Proteomics and the Identification of Serum Biomarkers in a Mouse Model of Oral Squamous Cell Carcinoma

by

Marcin Aleksander Mlynarek. MDCM'02

Department of Otolaryngology – Head and Neck Surgery
McGill University, Montreal, Canada

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science: Thesis in Otolaryngology.
Thesis submitted in February 2007.

Your file Votre référence
ISBN: 978-0-494-32847-7
Our file Notre référence
ISBN: 978-0-494-32847-7

Table of Contents

Abstract (English)

Objective: To establish a clinically relevant model for the identification of protein serum biomarkers for oral squamous cell carcinoma, and to identify specific candidate proteins.

Methods: Samples of oral cancer and adjacent normal tissue were obtained and were transplanted orthotopically into tongues of immunocompromized mice. When the mice lost 20% of their weight, they were sacrificed by exsanguinations. The serum was analyzed by two separate protocols: DIGE/MALDI and MudPIT/LC/ESI. Preliminary validation was conducted on an established cancer marker.

Results: We identified over one hundred proteins as being differentially expressed between control and cancer-bearing mice ($p<0.05$); including EGFR, cytokeratin 10, gelsolin, titin, vitronectin, retinoblastoma protein family, bullous pemphigoid antigen, and clusterin.

Conclusion: We report a proteomic approach for the identification of serum biomarkers of oral cancer using an orthotopic mouse model. We identified several proteins that can be exploited as potential markers for diagnosis of oral squamous cell carcinoma.

Résumé (français)

Objectifs: Établir un modèle cliniquement significatif pour l'identification dans le sérum de biomarqueurs protéiques liés au carcinome à cellules squameuses de la cavité buccale. Identifier des protéines candidates spécifiques pour ce biomarquage.

Méthodes: Des échantillons de lésions cancéreuses de la cavité buccale et de tissus adjacents normaux ont été obtenus et transplantés sur des langues de souris immunosupprimées. Lorsque les souris perdaient 20% de leur masse corporelle, elles étaient sacrifiées par exsanguination. Le sérum était alors analysé par deux protocoles distincts: le DIGE/MALDI et le MudPIT/LC/ESI. La validation préliminaire était effectuée en utilisant un marqueur de cancer reconnu.

Résultats: Nous avons recensé plus de cent protéines s'exprimant de façon différentielle entre les souris porteuses de tumeurs cancéreuses et les souris contrôles ($p<0.05$), incluant le EGFR (récepteur du facteur de croissance épidermique), la cytokératine 10, la gelsoline, la titine, la vitronectine, les protéines de la famille du rétinoblastome, l'antigène de la pemphigoïde bulleuse et l'apolipoprotéine J/clusterin.

Conclusion: Nous présentons une approche protéomique pour l'identification de biomarqueurs sériques liés au cancer de la cavité buccale en utilisant un modèle de souris. Nous avons identifié plusieurs protéines qui peuvent être exploitées comme biomarqueurs potentiels pour le diagnostique du carcinome à cellules squameuses.

Acknowledgements

I would like to thank Dr. Moulay A. Alaoui-Jamali, for his basic science supervision and support in all aspects of this project. I would also like to thank Dr. Michael P. Hier and Dr. Martin J. Black for their clinical supervision and their continued mentorship. I would like to acknowledge fellow resident Dr. Richard Balys for his help in the collection of the materials as well as the data. I would also like to acknowledge Su Jie and Yingjie Xu, for their help in performing some of the experiments at the Lady Davis Institute for Medical Research. I want to thank my wife Lisa Mlynarek, and my sister Julia Mlynarek for spending the time to proof read this thesis. Finally, I would like to thank Mrs. Susan James, Mr. Marcos DiFalco, and Mr. Leonid Kriazhev, from the Genome Quebec Innovation Centre: Proteomics Service Platform, for technical support and valuable advice on the proteomic analysis. This work is Supported by McGill University Head and Neck Cancer Fund, and in part by the Canadian Institutes for Health Research, the Canadian Breast Cancer Research Alliance, and the Cancer Research Society.

List of Abbreviations

DIGE: Differential Gel Electrophoresis

EBV: Epstein Barr Virus

EGFR: Epidermal Growth Factor Receptor

ELISA: Enzyme Linked Immunosorbent Assay

ESI: Electrospray Ionization

FT: Fournier Transform

HNSCC: Head and Neck Squamous Cell Carcinoma

HPLC: High Performance Liquid Chromatography

HPV: Human Papilloma Virus

ICAT: Isotope Coded Affinity Tag

iTRAQ: Proteomic labeling technique for quantitative analysis

IHC: Immunohistochemistry

LC: Liquid Chromatography

LMW: Low Molecular Weight

MALDI: Matrix Assisted Laser Desorption Ionization

MCAT: Mass Coded Affinity Tag

MudPIT: Multidimensional Protein Identification Technology

MS: Mass Spectrometry

MS/MS: Tandem Mass Spectrometry

m/z: mass to charge ratio of a particle

OSCC: Oral Squamous Cell Carcinoma

RAG2/γ(c): Immunocompromised mouse lacking Recombinase Activating 2 Gene

SELDI: Surface Enhanced Laser Desorption Ionization

TOF: Time Of Flight

Q or q: Quadrupole

QTRAP: Hybrid of a linear ion trap and a triple quadrupole mass spectrometer

2-DE: Two Dimensional Electrophoresis

1. Introduction

Head and neck squamous cell carcinoma (HNSCC) is diagnosed in an estimated 500,000 people each year worldwide.[1] Despite improvements in diagnostic technology and novel therapy regimens, HNSCC remains the sixth most common cause of cancer deaths in the world, and the 5-year survival rates have not changed significantly, remaining at approximately 50%.[2,3] Presence of cervical lymph node metastasis is presently the most significant prognostic factor of patient survival. [4] No serum biomarker for HNSCC is currently available.

Oral squamous cell carcinoma, a subgroup of HNSCC, is a relatively common and devastating disease, mainly among smokers. It begins as a focal overgrowth of epithelial stem cells near the basement membrane. The factors that provoke this change include tobacco, alcohol, dietary factors, oral hygiene and viruses, such as HPV and EBV.[5] The neoplastic process then goes through progressive genetic and pathological phases: normal epithelium, to hyperplasia, to dysplasia, to carcinoma in situ, and finally to invasive carcinoma. HNSCC is unique in that it remains locoregional for a long time and visceral metastasis develop only in the later stages of the disease.[6] Even though attempts have been made to create a molecular progression model,[7] the precise nature of the genetic and protein alterations occurring at each step remains unknown.

Presently, diagnosis and follow-up of patients with oral cancer is based on physical exam and different imaging modalities. These approaches are neither specific nor sensitive, and many patients are diagnosed only late in progression of the disease. Often, patients require morbid treatments, such as radical surgery [Figure 1], brachytherapy [Figure 2], or high dose radiotherapy, and relapses are frequent.

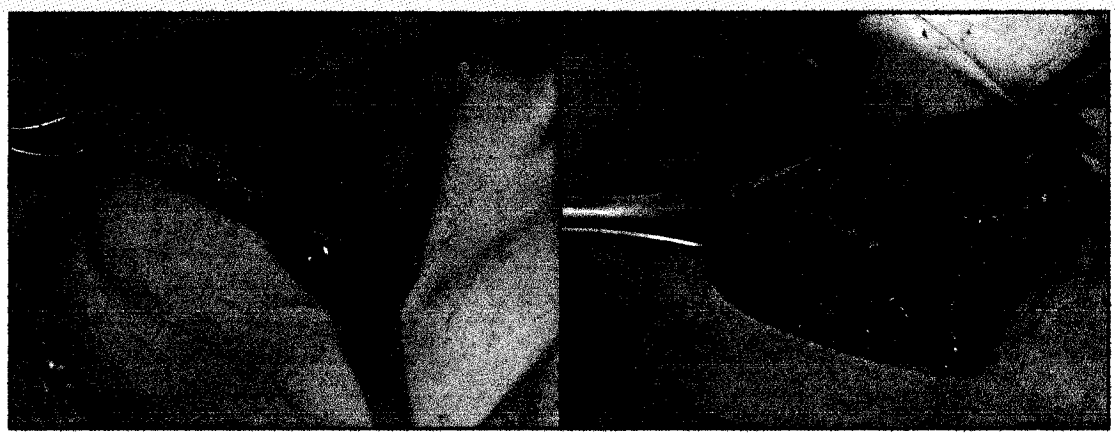

Figure 1. Intraopertive images of a (A) partial glossectomy and (B) neck dissection.

Figure 2. Intra-operative insertion of brachytherapy catheters inserted though the oral cavity (B) and exiting in the neck (A)

Like most solid tumors, HNSCC is dependent on a myriad of intracellular and intercellular transduction signals occurring in the tumor microenvironment between the heterogeneous tumor mass and the surrounding stromal and inflammatory cells.

Presently, clinically useful biomarkers are lacking for this complex disease. Identifying key biomarkers involved in disease progression will lead us to a greater

understanding of the biology of oral cancer, and will have enormous implications for the prognosis, diagnosis and treatment of this disease.

The objectives of this study were to first establish a reproducible proteomic platform for identifying human serum proteins in our mouse model of oral squamous cell carcinoma, and second to identify specific candidate proteins that can serve as potential biomarkers for the human disease.

2. Literature summary

a. Genomics studies on HNSCC

Numerous genetic analyses of primary tumors,[8,9] surgical margins,[10] and body fluids such as saliva [9,11,12] have been reported. A previous study, performed by our team of researchers, has found over 213 gene modifications in head and neck squamous cell carcinoma.[8] Our lab has also studied the genetic transformations in the relationship between HPV type 16 and HNSCC.[11] Despite this active research, no reliable genetic or protein marker has been proven to be useful for early diagnosis of head and neck squamous cell carcinoma. The known serologic tumor markers (such as CEA, CA 19-9, SCC, TK, and dTTPase) are not significant in HNSCC patients.[13] Tumorgenesis of head and neck cancer, like most other cancers, is not only dependant on the diseased tissue, but on a complex interaction between the tumor and its surrounding microenvironment.[14] Therefore, a combination of genetic profiling techniques and serum proteomics is needed to identify clinical biomarkers for HNSCC.[6,15]

b. Proteomics

Even though the proteome is a complete description of all the proteins encoded by the genome,[16] the information contained in that genome is not necessarily the same as the information in the corresponding proteome.[17] As the DNA is transcribed into RNA and then translated into protein, a number of transcriptional, translational and post-translational modifications can occur [Figure 3].[18] Therefore genetic profiling can not

predict accurately the function of marker proteins, thus proteomics holds great promise to fill this gap.

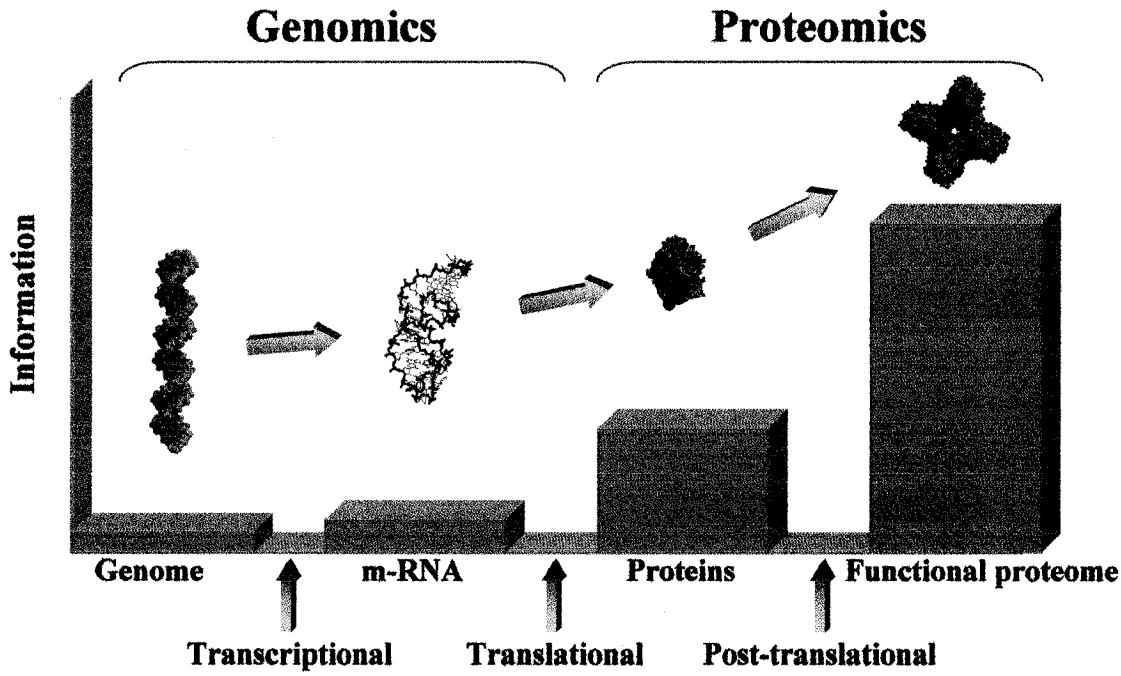

Figure 3. Information contained in the genome is not necessarily the same as the information in the corresponding proteome.

i. Basic proteomic techniques

The identification of cancer biomarkers using proteomic technology involves different steps. First, the large and abundant proteins are removed in order to permit analysis of the low molecular weight proteins. Second, the proteins are digested in a predictable manner and separated. Third, the resulting peptides are introduced into a mass

spectrometer, which separates them by their mass-to-charge ratios and produces a mass spectrum of a sample, unique to its composition. Typically a mass spectrometer is comprised of three parts: an ion source, a mass analyzer, and a detector. Finally, the individual spectra are compared to known peptide spectra from various databases. The following is a more detailed explanation of these various steps, including the methods for 1) large protein removal, 2) protein digestion, 3) protein separation, and 4) protein ionization, as well as a description of 5) mass analyzers, 6) quantitation, and 7) proteomic databases.

1. Large protein removal

Most biomarkers are small proteins that are relatively sparse in serum.[19] Therefore, when searching for new biomarkers, it is important to focus on this portion of the serum proteome.[16,20]

Albumin, IgG and transferrin comprise 80% of the total protein mass. Thus, the difficulty in identifying expected low molecular weight markers is that these large and abundant proteins overshadow the small ones. Currently, the first step in the search for new biomarkers is the removal of the abundant high molecular mass proteins (>30 kDa) such as albumin, transferrin, thyroglubulin and the immunoglobulins.[21,22]

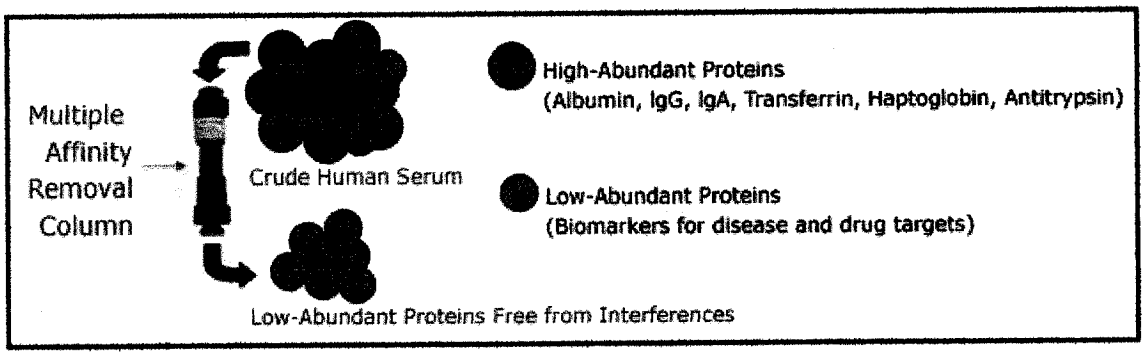

Figure 4. Removal of high abundant proteins using the Agilent multiple affinity removal system. Adapted from the Agilent website.

There are various systems designed for the separation of these high molecular mass proteins. The Agilent multiple affinity removal system removes the high-abundant proteins (from 3 to 6 proteins) from mouse serum simultaneously and reproducibly [Figure 4]. It depletes >98-99% of the three targeted proteins.[23] Our experiments were performed using the Agilent system which removed the three most abundant proteins (albumin, IgG and Transferrin). Denaturing using acetonitrile containing 0.1% trifluoroacetic acid is another way to precipitate the abundant proteins.[24]

Unfortunately, some marker proteins exist predominately in a bound phase, and the circulating large molecular weight carrier proteins act as a reservoir. Although affinities between the carrier proteins and the small markers are weak, removing and discarding these "contaminating" large proteins will lead to a loss of some possible markers.[25] Efficient capture of the carrier proteins and specific elution of the low abundant biomarkers will yield the greatest amount of diagnostic information.[26] Liotta et al propose agents that can be placed into the circulation to act as 'molecular mops' that soak up and amplify the biomarkers that exist [Figure 5].[27]

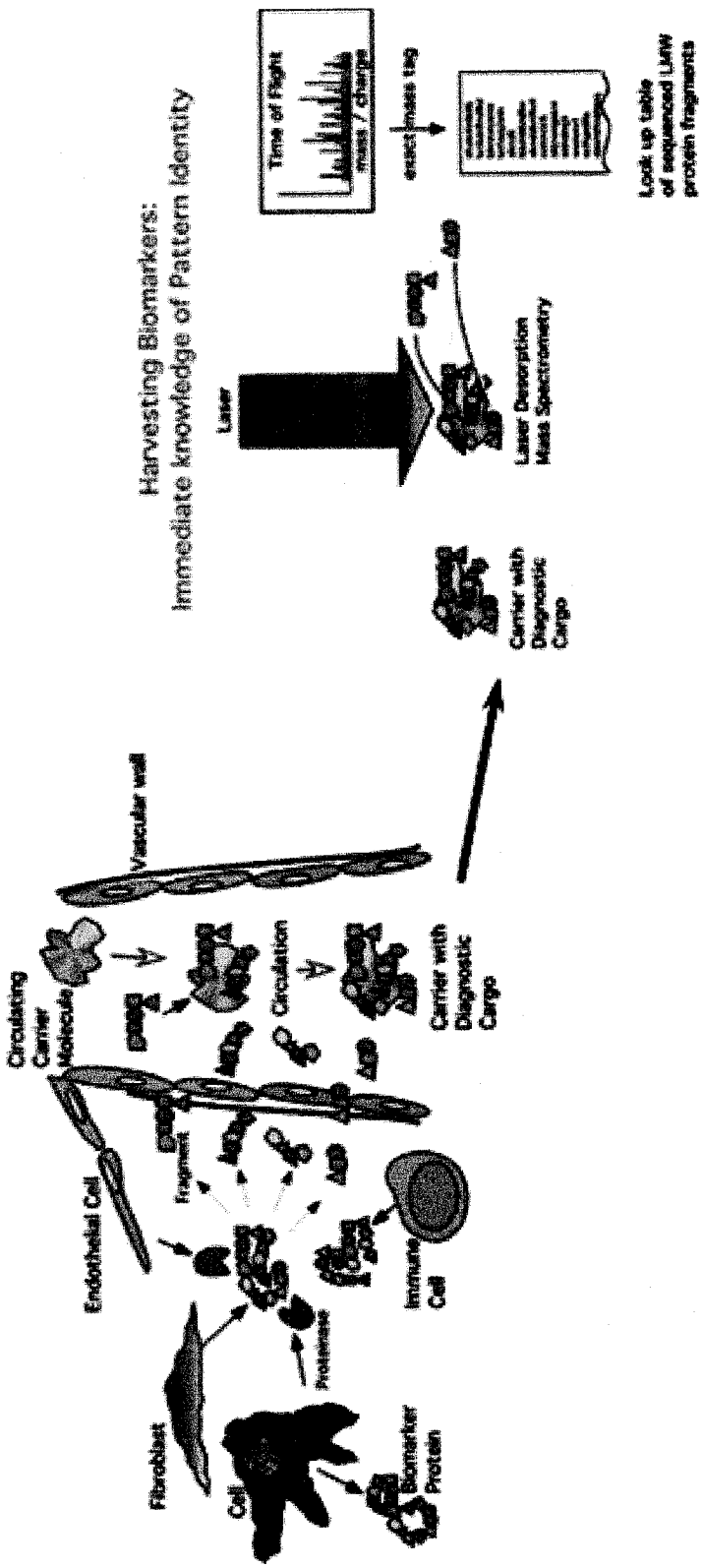

Figure 5. Biomarker amplification and harvesting by carrier molecules.[27]

2. Protein digestion

As the mass of a protein increases, the mass spectrometry measurement error increases, the efficiency of ionization decreases, and the sensitivity of small protein detection diminishes.[16] Therefore, enzymatic fragmentation of the separated proteins usually follows the first large protein removal step in biomarker search.[21] This involves digestion of the protein with a sequence-specific endoprotease, such as trypsin. The fragmentation occurs in a predictable manner along the peptide backbone.[28] Because these enzymes cleave proteins at well-defined positions (i.e. they are sequence specific), a peptide-mass fingerprint pattern will be generated that can be predicted from protein-sequence information provided by the Human Genome Project.[16]

Analysis of peptide fragments of proteins permits sensitive detection and accurate mass measurement. Smaller peptides are much easier to elute from gels as compared to full proteins.[29,30]

3. Protein separation

The third step in biomarker identification is separation of proteins from the complicated mixtures. Several techniques are available for this step: one and two dimensional gel electrophoresis, as well as liquid chromatography. One and two-dimensional gel electrophoresis (1-DE & 2-DE) have been the standard discovery tool in proteomics.[31] Proteins are first separated using a standard gel. Using a 1D gel, the separation occurs by molecular weight (MW) only; using a 2D gel, the separation occurs by both molecular weight and pH. The "pH point" is the point at which the overall charge of a particular protein becomes zero, so that the protein stops moving alongside the

electric field. The selected spots are then excised from the gel, digested and analyzed by mass spectroscopy [Figure 6].

This method suffers from several disadvantages including a low sensitivity to identify low molecular weight proteins, a low dynamic range (10^3), and the inability to resolve/solubilize proteins at the extremes of pH, size and hydrophobicity.[28,32] It is important to remember that the low molecular weight range (<15 000 Da) of the serum proteome is a rich source of previously undiscovered biomarkers. These markers exist below the range of detection achieved by conventional two-dimensional gel electrophoresis.[33] When combined with prefractionation techniques, however, Pieper et al were able to identify 325 different serum proteins.[34]

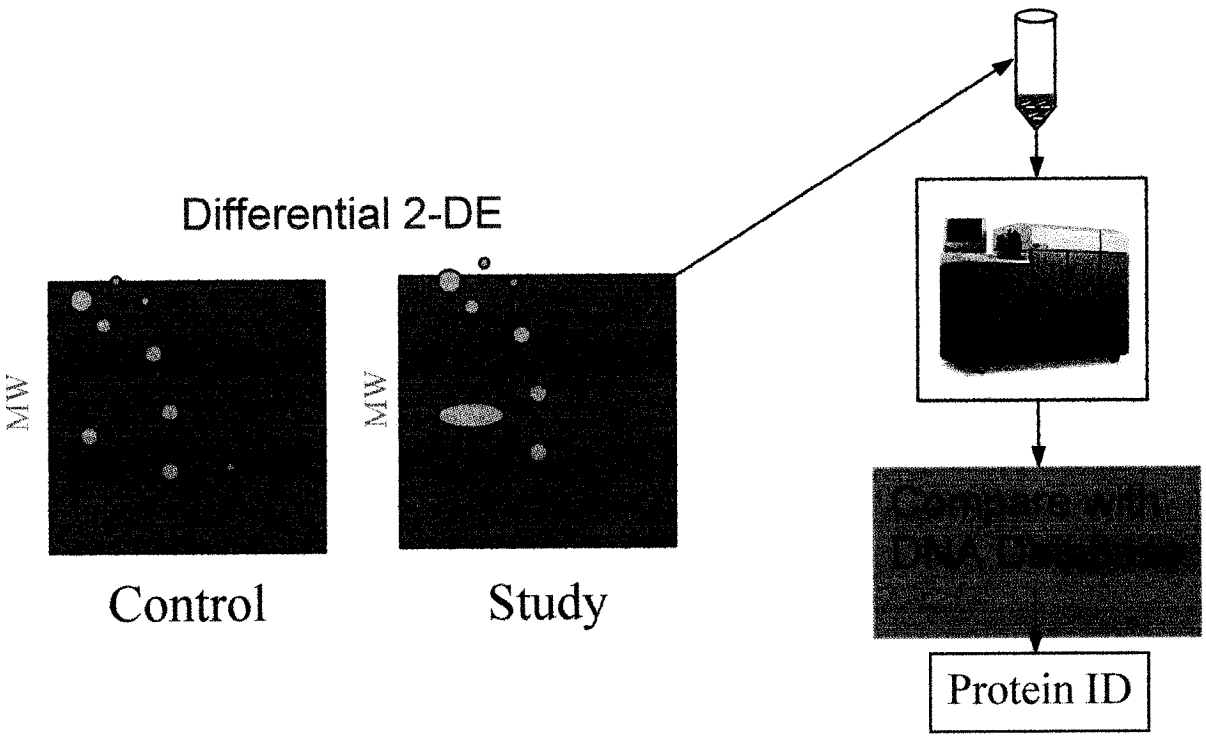

Figure 6. Conventional proteomic analysis using 2-Dimentional Gel Electrophoresis.

Liquid chromatography (LC) separates the ions or molecules that are dissolved in a solution, such as serum. It is a separation technique in which the mobile phase is a liquid. Therefore, as opposed to the gel based separation techniques, LC allows the separation of peptides directly from the solvent. Differences in ion-exchange, absorption, size and partitioning, determine the transit time of solutes through a column or a plane, and thus allow the separation of the mixture components [Figure 7].

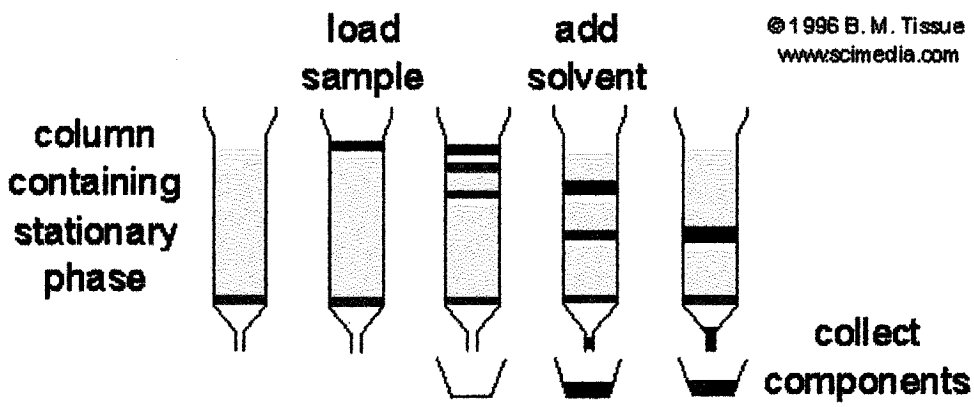

Figure 7. Liquid chromatography separation. Adapted from Science Hypermedia web site.

High-performance liquid chromatography (HPLC) is a form of LC, used to separate compounds that are dissolved in solution. Utilizing very small particles and a high pressure, HPLC forces the solution, or mobile phase, through a column, sometimes referred to as the stationary phase. The interaction of the solute with mobile and stationary phases can be manipulated through different choices of both solvents and stationary phases. As a result, HPLC acquires a high degree of versatility not found in

other chromatographic systems and it has the ability to easily separate a wide variety of chemical mixtures. HPLC also provides higher resolution and faster analysis time.[35]

4. Protein ionization

Ionization of the proteins is the fourth step in biomarker identification. Various ionization methods, outlined below, have been described in the literature including: matrix-assisted laser desorption ionization (MALDI); surface-enhanced laser desorption/ionization (SELDI); and electrospray ionization (ESI). All three are sensitive to the picomole-to-femtomole range that is required for application to biological samples.[16]

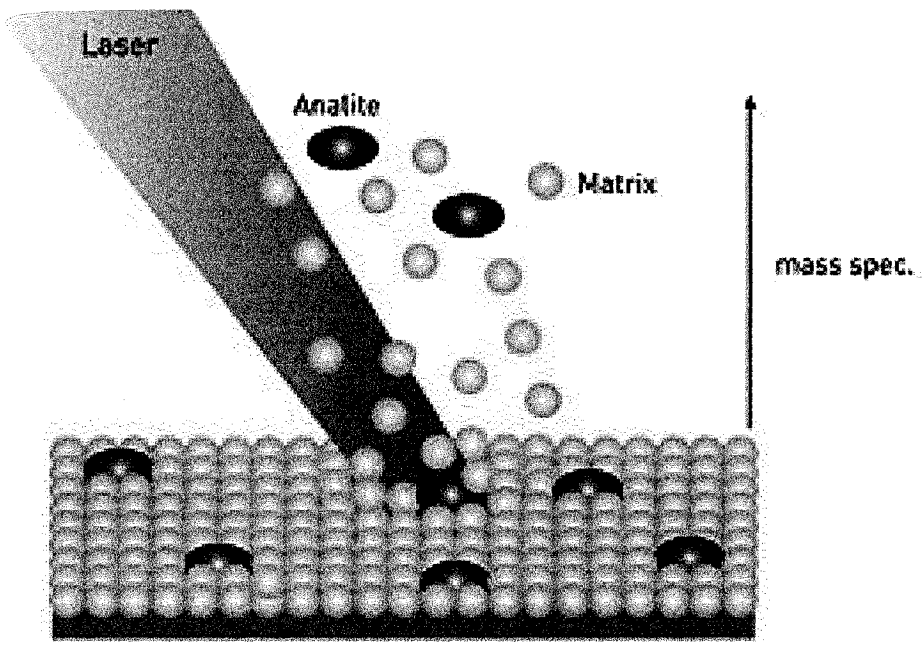

Figure 8. Matrix-assisted laser desorption/ionization (MALDI). Adapted from Eurogentec website.

Matrix-assisted laser desorption/ionization (MALDI) was first described in the 1980s.[36] It uses laser pulses to ionize the peptides out of a dry, crystalline matrix, as follows: 1) serum proteins bind to a chromatographic solid surface, referred to as a probe, and the unbound proteins are washed away; 2) the immobilized serum proteins are overlaid with a coating of an energy absorbing chemical matrix which then crystallizes; 3) the entire probe is inserted into a vacuum chamber; 4) a laser beam tuned to the excitation wavelength of the matrix, is then directed at each spot [Figure 8]; 5) the matrix serves as an energy transfer medium for protein ionization whereby the kinetic energy from the laser causes protons from the matrix to be transferred to the peptides resulting in positive ions; 6) this energy transfer causes the proteins to fragment, ionize, and vaporize into a gas cloud; 7) the gas cloud is then accelerated into a mass analyzer electrostatically.[16,37] It is very sensitive, simple to use, has an excellent mass accuracy and high resolution.[38]

Surface enhanced laser desorption/ionization (SELDI) was first described in the 1990s.[39] The steps used in SELDI are the same as the ones used in MALDI, with one addition: a protein-chip is used instead of a probe [Figure 9]. Protein-chips employ selective surfaces to capture only a fraction of proteins from a complex mixture in biological samples.[16,40] The disadvantages of using the SELDI technique to identify serum biomarkers include the need for a very large number of patients and controls and very high cost. Moreover, the protein concentrations, found using SELDI, are in the μg/ml range, much higher than known biomarkers such as the prostate specific antigen (PSA).[41]

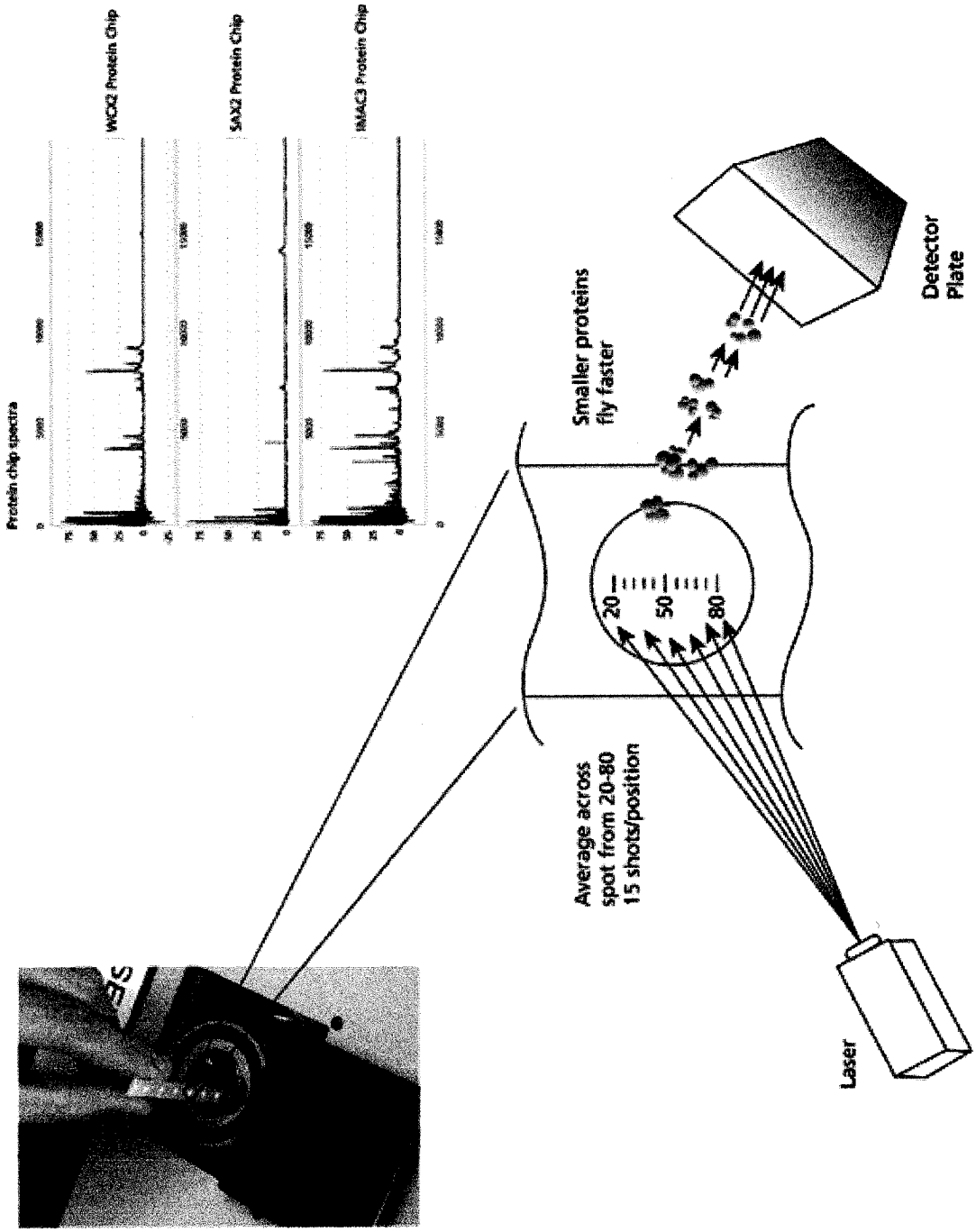

Figure 9. Surface enhanced laser desorption/ionization, SELDI.[20]

Electrospray ionization (ESI) was first described by John B. Fenn in the 1980s,[42] for which he won the 2002 Nobel Prize in chemistry, along with Koichi Tanaka for his work on laser desorption. ESI ionizes the peptides directly out of liquid solution and it is the preferred method for analysis of complex mixtures.[38] The steps include: 1) ion gas cloud created directly from the sample solution; 2) passage of the gas through a needle held at a high voltage produces highly charged droplets; 3) these droplets are electrostatically driven through nitrogen gas, air, heat, solvents, or other drying agents to evaporate water and solvents [Figure 10]; 4) as droplets decrease in size, the surface charges are deposited onto the peptides and proteins.[16,29]

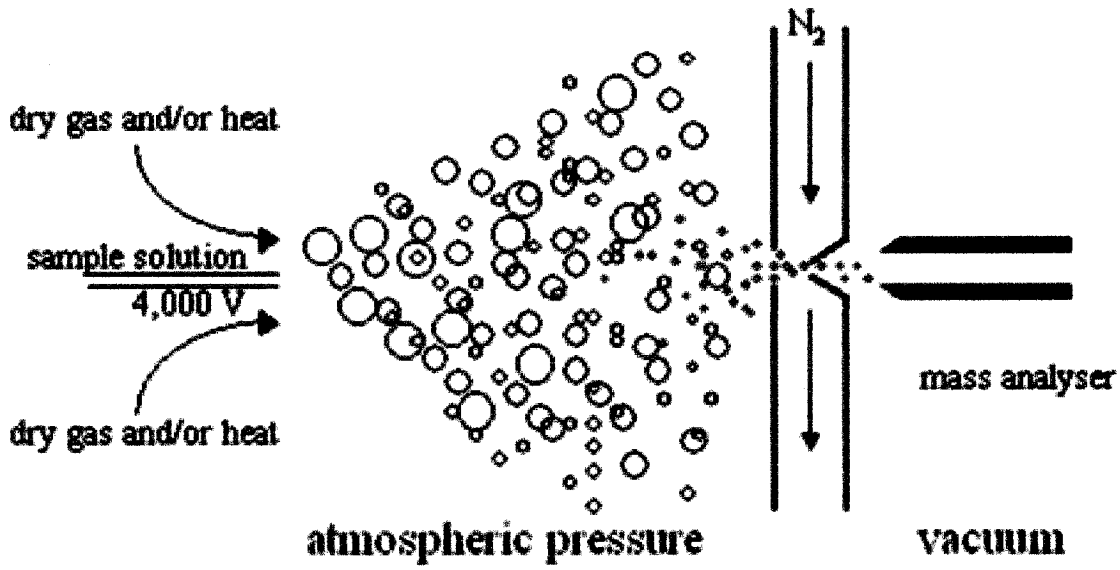

Figure 10. Electrospray ionization, ESI. Adapted from University of Wales - Aberystwyth website.

The numerous technical steps utilized in each of these methods are still in the process of being refined, and a debate continues as to which one of these techniques is best suited for biomarker identification. MALDI is similar to SELDI, but without the surface pre-selection or enrichment steps. The inherent limitation of SELDI, not observed with MALDI, rests in the selective nature of specific protein-chip surfaces. Thus, MALDI is less likely to be affected by technical artefacts, and it might be better for detection of biomarkers.[43] Moreover, both MALDI and SELDI have an advantage over ESI in the fact that they have a higher tolerance for salts, and thus are better suited for the examination of biological samples such as serum.[16] Nevertheless, ESI is more gentle than MALDI/SELDI. This kinder and gentler ionization method, with much less fragmentation, permits the formation of multicharged ion species which can be adequately detected in the more limited mass to charge ratio (m/z) range.[16] MALDI-time-of-flight (TOF) combination has a practical mass limit between 150,000 and 300,000 Da, whereas ESI-Quadrupole combination has a mass limit of 70 Da.[44]

Multidimensional protein identification technology (MudPIT) is a new method of analysis allowing in-line fractionation and mass spectroscopy to identify all proteins in a complex mixture.[30,32,45] It combines multidimensional liquid chromatography with electrospray ionization and tandem mass spectrometry. The multidimensional HPLC column allows ionization from the tip of the column directly into the tandem mass spectrometer.[30] This technique has exquisite sensitivity, reported to be in the femtomolar (10^{-15}) range, with a dynamic range of 10,000:1 for a complex mixture of peptides of tens of thousands of components, and is reproducible to within 0.5%.[30] A principle problem for analysis of complex protein mixtures is the dynamic range or quantity differences

among the proteins. With a dynamic range of 10 000:1, MudPIT allows for the detection and identification of a protein at 100 copies/cell in the background of a protein at 1,000,000 copies/cell or within a five orders of magnitude difference in concentration.[30] Nevertheless, although it is excellent in identifying proteins which are present in one solution, but not in another, it's major weakness is in identifying quantitative differences in protein expression.[32] Most agree that MudPIT should be used complementarily to 2-dimnetional gel electrophoresis when searching for biomarkers.[30]

5. Mass analyzers

The last step in biomarker identification is passing the peptides though mass analyzers. The different types of mass analyzers used in proteomics include: time-of-flight (TOF), quadrupole (q or Q), ion trap, or fourier transform (FT) [Figure 11].

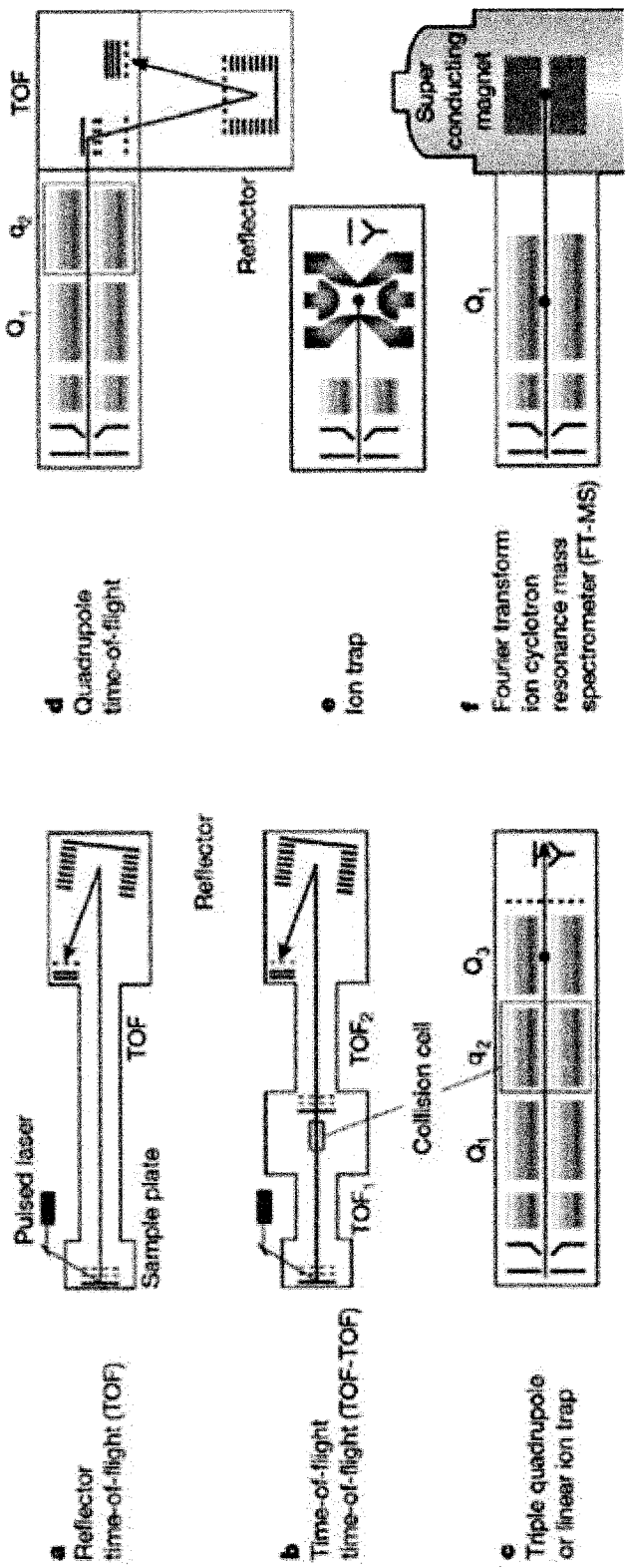

Figure 11. Examples of different mass analyzers and combinations.[38]

The TOF (time of flight) mass analyzer is the most widely used. It is affected by the mass (m) of the particle and the charge (z) it bears (m/z ratio). The detector plate records the intensity of the signal at a given m/z value [Figure 12]. The mass to charge value of each ion is estimated from the time it takes for the launched ion to reach the electrode; small ions travel faster. The result is a time-of-flight distribution of the peptides comprising the mixture. Consequently, the spectrum provides a "time-of-flight" signature of ions ordered by size with the different peaks in the spectrum corresponding to different m/z protein species. TOF is usually coupled with MALDI. It has a simple design, a low resolution, and is good for the detection of proteins <20 kDa.[16,29,46]

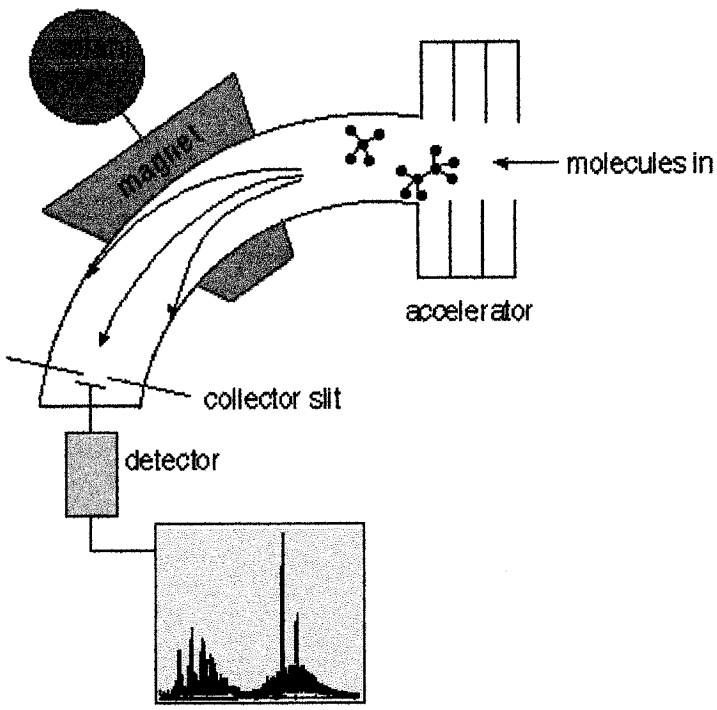

Figure 12. Example of a time-of-flight (TOF) mass spectrometer. Adapted from the University of Wisconsin – Madison website.

Quadrupole (q) mass analyzer is constructed from four parallel metal rods. Direct current (DC) and radio frequency (RF) voltages are applied to these rods, and create selective magnetic fields that control which ions pass through to the detector [Figure 13]. For a particular combination of DC to RF voltages, only ions of a specific m/z value can pass through.[16]

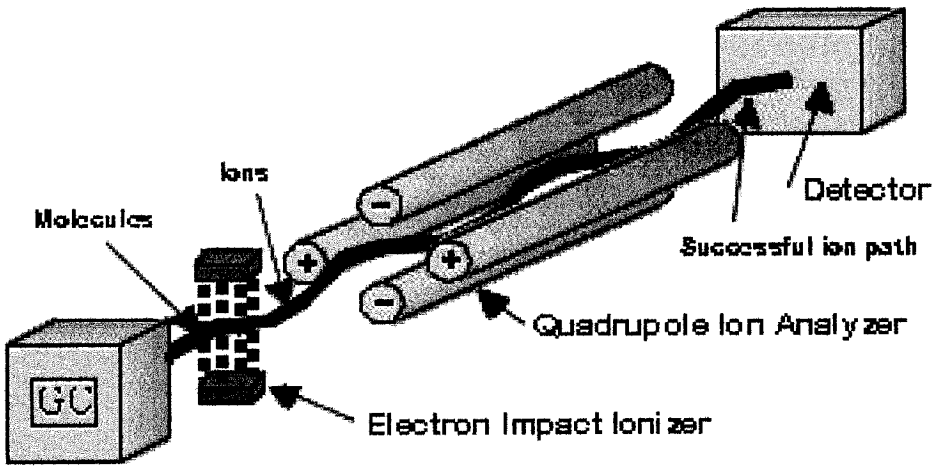

Figure 13. Examples of a quadrupole mass analyzer. Adapted from the University of Arizona website.

Using multiple analyzers together is called tandem mass spectrometry (MS/MS). Different combinations of analyzers can be used such as: QqTOF, TOF-TOF, linear ion traps [Figure 11]. Peptides sub-fragment in a reproducible way, resulting in a pattern somewhat like the sequence-ladder pattern obtained in DNA sequencing.[16] Tandem mass spectrometer not only measures the masses of the ions, but is also able to select individual ions, fragment them (usually by energetic collision with an inert gas) and then measure the masses of the resulting fragment ions. The most important benefit of MS/MS data is probably that the unit of identification is a single peptide rather than a group of peptides.[28,29]

6. Quantitation

Detecting quantitative differences across protein mixtures remains one of the limiting factors in proteomics. Quantitative differences can be found with the dilution of a stable isotope.[29] There are isotope-tags specific to sulphydryl groups, amino groups, phosphate ester groups, N-linked carbohydrates, as well as active sites for serine and cysteine hydrolases.[38]

Protein dye staining with fluorescent dyes,[47] or silver,[48-49] are used in gel-based technologies. Silver is the method used most often with 2D gel electrophoresis.[50] A newer method, called Differential Gel Electrophoresis (DIGE) involves tagging the two comparison protein solutions with different fluorescent dyes (Cy 2, Cy 3 or Cy 5), then mixing the solutions and running them on the same gel.[51] Mass and charge-matched fluorescent cyanine dyes undergo nucleophilic substitution with ε-amine groups of lysine residues of proteins. Because mass mapping requires an essentially purified target protein, the technique is commonly used in conjunction with prior protein fractionation. DIGE is essentially the same as 2D gel electrophoresis, just both (control and experimental) samples are run on the same gel. The system suffers the same disadvantages as standard 2-DE, most concerning for us is the lower sensitivity and lower resolution of low molecular weight proteins as compared to mass spectroscopy techniques.[47] Other problems associated with this technology include: possible multiple dye additions on each protein; only about 5% of protein is labelled; and it is less sensitive than silver stain.[47] The advantages over the regular (silver stained) gel include: less variability in spot positions, same proteins from both samples come on exactly the same position. The disadvantages greatly downplay the practical outcome from DIGE.

Chemically derived tags such as ICAT (isotope-coded affinity tag composed of a reactive group for the amino acid cysteine)[52] and MCAT (mass-coded affinity tag targets lysine residues, N-acetoxysuccinimide, succinic anhydride)[53] can be used in liquid mixtures. Alternatively, performing an endoproteolytic digest in the presence of water containing ^{18}O can be performed. This substitutes both oxygens on the caboxylic acid with a heavy form.[28,54] Applied Biosystems recently developed four new labelling reagents for mass spectroscopy called iTRAQ. Although we could not find publications using these reagents, they have been tested at the McGill Center for Proteomics and Genomics and show great promise in improving quantitative analysis of MS/MS.

7. Proteomic databases

The final step in the biomarker identification is the search through different proteomic databases [Figure 14]. The individual spectra can be used to search against predicted peptide spectra from the databases using different search algorithms such as Sequest [55] or Mascot.[56,57] Shotgun proteomics implies direct and rapid analysis of the entire protein complement within a complex protein mixture.[32]

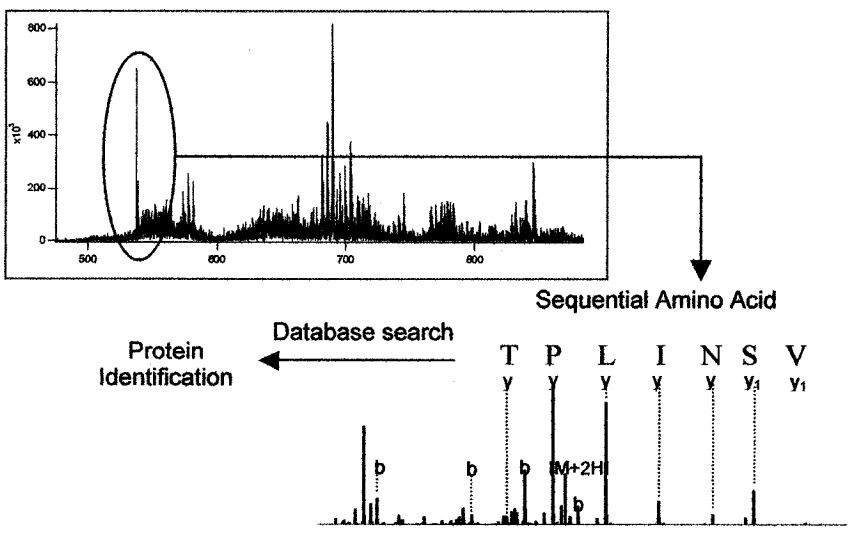

Figure 14. Protein identification by mass spectrometry.

ii. Proteomic studies and the search for cancer biomarkers

Proteomics, which is the large-scale study of proteins, particularly their structures and functions, has many applications in all fields of medicine. Because it studies the end products of genetics, it is the key to the discovery of new biomarkers in cancer research.[16,17] The study of proteins dates back to the 1970s, when databases of proteins were created using two-dimensional gel electrophoresis. However, it wasn't until the 1990s and the emergence of biological mass spectrometry that the field of proteomics exploded.[29] Some of the most exciting advances in the past three years have used surface enhanced laser desorption/ionization (SELDI) mass spectroscopy. It has been successfully used in the search for serum markers of various types of cancers, such as: ovarian,[58-62] pancreatic,[63] prostatic,[64] breast,[65] lung,[66] liver,[67] colon.[68] Using this technology, a study of ovarian cancer obtained results having 100% sensitivity and 95% specificity.[59]

Nevertheless, SELDI has some disadvantages in the biomarker hunt including the need for a very large number of patients and controls, as well as a very high operational cost. Moreover, the protein concentrations, found using SELDI, are in the μg/ml range.[41] The known biomarkers, such as the prostate specific antigen (PSA), are relatively sparse in serum.[61] Therefore, when searching for a previously undiscovered biomarker, it is important to focus on the low abundant range the serum proteome.[16,60] Matrix-assisted laser desorption/ionization (MALDI), on the other hand, is less likely to be affected by technical artefacts when compared to SELDI. As discussed previously, the inherent limitation of SELDI, not observed with MALDI, rests in the selective nature of specific ProteinChip surfaces. Proteins that do not bind to the chip are washed away, potentially

losing very important information. Therefore, some propose MALDI to be a better technique for detection of cancer markers.[43]

It is important to remember, however, that cancer is not confined only to the diseased cell. The microenvironment of the tumor-host interface, the surrounding stromal and vascular compartments, play a very important role in the tumorgenesis.[14] It is therefore extremely important to take into account all these aspects when hunting for a new biomarker.

iii. Proteomic studies on HNSCC

To date, at least twelve published studies reported proteomic technology to identify biomarkers for HNSCC.[43,49,69-78] Four have used surface-enhanced laser desorption/ionization (SELDI) technology,[69-72] two matrix-assisted laser desorption/ionization (MALDI),[49,73] and one liquid chromatography and electrospray ionization (LC/ESI).[74] Most of these studies reported differential protein expression levels in carcinoma tissues as compared with their paired normal mucosa.[49,69-74] Four studies have used mass spectrometry to analyze serum profiles of HNSCC patients.[43,75-77] These studies found that particular protein peaks had sensitivities between 68 and 83% and specificities between 73 and 90%, but no specific proteins were reported. Moreover, it was mentioned that the most dominant data points had very low intensity and can be easily mistaken for background noise. In contrast, a study by Gourin et al. reported that proteomic analysis of serum protein profiles could distinguish patients with HNSCC from controls with a high degree of sensitivity and specificity.[77] One study has used mass spectrometry to identify specific serum proteins in a mouse model of tongue cancer, and

identified the squamous cell carcinoma antigen 1 as the only protein over-expressed in serum from tumor-bearing mice.[78] The following is a summary of each of these studies.

Baker et al. analyzed tissues from five tongue squamous cell carcinoma patients, using ESI/LC/MS/MS.[74] They reported several protein differences between normal and tumor tissues including: heat shock proteins (HSP70 & HSP90) up-regulated in cancer, keratin 13 down-regulated, Wnt-6 and Wnt-4 identified in both normal and tumor, and placental growth factor (PIGF) found only in tumors.

He et al. studied ten oral tongue squamous cell cancer patients, using 2-DE, MALDI-TOF MS.[49] They found several proteins up-regulated in cancer tissues: myosin heavy chain 1 (MHC1), galectin 1, tropomyosin Mγ isoform, heat shock proteins (HSP60 & HSP27), nuclear fragile X mental retardation protein (FMRP), and calgranulin B. The down-regulated proteins included myosin light chains (MLC), tropomyosin β chain, ATP synthase β chain, and αB-crystalline.

The same group also studied ten buccal squamous cell cancer patients using 2-DE, MALDI-TOF MS.[73] Only three proteins were identified to have similar trends of alteration in both tongue[49] and buccal carcinoma: crystalline-B being down-regulated, whereas HSP27 and mitochondrial ribosomal protein (MRP)-L13 were up-regulated in cancer tissues (MRP-L13 being the same protein as FMRP). Up-regulated proteins in buccal cancer included: glycolytic enzymes, heat-shock proteins, tumor antigens, cytoskeleton proteins, enzymes involved in detoxification and anti-oxidation systems, and proteins involved in mitochondrial and intracellular signalling pathways. Finally, they suggested several candidate proteins for further HNSCC marker analysis: SCC antigen, G protein (GNBP), glutathione S-transferase (GST), manganese superoxide

dismutase (MnSOD), annexins, voltage-dependent anion channel (VDAC), cyclophilin A (CyP-A), stratifin and galectin 7.

Von Eggeling et al reported on six HNSCC cases using SELDI. They detected a protein of 8670 Da in tumor extracts of five of six HNSCC cases and not in matched normal tissues.[69]

Wu et al. studied two matched HNSCC cell lines derived from either the primary tumor or lymph node metastasis, using SELDI.[70] They identified the up-regulation of two membrane-associated proteins (annexin I and annexin II) and glycolytic protein enolase-α, as well as the down-regulation of calumenin precursor in the metastatic cell line.

Melle et al. analyzed fifty seven pharyngeal cancer tissues, using SELDI.[71] They identified calgranulin A and B, and calgizzarin as potential markers and suggested that the triad of microdissection, SELDI and immunohistochemistry (IHC) should be used in the identification of biomarkers.

Roesch et al. used cDNA microarrays, qRT-PCR and SELDI-TOF MS to study differential expression of calcium-binding proteins of the S100 and the annexin protein families.[72] They found algranulins A and B, and annexins 1 and 2 to be down-regulated at both mRNA and protein levels.

Three studies have used mass spectroscopy technology to analyze serum profiles of head and neck squamous cell carcinoma patients. [43,75-77] All of these studies included large numbers of patients, and found that particular protein peaks had sensitivities between 68 and 83% and specificities between 73 and 90%. However, these differences were not statistically significant. Moreover, the most dominant data points had very low intensity and could have been easily mistaken for background noise.

Soltys et al studied 113 HNSCC and 104 controls using SELDI.[43] They reported 65 significant data points for discrimination of normal from cancer profiles, with sensitivity of 68%, and specificity of 73%.

Sidransky et al studied 99 HNSCC patients and 143 controls using MALDI.[76] They found that total protein levels, particularly ten individual m/z peaks, from 5 to 111 kd, were higher in cancer patients with a sensitivity of 70% and a specificity of 90%. However these differences were not statistically significant.

Wadsworth et al performed serum proteomic profiling using SELDI on 99 HNSCC, 25 smokers, and 102 controls.[75] Several proteins with masses ranging from 2778 to 20800 Da were differentially expressed with a sensitivity of 83.3%, and a specificity of 90%. They also found a peak with an average mass of 10068 Da in sera from HNSCC patients and identified it as metallopanstimulin-1 (MPS-1) based on mass.

One study has used 2-DE, MALDI mass spectrometry to identify serum proteins in a mouse model of tongue cancer.[78] They compared serum of mice implanted with tumor to those injected with phosphate-buffered saline. Using this technology they found only one protein to be over-expressed in tumor-bearing mice: squamous cell carcinoma antigen 1.

In a previous study, Balys et al. developed a RAG2/γ(c) immunocompromised mouse model able to reliably engraft human oral cancer, as well as normal human tongue tissue.[79] This mouse model had several advantages for biomarker proteomic discovery, including: a known mouse genetic background, which was alike in all respects except for engraftment of the cancer tissue or normal tongue tissue from the same patient; the ability to obtain multiple samples from the same mice to demonstrate accuracy of the findings;

and the ability to differentiate proteins released from cancer tissue versus its host/tumor microenvironment. In the latter case, homology search allowed discrimination between distinct conserved regions between mouse and human proteins identified. Human proteins could only have arisen from the cancer, while the mouse proteins reflected host and tumor microenvironment response.

Numerous studies have reported on the orthotopic transplantation of oral cancer into immunocompromised mice. Six papers have reported successful take (50-100%) of human oral cancer xenograft into immunocompromised mice.[80-85] Local invasion and metastasis often occurred when cancer cells were transplanted orthotopically, and some metastasis models have been established.[86] Moreover, oral SCC cell lines and tissues have been reported to invade and metastasize to neck lymph nodes when transplanted into tongues of nude mice.[85] The RAG2γ(c) knockout mouse has been suggested as the first universal recipient, able to take virtually any tumor or normal tissue due to an absence of B cells, T cells, and natural killer cells.[87,88]

The objective of this study was to first establish a reproducible proteomic platform for identifying serum proteins in a mouse model of oral squamous cell carcinoma that can specifically predict human disease, and second, to identify specific candidate proteins that can serve as potential biomarkers for the human disease.

3. Methodology

a. Ethics

Ethics approval for the use of human tissue was obtained from the Research Ethics Committee at the Jewish General Hospital (protocol no. 04-082). Animal ethics was approved by the Animal Care Committee at the Lady Davis Institute for Medical Research (protocol no. 5018). [Appendix b]

b. Mouse model

Two patients with a planned surgical resection of an oral squamous cell cancer were informed of the study and asked to participate. During their operation, two biopsies were obtained: 1) of oral cancer; 2) of adjacent normal tissue. Tumor tissue was cut into 0.5 mm^3 pieces and implanted surgically into the tongues of five RAG-2/γ(c) immunocompromised mice. Normal tissue was confirmed to be free of cancer using standard pathological techniques by an experienced head and neck pathologist, and it was cut into $0.5mm^3$ and $0.7cm^3$ pieces. The smaller pieces were implanted into the tongues of five RAG-2/γ(c) mice and the larger pieces were implanted in a subcutaneous pocket. Our preliminary experiments demonstrated superior results when the sacrificed mice showed a tongue tumor diameter between 0.65 and 0.8cm (average of 0.7cm). To provide the best possible control, we implanted this amount of normal tissue into the control mice subcutaneously, as well as a small piece into the tongue to account for any localized inflammatory changes in the tongue tissue.

Post-operatively, the mice were kept under identical conditions. When the tumor-bearing mice lost 20% of their weight or showed signs of discomfort, they were

sacrificed as were their respective control mouse. Our previous study has demonstrated the most reliable method of obtaining a pure serum sample (no hemolysis) is through a puncture of the inferior vena cava. The serum was then separated into 50 µl aliquots with protease inhibitor cocktail tablets (Roche Diagnostics) and immediately frozen at -80°C until the time of proteomic analysis. The most abundant serum samples from each patient (one from a tumor mouse and one from a control mouse) were chosen as the representative samples, and sent for proteomic analysis [Figure 15].

In parallel, tumor tissues were used to establish matched cell lines using enzymatic digestion as our laboratory has described previously.[8,89] From each patient's cancer tissue, we isolated a fibroblast-free cancer cell population, denoted as OSCC1 and OSCC2, for patient 1 and patient 2, respectively. These cells were maintained in RPMI medium supplemented with 5% fetal calf serum and 50 units/ml of penicillin-streptomycin antibiotics. Tumors induced by these cells into tongue of RAG-2/γ(c) mice were confirmed to maintain the same pathologic, tumorigenic and invasive potential as their original human tissue.[79]

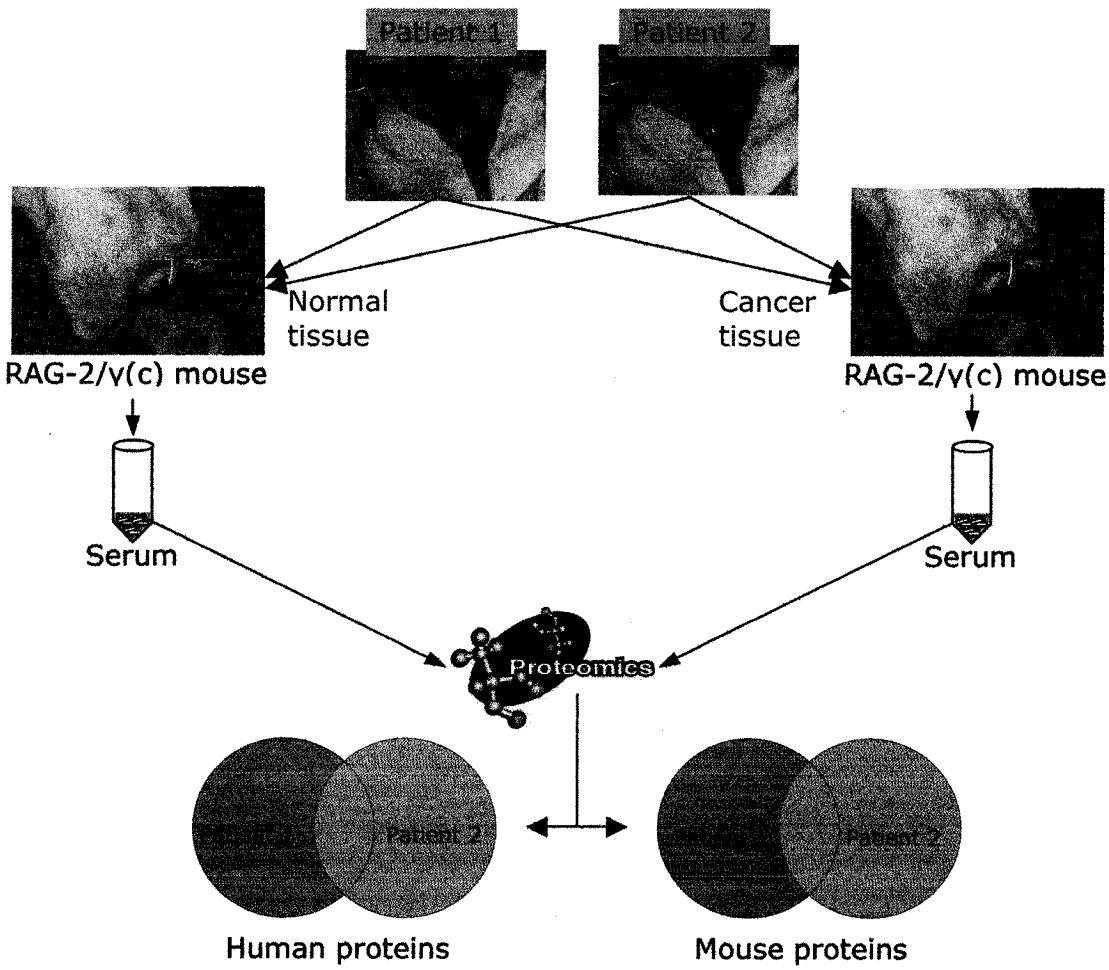

Figure 15. Summary of the methods used.

c. Proteomic techniques

The Agilent multiple affinity removal system was used to deplete the high abundant serum proteins. The depleted mouse serum was processed using an automated workstation and analyzed using a two-arm proteomic approach [Figure 16]. Each arm used a different method of protein identification and quantification, combined in a complimentary fashion to allow the highest probability of identification of proteins from cancer versus host and tumor microenvironment.

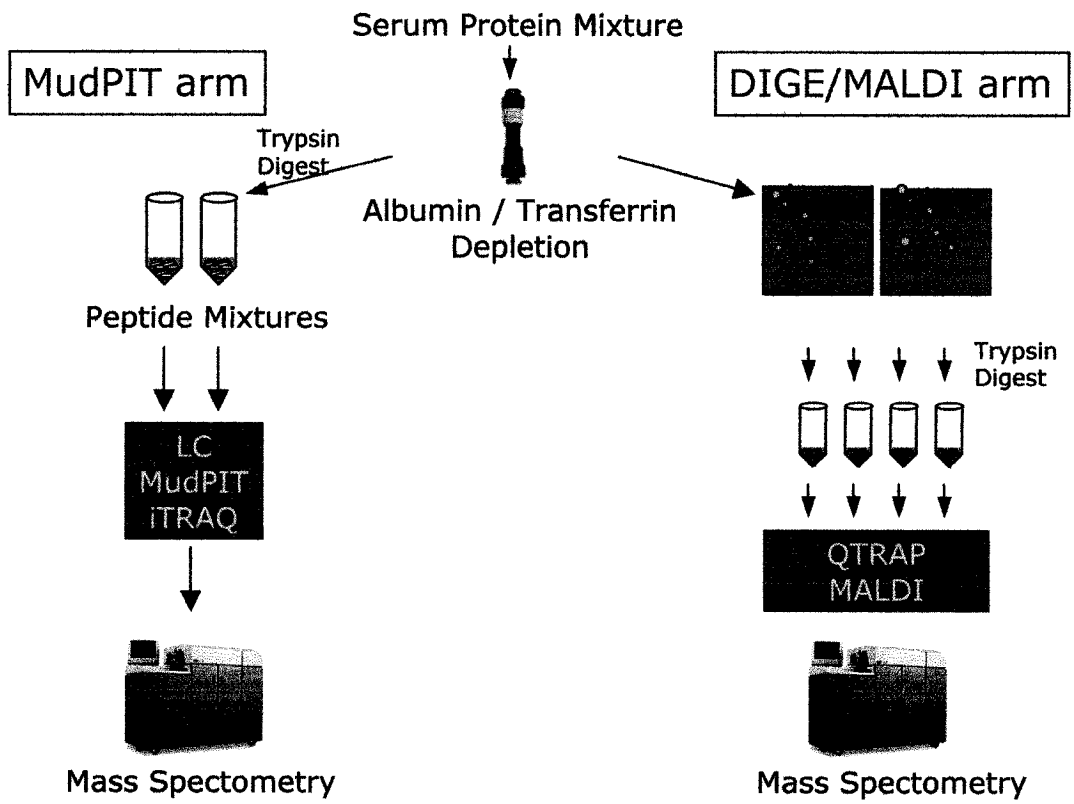

Figure 16. Two Arm Proteomic Model.

The first arm of our model was named DIGE/MALDI. It involved two-dimensional gel electrophoresis using Cy fluorescent dyes (DeCyder Differential Analysis Software v5.0, Amersham Biosciences). It tagged the two comparison protein solutions with different fluorescent dyes (Cy 2, Cy 3 or Cy 5) then mixed the solutions and processed them on the same gel.[51] The two-dimensional gels were analyzed using DyCyder software (Ettan DIGE, GE Healthcare), which permitted comparison across multiple gels. Semi-quantitation was done with the Phoretix 2004 Image Analysis program. Consistent candidate proteins, which reliably showed increased expression across samples, were cut from the gels, in-gel trypsin digested, and analyzed by matrix-assisted laser desorption ionization (QTRAP MALDI) tandem mass spectroscopy for identification.

The second arm, was named MudPIT (Multidimensional-Protein-Identification-Technology), combining LC/ESI/MS/MS with database searching to identify proteins in a complex solution.[30,32,90] This arm involved trypsin digestion, iTRAQ labelling (iTRAQ Reagents, Applied Biosystems), automated in-line two-dimensional column liquid chromatography (LC), electrospray ionization (ESI), and tandem mass spectroscopy (MS/MS). The trypsin digested mixtures were labelled with different iTRAQ reagents to allow detection of relative quantitative differences between samples. The elution was passed directly into an electrospray ionization (ESI) mass spectroscopy unit. Here the individual peptides were initially characterized by their mass to charge ratio. Identified proteins were then analyzed sequentially by searching human and mouse peptide databases (Mascot & ProQuant) of mass spectra.[56,57] Proteins of human origin must have originated from the tumor, while proteins of mouse origin must have originated from the host-tissue microenvironment.

d. Marker (EGFR) validation

Epidermal growth factor receptor (EGFR) validation was performed using ELISA, Western blot and immunostaining.

Enzyme linked immunosorbent assay (ELISA) (Oncogene Science, Bayer Corporation, Cambridge, UK) was used to measure the levels of circulating serum EGFR and EGFR released in conditioned serum-free cell culture medium. This sandwich type immunoassay uses a mouse monoclonal capture antibody against the extracellular domain of EGFR immobilized onto a microtiter plate, and an alkaline phosphatase-labeled mouse monoclonal antibody as detector specific for the electron capture detection of human EGFR. The procedure was conducted according to the manufacturer specifications. Briefly, 100ul of standards and samples diluted at 1:50 with PBS were added to the antibody-coated ELISA microtiter wells and incubated for 1.5h at 37. After washing, 100ul of alkaline phosphatase-labeled mouse monoclonal detector EGFR antibody was added for 30min at room temperature. After multiple washes to remove unbound antibody-enzyme reagent, a substrate solution containing Bluephos substrate was added for 1h at room temperature. The reaction was terminated by adding 100ul of a stop solution. Colorimetric quantification was performed using a multireader spectrophotometer at 620nm. The results were expressed in ng/ml based on a standard curve using human recombinant EGFR at concentrations ranging from 0.1 to 50ng/ml with a detection limit of 0.55 ng/ml. Statistical significance among groups was assessed by Student's t-test.

Western blot assay was described earlier.[91] Briefly, 50ug of a protein cell extract was prepared, electrophoresed on polyacrylamide gels, and the separated proteins were

transferred to nitrocellulose and detected with 10ug/ml of biotinylated EGFR antibody (clone13, Transduction Labs) followed by peroxidase-conjugated streptavidin by chemiluminescence. Blots were subsequently stripped and immunoblotted with monoclonal anti-GAPDH antibody.

Immunohistochemistry was performed with tumors fixed in 10% neutral buffered formalin and embedded in paraffin wax. Sections were immunostained using an anti-EGFR (clone13, Transduction Labs). Immunostaining was performed on 5-μm thick sections as previously described.[21] Sections were counterstained with Harris' Hematoxylin and mounted. All sections were analyzed by conventional light microscopy and digital photography (Leitz Aristoplan).

4. Results

Our study focused on tumor versus matched normal tissue from two patients with a defined pathology and distinct invasive property, based on clinical observations [Table 1]. Clinically, the cancer from patient 1 was highly invasive, while the cancer from patient 2 was less aggressive. This was correlated with the intrinsic invasive capacity of the isolated matched cancer cells as determined in vitro using the Boyden chamber matrigel assay.

	Patient 1	Patient 2
Age	45	74
Gender	Male	Male
TNM classification	T4N2b	T3N1
Surgery	Total glossectomy, laryngopharyngectomy, bilateral neck dissection	Hemi-glossectomy, bilateral neck dissection
Pathology	Poorly differentiated SCCa, with vascular, lymphatic and perineural invasion	Moderately to poorly differentiated SCCa, without vascular and perineural invasion
Previous Treatment	Chemotherapy + Radiotherapy	Surgery x 2
Past Medical History	None	None
Medications	None	Percacet
Habits	Smoker, Alcohol drinker	Smoker, Alcohol drinker
Invasiveness*	++++	++
Cell line**	OSCC1	OSCC2
In vitro Invasiveness***	38 %	4 %

Table 1. Patient characteristics.
*based on clinical characteristics
**cell lines isolated from the 2 human patients
***based on Boyden chamber invasion assay[21]

a. Mouse model

Implantation of oral cancer tissues into RAG-2/γ(c) mice was successful in 80% of cases [Figures 17 & 18]. Normal human tongue tissue implantation was successful in 30% of cases [Figure 19]. Analysis by an experienced head and neck pathologist confirmed that the original human patient tumors had the same histology as the corresponding tumor tissues obtained from the mice.

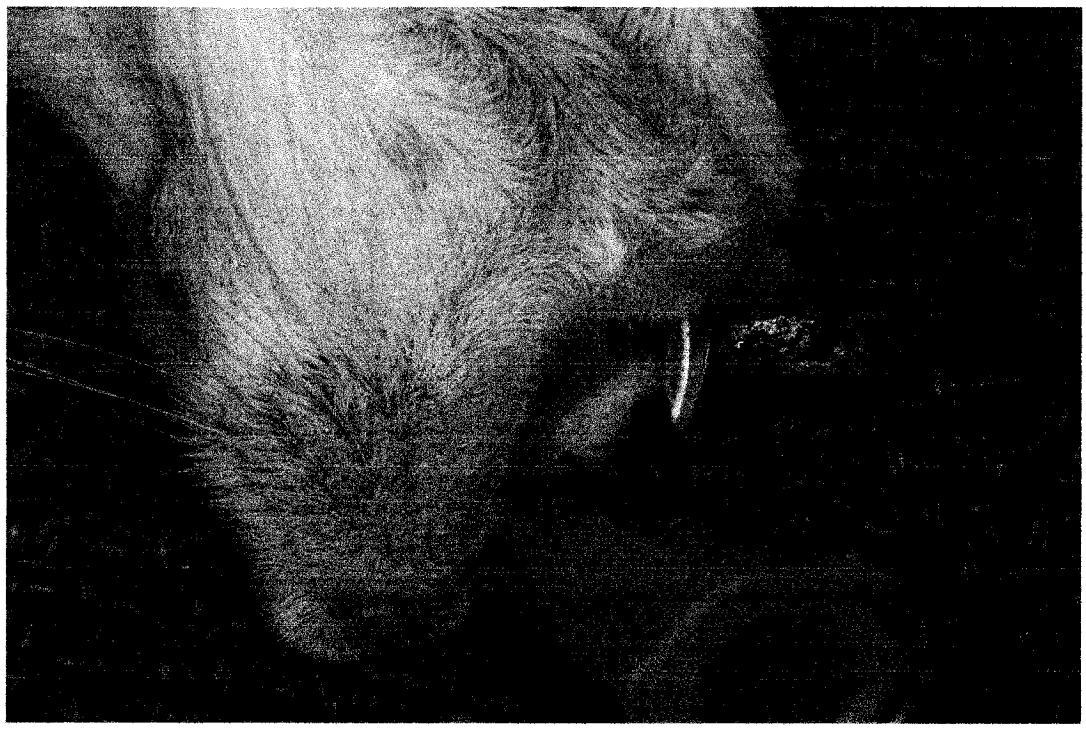

Figure 17. Orthotopically transplanted squamous cell carcinoma in an immunocompromised RAG-2/γ(c) mouse.

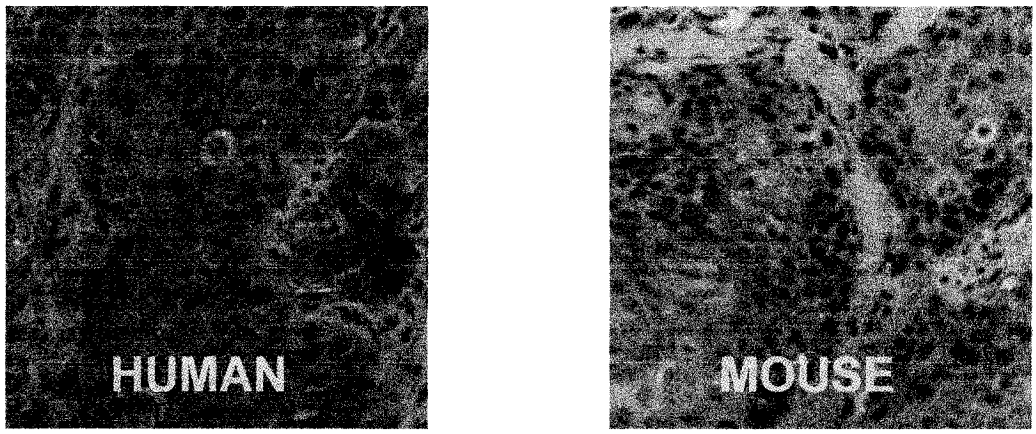

Figure 18. Histology of the squamous cell carcinoma found in the mice was identical to original patients' tumors.

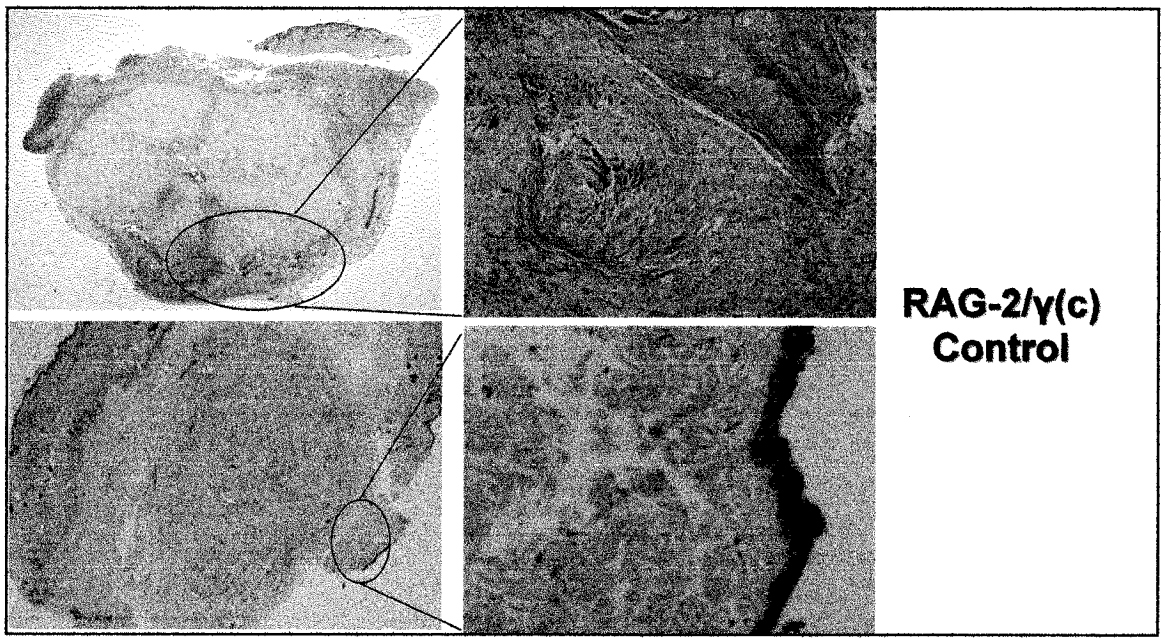

Figure 19. Histology showing normal human tongue surviving in the tongues of immunocompromised mice

b. Albumin depletion

Using the Agilent multiple affinity removal system, we were able to remove an average of 75% of total proteins, including 95% of albumin and 95% of transferrin [Figure 20].

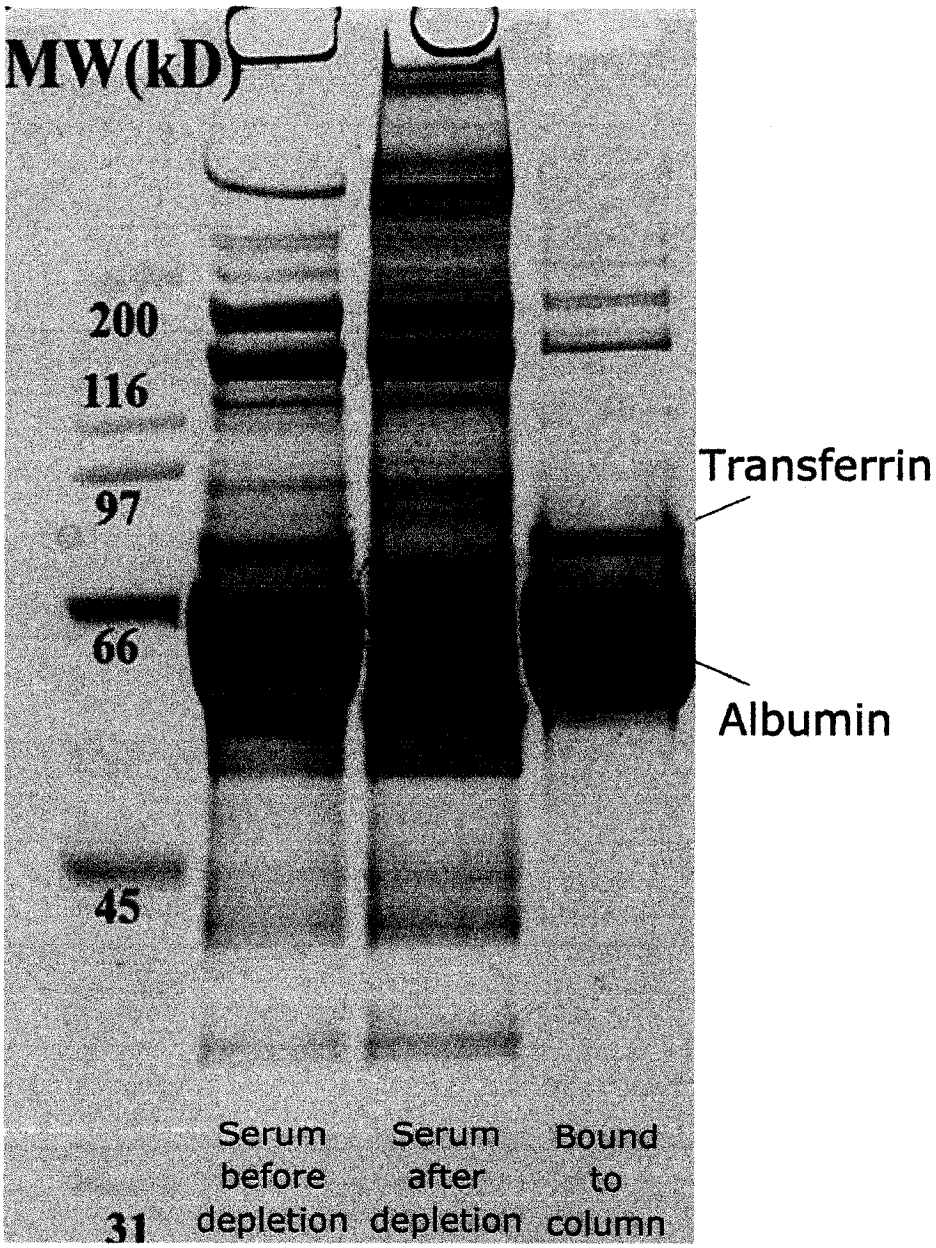

Figure 20. Results from the protein depletion column, showing removal of 75% of total proteins, including 95% of albumin and 95% transferrin.

c. DIGE / MALDI

Two-dimensional gel electrophoresis revealed 50 distinct proteins in the first patient and 75 proteins in the second patient [Figure 21]. The majority of identified protein spots were distinct and easy to cut out. However, even after the depletion of the high abundant proteins there were some large spots on the gel, most likely a representation of protein agglomerations.

All the spots were excised; trypsin digested and analyzed using QTRAP MALDI / MS / MS. The proteomic analysis of the spots from the first patient revealed only two proteins being differentially expressed (found in different quantities) in the cancer bearing mice as compared to the corresponding normal tissue bearing mice. Both of these proteins were down-regulated (found in lesser quantity in the tumor bearing mice) by at least two fold, and both were mouse proteins. Serum analysis of the second patient revealed 20 proteins to be differentially expressed. Seven proteins were found to be down-regulated in the tumor by at least two fold: one human and six mouse proteins. Thirteen proteins were found to be up-regulated (found in higher quantity in the tumor bearing mice): five human and eight mouse proteins.

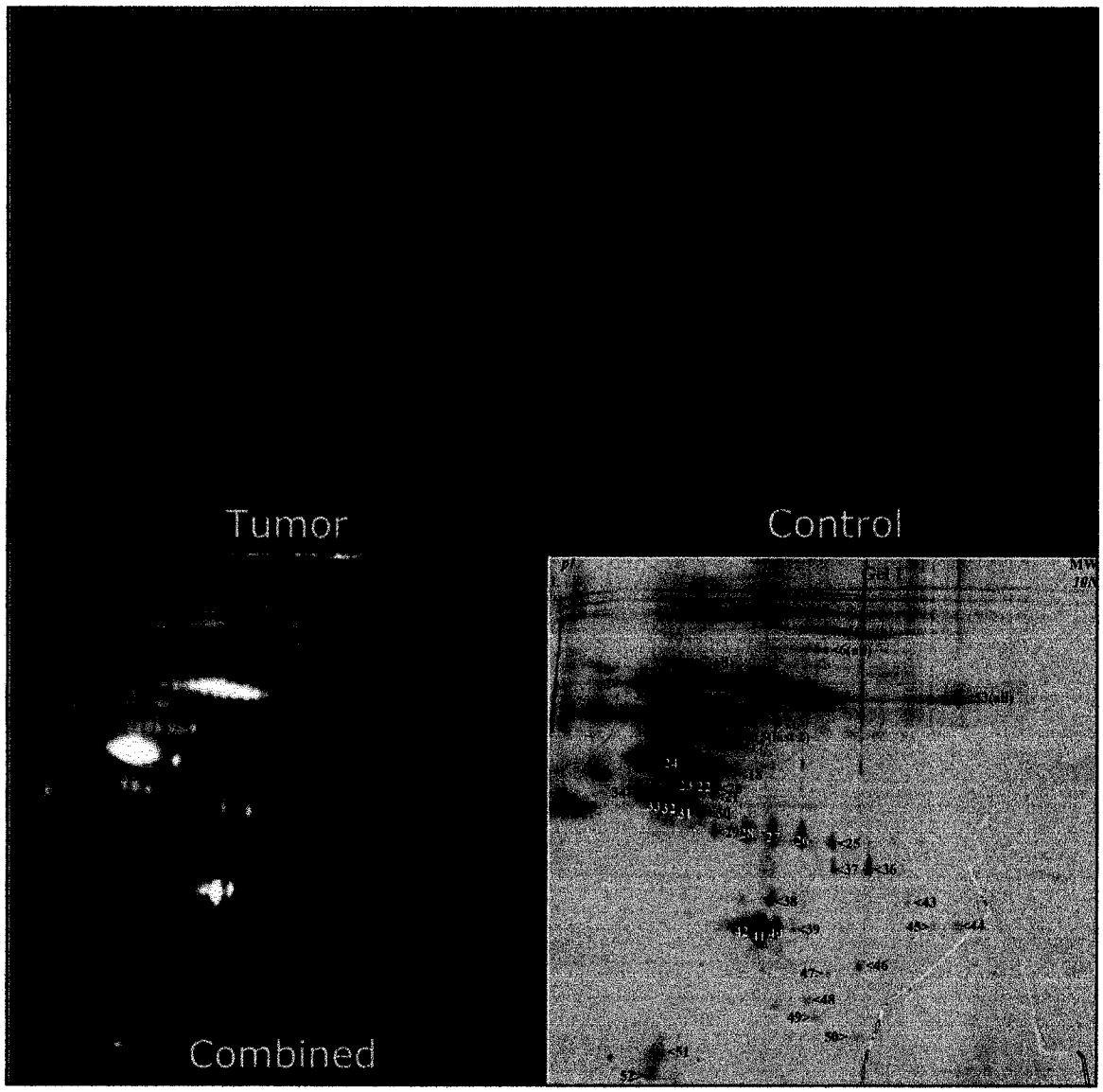

Figure 21. Results from the 2-Dimentional gel electrophoresis of serum from mice bearing oral cancer tissue and matched normal tongue tissue from patient 2.

d. MudPIT

iTRAQ labelling was found to be 90% successful in tagging proteins. Proteins from the tumor mice were tagged at position 117; proteins from the control mice were tagged at position 114.

Using the ProQuant search engine, we were able to identify, within a 95% confidence interval, 762 different spectra, 434 distinct peptides, 38 proteins that were expressed differentially in the tumor and the control mice in first patient [Table 2]. Fourteen of the proteins were found to be human, with eight being up-regulated and six down-regulated in the cancer bearing mice. ProQuant analysis of the second patient revealed 1862 different spectra, 1172 distinct peptides and 111 proteins within a 95% confidence interval [Table 3]. Twelve of these proteins were human, with six being up-regulated, and six down-regulated.

Confidence (ProtScore) Cutoff	Proteins Identified	Proteins before Grouping	Distinct Peptides	Spectra Identified	% of Total Spectra
>99 % (2.0)	23	122	222	547	13.20
>95 % (1.3)	38	342	434	762	18.39

Table 2. Summary of the ProQuant MudPIT results from first patient.

Confidence (ProtScore) Cutoff	Proteins Identified	Proteins before Grouping	Distinct Peptides	Spectra Identified	% of Total Spectra
>99 % (2.0)	84	425	1106	1785	24.44
>95 % (1.3)	111	902	1172	1862	25.49

Table 3. Summary of the ProQuant MudPIT results from second patient.

The Mascot search engine identified 181 proteins (60 human: 22 up-regulated, 28 down-regulated, 2 unchanged, 8 un-tagged) in the first patient and 102 (7 human: 6 up-regulated, 0 down-regulated, 1 un-tagged) proteins in the second patient as being differentially expressed between control and cancer-bearing mice ($p<0.05$).

Complete lists of the differentially expressed human and mouse proteins and their accession numbers are listed in appendix a.

e. Human proteins

Three human proteins, gelsolin, epidermal growth factor receptor (EGFR) and cytokeratin 10 were identified using the various proteomic methods and search engines. Only gelsolin was identified in serum of mice implanted with oral cancer tissues from both patients [Figure 22]. It was found to be down-regulated in serum from mice implanted with tissue from patient 1, but up-regulated in serum from mice implanted with tissue from patient 2. EGFR and cytokeratin 10 were both identified to be up-regulated in serum from mice implanted with tissue from patient 2. Other interesting human proteins found to be significantly affected between control and cancer-bearing mice included bullous pemphigoid antigen 1 (down-regulated in serum from mice implanted with tissue from patient 1), retinoblastoma associated factor 600 (down-regulated in serum from mice implanted with tissue from patient 2), titin, and BRCA2 (both up-regulated in serum from mice implanted with tissue from patient 1).

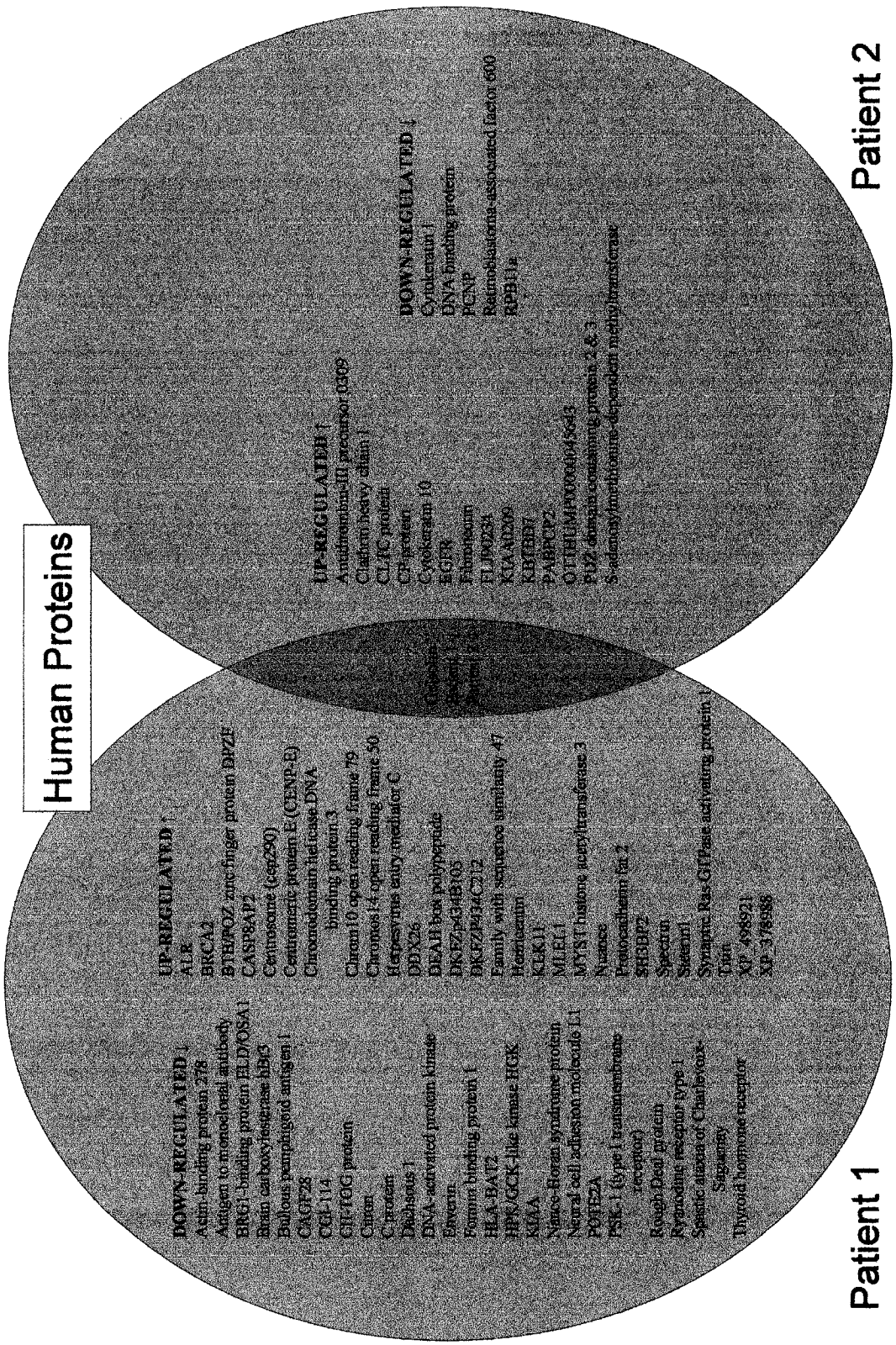

Figure 22. A Venn diagram summary of differentially expressed human proteins.

f. Mouse proteins

Twenty-eight of the 75 recognized mouse proteins were identified using the various proteomic methods and search engines. These proteins included several inflammatory and non-inflammatory-associated factors, including: alpha-1 protease inhibitor 2, apolipoprotein, BRCA2, clusterin, contraspin, fetuin, gelsolin, haptoglobin, hemopexin, histidine-rich glycoproteins, kininogen, serine (or cysteine) proteinase inhibitor, serpina1, serum amyloid P, titin, vitronectin, and vitamin D associated proteins. Twenty-four mouse proteins were identified in serum from mice implanted with tissues from both patients [Figure 23]. Proteins that were down-regulated in both patients included: alpha-1 protease inhibitor 2, DOM1, fetuin, haptoglobin, and orosomucoid. Proteins that were up-regulated in both patients include apolipoproteins, complement component C4, contraspin, histidine-rich glycoprotein, kininogen, PZP protein, RIKEN cDNA, vitronectin, and zinc associated protein. Proteins found in both patients but with different expressions included hemopexin, murinoglobulin, serpina1, carboxylesterase, complement component C3, Es1, gelsolin, serine proteinase inhibitor, serum amyloid P, and MHC class III.

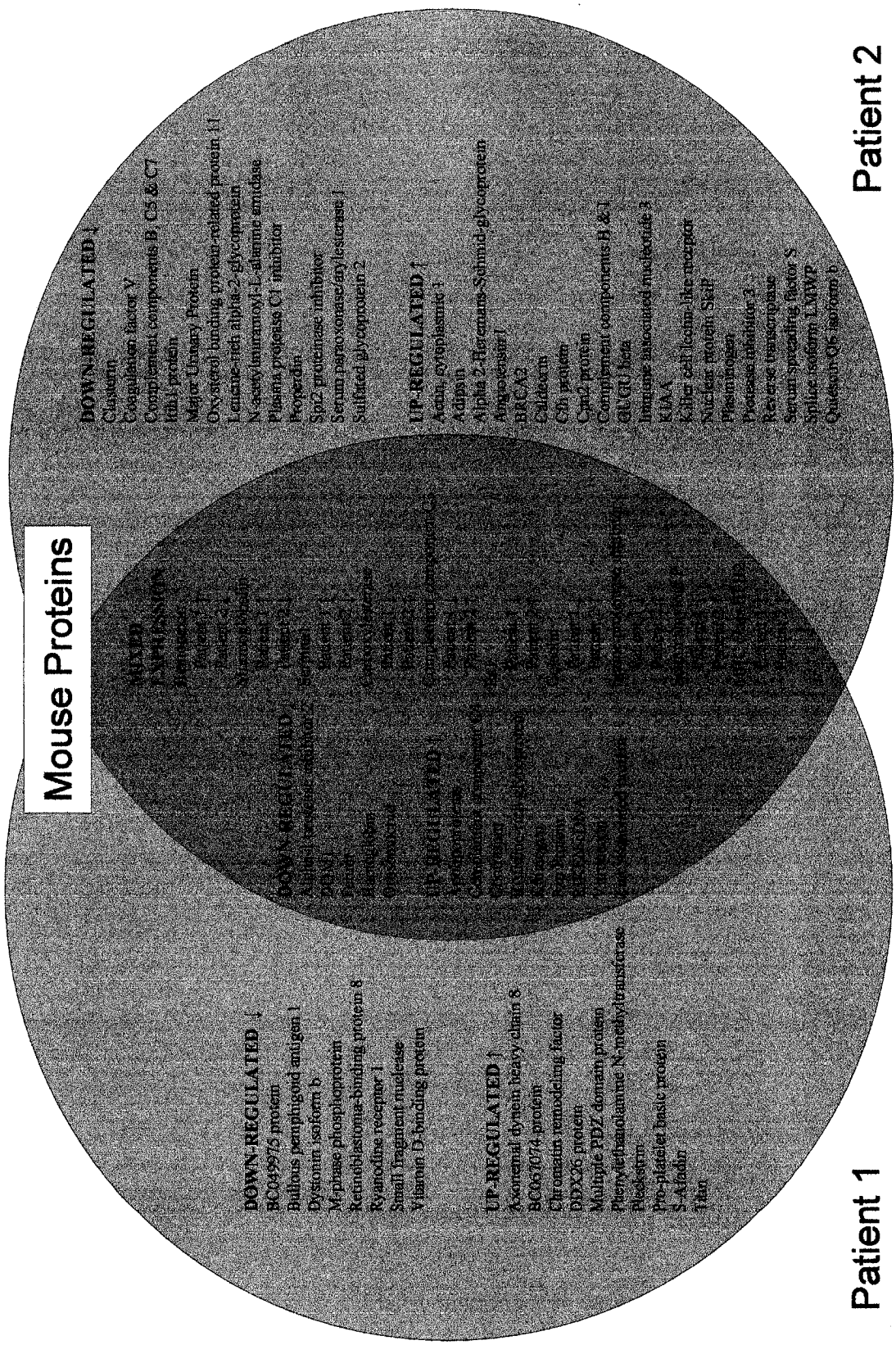

Figure 23. A Venn diagram summary of differentially expressed mouse proteins.

g. Marker validation

To validate the overall proteomic technology used in this study, we selected EGFR, which has been reported to be altered in many cancers including HNSCC. EGFR was found to be up-regulated in serum of mice bearing the cancer tissue from patient 2 [Figure 24].

Figure 25A shows that serum concentration of the extracellular domain of EGFR was significantly higher ($p<0.05$) in serum from mice implanted with tissue from patient 2 compared to serum from mice implanted with tissue from patient 1 or from normal tongue. Interestingly, the levels of circulating serum EGFR from mice bearing oral cancer tissues, correlated with the levels of the extracellular domain of EGFR secreted in conditioned cell culture medium [Figure 25B], as well as with the expression of EGFR in matched cell lines [Figure 25C] and cancer tissues [Figures 25D].

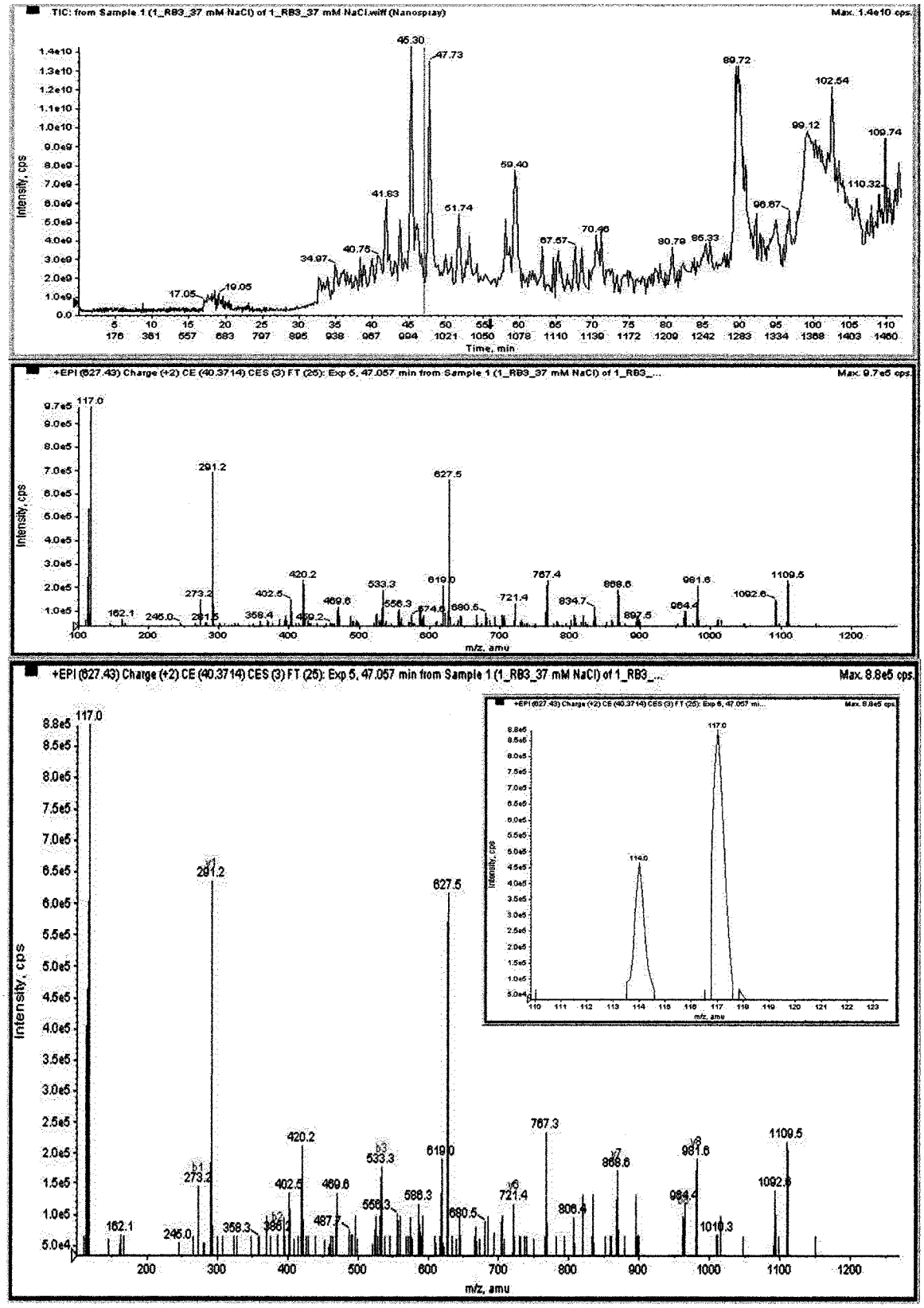

Figure 24. A representative MS result showing EGFR peptide and EGFR up-regulation in the cancer bearing mouse (117: tumor, 114: control).

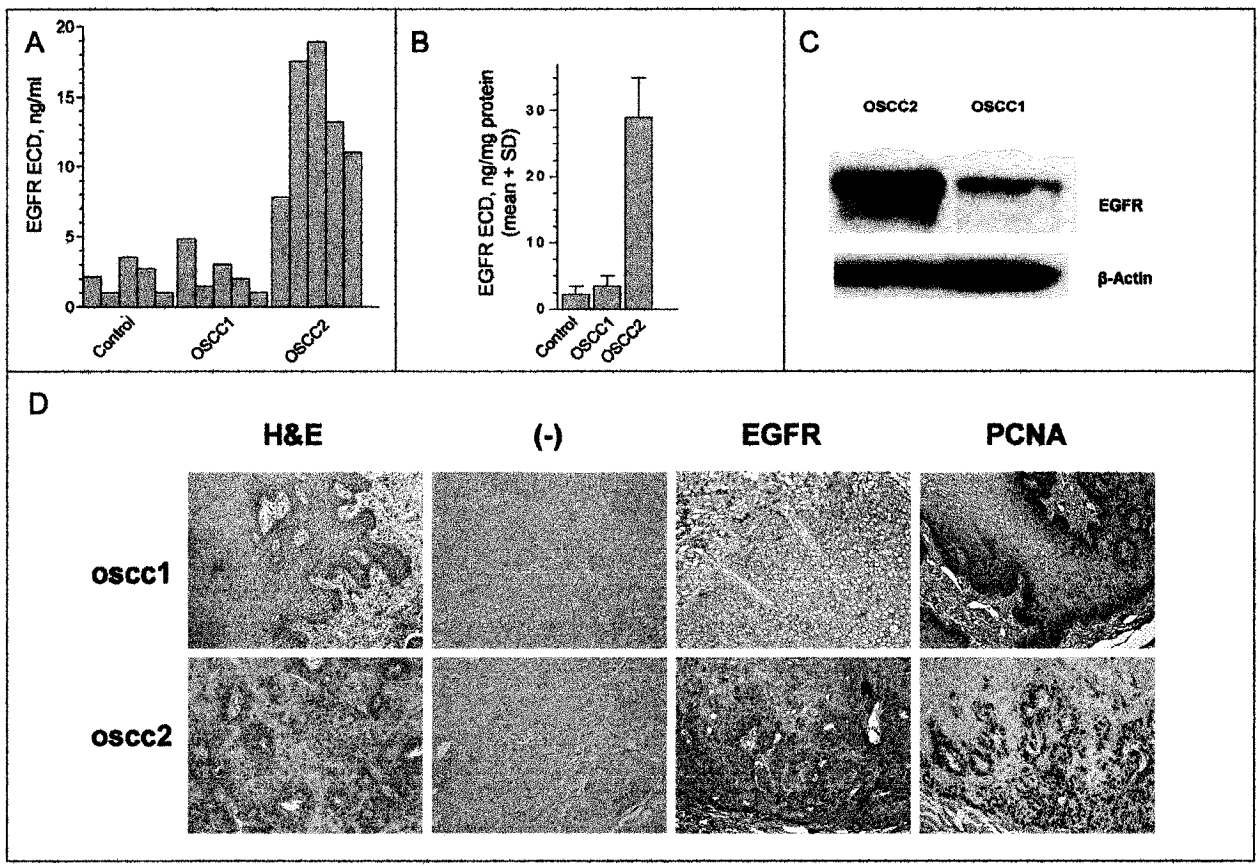

Figure 25. Validation of EGFR. (A) ELISA in the mouse serum and (B) in cell culture conditioned media, (C) Immunoblot analysis, (D) Immunohistochemical staining.

5. Discussion

Several proteomic techniques and instrumentations have been developed to facilitate discovery of cancer biomarkers. In this study, we used DIGE/MALDI and LC/ESI technology on serum from mice implanted orthotopically with human oral cancer or matched normal tongue tissue. We have reported earlier that this orthotopic xenograft model mimics human disease in term of pathology and tumor aggressiveness.[79]

a. Mouse model

Numerous studies have reported on the orthotopic transplantation of oral cancer into immunocompromised mice, with successful take varying between 50 and 100%. As mentioned previously, the RAG2γ(c) knockout mouse has been suggested as the first universal recipient.[87,88] In a previous study, our group was the first to report a successful engraftment of normal human tongue tissue into an immunocompromised mouse.[79] This study confirmed those results, with 80% successful implantation of oral cancer tissues into RAG-2/γ(c) mice, and 30% successful implantation of normal human tongue tissue.

b. Albumin depletion

Most known biomarkers are small and relatively sparse in serum. The study of these low abundant proteins is made difficult because the large and abundant proteins overshadow the small ones (Albumin, IgG and transferrin comprise 80% of the total protein mass). Therefore, when searching for undiscovered biomarkers, it is important to use methods that detect the small serum proteins.[16,60] At the moment, the first step in the

search for new biomarkers is the removal of abundant proteins such as albumin, transferrin, thyroglubulin and the immunoglobulins.[21,22] The Agilent multiple affinity removal system, used in our study, removed albumin, IgG and transferrin from mouse serum simultaneously and reproducibly. Reportedly, it depletes >98-99% of the three targeted proteins.[23] However, we found the depletion to be only about 95%. Moreover, we found that even the remaining 5% of these large proteins had an important overshadowing effect. Albumin remained the predominant protein in both DIGE and MudPIT arms, clearly overshadowing the smaller proteins.

More importantly, some marker proteins are known to exist predominately in a bound phase, where the circulating large carrier proteins act as a reservoir. Thus, albumin depletion can indirectly remove important small molecular weight proteins bound to albumin. Although affinities between the carrier proteins and the small markers are weak, removing and discarding these "contaminating" large proteins can potentially lead to a loss of some markers. Efficient capture of the carrier proteins and specific elution of the biomarkers will yield the greatest amount of diagnostic information.[26] Liotta et al have proposed agents that can be placed into the circulation to act as 'molecular mops' that soak up and amplify the biomarkers that exist.[27] Unfortunately these are not yet available commercially. Thus, the removal of the high abundance proteins remains the most important limiting factor in biomarker search.

c. Proteomic findings

Our results demonstrated that the LC/ESI (MudPIT) was much more sensitive as compared to two-dimensional gel electrophoresis (DIGE/MALDI). The DIGE/MALDI arm revealed only 50 and 75 proteins, and only 4 potential human proteins. The MudPIT arm, on the other hand, identified 38 proteins (14 human) from the first patient and 127 proteins (12 human) from the second patient with the ProQuant search engine. While the Mascot search engine was even more sensitive, identifying 181 (74 human) proteins and 102 (7 human) proteins in the two patients. The quantitative differences were also identified more sensitively by iTRAQ then with Cy dyes. iTRAQ labelling was found to be 90% successful, and picking up small differences between samples, whereas the Cy dyes were only able to pick up differences of two fold or greater.

Using the various proteomic methods and search engines, we identified over one hundred proteins as being differentially expressed between control and cancer-bearing mice. Several candidate proteins were identified as being selectively associated with oral cancer or contributed by the host [summarized as venn diagnrams in Figures 22 & 23]. Detection of several markers previously reported in the literature as potential cancer biomarkers, support that the observed variations in protein expression between mice implanted with tissues from the two patients are likely due to the disease phenotypic and genotypic characteristics.

Epidermal Growth Factor Receptor (EGFR)

Proteomic results with circulating EGFR, which is upregulated in serum from mice bearing oral cancer from patient 2 was confirmed using an ELISA assay for serum, as

well as for conditioned cell culture medium, and correlated with endogenous level of EGFR in matched cancer cell lines and cancer tissues [Figure 25]. This would indicate that high circulating EGFR is likely the result of proteolytically cleaved extracellular binding domain of EGFR from cell surface of cancer cells, as has been extensively reported in a variety of cancers, including oral cancer. Currently, EGFR is being tested as an alternative therapeutic target for head and neck cancer. [92-96]

Clusterin and Cytokeratin 10

Other serum proteins reported earlier include clusterin and cytokeratin 10. Clusterin is a ubiquitous secretory sulfated glycoprotein implicated in cell aggregation, inhibition of complement-mediated cytotoxicity, lipid transport, and anti-apoptotic functions. It has been reported as a potential biomarker for prostate cancer recurrence.[97] Cytokeratin 10 is an intermediate filament protein of the epithelial cells involved in cell motility and cell differentiation; cytokeratins have been identified as squamous cell carcinoma progression prognostic markers.[98-99]

Gelsolin and Vitamin D-binding protein

Furthermore, our study identified several proteins known to play a role in acute phase response. These include gelsolin, and group-specific component protein also known as vitamin D-binding protein. Both of these proteins are abundant components of normal human plasma that bind to G-actin with high affinity, thus preventing actin filament repolymerization, increasing their clearance from circulation, and preventing deleterious effects of long cytoskeletal polymers released during normal or pathological

cell death.[100] Gelsolin has been implicated in EGFR-associated cell motility, and is regulated by osteopontin and integrins, which play a role in cancer metastases. In addition gelsolin regulates cancer-associated signaling by interacting with proteins such as Src and PI-3K.[101,102] Previous studies report gelsolin to be up-regulated in lung cancer cells,[103] and renal cancer cells;[104] and down-regulated in ovarian cancer cells,[105] and in serum of patients with pancreatic cancer.[106] Nevertheless, the correlation between serum gelsolin and the susceptibility to inflammatory reactions versus cancer are still debated. In our study, human gelsolin, as well as mouse gelsolin were both down-regulated in patient 1, and up-regulated in patient 2. The fact that gelsolin was identified in both human and mouse forms suggest that gelsolin is contributed by cancer cells as well as the inflammatory host response. In the absence of validation, we cannot rule out bias due to homology overlaps. The fact that Gelsolin was down-regulated in one patient and up-regulated in the other, leads us to think that this protein is not related to the carcinogenesis of HNSCC, however more tests are needed to confirm this hypothesis. With regards to vitamin D binding proteins, we have reported previously that vitamin D3 analogs have therapeutic benefit for head and neck carcinoma.[107]

Retinoblastoma protein family, Bullous pemphigoid antigen 1, and Titin

Other human and mouse proteins found in serum from mice implanted with oral cancer tissues from both patients included retinoblastoma protein (pRB) family, bullous pemphigoid antigen 1 (BPAG1), titin, and BRCA2. Retinoblastoma protein family was identified with the human retinoblastoma associated factor 600 and the mouse retinoblastoma binding protein 8. These are implicated in regulation of cellular

proliferation and differentiation.[108-109] BPAG1 is a hemidesmosomal transmembrane component important in keratinocyte adhesion, motility, differentiation and proliferation.[110] Titin is a giant elastic protein important in muscle function and development. It has recently been reported as a possible biomarker for lung adenocarcinoma found in plasma and pleural effusions.[108]

BRCA2

An unexpected finding was the elevation of the breast cancer susceptibility protein BRCA2 and other genes never reported in oral cancer. BRCA2 is a large nuclear protein involved in DNA repair, cell cycle control, and genomic stability. Germline mutations in the BRCA2 gene are associated with breast and ovarian cancer.[111] In our study, the human form of this protein was up-regulated in serum from mice implanted with oral cancer tissue from patient 1, and the mouse homolog was up-regulated in serum from mice implanted with oral cancer tissue from patient 2. Further studies are needed to confirm the utility of this marker.

Pregnancy zone protein, Proteinase inhibitors and Vitronectin

Of significance to the host-tumor microenvironment are the 20 mouse proteins identified using the various proteomic methods and search engines. These mouse proteins likely represent a host response to the cancer. In light of the critical role of tumor microenvironment in tumor development and progression, we believe these proteins are potential biomarkers that need to be carefully examined. Among these proteins, the pregnancy zone protein (PZP), a plasma protein thought to be involved in

immunosuppressive effects of T-cells and in the production of interleukin-2 (IL-2), was the most significantly expressed mouse protein, and was found to be clearly up-regulated in our study. PZP levels have been found to be increased in the sera of women with ovarian cancer[112] and breast cancer.[113] Interestingly, all the tissues used in our study were obtained from male human subjects, and all the mice used in the study were male. Proteinase inhibitors, such as serine (or cysteine) proteinase inhibitor, alpha-1 protease inhibitor 2, Serpina1 and DOM1, were found in high abundance in all samples. All four proteins have a significant homology, and the majority were found to be up-regulated. Nevertheless, it is difficult to draw conclusions from these findings, because all the serum samples have been trypsin digested prior to the mass spectral analysis. Vitronectin, which was also up-regulated in our study, is an extracellular matrix protein that alters the strength of cellular adhesions, and has recently been reported to bind proteins important in the carcinogenesis of lingual carcinoma. [114-115]

6. Conclusion

This thesis describes the first proteomic in-vivo model of oral cancer for the identification of low abundant serum biomarkers. Using this novel approach, over 100 proteins were found to be differentially expressed between control and cancer bearing mice. Several candidate proteins were identified as being selectively associated with oral cancer, and were found to be significant ($p<0.05$).

Only time will tell whether any of these identified proteins will be clinically useful as a biomarker for head and neck squamous cell carcinoma. The most sensitive marker would be able to identify the most earliest stages of the disease, as well as objectively assess response to treatment, and predict recurrence of the tumor. This study focused on advanced squamous cell carcinoma cell lines, and it remains to be seen whether the same proteins are expressed in the early forms of the disease. Moreover, the most significant barrier to getting a measurable and reproducible marker for head and neck cancer is the fact that it is a multifactorial disease. Without a single identifiable etiology for this type of cancer, it is very unlikely that one protein will be able to serve as biomarker. Most probably we will need to identify multiple proteins that will be able to serve together as markers of HNSCC. However, the significance of this study is that it provides a reliable new tool for the identification of biomarkers, and it opens the door for subsequent studies.

The proteomic approach reported here addresses a step toward individualized proteomic screening and it identifies several potential cancer and host associated biomarkers for HNSCC. The current challenge with the resultant enormous data sets is the appropriate validation and correlation with the disease in a large number of patients.

Nevertheless, the combined use of proteomic technology and clinically relevant cancer models is a promising approach for the identification of protein markers for head and neck cancer.

7. References

1. Jemal A, Murray T, Samuels A, et al. Cancer Statistics, 2003. CA Cancer J Clin 2003;53(1):5-26
2. Forastiere A, Koch W, Trotti A, et al. Head-and-neck cancer. N Eng J Med 2001;345:1890-1900
3. Reid BC, Winn DM, Morse DE, et al. Head and neck in situ carcinoma: incidence, trends, and survival. Oral Oncology 2000;36:414-420
4. Mamelle G, Pampurik J, Luboinski B, et al. Lymph node prognostic factors in head and neck squamous cell carcinomas. Am J Surg 1994;168:494-498
5. Daftary DK, Murti PR, Bhonsle RB, et al. Risk factors and risk markers for oral cancer in high incidence areas of the world. In: Johnson NW, editor. Oral cancer, vol 2. Cambridge: Cambridge University Press; 1991: p.29-63
6. Timar J, Csuka O, Remenar E, et al. Progresión of head and neck squamous cell cancer. Cancer Metastasis Rev 2005;24:107-127
7. Califano J, van der Riet P, Westra W, Nawroz H, Clayman G, Piantadosi S, Corio R, Lee D, Greenberg B, Koch W, Sidransky D. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res 1996;56:2488-2492
8. Al Moustafa AE, Alaoui-Jamali MA, Batist G, et al. Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells. Oncogene 2002;21:2634-2640
9. Kim MM, Califano JA. Molecular pathology of head-and-neck cancer. Int J Cancer 2004;112:545-553

10. Nathan CAO, Amirghahri N, Rice C, et al. Molecular analysis of surgical margins in head and neck squamous cell carcinoma patients. Laryngoscope 2002;112:2129-2140

11. Franzmann EJ, Reategui EP, Carraway KL, et al. Salivary soluble CD44: a potential molecular marker for head and neck cancer. Cancer Epidemiol Biomarkers Prev 2005;14(3):735-739

12. Tavassoli M, Brunel N, Maher R, Johnson NW, Soussi T. p53 antiobodies in the saliva of patients with squamous cell carcinoma of the oral cavity. Int J Cancer 1998;78:390-391

13. Walther EK, Dahlmann N, Gorgulla HT. Tumor marker in the diagnosis and follow-up of head and neck cancer: role of CEA, CA 19-9, SCC, TK, and dTTPase. Head Neck 1993;15:230-235

14. Liotta LA, Kohn EC. The microenvironment of the tumour-host interface. Nature 2001;411:375-379

15. Juan HF, Chen JH, Hsu WT, Huang SC, Chen ST, Yi-Chung Lin J, Chang YW, Chiang CY, Wen LL, Chan DC, Liu YC, Chen YJ. Identification of tumor-associated plasma biomarkers using proteomic techniques: from mouse to human. Proteomics 2004;4(9):2766-2775

16. Rosenblatt KP, Bryant-Greenwood P, Killian JK, et al. Serum proteomics in cancer diagnosis and management. Annu Rev Med 2004;55:97-112

17. Yeatman TJ. The future of cancer management: translating the genome, transcriptome, and proteome (educational review). Annals of Surgical Oncology 2003;10(1):7-14

18. Verma M, Wright GL Jr, Hanash SM, et al. Proteomic approaches within the NCI

early detection research network for the discovery and identification of cancer biomarkers. Ann NY Acad Sci 2001;945:103-115

19. Petricoin E, Wulfkuhle J, Espina V, Liotta LA. Clinical Proteomics: Revolutionizing disease detection and patient tailoring therapy. J Proteome Res 2004;3:209-217
20. Petricoin EF, Liotta LA. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. Curr Opin Biotechnol 2004;15(1):24-30
21. Kantor AB. Comprehensive phenotyping and biological marker discovery. Dis Marker 2002;18:91-97
22. Rothemund DL, Locke VL, Liew A, et al. Depletion of the highly abundant protein albumin from human plasma using the Gradiflow. Proteomics 2003;3(3):279-287
23. Barrett W. Removal of multiple high-abundant proteins from mouse plasma using the agilent multiple affinity removal system for mouse. Agilent Technologies, publication 5989-1347EN
24. Chertov O, Biragyn A, Kwak LW, et al. Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass spectrometry. Proteomics 2004;4(4):1195-203
25. Granger J, Siddiqui J, Copeland S, et al. Albumin depletion of human plasma also removes low abundance proteins including the cytokines. Proteomics 2005;5(18):4713-4718
26. Mehta AI, Ross S, Lowenthal MS, et al. Biomarker amplification by serum carrier protein binding. Dis Marker 2002;18:1-10
27. Liotta LA, Ferrari M, Petricoin E. Written in blood. Nature 2003;425:905
28. McDonald WH, Yates 3rd JR. Shotgun proteomics and biomarker discovery. Dis

Markers 2002;18:99-105

29. Pandey A, Mann M. Proteomics to study genes and genomes. Nature 2000;405:837-846
30. Wolters DA, Washburn MP, Yates JR. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 2001;73:5683-5690
31. Hanash SM. Biomedical applications of two-dimensional electrophoresis using immobilized pH gradients: current status. Electrophoresis 2000;21:1202-1209
32. Wu CC, MacCoss MJ. Shotgun proteomics: Tools for the analysis of complex biological systems. Current Opinion in Molecular Therapeutics 2002;4(3):242-250
33. Tirumalai RS, Chan KC, Prieto DA, et al. Characterization of the low molecular weight human serum proteome. Mol Cell Proteomics 2003;2:1096-1103
34. Pieper R, Gatlin CL, Makusky AJ, et al. The human serum proteome: Display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. Proteomics 2003;3:1345-1364
35. IUPAC Compendium of Chemical Terminology 2nd Edition 1993;65:825
36. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 1988;60:2299-2301
37. Hillenkamp F, Karas M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. Methods Enzymol 1990;193:280-295
38. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature 2003;422:198-207
39. Hutchens TW, Yip T-T. New desorption strategies for the mass spectrometric

analysis of macromolecules. Rapid Commun Mass Spectrom 1993;7:576-580

40. Merchant M, Weinberger SR. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis 2000;21:1164-1177
41. Diamandis E. Analysis of Serum Proteomic Patterns for Early Cancer Diagnosis: Drawing Attention to Potential Problems. Journal of the National Cancer Institute, 2004;96(5):353-356
42. Fenn JB, Mann M, Meng CK, et al. Electrospray ionization for mass spectrometry of large biomolecules. Science 1989;246:64-71
43. Soltys SG, Le QT, Shi G, et al. The use of plasma surface-enhanced laser desorption/ionization time-of-flight mass spectrometry proteomic patterns for detection of head and neck squamous cell cancers. Clin Cancer Research 2004;10:4806-4812
44. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem 2001;70:437-473
45. Lee DW, Parh SW, Park SY, et al. Effects of p53 or p27 overexression on cyclooxygenase-2 gene expression in head and neck squamous cell carcinoma cell lines. Head Neck 2004;26:706-715
46. Posadas EM, Simpkins F, Liotta LA, et al. Proteomic analysis for the early detection and rational treatment of cancer – realistic hope? Ann Oncol 2005;16:16-22
47. Shaw J, Rowlinson R, Nickson J, et al. Evaluation of saturation labeling two-dimensional difference gel electrophoresis fluorescent dyes. Proteomics 2003;3:1181-1195

48. Tonge R, Shaw J, Middleton B, et al. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics 2001;1:377-396

49. He QY, Chen J, Kung, HF, Yuen APW, Chiu JF. Identification of tumor-associated proteins in oral tongue squamous cell carcinoma by proteomics. Proteomics 2004;4:271-278

50. He QY, Lau GK, Zhou Y, et al. Serum biomarkers of hepatitis B virus infected liver inflammation: a proteomic study. Proteomics 2003;3(5):666-674

51. Alban A, David SO, Bjorkesten L, et al. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 2003;3(1):36-44

52. Gygi SP, Rist B, Gerber SA, et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 1999;17(10):994-999

53. Cagney G, Emili A. De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abyndance tagging. Nat Biotechnol 2002;20(5):163-170

54. Stewart II, Thomson T, Figeys D. ^{18}O labelling: a tool for proteomics. Rapid Commun Mass Spectrom 2001;15(24):2456-2465

55. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 1994;5:976

56. Perkins DN, Pappin DJ, Creasy DM, et al. Probability-based protein identification by searching databases using mass spectrometry data. Electrophoresis 1999;20(18):3551-

3567

57. Creasy DM, Cottrell JS. Error tolerant searching of uninterpreted tandem mass spectrometry data. Proteomics 2002;2:1426-1434

58. Jones MB, Krutzsch H, Shu H, et al. Proteomic analysis and identification of new biomarkers and therapeutic targets for invasive ovarian cancer. Proteomics 2002;2(1):76-84

59. Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to indentify ovarian cancer. Lancet 2002;359:572-577

60. Petricoin EF, Liotta LA. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. Curr Opin Biotechnol 2004;15(1):24-30

61. Petricoin E, Wulfkuhle J, Espina V, Liotta LA. Clinical Proteomics: Revolutionizing disease detection and patient tailoring therapy. J Proteome Res 2004;3:209-217

62. Zhang Z, Bast RC Jr, Yu Y, et al. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Cancer Research 2004;64:5882-5890

63. Koopmann J, Zhang Z, White N, et al. Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. Clinical Cancer Research 2004;10:860-868

64. Li J, White N, Zhang Z, et al. Detection of prostate cancer using serum proteomics pattern in a histologically confirmed population. J Urol 2004;171:1782-1787

65. Rui Z, Jian-Guo J, Yuan-Peng T, et al. Use of serological proteomics methods to find biomarkers associated with breast cancer. Proteomics 2003;3:433-439

66. Yanagisawa K, Shyr Y, Xu BJ, et al. Proteomic patterns of tumour subsets in

nonsmall-cell lung cancer. Lancet 2003;362(9382):433-439

67. Melle C, Kaufmann R, Hommann M, et al. Proteomic profiling in microdissected hepatocellular carcinoma tissue using ProteinChip technology. Int J Oncol 2004;24(4):885-891

68. Shiwa M, Nishimura Y, Wakatabe R, et al. Rapid discovery and identification of a tissue-specific tumor biomarker from 39 human cancer cell lines using the SELDI ProteinChip platform. Biochem Biophys Res Commun 2003;309(1):18-25

69. von Eggeling F, Davies H, Lomas L, et al. Tissue-specific microdissection coupled with ProteinChip array technologies: applications in cancer research. BioTechniques 2000;29:1066-1070

70. Wu W, Tang X, Hu W, Lotan R, Hong WK, Mao L. 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;19:319-326

71. Melle C, Ernst G, Schimmel B, et al. Biomarker discovery and indentification in laser microdissected head and neck squamous cell carcinoma and proteinchip technology, two-dimensional gel electrophoresis, tandem mass spectrometry, and immunohistochemistry. Mol Cell Proteomics 2003;2:443-452

72. 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

73. Chen J, He QY, Yuen APW, et al. Proeomics of buccal squamous cell carcinoma: The involvement of multiple pathways in tumorigenesis. Proteomics 2004;4:2465-2475

74. 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

75. Wadsworth JT, Somers KD, Stack BC, et al. Identification of patients with head and neck cancer using serum protein profiles. Arch Otolaryngol Heand Neck Surg 2004;130:98-104

76. Sidransky D, Irizarry R, Califano JA, et al. Serum protein MALDI profiling to distinguish upper aerodigestive tract cancer patients from control subjects. J Natl Cancer Inst (Bethesda)2003;95:1711-1717

77. Gourin CG, Xia ZS, Han Y, et al. Serum protein profile analysis in patients with head and neck squamous cell carcinoma. Arch Otolaryngol Head Neck Surg 2006;132:390-397

78. Huang X, Wei Y, Li L, et al. Serum proteomics study of the squamous cell carcinoma antigen 1 in tongue cancer. Oral Oncol 2006;42:26-31

79. Balys R, Alaoui-Jamali M, Hier M, et al. Clinically relevant oral cancer model for serum proteomic eavesdropping on the tumour microenvironment. J Otolaryngol 35(3):1-9, 2006

80. Ji ZW, Oku N, Umeda M, Komori T. Establishment of an oral squamous cell carcinoma cell line (NOS-1) exhibiting amplification of the erb B-1 oncogene and point mutation of 53 tumor suppressor gene: its biological characteristics and animal model of local invasion by orthotopic transplantation of the cell line. Oral Oncology 2001;37:386-392

81. Kurokawa H, Katsube K, Podyma k, Ikuta M, Iseki H, Nakajima M, Akashi T,

Omura K, Takagi M, Yanagishita M. Heparanase and tumor invasion patterns in human oral squamous cell carcinoma xenografts. Cancer Sci 2003;94(3):277-285

82. Moriyama M. Development of diffuse invasive (grade 4D) human oral squamous cell carcinoma model in severe combined immunodeficiency mice: microangioarchitectural analysis and immunohistochemical study. Oral Oncology 1999;35:395-400

83. Qiu C, Hong W, Huajun H, Wiu W. A cervical lymph node metastatic model of human tongue carcinoma: serial and orthotopic transplantation of histologically intact patient specimens in nude mice. J Oral Macillofac Surg 2003;61:696-7000

84. Reuther T, Kubler AC, Staff CJ, et al. The RAG 2 Mouse Model for Xenografted Human Oral Squamous Cell Carcinoma. Laboratory Animal Science 2002;41(2):31-35

85. Umeda M, Yokoo S, Komori T, et al. Experimental model of invasion and metastasis by orthotopic transplantation of oral squamous and adenoid cystic carcinomas into the tongue of nude mice. British J of Oral and Maxfac Surg 2001;39:376-380

86. Kawashiri S, Kumagai S, Kojima K, et al. Development of a new invasion and metastasis model of human oral squamous cell carcinoma. Eur J Cancer B Oral Oncol 1995;31(4):216-221

87. Cao X, Shores EW, Hu-Li J, et al. Defective lymphoid development in mice lacking expression of the common cytokine receptor g chain. Immunity 1995;2:223-38

88. Greenberg LH, Slayden OD. Human endometriotic xenografts in immunodeficient RAG-2/γ(c) KO Mouse. American Journal of Obstetrics and Gynecology 2003;190:1788-1796

89. Zeng-Rong N, Paterson J, Alpert L, et al. Elevated DNA repair capacity is associated with intrinsic resistance of lung cancer to chemotherapy. Cancer Res 1995;55:4760-4764

90. Lee CL, Hsiao HH, Lin CW, et al. Strategic shotgun proteomics approach for efficient construction of an expression map of targeted protein families in hepatoma cell lines. Proteomics 2003;3:2472-2486

91. Benlimane N, He Q, Jie S, et al.FAK signaling is critical for ErbB-2/ErbB-3 receptor cooperation for oncogenic transformation and invasion. J Cell Biol 2005;171(3):505-16

92. Ang KK, Berkey BA, Tu X, Zhang et al. Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma. Cancer Res 2002;62:7350-7356

93. Shintani S, Li C, Mihara M, et al. Gefitnib ('Iressa'), and epidermal growth factor receptor tyrosine kinase inhibitor, mediates the inhibition of lymph node metastasis in oral cancer cells. Cancer Lett 2003;201:149-155

94. Thomas SM, Coppelli FM, Wells A, et al. Epidermal growth factor receptor-stimulated activation of phospholipase Cgamma-1 promotes invasion of head and neck squamous cell carcinoma. Cancer Res 2003;63:5629-5635

95. Zhang Q, Thomas SM, Xi S, et al. Src family kinases mediate epidermal growth factor receptor ligand cleavage, proliferation, and invasion of head and neck cancer cells. Cancer Res 2004;64:6166-6173

96. Ford AC, Grandis JR. Targeting epidermal growth factor receptor in head and neck cancer. Head Neck 2003;25(1):67-73

97. Pins MR, Fiadjoe JE, Korley, et al. Clusterin as a possible predictor for biochemical recurrence of prostate cancer following radical prostatectomy with intermediate Gleason scores: a preliminary report. Prostate Cancer Prostatic Dis 2004;7(3):243-8

98. Fillies T, Werkmeister R, Packeisen J, et al. Cytokeratin 8/18 expression indicates a poor prognosis in squamous cell carcinomas of the oral cavity. BMC Cancer 2006;13:6-10

99. Oliveira MC, Silveira EJ, Godoy GP, et al. Immunohistochemical evaluation of intermediate filament proteins in squamous papilloma and oral verrucous carcinoma. Oral Dis 2005;11(5):288-92

100. Lee WM, Galbraith RM. The extracellular actin-scavenger system and actin toxicity. N Engl J Med 1992;326(20):1335-41

101. Chellaiah MA, Biswas RS, Yuen D, et al. Phosphatidylinositol 3,4,5-trisphosphate directs association of Src homology 2-containing signaling proteins with gelsolin. J Biol Chem 2001;276(50):47434-44

102. Bucki R, Georges PC, Espinassous Q, et al. Inactivation of endotoxin by human plasma gelsolin. Biochemistry 2005;44(28):9590-7

103. Yang J, Tan D, Asch HL, et al. Prognostic significance of gelsolin expression level and variability in non-small cell lung cancer. Lung Cancer 2004;46(1):29-42

104. Flanigan R. Using tumor markers to predict the survival of patients with metastatic renal cell carcinoma. Urol Oncol 2006;24(1):81

105. Noske A, Denkert C, Schober H, et al. Loss of Gelsolin expression in human ovarian carcinomas. Eur J Cancer 2005;41(3):461-9

106. Orchekowski R, Hamelick D, Li L, et al. Antibody microarray profiling reveals

individual and combined serum proteins associated with pancreatic cancer. Cancer Res 2005;65(23):11193-202

107. Enepekides DJ, Black MJ, White JH. The independent and combined effects of RAR-, RXR-, and VDR-selective ligands on the growth of squamous cell carcinoma in vitro. J Otolaryngol 1999;28(2):83-9

108. Tyan YC, Wu HY, Su WC, et al. Proteomic analysis of human pleural effusion. Proteomics 2005;5(4):1062-74

109. Wikman H, Kettunen E, et al. Regulation of the G1/S phase of the cell cycle and alterations in the RB pathway in human lung cancer. Expert Rev Anticancer Ther 2006;6(4):515-30

110. Kaneko T, Tamai K, Matsuzaki Y, et al. Interferon-gamma down-regulates expression of the 230-kDa bullous pemphigoid antigen gene (BPAG1) in epidermal keratinocytes via novel chimeric sequences of ISRE and GAS. Exp Dermatol 2006;15(4):308-14

111. Antoniuo A, Pharoah PD, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet 2003; 72(5);1117-30

112. Bhattacharya M, Barlow JJ. Tumor markers for ovarian cancer. Int Adv Surg Oncol 1979;2:155-76

113. Zorina VV, Levchenco VG, Zorina RM, et al. Pregnancy-Associated Plasma Protein-A, Alpha-2-Macroglobulin, Pregnancy Zone Protein and Their Complexes with IgG in Sera of Healthy Non-Pregnant and Pregnant Woman, and Patients with

Breast Cancer. Russ J Immunol 2001;6(1):71-76

114. Thomas GJ, Nystrom ML, Marshall JF. Alphavbeta6 integrin in wound healing and cancer of the oral cavity. J Oral Pathol Med 2006;35(1):1-10

115. Wang J, Guo F, Wei H, et al. Expression of urokinase-type plasminogen activator receptor is correlated with metastases of lingual squamous cell carcinoma. Br J Oral Maxillofac Surg 2005;Dec 12:[Epub ahead of print]

8. Appendices

a. Complete protein tables

Patient 1				
Human Protein Name	**Accession**	**Difference**	**Mascot score**	**Spot #**
None				
Mouse Protein Name	**Accession**	**Difference**	**Mascot score**	**Spot #**
Serum amyloid P	gi|38174334	Downregulated x 2	177	2
Haptoglobin	gi|8850219	Downregulated x 2	174	1
Patient 2				
Human Protein Name	**Accession**	**Difference**	**Mascot score**	**Spot #**
Cytokeratin 10	gi|547749	Upregulated x 2	193	7
Clathrin heavy chain 1	gi|32451593	Upregulated x 2	142	5
CLTC protein	gi|30353925	Upregulated x 2	142	5
S-adenosylmethionine-dependent methyltransferase	gi|22749301	Upregulated x 2	53	29
DNA binding protein	gi|15706422	Downregulated x 2	51	8
KIAA0404 protein	gi|31565581	Upregulated x 2	44	7
Mouse Protein Name	**Accession**	**Difference**	**Mascot score**	**Spot #**
Hemopexin	gi|23956086	Downregulated x 2	1041	8
Complement C3 precursor (HSE-MSF)	gi|1352102	Upregulated x 2	871	53
Serine (or cysteine) proteinase inhibitor	gi|18252782	Upregulated x 2	677	13
Pzp protein	gi|34785996	Upregulated x 2	520	29
Gelsolin	gi|90508	Downregulated x 2	444	6
Alpha-1 protease inhibitor 2	gi|191844	Downregulated x 2	442	30
DOM1	gi|21322147	Upregulated x 2	436	29
Serpina1a protein	gi|15012149	Downregulated x 2	413	31
Major Urinary Protein	gi|13276755	Downregulated x 2	400	51
Haptoglobin	gi|41019125	Downregulated x 2	370	34
Plasminogen	gi|31982113	Upregulated x 2	325	5
Zinc-alpha-2-glycoprotein 1	gi|29838430	Upregulated x 2	209	22
Contraspin	gi|54173	Upregulated x 2	204	13
Contraspin	gi|54173	Upregulated x 2	64	53

Mascot Score (confidence):
>44 ($p<0.05$)

Table 4. DIGE/MASCOT results: human and mouse proteins identified.

Patient 1			
Human Protein Name	**Accession**	**117:114**	**ProtScore**
Unnamed protein	gi|51476390	0.2824	2.21
Gelsolin	gi|4504165	0.7737	2.06
SH3BP2	gi|18605724	1.2401	2.02
HVEC (herpesvirus entry mediator C)	gi|10312085	1.1793	2.02
DDX26	gi|33872145	1.3198	2.01
DKFZp434B105.	gi|7512527	1.0485	2.01
KLK11	gi|37183146	1.1445	2.00
POTE2A	gi|51460457	0.8027	1.72
Neural cell adhesion molecule L1	gi|35010	0.8313	1.40
Chromosome 14 open reading frame 50	gi|46362563	1.0842	1.40
XP_498921	gi|51460967	1.4463	1.40
XP_378988	gi|42662245	1.5646	1.40
Patient 2			
Human Protein Name	**Accession**	**117:114**	**ProtScore**
Cytokeratin 1	P04264	0.8862	4.03
Cytokeratin 10	P13645	0.9850	2.15
Retinoblastoma-associated factor 600	Q5T4S7	0.4724	2.06
Epidermal growth factor receptor	Q68GS6	0.9234	2.01
Antithrombin III precursor 0309	P01008	9.0619	2.15
PDZ domain containing protein 2 & 3	O15018	1.1455	2.01
FLJ90234	Q8NCI4	1.0228	2.00
PABPCP2	Q6NV95	1.8034	2.00
KIAA1821	Q969D8	1.0766	1.52
RPB11a	Q9H1A6	0.7223	1.10
PCNP	Q96CU3	0.6379	1.05
KBTBD7	Q6FI68	1.2016	1.02

117:114 ratio:
- 117: tumor, 114 control
- 117:114 > 1: upregulated
- 117:114 < 1: downregulated

ProtScore (confidence):
- >99 (2.0)
- >95 (1.3)
- >90 (1.00)

Table 5. MudPIT/PROQUANT results: human proteins identified.

Patient 1			
Mouse Protein Name	**Accession**	**117:114**	**ProtScore**
Contraspin	gi|54173	0.8314	7.53
Pzp protein	gi|34785996	1.1465	6.16
Apolipoprotein A-II	gi|7304897	0.7702	6.02
Kininogen	gi|12643495	1.1955	5.78
Hemopexin	gi|23956086	1.3851	5.40
Fetuin	gi|2546995	1.3499	3.41
Serine (or cysteine) proteinase inhibitor	gi|6678087	0.7924	2.60
RIKEN cDNA 1300017J02	gi|18204720	0.9940	2.56
C4 complement	gi|387438	1.5661	2.23
Gelsolin	gi|90508	0.9738	2.07
Apolipoprotein C-III	gi|15421856	1.0083	2.04
Chromatin remodeling factor	gi|13442965	1.1318	2.02
DDX26 protein	gi|33872145	1.3198	2.01
Unnamed protein product	gi|12846768	1.4405	2.00
Pro-platelet basic protein	gi|12963823	1.2074	2.00
Titin	gi|51706225	1.3423	1.60
Histidine-rich glycoprotein	gi|16716461	1.3708	1.44
Unnamed protein product	gi|12846768	0.7710	1.40
Vitronectin	gi|15215172	1.3181	1.40
Phenylethanolamine N-methyltransferase	gi|383548	1.8083	1.40
Zinc finger protein 592 (Zfp-592)	gi|48428710	1.0277	1.32
Patient 2			
Mouse Protein Name	**Accession**	**117:114**	**ProtScore**
Complement component 3	Q80XP1	1.2859	103.18
Pzp protein	Q6PEM2	1.1856	80.93
Murinoglobulin 1	Q80XE6	0.9471	62.90
Fibronectin	P11276	1.2034	56.74
C4a anaphylatoxin (complement C4 precursor)	P01029	1.0362	48.36
Hypothetical protein	Q6P5C8	0.9711	49.70
Hemopexin	Q91X72	0.6940	41.29
Contraspin	Q62257	0.7631	31.05
Liver tumor cDNA, RIKEN	Q8C7G9	1.0778	29.23
Apolipoprotein A-I (ApoA-I)	Q00623	1.0942	25.73
Complement factor H	P06909	0.9585	23.86
PES-N (liver carboxylesterase N precursor)	P23953	1.2331	23.80
(O08677) Splice isoform LMWP	O08677-2	1.4196	20.75
C3/C5 convertase (complement factor B precursor)	P04186	0.7218	20.10
Angiostatin (plasminogen precursor)	P20918	1.0795	16.56
Apolipoprotein A-IV	Q01488	0.5178	16.55
DBP (vitamin D-binding protein precursor)	P21614	0.8562	16.31
Fetuin-A (alpha-2-HS-glycoprotein precursor)	P29699	0.9517	14.68
Histidine-rich glycoprotein	Q9ESB3	1.0452	13.03
Hemolytic complement C5a anaphylatoxin	P06684	0.9521	12.80
Alpha-2-plasmin inhibitor	Q61247	0.9799	11.21
C1 Inh (plasma protease C1 inhibitor precursor)	P97290	0.6874	11.08
DKFZp459F2310	Q5NVH5	0.0229	10.19
Apolipoprotein H	Q01339	1.1182	10.16
Apob protein	Q8CGG8	1.8595	9.86
Spi2 proteinase inhibitor	Q62258	0.9134	9.07
Alpha-albumin (afamin precursor)	O89020	1.0051	8.61

Antithrombin-III	P32261	1.2386	8.57
Serum amyloid P (SAP)	P12246	1.7098	8.13
Gelsolin	Q68FP1	0.7052	8.13
Serum spreading factor S (vitronectin precursor)	P29788	1.0270	7.55
Prothrombin	P19221	0.9632	7.40
AGP1 Orosomucoid1 (alpha-1-acid glycoprotein 1 prec)	Q60590	0.8300	6.84
Sulfated glycoprotein 2 (Clusterin precursor)	Q06890	0.6280	6.68
GUGU beta	Q6YJU1	1.3983	6.68
Apolipoprotein E	P08226	1.1703	4.17
Apolipoprotein C-III	P33622	2.1296	4.00
AGP 2 Orosomucoid 2 (alpha-1-acid glycoprotein 2 precursor)	P07361	0.5844	4.64
Carboxypeptidase N, polypeptide 1	Q91WM9	0.9266	4.00
Angiotensin I (angiotensinogen precursor)	P11859	1. 8828	4.00
N-acetylmuramoyl-L-alanine amidase	Q8VCS0	0.7445	4.00
Itih1 protein	Q8K159	0.8160	3.52
C3B/C4B inactivator (complement factor I precursor)	Q61129	1. 2021	3.48
Plasma glutathione peroxidase (GSHPx-P)	P46412	1. 6102	3.21
Urinary bladder cDNA, RIKEN	Q9D239	1. 0587	3.10
Serum paraoxonase/arylesterase	P52430	1. 175	2.62
(P03953) Splice isoform 2 of P03953	P03953-2	1. 1898	2.57
Leucine-rich alpha-2-glycoprotein	Q91XL1	0.6742	2.15
Zinc-alpha-2-glycoprotein	Q64726	0.7148	2.06
Coagulation factor V	O88783	0.9661	2.02
Kininogen, LMW	Q7M084	1.382	4.00
(Q7TQH0) Splice isoform 2	Q7TQH0-2	0.3670	2.00
C4b-binding protein alpha-chain	Q80SX2	0.3052	2.00
Killer cell lectin-like receptor	Q75XR6	3.0951	2.00
ES cells cDNA, RIKEN	Q9CWS5	0.3285	5.53
Spectrin, non-erythroid alpha chain	P16086	0.0426	1. 54
Oxysterol binding protein-related protein 11	Q8CI95	0.5678	1.53
BRCA2	Q95143	0.7773	1.11
Immune associated nucleotide 3	Q8R379	1.5373	1.10
(O95789) Splice isoform 2 of O95789	O95789-2	0.0582	1.00

117:114 ratio:
- 117: tumor, 114 control
- 117:114 > 1: upregulated
- 117:114 < 1: downregulated

ProtScore (confidence):
- >99 (2.0)
- >95 (1.3)
- >90 (1.00)

Table 6. MudPIT/PROQUANT results: mouse proteins identified.

Patient 1			
Human Protein Name	**Accession**	**117:114**	**Mascot score**
Bullous pemphigoid antigen 1	gi|55960095	0.8462	97
Chromodomain helicase DNA binding protein 3	gi|52630324	1.2667	79
ALR	gi|2358287	2.8571	73
Ataxia/oculomotor apraxia protein 2	gi|38195410	1.0000	73
Dachsous 1	gi|16933557	0.5714	72
BRG1-binding protein ELD/OSA1	gi|18568414	0.9263	72
Titin	gi|407139	1.2491	70
Brain carboxylesterase hBr3	gi|6009628	0.1974	69
Nuance	gi|17016967	1.1581	68
MLEL1	gi|11526793	1.2215	68
CASP8AP2	gi|55958516	1.9351	67
Family with sequence similarity 47, member C	gi|57163103	1.1140	67
Spectrin	gi|4507191	1.0187	66
KIAA	gi|20521808	3.0833	64
Centrosome (cep290)	gi|51890223	1.2258	64
CGI-114	gi|4929697	0.9112	62
HPK/GCK-like kinase HGK	gi|4322936	0.8632	62
Centromeric protein E (CENP-E)	gi|41149911	1.4750	62
C protein	gi|36501	0.9591	61
DEAH box polypeptide	gi|4503297	1.7489	61
DKFZP434C212	gi|51093832	1.8950	60
MYST histone acetyltransferase (monocytic leukemia) 3	gi|5803098	2.8043	60
Actin-binding protein (ABP-278)	gi|3282771	0.5916	60
PSK-1 (type I transmembrane receptor)	gi|6018464	0.8469	60
antigen to monoclonal antibody Ki-67	gi|55958671	0.6447	59
Rough Deal (centromere / kinetochore) protein	gi|7661960	0.7973	57
HLA-B associated transcript-2 (BAT2)	gi|38173707	0.5052	57
Formin binding protein 1 (FNBP1)	gi|41581463	0.8471	55
BRCA2	gi|16116616	2.4870	55
Ryanodine receptor type 1	gi|107631	0.8604	55
Nance-Horan syndrome protein	gi|37789887	0.5079	55
Protocadherin fat 2	gi|7407144	1.1687	54
Chromosome 10 open reading frame 79	gi|51339293	1.4589	54
CH-TOG protein (Colonic and hepatic tumor over-expressed protein)	gi|3121951	0.4203	54
DNA-activated protein kinase	gi|1362789	0.9793	54
Citron	gi|32698688	0.8503	54
WD repeat domain 17 isoform 1	gi|31317311	1.0000	53
Steerin1	gi|27526775	2.6133	53
Spastic ataxia of Charlevoix-Saguenay (sacsin)	gi|55957308	0.1365	52
Synaptic Ras GTPase activating protein 1	gi|56203648	1.0239	52
Thyroid hormone receptor-associated protein complex component TRAP220	gi|3319290	0.6849	52
Hemicentin	gi|31083306	1.1437	52
CAGF28	gi|2565046	0.0659	52
Enverin	gi|20978308	0.8542	52
BTB/POZ zinc finger protein DPZF	gi|13386602	1.9050	51
Patient 2			
Human Protein Name	**Accession**	**117:114**	**Mascot score**
Keratin 10	gi|186629	Not tagged	110

CP protein	gi|47125416	1.4444	73
Fibronectin	gi|51476292	2.0682	72
Gelsolin	gi|90508	1.3158	60
EGFR	gi|757924	1.3462	59
KIAA	gi|50510415	1.7169	51
OTTHUMP00000045643	gi|7263925	2.5862	50

117:114 ratio:
- ☐ 117: tumor, 114 control
- ☐ 117:114 > 1: upregulated
- ☐ 117:114 < 1: downregulated

Mascot Score (confidence):
- ☐ >48 ($p<0.05$)

Table 7. MudPIT/MASCOT results: human proteins identified.

Patient 1			
Mouse Protein Name	**Accession**	**117:114**	**Mascot score**
Contraspin	gi|54173	2.0100	644
Apolipoprotein A-I	gi|6753096	2.7885	628
Pzp protein	gi|34785996	2.8462	548
Murinoglobulin 1	gi|31982171	3.8137	468
Complement C3 precursor (HSE-MSF)	gi|1352102	0.5349	436
Serpina1b protein	gi|15277553	1.6111	433
Alpha-1 protease inhibitor 2	gi|191844	0.9364	410
Serine (or cysteine) proteinase inhibitor	gi|6678085	0.7937	410
DOM1	gi|21322147	0.3850	401
Apolipoprotein A-IV	gi|109575	1.7053	266
Unnamed protein product Apoa4	gi|14789706	1.4307	266
Apolipoprotein E	gi|192005	1.2006	126
Es1 protein	gi|22135640	0.1372	108
Hemopexin	gi|1881768	1.1687	92
Complement 3	gi|192281	0.5371	90
Dystonin isoform b	gi|19882221	0.6711	88
Fetuin	gi|2546995	0.6024	83
Sex-limited protein, complement component 4	gi|90402	0.8782	71
BC049975 protein	gi|51767018	0.0704	69
RIKEN cDNA 4921517L17	gi|51706473	1.0187	69
Ataxia/oculomotor apraxia protein 2	gi|57092031	1.0000	66
Bullous pemphigoid antigen 1 isoform 1	gi|57094564	0.6746	65
S-Afadin	gi|2555013	7.8776	64
Fibronectin 1	gi|46849812	0.5682	64
Small fragment nuclease, Smfn	gi|13097411	0.9191	62
Titin	gi|51706225	1.3067	62
Multiple PDZ domain protein	gi|17225415	1.3182	61
Zinc finger protein 288	gi|57109650	1.9949	61
Axonemal dynein heavy chain 8	gi|14335446	1.7143	56
Pleckstrin	gi|29243976	1.2251	56
Retinoblastoma-binding protein 8 (RBBP-8)	gi|12643796	0.8534	55
MHC class III H2 sex-limited protein	gi|387476	0.8721	54
BC067074 protein	gi|51767918	1.3196	54
Carboxylesterase MH1	gi|14331135	0.1927	51
M-phase phosphoprotein	gi|30144662	0.3211	51
Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor)	gi|134134	0.7929	51
Patient 2			
Mouse Protein Name	**Accession**	**117:114**	**Mascot score**
Pzp protein	gi|34785996	2.2840	1135
Alpha-2 macroglobulin MUG1	gi|109550	1.2533	1058
Complement C3 precursor (HSE-MSF)	gi|1352102	1.9747	977
Contraspin	gi|54173	0.5763	840
Cp protein	gi|38614350	1.7685	547
Es1 protein	gi|22135640	3.6542	508
Serine (or cysteine) proteinase inhibitor	gi|6678083	1.3992	490
Alpha-1 protease inhibitor 2	gi|191844	1.4359	487
DOM1	gi|21322147	2.9328	477
Alpha-1 antitrypsin	gi|309079	1.3934	447
Murinoglubulin 4	gi|47169562	1.2467	266
C4 complement	gi|387438	1.3732	382

Fibronectin	gi|1181242	3.8713	376
Apolipoprotein A-I	gi|2145141	5.2603	361
Apolipoprotein A-IV	gi|91885	0.7369	358
Apoa4	gi|12836356	0.7306	358
MHC class III H2 sex-limited protein	gi|387476	1.2459	352
Fibronectin 1	gi|46849812	0.7723	333
Serpina1d protein	gi|18256880	0.4047	304
Hemopexin	gi|1881768	0.4508	244
Histocompatibility 2, complement component factor B	gi|6996919	1.4322	241
Afamin	gi|21553101	3.0625	240
Kininogen	gi|12643495	2.2182	224
Plasminogen	gi|200403	1.8708	214
Clusterin	gi|7304967	0.5561	206
Carboxylesterase 1	gi|10946842	1.9698	193
Complement component factor h	gi|19072788	1.4223	167
Histidine-rich glycoprotein	gi|11066003	1.9289	158
Complement C3 precursor	gi|544053	1.9548	157
Beta-2 glycoprotein I	gi|1938223	1.5198	154
Serum paraoxonase/arylesterase 1 (PON 1)	gi|1709718	0.7647	148
Apolipoprotein B	gi|63629958	0.1901	146
Macroglobulin alpha2	gi|224053	1.9009	145
Serum amyloid P (SAP)	gi|200924	3.4394	139
Alpha 2 HS-glycoprotein	gi|248764	1.6812	139
Alpha 2-Heremans-Schmid-glycoprotein	gi|2511777	1.8824	132
Orosomucoid 2	gi|6754950	0.7249	129
C4/Slp hybrid	gi|220354	0.4129	126
Vitamin D-binding protein	gi|193446	1.3558	124
antithrombin III	gi|179161	0.7329	121
Complement component 7	gi|63672487	0.4426	117
Unnamed protein product	gi|53814	0.7095	114
Apolipoprotein C-III	gi|15421856	5.0395	110
Fetub protein	gi|17390796	2.4060	110
RIKEN cDNA 1300017J02	gi|18204720	1.0917	104
Alpha-1-acid glycoprotein 2A	gi|109546	1.0880	104
Angiostatin	gi|38000582	1.6638	100
Serum amyloid A 4	gi|6755398	0.9119	87
Vitronectin	gi|202372	1.2876	85
Gelsolin	gi|28916693	0.5831	74
Quiescin Q6 isoform b	gi|12963609	1.2379	72
Adipsin	gi|673431	3.0394	64
Proapolipoprotein	gi|178775	1.8861	61
Coagulation factor V	gi|6679731	1.6754	58
B-factor, properdin	gi|47059181	0.4197	54
Chymotrypsin C (caldecrin)	gi|55585960	1.1661	54
Actin, cytoplasmic 1 (Beta-actin)	gi|60389477	1.6033	53
mKIAA0209 protein	gi|50510415	1.5673	51
Nuclear protein SkiP	gi|57090307	1.0572	51
Liver carboxylesterase 1	gi|62510567	3.4954	50
Reverse transcriptase	gi|1174092	1.3907	50
Cfh protein	gi|20071242	1.7009	50
Protease inhibitor 3	gi|227259	2.1461	49
Cpn2 protein (Carboxypeptidase N subunit 2)	gi|19388017	1.2534	49
BRCA2	gi|1743860	1.9949	47

117:114 ratio:
- ☐ 117: tumor, 114 control
- ☐ 117:114 > 1: upregulated
- ☐ 117:114 < 1: downregulated

Mascot Score (confidence):
- ☐ >51 ($p<0.05$)

Table 8. MudPIT/MASCOT results: mouse proteins identified.

b. Ethics consent and approval

SMBD-Hôpital général juif – Comité de l'éthique de la recherche

SMBD - Jewish General Hospital

Otolaryngology

Principal Investigator:
Dr. Moulay Alaoui-Jamali
Co-Investigators:
Dr. M. Black
Dr. M. Hier
Dr. A. Mlynarek
Dr. R. Balys

Addressograph

Consent form

Proteomics in Oral Squamous Cell Carcinoma

We invite you to take part in a research study because you have an early cancer which will be removed surgically. Through this research, your tissue will help us identify a test to diagnose cancer earlier and improve patient survival.

What you should know about a research study

We give you this consent so that you may read about the purpose, risks and benefits of this research study.

Routine care is based upon the best known treatment and is provided with the main goal of helping the individual patient. The main goal of research studies is to gain knowledge that may help future patients.

- We cannot promise that this research will benefit you.
- You have the right to refuse to take part, or agree to take part now and change your mind later on.
- Whatever you decide, it will not affect your regular care.
- Please review this consent form carefully and ask any questions before you make a decision.
- Your participation is voluntary.

1 - Why is this research being done?
To develop a blood test this will tell if a patient has, or does not have, squamous cancer of the head and neck.

2 - What is the purpose of this study?

Squamous cell cancer of the head and neck is the sixth most common cancer. The most important factor that affects a patient survival is the size and the extent of the tumor at the time of diagnosis. Simply put, the earlier the cancer is found, the better the likelihood of a cure. There are many patients in which the diagnosis is difficult due to the lack of symptoms (the patient doesn't know anything is wrong) and the location of the t
umor (you can't see it). Patients often have to undergo general anaesthetic for an examination and biopsy in the operating room, many of which come back as not being cancer. These patients have a delay in the diagnosis, allowing more time for the tumor to grow before treatment can be started.

The purpose of this research is to develop a blood test which will tell us if a patient has or doesn't have cancer. For some types of cancer these tests already exist; however, none exists for head and neck cancer. This research will help to diagnose this cancer at its earliest stage so we can offer more of our patients a cure.

3 - What will happen to you if you take part in the study?

Before your operation a single blood test will be performed (the same as any blood test you have recently received). This is frozen and stored in the laboratory of Dr. Moulay Alaoui-Jamali for the duration of this study (about two years). It will be used to confirm the results of our study by looking for any newly discovered proteins that are related to oral cancer in your blood sample. If they are present it will confirm that this protein can be used to screen for cancer. This will be the only test that will be done on your blood sample and, after this test is done, your blood sample will be destroyed. The blood samples will be coded so that only Dr. Alaoui-Jamali will know which sample belongs to which patient.

During your operation, we would like to take a small piece of the cancer (about 3mm) and a small piece of normal tissue beside the cancer (also about 3mm) and a blood sample from you. The tissue samples will be transplanted into immunocompromised mice. These mice do not have a functioning immune system and therefore your tissue will grow in the mice as though it were their own. The tumors and normal tissue will be allowed to grow in these mice for a maximum of 8 weeks. At this time a blood test will be taken from the mice and analyzed for the proteins. We hope to identify proteins in the blood of mice which have the cancer transplanted but are not found in the blood of mice which have your normal tissue transplanted.

The tissue transplanted into the mice will also be taken out of the mice and grown in culture dishes. There will be no identifying information or numbering system that could link you to this tissue to ensure confidentiality. The anonymized tissue will be grown and stored in the laboratory of Dr. Alaoui-Jamali indefinitely. Nobody has successfully cultured early oral squamous cell carcinoma. This tissue will be used in future studies to learn more about early cancers and to test new drug therapies for cancer.

4 - What are the possible risks and discomforts?

Because tissue is being taken in your planned surgery, there are no extra risks you are taking. This will not affect your chances of a complete cure from your cancer. The adjacent normal tissue taken is very small (5mm) and you will have no increased discomfort. We will also take a single blood sample just before the surgery starts. This is the same any of the blood tests you have received in preparation for your surgery. This blood test will be used to confirm the results of this study.

5 - What are the possible benefits?
There are no medical benefits to you from your taking part in this study. However, your participation in this study may increase knowledge of head and neck cancers and might help other people in the future.

6 - If you have any questions or problems, whom can you call?
If you have any questions about the research now or later, or if you think you have had a research-related injury, you should call Dr. M Black or Dr. M Hier at the Jewish General Hospital Department of Otolaryngology: 340-8222 ext 5985. If you cannot reach him or her, or if you have questions about your rights as a research participant, you may call the Jewish General Hospital patient representative, Ms. Laurie Berlin, at (514) 340-8222 ext. 5833.

7 - What information will be kept confidential?
We will keep all research records that identify you private to the extent allowed by law. However, someone from the Jewish General Hospital Research Ethics Committee, and Dr. Black and Dr. Hier may inspect and/or copy the records that identify you. Results of the study may be published; however, we will keep your name and other identifying information private. You may be contacted by Dr. Black or Dr. Hier in the future in the event of a major discovery that could affect the lives of future patients. Again, any further involvement at that time will be completely voluntary.

8 – Voluntary Participation
Your participation is voluntary. You may decide not to participate or to withdraw from the study at any time. In the event that you withdraw or your participation in this study is ended, all data collected for the purpose of this study may be used to preserve the scientific integrity of the study. You can decide, at any time, to have any identifying information removed from the study.

10 - What else do you need to know?
There are no costs and no payment for your participation in this study. Your participation is voluntary.
You do not give up any of your legal rights by signing this form. We will give you a copy of this consent form and a copy will be placed in your medical chart.

Would you like to be contacted for future research in head and neck cancer? Please circle:

YES **NO**

CONSENT

Signature:_____Date:_____

Nameof Participant:_____

Consent form administered and explained in person by:

Signature:_____Date:_____

Name and title:_____

Signature of witness: (if required)

Signature: _____

HÔPITAL GÉNÉRAL JUIF
SIR MORTIMER B. DAVIS
JEWISH GENERAL HOSPITAL
HÔPITAL D'ENSEIGNEMENT DE L'UNIVERSITÉ MCGILL
A MCGILL UNIVERSITY TEACHING HOSPITAL

BUREAU D'ÉTHIQUE DE LA RECHERCHE
RESEARCH ETHICS OFFICE

JACK MENDELSON, M.D., DIRECTOR
BUREAU/ROOM G-142
TEL.: (514) 340-7940
FAX: (514) 340-8222 - 2390
E-MAIL: jack.mendelson@mcgill.ca

FRANCA CANTINI, M.Sc.N., RESEARCH ETHICS OFFICER
BUREAU/ROOM A-925
TEL.: (514) 340-8222 #2445
FAX: (514) 340-7951
E-MAIL: fcantini@lab.jgh.mcgill.ca

October 14, 2005

Dr. Moulay Aloui-Jamali
Otolaryngology
SMBD-Jewish General Hospital

SUBJECT: Protocol #04-082 entitled "Proteomics in Oral Squamous Cell Carcinoma"

Dear Dr. Aloui-Jamali,

Thank you for submitting the following documents pertaining to the above-mentioned protocol to the Research Ethics Office for review of your Continuing Review Application:

- Protocol
- English consent form (September 1, 2005)
- Amendment #1 (dated September 15, 2005)

The Research Ethics Committee of the SMBD-Jewish General Hospital (Federalwide Assurance Number: 0796) is designated by the province (MSSS) and follows the published guidelines of the Tri-Council Policy Statement, 1998 (with 2000, 2002 updates), in compliance with the "Plan d'action ministériel en éthique de la recherche et en intégrité scientifique" (MSSS, 1998), the membership requirements for Research Ethics Boards defined in Part C Division 5 of the Food and Drugs Regulations; acts in conformity with standards set forth in the United States Code of Federal Regulations governing human subjects research, and functions in a manner consistent with internationally accepted principles of good clinical practice.

We are pleased to inform you that expedited re-approval for the above-mentioned protocol, as well as the English consent form (September 1, 2005) is granted for a period of **one year**. For quality assurance purposes, you must use the approved REO stamped consent form when obtaining consent by making copies of the enclosed ones.

Please be informed that this study proposal will be presented for corroborative approval at the next meeting of the Committee on November 21, 2005.

Expedited Re-Approval Date: **October 14, 2005**
Expiration date of Expedited Re-Approval: **October 13, 2006**

...2

c. Associated Publications

The data in this thesis have been accepted for publication in 2 original publications:

Mlynarek AM, Balys RL , Jie S, Hier MP, Black MJ, Alaoui-Jamali MA. A Cell Proteomic Approach for the Detection of Secretable Biomarkers of Invasiveness in Oral Squamous Cell Carcinoma. In Press: Archives of Otolaryngology – Head and Neck Surgery

Mlynarek AM, Balys RL, Jie S, Xu Y, Hier MP, Black MJ, DiFalco M, Alaoui-Jamali MA. A Serum Proteomic Approach for the Identification of Serum Biomarkers Contributed by Oral Squamous Cell Carcinoma and Host Tissue Microenvironment. In Press: Molecular and Cell Proteomics