Quantification of Serum Proteins of Metastatic Oral Cancer Patients Using LC-MS/MS and iTRAQ Labeling

Lifeng Zhang, Jiang Jiang, Martha Arellano, Lei Zhang, Xinmin Yan, David T. Wong and Shen Hu*

School of Dentistry and Dental Research Institute, University of California, Los Angeles CA 90095, USA

Abstract: Metastasis is a critical event in oral squamous cell carcinoma (OSCC) progression. In this study, we have performed quantitative analysis of serum proteins from non-metastatic (lymph-node metastasis free) and metastatic OSCC patients using liquid chromatography-tandem mass spectrometry (LC-MS/MS) with iTRAQ labeling (isobaric tagging for relative and absolute quantitation). To eliminate highly abundant proteins, the serum samples were initially separated by SDS-PAGE and only low abundant protein bands were excised for subsequent in-gel tryptic digestion. The resulting peptides were then extracted from each sample gels and labeled with iTRAQ reagent 114 (control), 116 (non-metastatic) and 117 (metastatic), respectively. Afterwards, the labeled samples were combined and subjected to LC-MS/MS analysis using linear ion trap (LIT) MS with pulsed Q collision induced dissociation (PQD). A total of 64 proteins were identified and quantified by this approach. Our study showed that iTRAQ labeling and LIT-MS with PQD is a valuable approach to quantification of serum proteins. We also demonstrated the presence of differentially expressed serum proteins between non-metastatic and metastatic OSCCs that may be further validated as biomarkers for metastatic OSCC. However, in order to comprehensively quantify low abundant serum proteins, a more efficient approach is needed to deplete highly abundant proteins prior to quantitative serum proteome analysis of OSCC.

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is the sixth most common cancer worldwide. Despite the tremendous improvements in surgery, radiotherapy and chemotherapy, the prognosis for patients with OSCC is more or less unchanged for the past 3 decades. This is because oral cancers are often diagnosed at late stage when the disease has metastasized from the primary tumor site [1-3]. Improvement in patient survival requires an increased understanding of tumor metastasis so that aggressive tumors can be detected early in the disease process and targeted therapeutic interventions can be developed. This suggests an imperative need for developing novel biomarkers for predicting metastatic potential of the patients with primary oral cancers. These biomarkers will certainly help to differentiate patients who clinically have no detectable disease but are potential candidates for lymph nodes metastasis and should have prophylactic neck dissection and/or adjuvant radiotherapy. Conversely, such a set of reliable biomarkers would also help avoid unnecessary surgery treatment for those individuals who are, or would be metastasis-free.

Quantitative proteomics using tandem MS with stable isotope labeling strategy represents an emerging technology for global quantification of protein levels in disease or perturbed biological samples. Recently, a variety of stable isotope reagents have been developed for relative quantification in proteomics, including ICAT (isotope-coded affinity tagging), SILAC (Stable isotope labeling with amino acids in cell culture), AQUA (absolute quantification using internal standard peptides), iTRAQ (isotope tagging for relative and absolute quantitation), etc. Most methods enable quantification in the full MS scan, and peptide identification based on subsequent fragmentation (MS/MS) of precursor ions, with the exception of iTRAQ, where both the identification and quantification are performed in the MS/MS scan [4]. The iTRAQ label attaches to the N-terminal amino group of peptides and the epsilon amino group of lysine. The labeled peptides fragment during MS/MS scans to produce the iTRAQ reporter ions (113-119, or 121 m/z). Therefore, quantification of peptides and corresponding proteins from up to 8 samples can be performed simply by comparing the intensities of these iTRAQ reporter ions. However, due to extremely low m/z of these reporter ions, there is limitation of using ion traps for iTRAQ applications under collision induced dissociation (CID) mode. Recently, a new dissociation technique called PQD, has been developed and implemented for linear ion trap mass spectrometers [5]. The PQD technique generates spectra qualitatively similar to CID, but it allows the observation of low m/z fragments that are usually excluded from CID. With the ability to trap and detect lower m/z product ions, PQD can be applied successfully to peptide quantification utilizing iTRAQ tags [6].

Serum/plasma proteomics is a very attractive approach to disease biomarker discovery because testing of biomarkers in blood is simple, safe and minimally invasive [7]. Compared to tissue biopsies, blood samples are easily accessible and therefore a large number of specimens can be enrolled for a clinical proteomic study. This allows adequate statistical power for a robust study design, and true signatures can be unveiled for disease detection. In this study, we have demonstrated quantitative mass spectrometry (MS) based on

*Address correspondence to this autohr at the Dental Research Institue and School of Dentistry, University of California, Los Angeles, CA 90095-1668, USA; Tel: 310-206-8834; E-mail: shu@dentistry.ucla.edu

1875-0397/08

linear ion trap MS and stable isotope labeling technique can reveal differentially expressed serum proteins between non-metastatic and metastatic OSCCs. The newly identified target proteins may be used for further validation on new patient samples to test their prediction value for metastatic OSCC.

MATERIALS AND METHODS

Sample Collection

The OSCC patients were recruited from the University of California-Los Angeles Medical Center and School of Dentistry. All patients received diagnosis of OSCC and had no prior treatment in the form of chemotherapy, radiotherapy, surgery, or alternative medicine. No patients had a history of prior malignancy, immunodeficiency, autoimmune disorders, hepatitis or HIV infection. Healthy control subjects were also recruited and completely matched for gender, age and ethnicity for this study. All subjects signed informed consent and were enrolled with an approved IRB at UCLA.

SDS-PAGE, in-Gel Tryptic Digestion and iTRAQ Labeling

The total protein concentration of 30 serum samples (10 healthy controls, 10 lymph-node metastasis free and 10 metastatic OSCCs) was measured using the 2-D Quant Kit (GE Healthcare, Piscataway, NJ). We then pooled the samples in each group, equally according to total amount of proteins, and precipitated the proteins using cold 10% TCA in acetone. The protein pellets were then dissolved in a buffer containing 7M urea, 2M thiourea, 20mM DTT, 1.2% CHAPS (w/v), 5% glycerol (v/v), 10% isopropanol (v/v), and 0.4% ASB-14 (w/v). Finally, the serum proteins (10 µg, each pooled sample) were separated using 12% NuPAGE gels (Invitrogen, Carlsbad, CA) and stained with Sypro Ruby (Bio-Rad, Fullerton, CA).

Selected protein bands (Fig. **1**) were excised using the ProteomeWorks™ Spot Cutter (Bio-Rad). In-gel tryptic digestion of the proteins was then performed and the resulting peptides were extracted with 0.1% TFA in 50% acetonitrile. The peptides from each pooled sample were combined, vacuum dried, re-dissolved in 10 µL of triethylammonium bicarbonate (TEAB) and labeled with isobaric tags (114, 116, and 117 for control, non-metastatic patients, and metastatic patients, respectively) according to the manufacturer's instruction manual (Applied Biosystems, Foster City, CA). All three labeled samples were finally combined and desalted with the cation-exchange cartridge system (Applied Biosystems).

LC-MS/MS and Database Searching

The combined peptide sample was analyzed using nano-LC (Eksigent Technologies, Dublin, CA) and linear ion trap MS (LTQ XL, Thermo-Fisher Scientific, Waltham, MA). Peptide separation was performed using the PepMap100 C18 column (3 µm, 100A°, 75 µm I.D., 15 cm length, Dionex, Sunnyvale, CA) with a 300-µm I.D. 3.5-mm-long C18 guard column (Dionex). Peptides were eluted using a linear gradient of 5–95% HPLC buffer B (5% H2O, 95% ACN containing 0.1% formic acid) over 60 min, followed by isocratic elution at 95% buffer B for 15 min with a flow rate of 400 nL/min across the capillary column. Electrospray (ESI) was performed at 2.0 kV using a PicoTip nanospray emitter (10-µm I. D., New Objective, Woburn, MA).

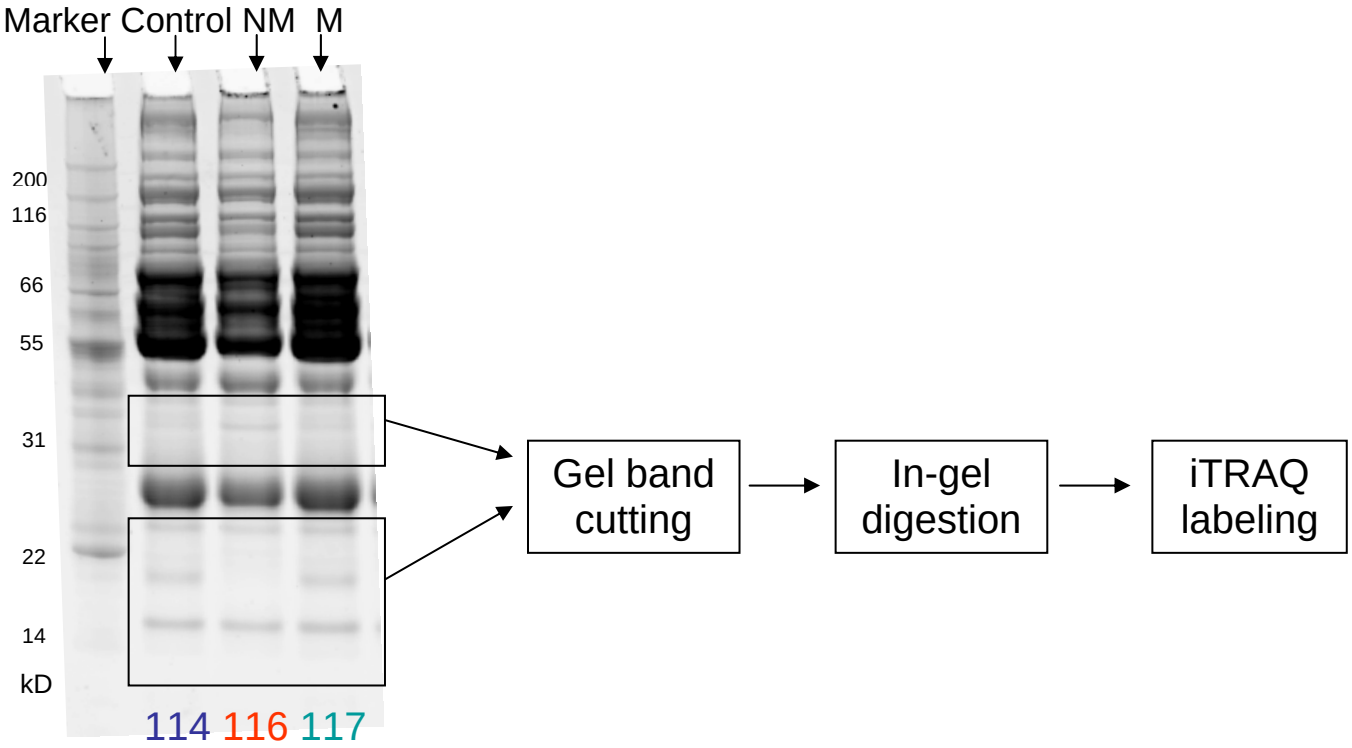


Fig. (1). SDS-PAGE of serum proteins in pooled samples of healthy control, non-metastatic (NM) and metastatic (M) OSCCs. Serum proteins were separated using a 12% NuPAGE gel and then stained with Sypro Ruby fluorescent stain. To avoid the highly abundant proteins, selected protein bands were excised from each lane for subsequent in-gel tryptic digestion. The resulting peptides were extracted and labeled with iTRAQ 114, 116 and 117, respectively.

Peptides were selected for MS/MS using pulsed Q collision induced dissociation (PQD) operating mode with a normalized collision energy setting of 42%. The ion trap was operated in a data-dependent mode, with one MS survey scan (400-1800 m/z) followed by five MS/MS scans for the five most abundant precursor ions in the MS survey scan. The m/z values selected for MS/MS were dynamically excluded for 20 s.

The obtained MS/MS spectra were searched against the International Protein Index database (IPI, Version 3.32) using the Sequest algorithm through the BioWorks 3.3 (Thermo-Fisher Scientific). Search parameters included differential mass shifts for methionine oxidation (+16 Da) and tyrosine (+144 Da), static mass shift (+144Da) at lysine and the N-terminus of all peptides, mass tolerance of +/-2.0 Da for precursor peptides, partial trypsin specificity and up to one internal missed cleavage. Proteins with minimum of two peptide hits were considered as positive identification.

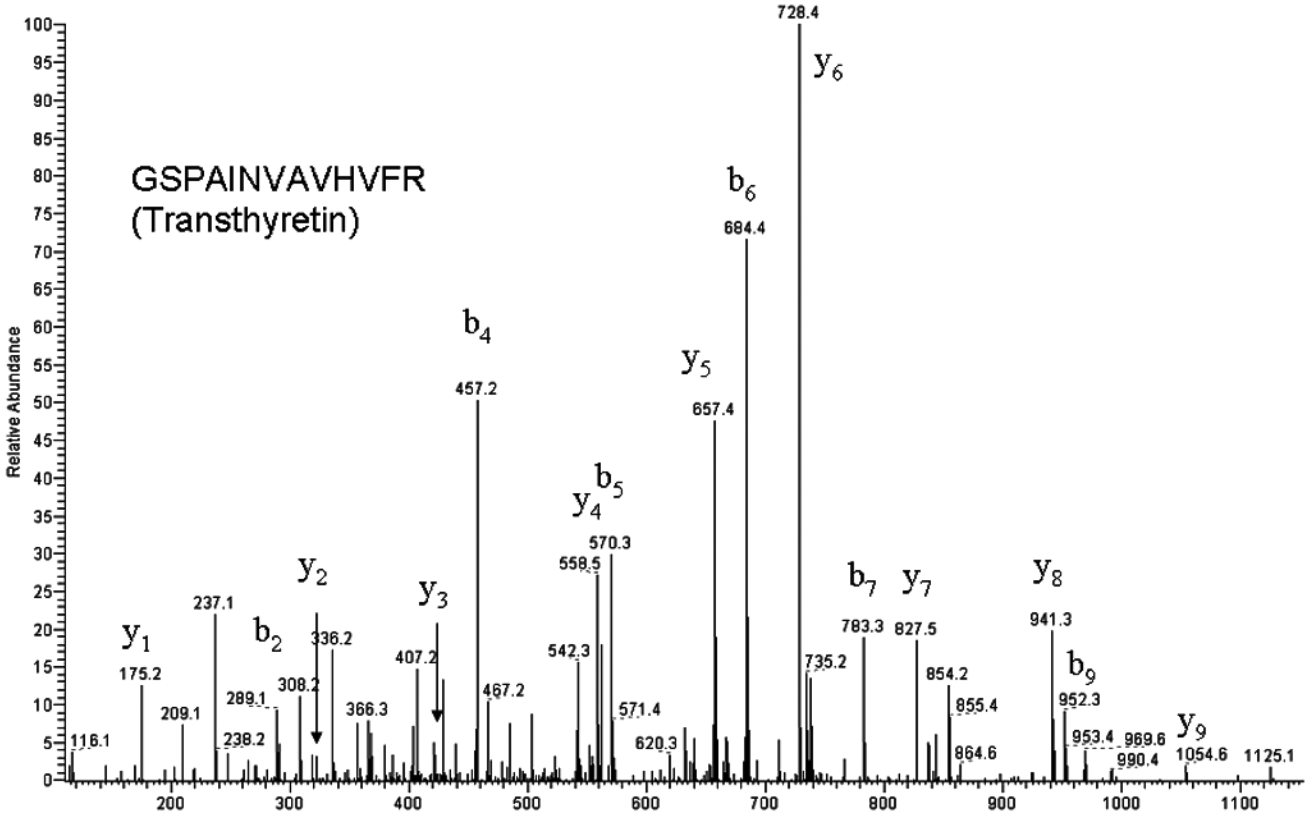
RESULTS AND DISCUSSION

The purpose of this study was to identify differentially expressed serum proteins among healthy control subjects, non-metastatic and metastatic OSCC patients. Since serum samples contain highly abundant proteins such as albumin, transferrin and immnoglobulins, we used SDS-PAGE initially to resolve the serum proteins from three pooled samples (healthy control, metastatic and non-metastatic OSCC) and only excised the low abundant protein bands for subsequent analysis (Fig. **1**). The proteins in gel bands were then digested and the resulting peptides from each gel were combined and labeled with iTRAQ 114, 116 and 117, respectively. The three labeled samples were finally combined and analyzed using capillary LC and linear ion trap MS under the PQD operation mode.

As an example, Fig. (**2**) depicts the PQD spectrum for a tryptic peptide, GSPAINVAVHVFR, originated from transthyretin (TTR). TTR is a thyroid hormone-binding protein and has been associated with lung and colon cancers previously [8-11]. Our study suggested that TTR was overexpressed in non-metastatic OSCC as compared to metastatic OSCC. The extended mass range obtained by PQD relative to CID allowed the detection of iTRAQ reporter ions as well as the fragment ions (b and y ions) (Fig. **2A** & **B**). Most of the abundant peaks in the spectrum were well matched, providing confident identification of the peptide (GSPAINVAVHVFR). The relative levels of the peptide among three serum samples could be directly calculated from the peak intensity of three iTRAQ reporter ions 114, 116 and 117. Fig. (**2C**) indicates the peak height ratios of iTRAQ 116 and 117 versus 114 for the peptide from four repetitive analyses. Significant variation in relative ratios was observed, which suggests that quantification may be biased based on the relative ratio from a single peptide and the use of average ratios for all identified peptides of a protein may be more accurate. Repetitive analysis of the same combined sample may also improve the accuracy of measuring relative protein levels among different samples.

In total, 85 proteins were identified from the three combined samples, however, only 64 (75%) of the proteins showed detectable levels of iTRAQ reporter ions (Table **1**).

A



(Fig. 2) contd….

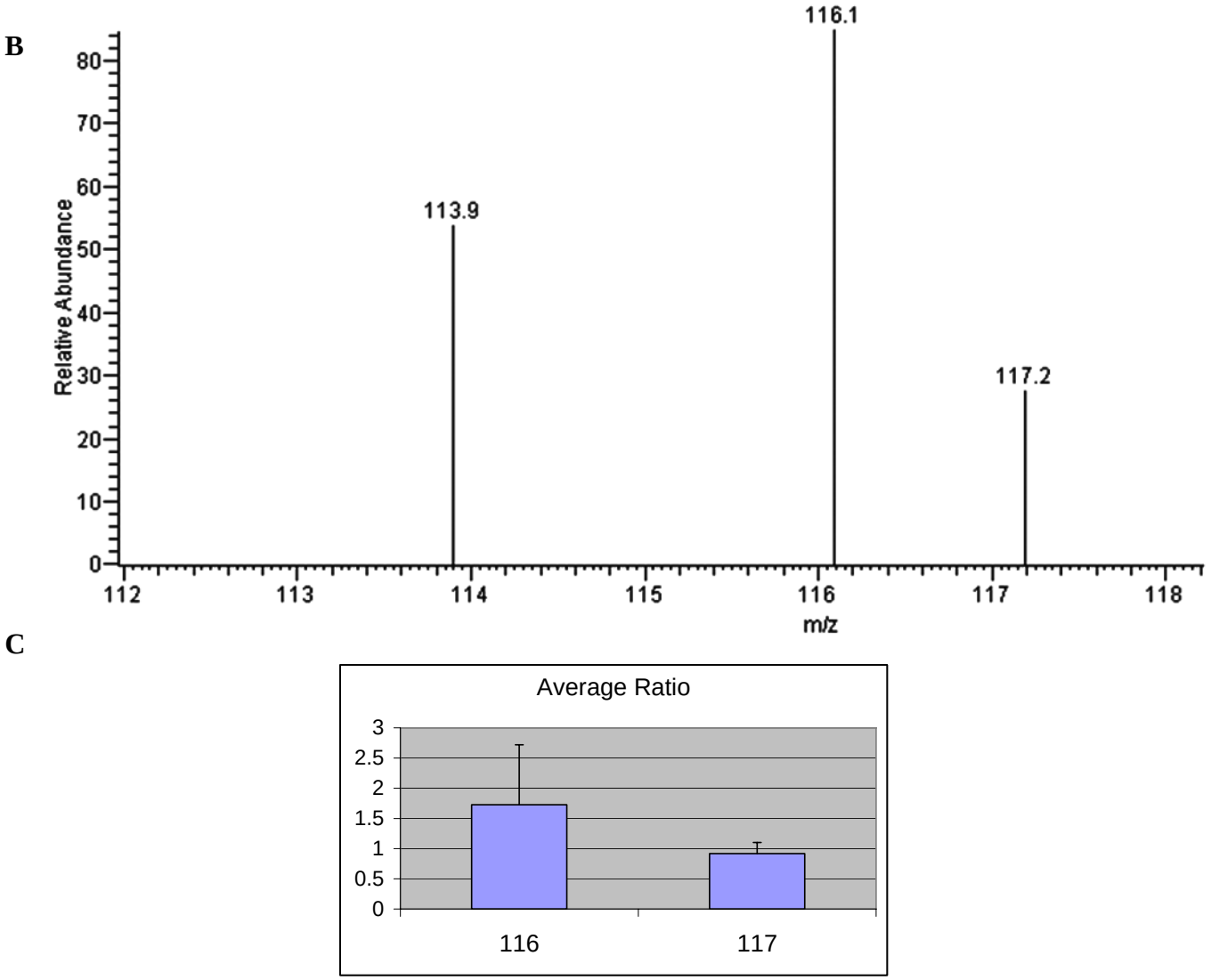


Fig. (2). Tandem MS spectrum for a tryptic peptide, GSPAINVAVHVFR, originating from transthyretin. Most abundant peaks in the MS/MS spectrum were well matched (Fig. **2A**), providing confident identification of the peptide. Fig. (**2B**) depicts the peaks for reporter ions 114, 116 and 117 reflecting the peptide's level among three samples. Fig. (**2C**) indicates the average ratio for reporter ions 116 and 117 versus reporter ion 114 from 4 repetitive analyses.

This may be due to inefficient iTRAQ labeling or inadequate sensitivity of LIT-PQD for iTRAQ reporter ions. More than half of the identified proteins were immunological proteins such as immunoglobulins and complement factors, which are abundant proteins in serum. Many proteins were found beyond the molecular weight range as observed on SDS-PAGE gels, suggesting that the identified proteins may be side chains of the intact proteins. Around 52 proteins were over-expressed in both non-metastatic and metastatic OSCCs compared to healthy controls. However, 27 proteins were found over-expressed and 37 were under-expressed in metastatic OSCC as compared to non-metastatic OSCC. One such protein over-expressed in metastatic *versus* non-metastatic cancer is hemicentin 1, which is a member of the fibulin family of extracellular matrix molecules. Fibulins are secreted glycoproteins functioning to modulate cell morphology, growth, adhesion and motility and therefore proposed to have both tumour suppressive and oncogenic activities [11, 12]. Invasion of cells through breakdown of the basement membrane is a crucial step during development and cancer metastasis. Recently Sherwood *et al.* have identified that hemicentin, matrix metalloproteinase and fat-like protocadherin are down-stream transcriptional targets of FOS-1 that promote cellular invasion [13]. Their study suggests that the accumulation of hemicentin in the basement membrane is required to promote basement membrane removal and invasion. Some other differentially expressed proteins such as HSP protein A14, MUC16 (CA125), ITIH2, and transthyretin were also found involved in human cancers (gastric, ovarian, breast or head and neck) in previous studies, which may be implicated as cancer biomarkers [14-17].

In summary, we have demonstrated that iTRAQ labeling and linear ion trap MS with PQD is applicable to quantification of serum proteins. The inclusion of SDS-PAGE for initial protein separation was able to eliminate some of the high-abundance proteins such as albumin, transferrin and macroglobulin. However, immunoglobulins and associated

Table 1. Quantification of Serum Proteins from Healthy Control (114), Non-Metastatic (116) and Metastatic (117) OSCCs

Protein name	MW	Accession No.	Average ratio (114:116:117)
AOC3 Membrane copper amine oxidase	84621	IPI00004457	0.48 : 0.87 : 1.00
APOD Apolipoprotein D	21275	IPI00006662	2.85 : 0.33 : 1.00
C4B;C4A complement component preproprotein	192732	IPI00418163	0.10 : 1.52 : 1.00
C7 protein	11359	IPI00642632	0.58:1.22:1.00
Complement C3 (Fragment)	187129	IPI00783987	0.47 : 0.88 : 1.00
Complement C4-A	192770	IPI00032258	0.10 : 1.52 : 1.00
Complement C4-B	192774	IPI00654875	0.10 : 1.52 : 1.00
Complement component C9	63173	IPI00022395	0.53 : 0.71 : 1.00
RalBP1 associated Eps domain containing protein 2 variant	57217	IPI00556669	ND:ND:1.00
F7 Factor VII active site mutant immunoconjugate	75553	IPI00382606	1.23:1:57:1.00
Haptoglobin	46723	IPI00641737	0.55:1.08:1.00
Hemicentin 1	613704	IPI00045512	0.59 : ND : 1.00
HP protein	38451	IPI00478493	0.58:125:1.00
HPR Isoform 1 of Haptoglobin-related protein	39007	IPI00477597	0.59:1.32:1.00
HPR Isoform 2 of Haptoglobin-related protein	43055	IPI00607707	0.77:1.20:1.00
HSPA14 protein	26982	IPI00152085	0.60 : 0.65 : 1.00
IAH1 protein (Fragment)	25634	IPI00432867	1.96:3.08:1.00
Ig kappa chain V-I region AG	11992	IPI00387022	0.58 : 0.70 : 1.00
Ig kappa chain V-I region Daudi	14235	IPI00385556	0.63 : 0.51 : 1.00
Ig kappa chain V-I region DEE	11661	IPI00387025	ND : 0.40 : 1.00
Ig kappa chain V-I region Gal	11814	IPI00387027	0.63 : 0.56 : 1.00
Ig lambda chain V-IV region Hil	11517	IPI00382440	ND : 0.58 : 1.00
IGHG1 protein	51395	IPI00448938	0.94:1:15:1.00
IGHG3 protein	57156	IPI00472345	1.47 : 1.49 : 1.00
IGHG4 protein	51986	IPI00550640	0.41:1.63:1.00
IGHM protein	52666	IPI00472610	0.93:1.16:1:00
IGHM protein	52586	IPI00761159	0.19:0.55:1:00
IGHV4-31 CDNA FLJ26266 fis, clone DMC05613	22208	IPI00442911	0.29:0.99:1.00
IGKC Ig kappa chain V-I region Walker	14069	IPI00024134	0.63 : 0.51 : 1.00
IGKC protein	25628	IPI00430808	ND:1.07:1.00
IGKV1-5 protein	26234	IPI00419424	ND:1.07:1.00
IGKV1-5 protein	25747	IPI00430820	0.14:0.99:1.00
IGKV1-5 protein	25853	IPI00478600	ND:1.07:1.00
IGKV2-24 protein	26228	IPI00440577	ND:1.07:1.00
IGL@ protein	24792	IPI00154742	0.59:1.30:1.00
IGL@ protein	24888	IPI00450309	0.62:1.04:1.00
IGL@ protein	25021	IPI00658130	0.46:0.99:1.00

(Table 1) contd….

Protein name	MW	Accession No.	Average ratio (114:116:117)
IGL@ protein	25148	IPI00719373	0.46:1.04:1.00
IGL@ protein	24823	IPI00744476	0.45:0.99:1.00
IGL@ protein	24857	IPI00745660	ND : 0.48 : 1.00
IGL@ protein	24961	IPI00829626	ND : 0.58 : 1.00
IGL@ protein	24855	IPI00829640	ND : 1.62 : 1.00
IGLC2;IGLV2-14;IGLC1;IGLC3 IGLV2-14 protein	24696	IPI00718819	0.47:1.00:1.00
IGLL1 Immunoglobulin lambda-like polypeptide 1	22963	IPI00013438	0.46 : 1.22 : 1.00
IGLV3-25 protein	24867	IPI00550162	0.47:1.00:1.00
IGLV4-3 protein	25977	IPI00382938	0.47:1.04:1.00
ITIH2 Inter-alpha-trypsin inhibitor heavy chain H2	106436	IPI00305461	0.93 : 1.56 : 1.00
MUC16 protein	1518181	IPI00646572	0.19 : 0.68 : 1.00
PTPN13 Isoform 2 of Tyrosine-protein phosphatase non-receptor type 13	256598	IPI00472958	1.32:ND:1.00
Putative uncharacterized protein	26216	IPI00550731	ND:1.07:1.00
Putative uncharacterized protein	20650	IPI00556287	0.27 : 0.79: 1.00
Putative uncharacterized protein	25015	IPI00784519	1.07:1:12:1.00
Putative uncharacterized protein	24712	IPI00785164	ND : 0.48 : 1.00
Hypothetical protein DKFZp686I04196	46061	IPI00399007	1.11:1.76:1.00
Putative uncharacterized protein DKFZp686H20196	52759	IPI00423466	0.94:1.15:1.00
Putative uncharacterized protein DKFZp686I15212	57019	IPI00418153	1.17 : 1.18 : 1.00
Putative uncharacterized protein DKFZp686N02209	52852	IPI00384938	0.94:1.15:1.00
Putative uncharacterized protein DKFZp686O01196	52612	IPI00423463	0.94:1.15:1.00
Putative uncharacterized protein DKFZp686P15220	51724	IPI00645363	0.22:0.80:1.00
Putative uncharacterized protein DKFZp781M0386	25002	IPI00784589	ND : 0.48 : 1.00
TTR 13 kDa protein	13155	IPI00646384	0.72:1.81:1.00
TTR Transthyretin	15887	IPI00022432	1.13:1.73:1.00
Uncharacterized protein ENSP00000366428	10497	IPI00259410	0.63 : 0.56 : 1.00
V2-7 protein	12441	IPI00747752	ND : 0.58 : 1.00

Note: ND, not detected; MW, molecular weight.

side chains were still observed at significantly abundant levels. We are currently testing immunoaffinity columns to deplete the 12 most abundant serum proteins, and expect to quantify a significant larger number of lower abundant proteins. Our study has also identified target proteins that can be used for further validation on new patient samples to test their prediction value for metastatic OSCC. Once robust biomarkers are developed in the future, a simple and valuable clinical tool will be available for the physicians to forecast the metastatic potential of primary oral cancers by monitoring the serum levels of these biomarkers, and therefore to make appropriate treatment decisions.

ACKNOWLEDGEMENTS

This work was supported by the PHS grant R21-CA122806, R03-DE017144 (S. Hu) and R01 DE17170 (D. Wong).

REFERENCES

[1] Brinkman, B.M.; Wong, D.T. *Curr. Opin. Oncol.*, **2006**, *18*, 228.
[2] Hu, S.; Wong, D.T. *Curr. Opin. Mol. Ther.*, **2007**, *9*, 467.
[3] Ye, H.; Wang, A.; Lee, B.; Yu, T.; Shen, S.; Peng, T.; Hu, S.; Crowe, D.; Zhou, X. *Cancer Genomics Proteomics*, **2008**, *5*, 85.
[4] Ross, P.L.; Huang, Y.N.; Marchese, J.N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Pur-

kayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D.J. *Mol. Cell Proteomics*, **2004**, *3*, 1154.
[5] Schwartz, J.C.; Syka, J.P.; Quarmby, S.T. The 53rd ASMS Conference on Mass Spectrometry, San Antonio, Texas, **2005**.
[6] Griffin, T.J.; Xie, H.; Bandhakavi, S.; Popko, J.; Mohan, A.; Carlis, J.V.; Higgins, L. *J. Proteome Res.*, **2007**, *6*, 4200.
[7] Hu, S.; Loo, J.A.; Wong, D.T. *Proteomics*, **2006**, *6*, 6326.
[8] Chatterji, B.; Borlak, J. *Proteomics*, **2007**, *7*, 3980.
[9] Maciel, C.M.; Junqueira, M.; Paschoal, M.E.; Kawamura, M.T.; Duarte, R.L.; Carvalho, Mda G.; Domont, G.B. *J. Exp. Ther. Oncol.*, **2005**, *5*,,31.
[10] Fentz, A.; Spörl, M.; Spangenberg, J.; Joachim, H.; Zornig, C.; Dörner, A.; Layer, P.; Juhl, H.; David, K.A. *Proteomics Clin. Appl.*, **2007**, *1*, 536.
[11] Gallagher, W.M.; Currid, C.A.; Whelan, L.C. *Trends Mol. Med*., **2005**, *11*, 336.
[12] Qing, J.; Maher, V.M.; Tran, H.; Argraves, W.S.; Dunstan, R.W.; McCormick, J.J. *Oncogene*, **1997**, *18*, 2159.
[13] Sherwood, D.R.; Butler, J.A.; Kramer, J.M.; Sternberg, P.W. *Cell*, **2005**, *21*, 951.
[14] Rohde, M.; Daugaard, M.; Jensen, M.H.; Helin, K.; Nylandsted, J.; Jäättelä, M. *Genes Dev.*, **2005**, *19*, 570.
[15] Isomoto, H.; Oka, M.; Yano, Y.; Kanazawa, Y.; Soda, H.; Terada, R.; Yasutake, T.; Nakayama, T.; Shikuwa, S.; Takeshima, F.; Udono, H.; Murata, I.; Ohtsuka, K.; Kohno, S. *Cancer Lett.*, **2003**, *198*, 219.
[16] Yin, B.W.; Lloyd, K.O. *J. Biol. Chem.*, **2001**, *276*, 27371.
[17] Megerian, C.A.; Pilch, B.Z.; Bhan, A.K.; McKenna, M.J. *Laryngoscope*, **1997**, *107*, 216.

Received: March 30, 2008 Revised: May 31, 2008 Accepted: June 2, 2008