surg 1997;123:412–419.
7. Kyzas PA, Loizou KT, Ioannidis JP. Selective reporting biases in cancer prognostic factor studies. J Natl Cancer Inst 2005;97: 1043–1055.
8. Lothaire P, de AE, Dequanter D, *et al.* Molecular markers of head and neck squamous cell carcinoma: promising signs in need of prospective evaluation. Head Neck 2006;28:256–269.
9. McShane LM, Altman DG, Sauerbrei W. Identification of clinically useful cancer prognostic factors: what are we missing? J Natl Cancer Inst 2005;97:1023–1025.
10. Forastiere AA, Goepfert H, Maor M, *et al.* Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer. N Engl J Med 2003;349:2091–2098.
11. Belbin TJ, Singh B, Barber I, *et al.* Molecular classification of head and neck squamous cell carcinoma using cDNA microarrays. Cancer Res 2002;62:1184–1190.
12. Head and Neck Cancer Group. Head and neck cancer: reduce and integrate for optimal outcome. Cytogenet Genome Res 2007 (in press).
13. Rheinwald JG, Beckett MA. Tumorigenic keratinocyte lines requiring anchorage and fibroblast support cultures from human squamous cell carcinomas. Cancer Res 1981;41:1657–1663.
14. Rangan SR. A new human cell line (FaDu) from a hypopharyngeal carcinoma. Cancer 1972;29:117–121.
15. Luo Q, Siconolfi-Baez L, Annamaneni P, *et al.* Altered protein expression at early stage rat hepatic neoplasia. Am J Physiol Gastrointest Liver Physiol 2007;292:G1272–G1282.
16. Du P, Angeletti RH. Automatic deconvolution of isotope-resolved mass spectra using variable selection and quantized peptide mass distribution. Anal Chem 2006;78:3385–3392.
17. Du P, Sudha R, Prystowsky M, *et al.* Data reduction of isotope-resolved LC-MS spectra. Bioinformatics 2007 (in press).
18. Shrout P, Fleiss J. Intraclass correlation: uses in assessing rater reliability. Psychol Bull 1979;86:420–428.
19. Belbin TJ, Gaspar J, Haigentz M, *et al.* Indirect measurements of differential gene expression with cDNA microarrays. Biotechniques 2004;36:310–314.
20. Belbin TJ, Singh B, Smith RV, *et al.* Molecular profiling of tumor progression in head and neck cancer. Arch Otolaryngol Head Neck Surg 2005;131:10–18.
21. Patel SG, Shah JP. TNM staging of cancers of the head and neck: striving for uniformity among diversity. CA Cancer J Clin 2005;55: 242–258.
22. 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;114:16–23.
23. Madan R, Brandwein-Gensler M, Schlecht NF, *et al.* Differential tissue and subcellular expressionof ERM proteins in normal and malignant tissues: cytoplasmic ezrin expression has prognostic signficance for head and neck squamous cell carcinoma. Head Neck 2006;28: 1018–1027.
24. Yarbrough WG, Slebos RJ, Liebler D. Proteomics: clinical applications for head and neck squamous cell carcinoma. Head Neck 2006;28: 549–558.
25. Diamandis EP. Serum proteomic profiling by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry for cancer diagnosis: next steps. Cancer Res 2006;66:5540–5541.
26. McShane LM, Altman DG, Sauerbrei W, *et al.* Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst 2005;97:1180–1184.
27. Zolg W. The proteomic search for diagnostic biomarkers: lost in translation? Mol Cell Proteomics 2006;5:1720–1726.
28. Seibert V, Ebert MP, Buschmann T. Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. Brief Funct Genomic Proteomic 2005;4:16–26.
29. Solassol J, Marin P, Maudelonde T, *et al.* Proteomic profiling: the potential of Seldi-Tof for the identification of new cancer biomarkers. Bull Cancer 2005;92:763–768.
30. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. Mol Cell Proteomics 2004;3:367–378.
31. Henderson NA, Steele RJ. SELDI-TOF proteomic analysis and cancer detection. Surgeon 2005;3:383–390, 422.
32. Roesch EM, Nees M, Karsai S, *et al.* Transcript and proteome analysis reveals reduced expression of calgranulins in head and neck squamous cell carcinoma. Eur J Cell Biol 2005;84:431–444.
33. 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;41:183–199.
34. Roesch-Ely M, Nees M, Karsai S, *et al.* Proteomic analysis reveals successive aberrations in protein expression from healthy mucosa to invasive head and neck cancer. Oncogene 2007;26:54–64.
35. Fang R, Elias DA, Monroe ME, *et al.* Differential label-free quantitative proteomic analysis of *Shewanella oneidensis* cultured under aerobic and suboxic conditions by accurate mass and time tag approach. Mol Cell Proteomics 2006;5:714–725.
36. Nakamura T, Dohmae N, Takio K. Characterization of a digested protein complex with quantitative aspects: an approach based on accurate mass chromatographic analysis with Fourier transform-ion cyclotron resonance mass spectrometry. Proteomics 2004;4:2558–2566.
37. Qian WJ, Jacobs JM, Camp DG, *et al.* Comparative proteome analyses of human plasma following *in vivo* lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry. Proteomics 2005;5:572–584.
38. Ru QC, Zhu LA, Silberman J, *et al.* Label-free semiquantitative peptide feature profiling of human breast cancer and breast disease sera via two-dimensional liquid chromatography-mass spectrometry. Mol Cell Proteomics 2006;5:1095–1104.
39. Silva JC, Denny R, Dorschel C, *et al.* Simultaneous qualitative and quantitative analysis of the *Escherichia coli* proteome: a sweet tale. Mol Cell Proteomics 2006;5:589–607.
40. Gaspari M, Verhoeckx KC, Verheij ER, *et al.* Integration of two-dimensional LC-MS with multivariate statistics for comparative analysis of proteomic samples. Anal Chem 2006;78:2286–2296.
41. Stewart II, Zhao L, Le BT, *et al.* The reproducible acquisition of comparative liquid chromatography/tandem mass spectrometry data from complex biological samples. Rapid Commun Mass Spectrom 2004;18:1697–1710.