

Potential Salivary Proteomic Markers of Oral Squamous Cell Carcinoma

CRESCENZIO GALLO¹, DOMENICO CIAVARELLA¹, ANDREA SANTARELLI²,
ELENA RANIERI³, GIUSEPPE COLELLA⁴, LORENZO LO MUZIO¹ and LUCIO LO RUSSO¹

¹Department of Clinical and Experimental Sciences, University of Foggia, Foggia, Italy;

²Department of Clinic Specialistic and Odontostomatological Sciences,
Polytechnic University of Marche, Ancona, Italy;

³Department of Medical and Surgical Sciences, University of Foggia, Foggia, Italy;

⁴Department of Head and Neck Pathology, Second University of Naples, Naples, Italy

Abstract. *Aim: The aim of the present study was to investigate the diagnostic and prognostic potential of proteomic signatures in saliva of patients with oral squamous cell carcinoma (OSCC). Materials and Methods: Data from SELDI-TOF mass spectrometry of saliva from 45 OSCC patients and 30 healthy controls were analyzed by means of univariate and multivariate statistical approaches, in order to identify proteomic OSCC signatures, reduce dimensionality and build models for discriminating between OSCC and controls, as well as predict nodal status. Results: The saliva proteome presents significant modifications in OSCC patients; some of them seem to be related to nodal involvement, and may be useful for knowledge advancement regarding oral carcinogenesis and definition of diagnostic and prognostic biomarkers. Our attempt to create a predictive model using different artificial neural networks (i.e. feed-forward (FF), radial basis function (RBF), vector quantization (VQ)) demonstrated that such biostatistical tools are powerful but not all network architectures have similar performance. RBF architecture showed the best diagnostic performance (91.89%), whereas FF had the best (77.27%) prognostic accuracy (distinguishing between N⁻ and N⁺). Discussion: Searching for potential biomarkers among differently expressed peptides is a challenge requiring for appropriate strategies that still remain to be defined. A number of factors may potentially impair results, e.g.: (i) a group's definition for adequate comparison; (ii) reduction of data dimensionality and selection of variables to be tested in predictive models; (iii) selection of the biostatistical tool for predictive models.*

Delay in diagnosis can still be considered a major cause of high morbidity and mortality of oral squamous cell carcinoma (OSCC). In fact, the majority (two-thirds) of OSCCs are diagnosed at an advanced stage (20), when prognosis is fairly poor (the overall 5-year relative survival rate of oral and pharyngeal cancer patients is approximately 59%) (17) and current standards-of-care (surgery and/or radiotherapy) often end-up with devastating consequences on the appearance and function of affected organs, causing a marked detriment on the quality of life, even in successfully-treated patients. This has medical and social implications since OSCC is the tenth most common malignancy in men, according to the worldwide cancer cases estimation (16), and its incidence trends are increasing in several countries (24). Thus, improving early diagnosis of oral cancer is pivotal in obtaining a better prognosis and quality of life of affected patients.

At present, early diagnosis of oral cancer is mainly based on clinical oral examination; various tools have been proposed as adjuvants for clinicians in identifying oral cancer or malignant transformation of potentially malignant oral lesions, but they are not completely validated by adequate clinical trials and a number of drawbacks still need to be addressed (26).

Thus, discovering new reliable markers for early OSCC and developing new diagnostic tools for its early and easy detection is a key issue in the field of oral medicine and oral pathology research. Under this point of view new diagnostic matrices (e.g. saliva) (2) are being investigated by means of new molecular technologies (e.g. proteomics).

In fact, the use of saliva as a preferential diagnostic matrix is being increasingly investigated (13, 14, 22) for the following reasons: (i) very easy and non-invasive collection (less risk for health professionals and a better psychological acceptance for the patient); (ii) easy storage and shipping; (iii) saliva contains proteomic markers for both oral and systemic diseases (7, 21); (iv) salivary proteome has good stability

Correspondence to: Dr. Lucio Lo Russo, Via Serro D'Annunzio, 55, 83050 Vallesaccarda (AV), Italy. Tel: +39 0881588085, Fax: +39 0881588081, e-mail: lucio.lorusso@unifg.it

Key Words: Saliva, oral cancer, proteomic, biomarkers, SELDI-TOF.

even at room temperature up to 48 h with the addition of protease inhibitors (23). These features, in the particular case of OSCC, make saliva testing particularly adequate to develop diagnostic tools to be used even in a non-laboratory setting, that may be of utmost importance in improving early diagnosis. Saliva screening by SELDI-TOF/MS (surface-enhanced laser desorption/ionization time of flight mass spectrometry), a high-throughput proteomic technology with sensitivity up to femtomole concentration particularly suitable for biomarkers discovery (15), particularly cancer biomarkers (1, 11, 25), has been recently proposed in a number of experimental studies to investigate changes in proteomic profiles in OSCC compared to other potentially malignant lesions such as oral leukoplakia (10), or in OSCC saliva samples collected either pre- and post-treatment (27).

Materials and Methods

In the current study, we utilised the same dataset as in our previous article (19), with a different and more robust statistical approach. Saliva collected from 45 patients affected by histopathologically-proven OSCC and 30 healthy controls (CTRL), matched for mean age, gender, smoking and alcohol consumption habits, was studied. All patients gave their consent for enrolment in the study and saliva collection.

Patients (33 men and 12 women) had a mean age of 60.3 years (range=26-80 years). The majority of patients had advanced-stage tumors, according to classification of the international union against cancer (UICC) (28): 48.9% had stage IV cancers, 11.1% presented with stage III tumors, while the remaining were stage II (20%) or stage I (20%) tumors. Saliva collection and processing methods, as well as instrument settings are detailed in our previous article (19).

Biostatistical analysis. A number of univariate and multivariate analyses were performed, according to the following specifications: a) Protein profiles of all samples were first analyzed by the Bio-Rad DataManager™ software (Ver 3.5) (Bio-Rad, Hercules, CA, USA). Differences between protein peaks intensities were tested using the Mann-Whitney test (level of significance: $p < 0.05$) in order to identify a list (list 1) of differently expressed mass peaks between OSCC cases and controls. Intensities of list 1 peaks were then exported to Wolfram Mathematica ver.9 software for further analysis: to this end, OSCC dataset was stratified in two groups: N⁺ and N⁻ according to the presence or absence of nodal metastases, respectively. b) Differential expression of list 1 protein peaks intensities was investigated by unpaired t-test in the following groups' comparisons: OSCC vs. CTRL, CTRL vs. N⁻, CTRL vs. N⁺ and N⁻ vs. N⁺; degree of freedom was determined in order to assess if t followed a Student distribution with identical degree of freedom and, then, the *p*-value calculated, assuming a 5% level of significance. For peaks with a statistical significant differential expression (list 2), fold-change (and its logarithm) was also calculated. c) Principal component analysis (PCA), an excellent method for reducing high-dimensional data and identifying outliers samples, was performed in order to identify, for each of the studied groups (*i.e.* CTRL, the subsets of N⁻ and N⁺), the principal and most variable components and their correlation (60% cut-off value was used) with list 1 peaks. Particular attention was given to assess the correlation of principal components with list 2

peaks. d) Peaks included in list 2 (*i.e.* with a significant differential expression) and having at least a 0.6 correlation coefficient with the principal component, were identified and selected as peaks of interest (list 3). e) Intensities of list 1 peaks were used to build a correlation matrix, by calculating the Spearman rank correlation coefficient (ρ) among intensity values. f) Starting from the correlation matrix (point e), a "correlation graph" was built connecting those peaks with a correlation coefficient above a given threshold ($\rho \geq 0.6$). In this graph, the numerical weights on the edges were correlation coefficients, while the nodes represented the peaks. Clustering was performed on peaks included in the graph in order to identify potential biological networks for the studied conditions (*i.e.* CTRL, N⁻ and N⁺). We named "communities" the obtained clusters, with many edges joining vertices of the same community and comparatively few edges joining vertices of different communities. A visual representation through community graphs was constructed. g) Clustering was also performed in order to visualize clusters formed by peaks of interest (list 3). In such a particular case, the same method in point f was used. This was represented by means of correlation graphs. h) Supervised and mixed-architecture artificial neural networks were also used to classify samples according to the following conditions: CTRL vs. OSCC, and N⁻ vs. N⁺. In particular, three network architectures were used: feed-forward (FF), radial basis function (RBF), vector quantization (VQ)(5), and results from the network architecture showing the best performance were taken into account. For each of them, classification attempts were performed using different network parameters and different sets of input variables (Table II) relative to the dataset constituted by 75 records (cases and CTRL). In particular, each set was constituted by intensities of the following peaks: (i) list 1 peaks; (ii) list 2 peaks; (iii) peaks belonging to list 2 with a significant differential expression in all examined conditions (*i.e.* CTRL, N⁻, N⁺), which will be addressed for simplicity as "common peaks"; (iv) list 3 peaks.

Classification between CTRL and OSCC was performed by randomly creating one training set (38 records: 15 CTRL and 23 OSCC) and one validation set (37 records: 14 CTRL and 23 OSCC), whereas classification between N⁻ and N⁺ was attempted by randomly dividing OSCC patients in one training set (23 records: 11 N⁻ and 12 N⁺) and one validation set (22 records: 11 N⁻ and 11 N⁺).

Results

Application of the Mann-Whitney test allowed for recognition of 74 mass peaks whose intensities were significantly different between controls and OSCC ($p < 0.05$). Differential expression analysis showed significance for 22 of those peaks (Table I).

Comparing CTRL with subset N⁻, differential expression analysis identified 15 significant peaks, eleven of which were below 9 kDa. Altered intensity of 16 significant peaks was identified comparing CTRL to N⁺; only four peaks were below 9 kDa. Six differently expressed peaks were also identified comparing N⁻ to N⁺ (3353, 3433, 4784, 6239, 8041, 13841), but, this time, just one was over 9 kDa.

When compared to CTRL, six peaks (*i.e.* 5235, 8041, 11064, 11948, 13287 and 27280) were significantly altered in both N⁻ group and N⁺ group; nine peaks (*i.e.* 3353, 3433, 3482, 4136, 5384, 6165, 6239, 6913, 8086) were selectively

Table I. Fold-change of proteomic peaks showing a significant differential expression.

Peak	Fold-change			
	OSCC vs. CTRL	CTRL vs. N ⁻	CTRL vs. N ⁺	N ⁻ vs. N ⁺
3353	-2.352	3.310	-	-2.057 *
3482	-2.636	3.369	-	-
3433	-	3.186	-	-2.189 *
4038	-2.058	-	2.146 *	-
4136	-	2.522 *	-	-
4784	-	-	-	-2.504
5134	2.468	-	-	-
5154	2.24	-	-	-
5235	3.319	-2.752	-2.378	-
5384	2.959	-2.186	-	-
6165	-	-2.952	-	-
6239	3.393	-4.010	-	2.863
6913	-3.208	3.192	-	-
7668	-	-	-2.644	-
7989	2.693	-	-	-
8041	4.116	-4.472	-2.088	3.162
8086	2.947	-2.557	-	-
10930	3.13*	-	-2.800	-
11002	2.842*	-	-2.661	-
11064	3.583	-2.474	-2.895	-
11164	3.147	-	-2.786	-
11948	-3.487	2.776	3.495	-
13287	-2.549*	2.195 *	2.828*	-
13746	-2.694*	-	3.244	-
13841	-	-	3.834	2.434 *
15531	3.027	-	-2.661	-
16807	-2.503	-	3.040	-
17127	-	-	3.200	-
27280	-2.36*	2.218	2.345	-

-: Differential expression was not significant. *Peaks having at least a 0.6 correlation coefficient with the principal component.

altered in N⁻ group; ten peaks were selectively altered in N⁺ group (*i.e.* 4038, 7668, 10930, 11002, 11164, 13746, 13841, 15531, 16807, 17127).

PCA showed that, considering OSCC and CTRL, 18 components account for 95.18% of total variance; 12 peaks, having at least a 0.6 correlation coefficient with the principal component, were identified; among them, 5 peaks, *i.e.* 10930, 11002, 13287, 13746, 27280 m/z peaks, were identified also by differential expression analysis, and, thus, included in the list of peaks of interest.

PCA showed that, considering N⁻ and CTRL, 18 components account for 96.56% of total variance; 16 peaks, having at least a 0.6 correlation coefficient with the principal component, were identified; among them, two peaks, *i.e.* 4136 and 13287 m/z peaks, were identified also by differential expression analysis, and, thus, included in the list of peaks of interest.

Considering N⁺ and CTRL, 18 components account for 96.56% of total variance; 14 peaks, having at least a 0.6 correlation coefficient with the principal component, were identified; among them, two peaks, *i.e.* 4038 and 13287 m/z peaks, were identified also by differential expression analysis, and, thus, included in the list of peaks of interest.

Considering N⁻ and N⁺, 17 components account for 96.18% of total variance; 19 peaks, having at least a 0.6 correlation coefficient with the principal component, were identified; among them, three peaks, *i.e.* 13841, 3433 and 3353 m/z peaks, were identified also by differential expression analysis.

Community graphs visualizing the clustering of peaks intensities, constructed using method indicated at point f of biostatistical analysis section, are shown in Figure 1. Details of clusters of peaks of interest, according to method indicated at point g of biostatistical analysis section, are shown in Figure 2.

Artificial neural networks showed different performances in classifying cases and controls according to the set of variables (*i.e.* peaks selected) for classification and the network architecture (Table II). The best classification performance, *i.e.* 91.89% of OSCC and CTRL, was obtained using the six common peaks; whereas, considering N⁻ and N⁺, the best classification was achieved by using the three peaks with a significant differential expression and having at least a 0.6 correlation coefficient with the principal component. Thus, a confirmation of the high association of differentially expressed peaks to the corresponding conditions was obtained.

Discussion

Since proteins are the ultimate products of genetic information and the final effectors of many cell functions, proteome profiling may be of great relevance in understanding pathogenetic mechanisms of disease, identifying reliable markers and providing important clues for targeted therapy. This is particularly true in oncology and interesting results have been reported for breast (6), prostate (3) and ovarian cancer (18).

Proteomic investigations regarding head and neck squamous cell carcinomas have shown that classification algorithms based on differently expressed serum proteins can distinguish patients with head and neck cancer from controls with a high degree of sensitivity (range=68%-83.3%) (29, 30) and specificity (range=76%-90%) (8, 30). Encouraging results have been also obtained in classifying pre-treatment and post-treatment serum samples from head and neck cancers patients (4, 9). Salivary proteomics seems even more promising; in fact, potential salivary biomarker peptides with sensitivity up to 90%, and specificity of 83% in detecting OSCC have been proposed (12). Recently, Shintani *et al.*



Figure 1. Clustering of peak intensities in the studied groups. Peaks with significant differential expression are highlighted: in green, peaks common to all groups; in yellow and blue peaks significant for N⁻ and N⁺, respectively.

identified a differently excreted cistatin S-1 fragment in OSCC samples collected either before and after surgical treatment (27), while He and colleagues described a proteomic pattern useful to discriminate OSCC from other pre-cancerous lesions, namely oral leukoplakia (10).

These findings seem to confirm the utility of SELDI screening of saliva samples in order to recognize useful biomarkers for non-invasive diagnosis of oral cancer. However, most studies have investigated OSCC without considering potential differences occurring at different stages of the tumor progression or associated with the occurrence of nodal metastasis.

In the current study we found that salivary proteome of OSCC patients is significantly different from healthy controls. In fact, 74 mass peaks have been identified, by SELDI-TOF analysis, with a significant different intensity in saliva of OSCC when compared to controls (Mann-Whitney test, $p < 0.05$). 22 out of those 74 peaks were characterized by a significant differential expression (Table I).

Now, since oral carcinogenesis, under a theoretical point of view, can be considered a continuum starting from a mutated epithelial cell which progresses towards malignant phenotype and then acquires a more aggressive behaviour (*i.e.* ability to invade and give metastases), we have considered the grouping of our samples in three categories, *i.e.* controls, N⁻ and N⁺, as representative of this continuum, and used it for our analysis.

Considering such a stratification for our samples, the different salivary proteome profiling is even more meaningful. In fact, our analysis found that there were peptides (*i.e.* 5235, 8041, 11064, 11948, 13287 and 27280) whose variation of concentration was differently altered when passing from CTRL to N⁻ and N⁺ groups, as well as peptides whose variation of concentration was selectively significant in N⁻ group (*i.e.* 3353, 3433, 3482, 4136, 5384, 6165, 6239, 6913, 8086) and N⁺ group (*i.e.* 4038, 7668, 10930, 11002, 11164, 13746, 13841, 15531, 16807, 17127).

Some of these selectively expressed peaks (*i.e.* 3433, 4136 and 6165 in N⁻ group; 7668, 13841, 17127 in N⁺ group) were not identified at differential expression analysis when considered N⁻ and N⁺ as a whole group (*i.e.* OSCC group), which seems to suggest that lack of stratification of OSCC patients according to nodal status, which is at present one of the most reliable prognosis predictor, might lead to miss some important markers. This seems to be also confirmed by PCA; in fact, in N⁻ group, the 3433 and 4136 peaks were correlated with the principal component with a $q > 0.6$.

Under a biological point of view, it is possible to speculate that common altered peptides among groups may reflect mechanisms deranged all throughout the continuum leading, or at least participating, to oral carcinogenesis, whereas selectively altered peptides may reflect mechanisms switched on and off in specific steps of the carcinogenesis.

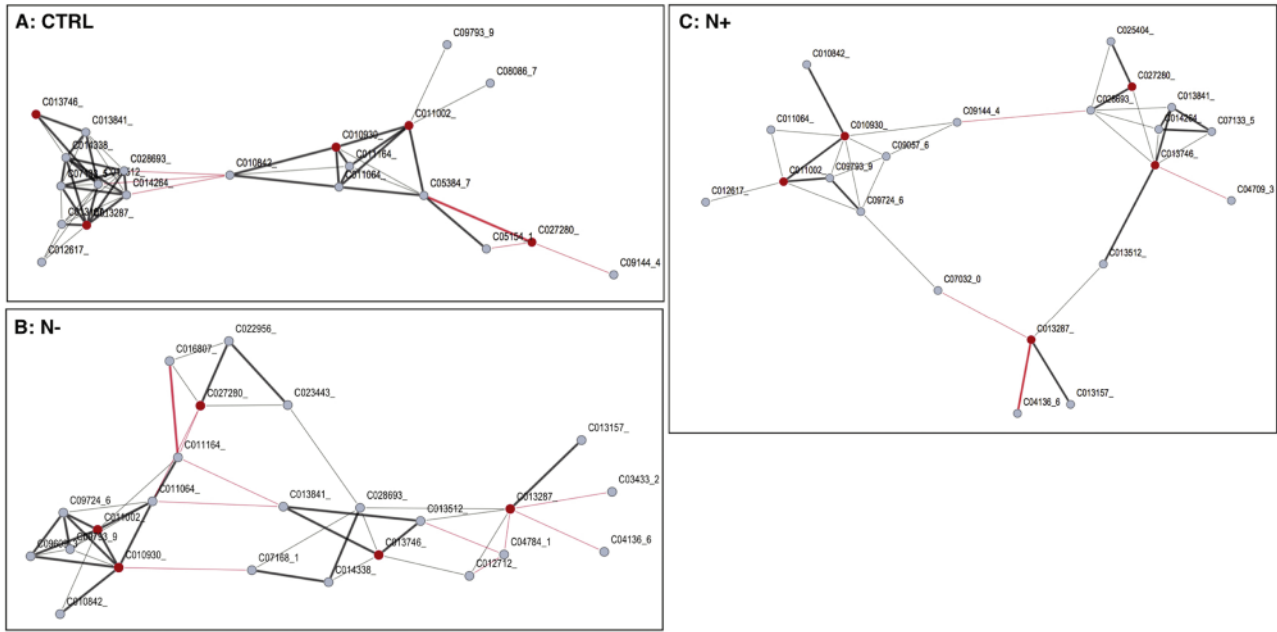


Figure 2. Clustering of peaks of interest. Only peaks directly connected to them and with a correlation coefficient (q) ≥ 0.6 have been considered. Red dot: Peak of interest; red line: negative correlation; gray line: positive correlation; heavy line: $q > 0.7$ (absolute value); light line: $0.6 < q < 0.7$ (absolute value).

Table II. Classification performance of different artificial neural networks: feed forward (FF), radial basis function (RBF), vector quantization (VQ).

	Training set (number of cases)	Validation set (number of cases)	Correct classification of validation set (% cases)					
			Set of variables used for classification					
			List 1 peaks (74 peaks)	Common peaks (6 peaks)	List 2 peaks (22 peaks)	List 2 peaks (6 peaks)	List 3 peaks (5 peaks)	List 3 peaks (3 peaks)
CTRL vs. OSCC								
Network architecture			RBF	RBF	RBF	-	FF	-
Network parameters (input-(hidden)-output)			74-(37)-2	6-(8)-2	22-(66)-2	-	5-(8)-2	-
CTRL	15	14	80%	80%	73.33%	-	60%	-
OSCC	23	23	90.91%	100%	95.45%	-	90.91%	-
Total	38	37	86.47%	91.89%	86.47%	-	78.38%	-
N⁻ vs. N⁺								
Network architecture			VQ	RBF	-	FF		FF
Network parameters (input-(hidden)-output)			74-(6)-2	6-(24)-2	-	6-(12)-2	-	3-(9)-2
N ⁻	11	11	27.27%	54.55%	-	54.55%	-	72.73%
N ⁺	12	11	63.64%	54.55%	-	36.36%	-	81.82%
Total	23	22	45.45%	54.55%	-	45.45%	-	77.27%

List 1 peaks: differently expressed mass peaks between OSCC cases and controls (Mann-Whitney test, $p < 0.05$). List 2 peaks: sub-group of list 1 peaks with a statistical significant differential expression in the specific condition. Common peaks: sub-group of list 2 peaks with a significant differential expression in all examined conditions (*i.e.* CTRL, N⁻, N⁺). List 3 peaks: sub-group of list 2 peaks having at least a 0.6 correlation coefficient with the principal component.

Interestingly, in N^- group the majority (11/15) of differentially expressed peptides maps in a mass range below 9kDa (3353-8086 Da), whereas in N^+ group twelve out of the 16 characterizing peptides fall in the mass range between 10930 and 27,280 Da. Thus, it seems that alteration of smaller molecules participates in early steps of carcinogenesis, whereas heavier molecules are involved in metastasis. This seems to also be confirmed by clustering analysis; in fact, in N^- it is possible to observe (Figure 1) the segregation of a three small molecules group (*i.e.* 3353, 3433, 3482), whose significance remains to be investigated.

Results of a previous study of our group (19) focusing in identifying markers capable to distinguish between early- and late-stage OSCC, have been confirmed by the present analysis. In fact, the 8041 and 6239 m/z peaks showed the highest fold change at differential expression analysis when comparing N^- vs. N^+ (Table I).

Our attempt to create a predictive model by means of different artificial neural networks architectures has clearly demonstrated that such biostatistical tools have great potential but, at the same time, not all network architectures have the same performance in correcting classification between cases and controls, as well as between presence or absence of nodal involvement. In addition, the process of data mining and selection of the variables to work with in the predictive model is of outstanding importance. To this end, we have used different strategies to reduce the dimensionality of our data, and we have tested the different set of variables output by those strategies. Under this point of view, the best diagnostic performance (91.89%) was achieved with a RBF neural network architecture (Table II) using the 6 peaks with a significantly altered differential expression in all our groups (*i.e.* CTRL, N^- , N^+); on the other hand, the best (77.27%) prognostic accuracy (capacity to distinguish between N^- and N^+) was obtained with a FF neural network architecture using a panel of three proteomic peaks with a significant differential expression between N^- and N^+ and a correlation coefficient >0.6 with the principal component.

Overall, our data confirm that saliva proteome presents several significant modifications in OSCC patients; in addition, some of these modifications seem to be related to nodal involvement, thus, theoretically following oral carcinogenesis and offering the opportunity for a better understanding over this phenomenon, as well as for discovering potential biomarkers useful for clinical purposes. Under this view, our results clearly indicate the challenge, at the biostatistical level, in selecting for the most appropriate predictive model, a fact that remains to be further investigated. Nonetheless, it appears that salivary proteomics is really worth further investigation and may lead to valuable results in the struggle against oral cancer.

References

- Engwegen JY, Helgason HH, Cats A, Harris N, Bonfrer JM, Schellens JH and Beijnen JH: Identification of serum proteins discriminating colorectal cancer patients and healthy controls using surface-enhanced laser desorption ionisation-time of flight mass spectrometry. *World J Gastroenterol* 12: 1536-1544, 2006.
- Fabian TK, Fejerdy P and Csermely P: Salivary Genomics, Transcriptomics and Proteomics: The Emerging Concept of the Oral Ecosystem and their Use in the Early Diagnosis of Cancer and other Diseases. *Curr Genomics* 9: 11-21, 2008.
- Fredolini C, Liotta LA and Petricoin EF: Application of proteomic technologies for prostate cancer detection, prognosis, and tailored therapy. *Crit Rev Clin Lab Sci* 47: 125-138, 2010.
- Freed GL, Cazares LH, Fichandler CE, Fuller TW, Sawyer CA, Stack BC, Jr., Schraff S, Semmes OJ, Wadsworth JT and Drake RR: Differential capture of serum proteins for expression profiling and biomarker discovery in pre- and posttreatment head and neck cancer samples. *Laryngoscope* 118: 61-68, 2008.
- Freeman JA and Skapura DM: Algorithms, Applications, and Programming Techniques. USA: Addison-Wesley Publishing Company, 1991.
- Gast MC, Schellens JH and Beijnen JH: Clinical proteomics in breast cancer: a review. *Breast Cancer Res Treat* 116: 17-29, 2009.
- Giusti L, Baldini C, Bazzichi L, Ciregia F, Tonazzini I, Mascia G, Giannaccini G, Bombardieri S and Lucacchini A: Proteome analysis of whole saliva: a new tool for rheumatic diseases--the example of Sjogren's syndrome. *Proteomics* 7: 1634-1643, 2007.
- Gourin CG, Xia ZS, Han Y, French AM, O'Rourke AK, Terris DJ and Adam BL: Serum protein profile analysis in patients with head and neck squamous cell carcinoma. *Arch Otolaryngol Head Neck Surg* 132: 390-397, 2006.
- Gourin CG, Zhi W and Adam BL: Proteomic identification of serum biomarkers for head and neck cancer surveillance. *Laryngoscope* 119: 1291-1302, 2009.
- He H, Sun G, Ping F and Cong Y: A new and preliminary three-dimensional perspective: proteomes of optimization between OSCC and OLK. *Artif Cells Blood Substit Immobil Biotechnol* 39: 26-30, 2011.
- He J, Gornbein J, Shen D, Lu M, Rovai LE, Shau H, Katz J, Whitelegge JP, Faull KF and Chang HR: Detection of breast cancer biomarkers in nipple aspirate fluid by SELDI-TOF and their identification by combined liquid chromatography-tandem mass spectrometry. *Int J Oncol* 30: 145-154, 2007.
- Hu S, Arellano M, Boontheung P, Wang J, Zhou H, Jiang J, Elashoff D, Wei R, Loo JA and Wong DT: Salivary proteomics for oral cancer biomarker discovery. *Clin Cancer Res* 14: 6246-6252, 2008.
- Hu S, Loo JA and Wong DT: Human saliva proteome analysis. *Ann NY Acad Sci* 1098: 323-329, 2007.
- Hu S, Yu T, Xie Y, Yang Y, Li Y, Zhou X, Tsung S, Loo RR, Loo JR and Wong DT: Discovery of oral fluid biomarkers for human oral cancer by mass spectrometry. *Cancer Genomics Proteomics* 4: 55-64, 2007.
- Issaq HJ, Veenstra TD, Conrads TP and Felschow D: The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochem Biophys Res Commun* 292: 587-592, 2002.

- 16 Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D: Global cancer statistics. *CA Cancer J Clin* 61: 69-90, 2011.
- 17 Jemal A, Siegel R, Ward E, Murray T, Xu J and Thun MJ: Cancer statistics, 2007. *CA Cancer J Clin* 57: 43-66, 2007.
- 18 Lee JM and Kohn EC: Proteomics as a guiding tool for more effective personalized therapy. *Ann Oncol* 21(Suppl 7): vii205-vii210, 2010.
- 19 Lo Russo L, Papale M, Perrone D, Ranieri E, Rubini C, Giannatempo G, Santarelli A, Colella G and L. LM: Salivary proteomic signatures of oral squamous cell carcinoma. *European Journal of Inflammation* 10: 61-70, 2012.
- 20 McGurk M, Chan C, Jones J, O'Regan E and Sherriff M: Delay in diagnosis and its effect on outcome in head and neck cancer. *Br J Oral Maxillofac Surg* 43: 281-284, 2005.
- 21 Miller CS, Foley JD, Bailey AL, Campell CL, Humphries RL, Christodoulides N, Floriano PN, Simmons G, Bhagwandin B, Jacobson JW, Redding SW, Ebersole JL and McDevitt JT: Current developments in salivary diagnostics. *Biomark Med* 4: 171-189, 2010.
- 22 Nagler RM: Saliva as a tool for oral cancer diagnosis and prognosis. *Oral Oncol* 45: 1006-1010, 2009.
- 23 Papale M, Pedicillo MC, Di Paolo S, Thatcher BJ, Lo Muzio L, Bufo P, Rocchetti MT, Centra M, Ranieri E and Gesualdo L: Saliva analysis by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF/MS): from sample collection to data analysis. *Clin Chem Lab Med* 46: 89-99, 2008.
- 24 Petersen P: The World Oral Health Report 2003: continuous improvement of oral health in the 21st century--the approach of the WHO Global Oral Health Programme; Community dentistry and oral epidemiology. 3-23, 2003.
- 25 Scarlett CJ, Saxby AJ, Nielsen A, Bell C, Samra JS, Hugh T, Baxter RC and Smith RC: Proteomic profiling of cholangiocarcinoma: diagnostic potential of SELDI-TOF MS in malignant bile duct stricture. *Hepatology* 44: 658-666, 2006.
- 26 Seoane Leston J and Diz Dios P: Diagnostic clinical aids in oral cancer. *Oral Oncol* 46: 418-422, 2010.
- 27 Shintani S, Hamakawa H, Ueyama Y, Hatori M and Toyoshima T: Identification of a truncated cystatin SA-I as a saliva biomarker for oral squamous cell carcinoma using the SELDI ProteinChip platform. *Int J Oral Maxillofac Surg* 39: 68-74, 2010.
- 28 Sobin LH and Wittekind CH: International union against cancer (UICC): TNM classification of malignant tumors. New York: Wiley, 2002.
- 29 Soltys SG, Le QT, Shi G, Tibshirani R, Giaccia AJ and Koong AC: 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 Res* 10: 4806-4812, 2004.
- 30 Wadsworth JT, Somers KD, Stack BC Jr., Cazares L, Malik G, Adam BL, Wright GL Jr. and Semmes OJ: Identification of patients with head and neck cancer using serum protein profiles. *Arch Otolaryngol Head Neck Surg* 130: 98-104, 2004.

Received September 14, 2015

Revised October 2, 2015

Accepted October 20, 2015