CANCER GENOMICS & PROTEOMICS 2: 353-364 (2005)

Tumour-associated Proteins in Oral Squamous Cell Carcinomas by Proteomics

SALAH OSMAN IBRAHIM[1], TALIA MIRON[2], MARIT KROHN[3], ASOKA N. AMARATUNGA[4], SAMAN WARNAKULASAURIYA[5] and ENDRE NORMANN VASSTRAND[6]

[1]*Department of Biomedicine, Section of Biochemistry and Molecular Biology, Jonas Lies Vei 91, 5009 Bergen, University of Bergen, Norway;*
[2]*Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel;*
[3]*Department of Cell Biology, Institute for Cancer Research, University of Oslo, Norway;*
[4]*Division of Oral Surgery, Dental School, University of Peradeniya, Sri Lanka;*
[5]*Guy's, King's and St. Thomas' Dental Institute, King's College, London, U.K.;*
[6]*Faculty of Oral Sciences –Periodontology, Arstadveien 17, 5009 Bergen, University of Bergen, Norway*

Abstract. *Antibody microarrays, two-dimensional electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (2-DE/MALDI-TOF-MS) were used to examine protein changes in 56 oral cancers (OCs)/normal controls (NCs) from Sudanese (41 OCs vs. 31 NCs) and Sri Lankan (15 OCs vs. 15 NCs) patients. Pools of extracted proteins were prepared and used for microarrays/2-DE/MALDI-TOF-MS. From 2-DE, protein spots (differentially- expressed) were cut and identified with peptide mass fingerprinting based on MALDI-TOF-MS, and the proteins were identified by submitting peptide mass profiles to the NCBInr database. By microarrays, 6 and 8 proteins demonstrated significant differences in their abundance values as differentially-expressed in OCs examined from Sudan and Sri Lanka, respectively. For some of the proteins found, like p56dok2 and NEK2, this is the first report in OCs. By MALDI-TOF-MS/2-DE, patterns of OCs/NCs were acquired and tumour-associated proteins, like psoriasin, calgranulin-B and glutathione transferase, were found to be altered in OCs compared to NCs. The proteins found in this work (by two different methods) represent a global protein change specific to OCs from two populations. This might indicates involvement of multiple pathways in the process of tumorgenesis; thus, multiple proteins should be simultaneously targeted in OCs. The finding of few common proteins might suggest involvement of different pathways, which may parallel differences in ethnicity and/or lifestyle.*

Correspondence to: Salah Osman Ibrahim, Department of Biomedicine, Section of Biochemistry and Molecular Biology, University of Bergen, Jonas Lies Vei 91, 5009 Bergen, Norway. Tel: +47-55-586423, Fax: +47-55-586360, e-mail: sosman@biomed.uib.no

Key Words: Antibody microarrays, mass spectrometry, protein profiling, matrix-assisted laser desorption, ionization time-of-flight mass spectrometry

Oral squamous cell carcinoma (OSCC) is a major health problem in developing countries, where a high incidence is found in people addicted to smokeless tobacco (1, 2). Of the 500,000 new cases of OSCCs reported annually, 62% occur in developing countries (1, 2). A high incidence of OSCCs is reported from Sri Lanka and other parts of Southeast Asia (1-3). In Africa, the highest incidence of OSCCs (excluding Asian-Africans) is reported from Sudan (4), being strongly attributed to the extensive use of oral snuff, locally called *toombak*, with over 6.5 million users (5, 6). *Toombak* is dipped and retained between the gums and lip or cheeks or floor of the mouth for prolonged periods, while quid is sometimes retained in the mouth during sleep (5, 6). *Toombak* contains high levels of tobacco-specific nitrosamines exceeding those reported in snuff from Sweden and the U.S.A. (7). In Sri Lanka, OSCC is linked to the habit of betel quid (BQ) chewing (3). The areca nut (*Areca catechu*), the major component of BQ, has been found to contain several alkaloids that give rise to nitrosamines, *e.g.*, 3-(methylnitrosamino) propionaldehyde (8, 9). Genomic and proteomic investigations of oral neoplasms in these populations are meagre.

Systematic analysis of protein profiles, "proteomics", paralleling that of genomics, is one of the recent technologies that is rapidly changing our approach to cancer research (10, 11). It represents a powerful tool to enhance the study of markers that can be useful in the search for tumour biomarkers, in screening, diagnosis, prognosis, monitoring therapeutic treatment and the prediction of disease recurrence (10, 11). At present, research in proteomics is

limited to serial processing of samples using high-resolution two-dimensional electrophoresis (2-DE) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) technologies for partially purified mixtures (10, 11). This approach can rapidly yield structural information on the identity of proteins that can be resolved from complex mixtures (10, 11). Recently, however, protein-based microarray technologies have been developed as screening tools (10, 11). The stability of immobilized proteins to surfaces has facilitated the use of antibody arrays in screening complex mixtures (10, 11). The use of protein microarray technology, together with 2-DE and MALDI-TOF-MS technologies, enables the identification of proteins from complex mixtures (10-12).

In OSCCs, there have been many studies related to the analysis of prognostic factors, the development of chemotherapeutic agents and the search for early detecting molecules, albeit that there are very few molecular markers clinically in use (13). The proteomic approach might lead to the molecular characterization of cellular events associated with progression, signalling and the developmental stages of this pathology (14-16). Proteomic approaches were applied to compare protein changes in OSCCs from Sudan and Sri Lanka, and to search for common biomarkers in OSCCs obtained from patients from the two countries.

Table I. *Patients' clinicopathological parameters in the oral squamous cell carcinomas examined from Sudan and Sri Lanka.*

Clinical feature	Sudan cases No. (%)	Sri Lanka cases No. (%)
Total number of cases	41	15
Gender		
Male, *no.* (%)	25 (61)	11 (73)
Female, *no.* (%)	16 (39)	4 (27)
Median age (yr)	57	59
Tobacco use[a], *no.* (%)	8 (20)	9 (60)
Alcohol use[b], *no.* (%)	8 (20)	9 (60)
Snuff dipping[c] or betel quid, *no.* (%)	13 (32)[c]	13 (87)[d]
Anatomic location, *no.* (%)		
Gingiva	4 (10)	0
Tongue/floor of mouth	15 (37)	2 (13)
Buccal	16 (39)	7 (47)
Lower lip	1 (2)	0
Hard palate/alveolus	5 (12)	6 (40)
Clinical stage, *no.* (%)		
Stage I-II	18 (44)	9 (60)
Stage III-IV	23 (56)	6 (40)
Tumour differentiation, *no.* (%)		
Well	17 (41)	6 (40)
Moderate	16 (39)	2 (13)
Poor	8 (20)	7 (47)

[a]Current or past history of use. [b]Current or past history of alcohol drinking. [c]Data indicates users of the Sudanese snuff (*toombak*). [d]Data indicates betel quid chewers.

Materials and Methods

Patients and tissue specimens. A total of 56 surgical, fresh biopsy specimens of OSCCs (41 from Sudan and 15 from Sri Lanka) with 46 paired normal controls (NCs) (31 from Sudan and 15 from Sri Lanka) were acquired from consecutive patients with previously untreated OSCCs operated on at the Departments of Oral and Maxillofacial Surgery of the Khartoum University Dental Teaching Hospital, Khartoum, Sudan, and the Maxillofacial Surgical Unit of the Dental Teaching Hospital, University of Peradeniya, Sri Lanka. In both Sudan and Sri Lanka, the studies were approved by their corresponding University Medical Ethical Committees. A common protocol was used for sample collection in both countries. NC samples were obtained either from the contralateral site to that of the tumour or from normal mucosa that was at least 4-5 cm away from the cancerous tissue. Following surgery, tissue samples (malignant and normal) were immediately submerged in the tissue storage and RNA stabilization solution, RNA*later*™ (Ambion, USA), shown to be suitable for protein stability and recovery (17), and then stored at –20˚C until protein extraction. The samples were transported to Bergen, Norway, on dry ice.

All tumours were histopathologically confirmed as SCC and were staged by the 1987 UICC staging system. The tumours were histologically graded into highly-, moderately- or poorly-differentiated according to Cawson and Eveson (18). To rule out protein expression alterations because of stromal cell contamination, it was confirmed histopathologically that each tumour specimen used for the protein assay contained ≥70% tumour tissue by analysis of the corresponding haematoxylin and eosin-stained sections. A representative histopathological picture for a pair of matched tissue samples (tumour *versus* normal control) from Sudan and Sri Lanka is shown in Figure 1. For all the OSCC patients, data on the clinicopathological parameters, history of cigarette smoking, alcohol drinking, snuff (*toombak*) dipping and betel quid chewing were available, as shown in Table I.

Protein extraction, antibody microarray and data analysis. For protein extraction, samples were removed from the RNA*later*, placed on a clean surface and ~25 mg of the tissue sample were cut out and washed 3-5 times for 30 min with cold phosphate-buffered saline. Following washing and gentle drying on tissue papers, the samples were homogenized in the presence of ~5 mg of alumina (Sigma, USA), followed by resuspension in non-denaturing extraction and labelling buffer (Clontech, USA). Insoluble tissue fragments were removed by centrifugation at 10,000xg for 30 min at 4˚C. The protein concentration had been assayed with bicinchoninic acid (BCA) using the BCA Protein Assay Reagent kit (Pierce Biotechnology, USA), and the protein samples were stored at –20˚C in aliquots. To improve the yield of intact proteins, a protease inhibitor cocktail for mammalian tissues (Sigma) was used. Following protein extraction, and due to lack of sufficient protein from the majority of the tumours and the normal controls to perform case by case analysis, protein samples from corresponding tumours and normal controls were pooled in equal quantities as biological averaging after grouping of the samples according to the country of origin (OSCCs *versus* NCs) and patients' habit of *toombak* (Sudan cases) use. Accordingly, 4

pools containing equal amounts of extracted total protein from OSCCs and NCs from Sudan and Sri Lanka, respectively, were prepared and used for the array experiment following the manufacturer's protocol. For the 2-DE-MALDI-TOF-MS, 6 pools (4 corresponding to Sudan cases grouped according to *toombak* use, and 2 corresponding to cases from Sri Lanka) containing equal amounts of extracted total protein from OSCCs and NCs were prepared.

Clontech antibody microarrays (printed on standard glass size of 75 x 25 x 1 mm) consisting of 500 highly specific and sensitive individual antibodies against human polypeptides spotted in duplicate (with two slides provided for each experiment), were purchased from BD Clontech Laboratories Inc, USA. The microarray detects a wide variety of proteins (both cytosolic and membrane-bound) representing a broad range of biological functions, including signal transduction, cell cycle regulation, gene transcription and apoptosis. The complete list of the array antibodies, including the Swiss-Port ID numbers of the target antigens, spotted on the array, plus the layout of the microarrays are available at Clontech's web site (http://bioinfo.clontech.com/abinfo/array-list-action.do). For the antibody array analysis, fresh protein pools (adjusted to 1.1 mg/ml), from OSCCs and NCs of cases from Sudan and Sri Lanka, were immediately used for separate labelling with Cy3 and Cy5 (Amersham Biosciences). Unbound dyes were removed by using disposable PD-10 desalting columns (Amersham Bioscience, Denmark). The pool labelled with Cy3/OSCCs was mixed with Cy5/NCs (100 μg each) and 10 μg was added to slide 1. For the reverse colour labelling, a mixture of the pool labelled with Cy5/OSCCs and that with Cy3/NCs was treated in a similar way for slide 2, following the manufacturer's protocol. The slides were washed, dried and scanned at 5-μm resolution with a GenePix 4000B scanner (Axon Instruments, USA), using appropriate gains on the photomultiplier tube to obtain the highest signal intensities without saturation. Cy5-labelled protein samples were scanned at 635 nm and Cy3-labelled proteins were scanned at 532 nm. The scanned images for Cy3 (green) and Cy5 (red) were then overlaid with GenePix Pro 4.0 software (Axon Instruments) and output from the scanner, gridding and spot finding were read. A text file with the signal intensity values (means of the ratios of Cy5/Cy3 for each polypeptide of both slides) and various other statistics for both channels was generated, and the outputs from the GenePix software were analyzed with Clontech software (as recommended by the manufacturer). Replicate ratio values, or ratio of ratios for each spot within each slide were averaged and an internally normalized ratio (INR) was calculated using the Microsoft Excel software program applying the formula "INR=√Ratio 1/Ratio 2", with Ratios 1 and 2 corresponding to slides 1 and 2, respectively. Ratio 1 = NCs-Cy5 relative fluorescent units/OSCCs-Cy3 relative fluorescent units and Ratio 2=OSCCs-Cy5 relative fluorescent units/NCs-Cy3 relative fluorescent units. This calculation normalizes for differences due to labelling efficiency and antibody-antigen binding affinity, greatly enhancing the precision and accuracy of the assay. The average INR was calculated for each pair of antibodies, and duplicate INR values that varied by more than 30% were discarded. INR values >1.7 or <0.6 were considered as true valid changes in protein abundance for this study. These thresholds were determined by taking the global INR value for each experiment (*i.e.*, calculated average INR of all spots) and multiplying by 1.7 and 0.6 to get the specific threshold INR values for each subsequent microarray analysis. This analysis was considered significant within the Excel software developed specifically for the antibody array by the manufacturer, and these INR values were selected to indicate, with high probability, that there was a true difference in protein abundance, as suggested by the manufacturer.

2-DE and MALDI-TOF-MS. For the 2-DE, 6 pools [4 for cases from Sudan (after grouping according to *toombak* use) and 2 for cases from Sri Lanka] containing equal amounts of extracted total proteins were used. 2-DE was performed essentially as described by Rabilloud *et al.* (19). Protein samples were precipitated with cold 10% trichloroacetic acid (TCA) and the pellets were collected by centrifugation (10,000xg for 30 min at 4˚C), followed by washing with cold 5% TCA, cold acetone and drying by air. After solubilization of the protein samples in isoelectric focusing (IEF) buffer [containing 7M urea, 2M thiourea, 4% w/v CHAPS, 0.5% v/v Triton X-100, 20 mM DTT, bromophenol blue and 0.5% v/v carrier ampholytes (Amersham-Pharmacia Biotech, Sweden)], immobilized pH gradient (IPG) strips (18 cm, pH 3-10 non-linear, Amersham-Pharmacia Biotech) were rehydrated by the protein samples (overnight incubation at room temperature) in the Immobiline Drystrip Reswelling Tray (Amersham Pharmacia Biotech). The first-dimension IEF was carried out on a Multiphor II system at 20˚C using the Pharmacia EPS 3500 XL power supply in gradient mode, as described by the manufacturer (Amersham Pharmacia Biotech). Prior to the second dimension, the IPG strips were equilibrated for 15 min with the SDS-PAGE sample buffer [containing 6 M urea, 30% v/v glycerol, 2% SDS, 16 mM (2.5 mg/ml) DTT in 50 mM Tris-HCl, pH 8.8], followed by an alkylation step carried out by 15-min incubation in the presence of 245 mM (45 mg/ml) iodoacetamide in the sample buffer without DTT. The second dimensional SDS-PAGE was carried out in the Protean II XI (Bio-Rad, USA) apparatus at 20˚C using a 12.5% polyacrylamide gel, as described earlier (20, 21). In all cases, tumours and normal controls were run in pairs and electrophoresis was performed 2-3 times for each pair of samples to ensure reproducibility. Proteins were visualized by the SYPRO® Ruby protein staining kit (Bio-Rad), following the manufacturer's recommendations. The stained gels were directly scanned using a Fuji FLA-2000 phosphoimager, and image acquisition was done by Image reader and Image Gaugean software. Spot detection, quantification and matching were performed with Image Master 2D Platinum Version 5.0 software, according to the manufacturer's directions (Amersham Biosciences). Accordingly, gel images from the OSCCs and NCs were compared to each other in pairs and the normalized volume differences (% vol) were statistically calculated for each sample. Spots that were significantly different were selected for further analysis by MALDI-TOF-MS. From the SYPRO® Ruby-stained 2-DE, protein spots were excised (as duplicate/or triplicate) from tumours/or normal control 2-DE gels from Sudanese and Sri Lankan cases under stringent conditions (*i.e.* ionizing air, gloves) to avoid contamination (in particular with keratins), transferred to 1.5 ml Eppendorf tubes and submitted to the Proteomic Unit of the University of Bergen (PROBE), Norway, at (http://www.probe.uib.no/), for peptide mass fingerprinting (PMF) using standard protocols. The spots were cut, washed twice with 50% acetonitrile in 25 mM ammonium bicarbonate for 15 min, followed by drying in a speedovac for 20 min. The dehydrated gels were rehydrated for 30 min on ice in 30 μl of 1.25 ng/μl trypsin (sequence grade modified, Promega, USA) dissolved in 50 mM ammonium bicarbonate and then incubated at 37˚C overnight. Following centrifugation, acidification and

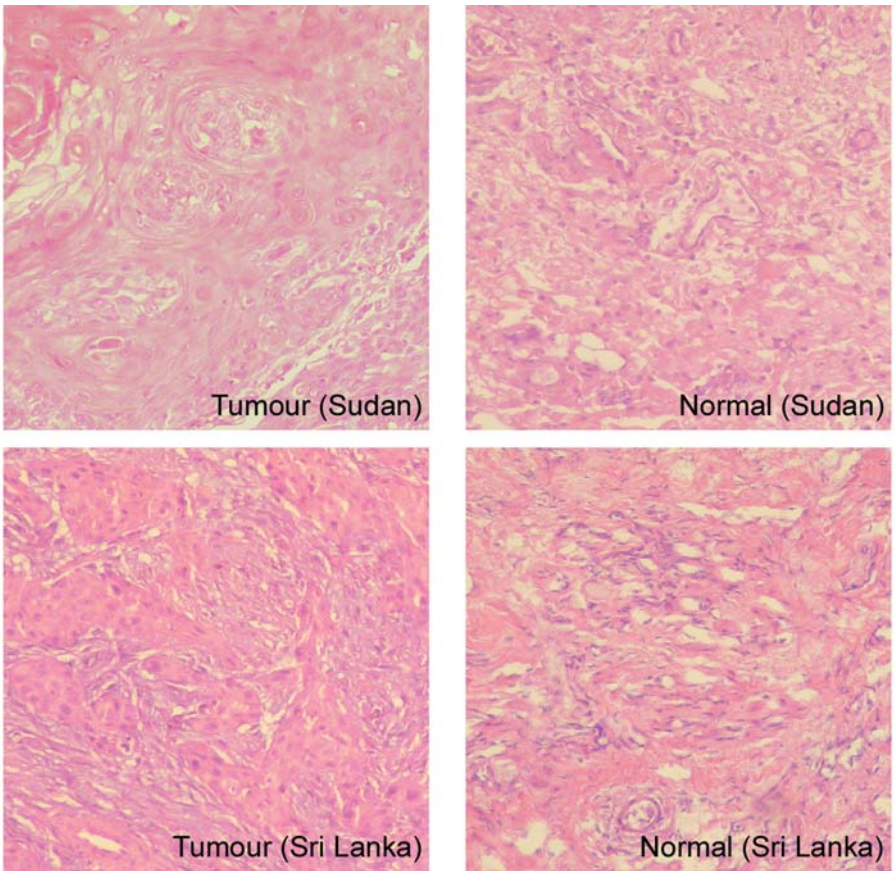

Figure 1. *Histopathological picture of a pair of matched tissue specimens of oral squamous cell carcinomas and normal controls from patients from Sudan and Sri Lanka.*

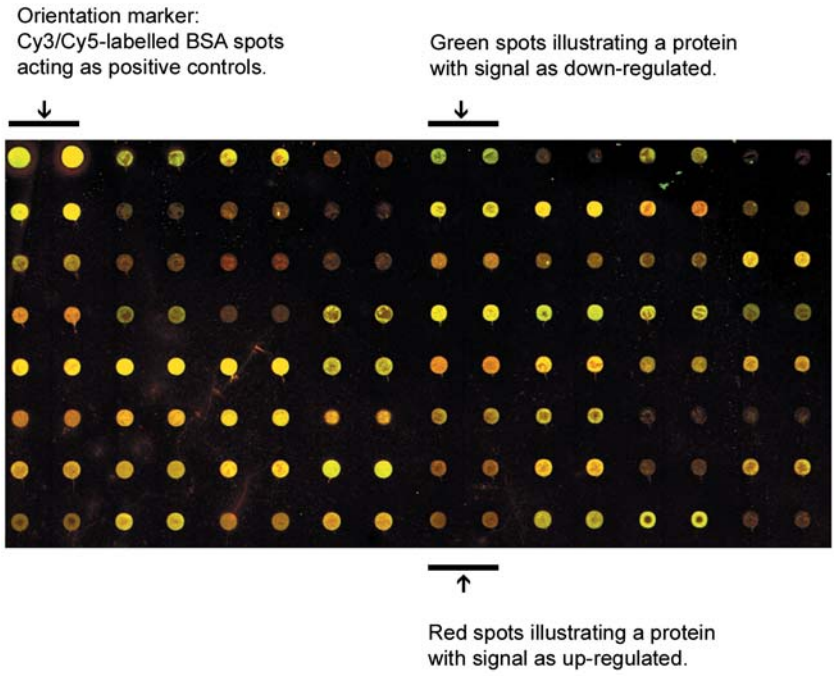

Figure 2. *Antibody microarray from one of the slides (Sudan) analyzed with GenePix software. Each spot represents an individual antibody spotted in duplicate. The yellow color indicates that antibodies label with equal amounts of Cy3 and Cy5 and are, therefore, yellow when scanned.*

Table II. *List of the differentially-expressed proteins within the antibody microarray in the oral squamous cell carcinomas examined from patients from Sudan and Sri Lanka.*

Antibody name (fold change)	Protein category/Basic function
Up-regulated proteins	
A) Sudan cases	
CAF-1 p150 (2.9)	Regulatory/chromatin assembly during DNA replication
Adenovirus 5E1A (2.1)	Transactivates early viral promoters
DFF 45 (1.9)	Enzyme/ DNA fragmentation factor
PMF-1 (1.9)	Transcription factor/coregulator of transcription of *SAT* gene
Fos (1.8)	Transcription factor/regulator for cellular mechanism
EBP50 (1.8)	Transport carrier/epithelial membrane transport
B) Sri Lankan cases	
NEK2 (2.4)	Enzyme/functioning as a serine/threonine kinase
TGF-β1 (2.4)	Signalling growth factor/stimulating and up-regulatory factor
p56dok2 (2.1)	Adaptor/insulin receptor ligand
Caveolin 1 (1.9)	Chaperone and stress/transforming suppressor activity
Psme3/PA28-g (1.9)	Chaperone and stress/activator subunit of the proteasome
Gap1m (1.9)	Enzyme/GTPase-activating protein
Neuroglycan C (1.8)	Adhesion/structural protein of neural transmembrane
Down-regulated proteins	
Sri Lankan cases	
pp120 src substrate (0.5)	Adhesion/cell structure and adhesion

supernatant collection, the gel spots were extracted once for 20 min at room temperature with 60% acetonitrile containing 0.1% trifluoroacetic acid (TFA) and the supernatants were pooled. After concentration of the peptide mixture (10- 20 μl) on a speedovac, 1 μl was mixed with an equal volume of α-cyano-4-hydroxycinnamic acid (CHCA) matrix, 10 mg/mL in 60% TFA, 0.1% trifluoroacetic acid, and 1.5 μl of the mixture was applied to MALDI-TOF-MS sample targets. The samples were analyzed on Autoflex (Bruker Daltonics, Germany), peptide mass spectra were recorded and a PMF search was done from the NCBInr database protein matching using MASCOT (www.matrixscience.com) database search algorithms. Trypsin autolytic peptides (m/z = 842.51, 1045.56 and 2211.10) were used to internally calibrate each spectrum and, when not detected (as in the cases with very little protein), an external standard was applied from a neighbouring sample on the target. Up to one missed tryptic cleavage was considered and a mass accuracy of 100 ppm was used for all tryptic-mass searches. Searches were performed without constraining the protein molecular weight or isoelectric point (pI), and allowed for carbamidomethylation of cysteine and partial oxidation of methionine residues. Only proteins found significant ($p<0.05$) according to MOWSE highest search score(s) and with 20 ppm or better were selected. Cases with lower mass accuracy and score matching were run in duplicate or triplicate to ensure an accurate analysis.

Results

Antibody microarray. After hybridization and washing, the antibody microarray slides were scanned immediately on a range of different laser intensities to ensure that none of the spots were saturated and that all data remained within a linear range. The slides were scanned at the appropriate wavelength first for Cy3 and then Cy5 and both images were generated by the GenePix Pro 4.0 software. In each slide, the control spots of BSA at the three corners hybridized strongly with both labels that determined the best laser power. The two blank spots at the fourth corner, which acted as negative controls, had no signal above that of the local background. A typical section of one of the slides is shown in Figure 2. On experimenting with a range of signal intensities, most spots were yellow, which indicated an equal hybridization of the two protein samples for the cases from each country independently. The arrows in Figure 2 indicate the BSA-positive control dots and two antibodies (each spotted in duplicate) that appeared green and red, indicating stronger (green) and weaker (red) hybridization by the normal protein compared to the cancer protein. When analyzed initially by GenePix Pro 4.0 software and then by Clontech software, over 80% of the proteins had INRs of 0.9–1.1. Only proteins that had INR ratios >1.7 or <0.6 (based on the dynamic range of the array) were considered significant within the Excel software developed specifically for the antibody array by the manufacturer. These INR values indicate with high probability that there is a true difference in protein abundance. Although INR ratio changes in the range of 1.7-1.5 can be accepted, but with less confidence, all proteins that showed these changes were ignored and only those of INR ratios >1.7 or <0.6 were selected as significant. In Table II, a list of the proteins found to be up-or down-regulated in the OSCCs examined from Sudan and Sri Lanka is shown.

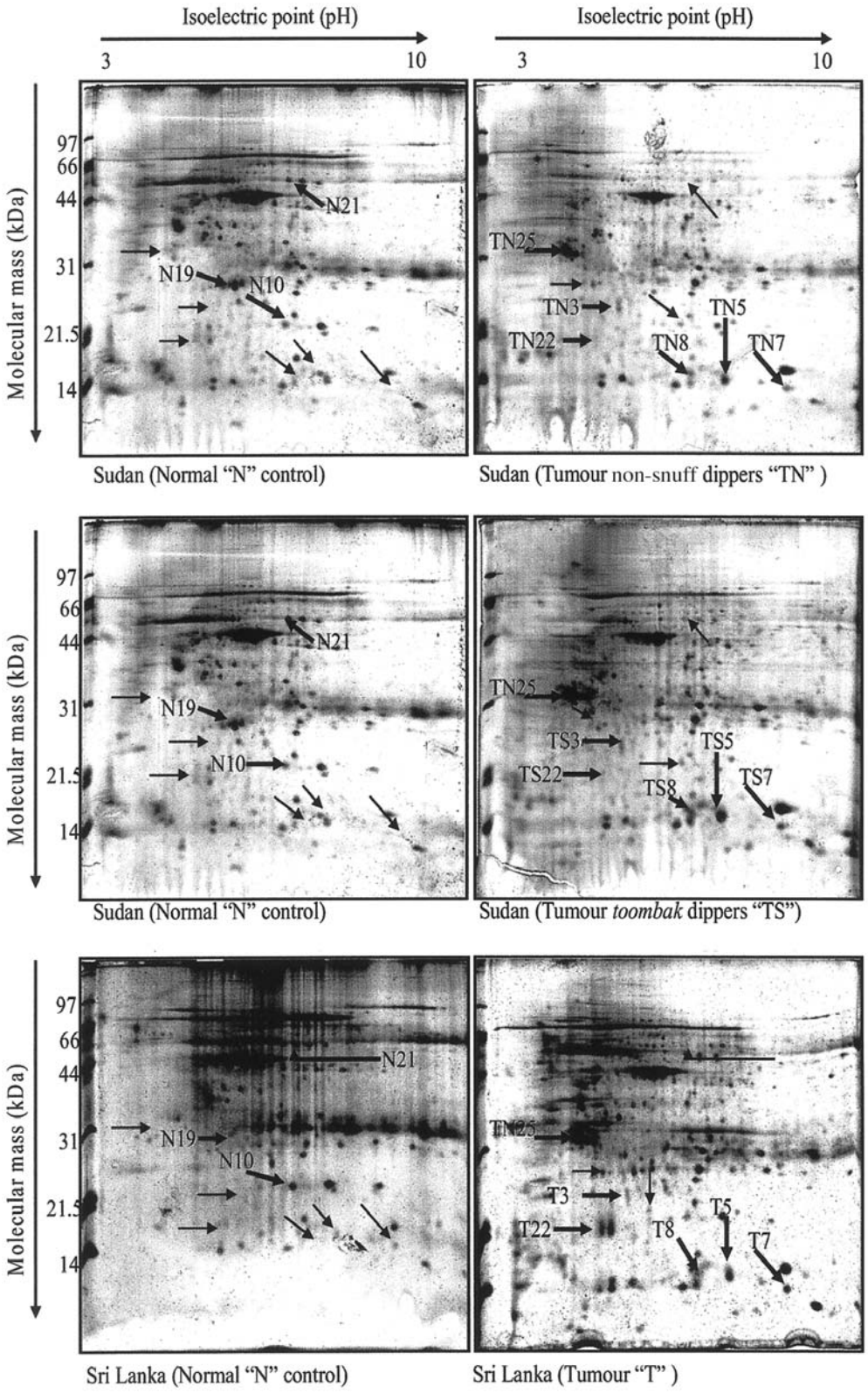

Figure 3. *Representative SYPRO-Ruby-stained 2-DE images of oral normal (N) controls and tumour (T) samples analyzed from patients from Sudan and Sri Lanka. Numbered matched spots identified by peptide mass fingerprinting are shown.*

Table III. *Results from the PMF search in the NCBInr database for the cases of oral squamous cell carcinomas and normal controls examined from patients from Sudan (toombak dippers and non-dippers) and Sri Lanka.*

Spot no./ID no./ Protein name	Peptides matched	Sequence coverage (%)	Mass accuracy/ Mass (kDa)/pI
TS5/4506773/S100A9	8	78	26 ppm/13/5.7
TS7/4389153/S100A7	8	70	24 ppm/11/6.3
TS8/4506773/S100A9	7	59	21 ppm/13/5.7
TN3/87564/Glutathione transferase	4	12	3 ppm/23/5.4
T22/27806861/Lactoglobulin	9	51	25 ppm/20/4.9
T25/34753/Stratifin	16	62	23 ppm/27/4.6
T25/38956/Epithelial cell marker protein 1	16	62	23 ppm/27/4.7
N10/182516/Ferritin light subunit	7	43	8 ppm/16/5.6
N19/17986273/Fast skeletal myosin	12	64	14 ppm/21/4.9
N21/223170/Fibrinogen	9	26	21 ppm/46/5.5
N21/31874240/Hypothetical protein [*Homo sapiens*]	8	21	22 ppm/45/5.7

Table IV. *Summary of the protein changes found in the oral squamous cell carcinomas and normal controls examined from patients from Sudan and Sri Lanka.*

Protein ID	Oral cancer (% vol)	Normal (% vol)	Fold change
S100A9	9.129335	2.67601	3.4
S100A7	4.65336	2.58579	1.8
S100A9	7.38872	2.92146	2.5
Glutathione transferase	1.08828	8.5769	–7.9
Lactoglobulin	0.871626	4.25236	–4.9
Epithelial cell marker protein 1 (Stratifin)	7.73702	5.50145	1.4
Ferritin light subunit	0.694093	4.95993	–7.1
Fast skeletal myosin	0.741926	4.38502	–5.9
Fibrinogen	0.482402	5.61916	–11.68
Hypothetical protein [*Homo sapiens*]	0.482402	5.61916	–11.68

2-DE PAGE and MALDI-TOF-MS. Proteins (400 μg, trichloroacetic acid precipitates), used for the 2-DE, were evenly distributed in the 18-cm gel with pH in the range of 3-10 and molecular masses of 14-97 kDa. About 400 protein spots were detected in each gel by Image Master software, and some were present in train spots, indicating that modified proteins or isoforms were separated by the 2-DE. In Figure 3, typical gel images for both tumours and normal controls for Sudanese and Sri Lankan cases are shown side by side. In Table III, data for the proteins (or the sum of the isoforms) that showed statistically significant differences are summarized with total normalized volumes [(% vol), using mean 100% and M.S.D. statistics] and fold differences found between OSCC and NC samples. Roughly, 4 proteins (or isoforms) were up-regulated and 6 proteins (isoforms) were significantly suppressed in the OSCCs.

MALDI-TOF-MS and protein identification. From the SYPRO® Ruby-stained 2-DE, excised protein spots were subjected to in-gel tryptic digestion followed by MALDI-TOF-MS measurement and database matching. In Figure 3, protein spots identified in tumours and normal controls obtained from patients from Sudan (*toombak* dippers and non-dippers) and Sri Lanka are illustrated. Table IV illustrates a summary of the proteins identified in the spots analyzed. Several isoforms gave an identical primary structure in the protein matching and were classified as one protein. Proteins which displayed an apparent change in their expression level in tumour samples from the two countries included S100A7, S100A9, glutathione transferase, lactoglobulin and epithelial cell marker protein 1 (stratifin), while those which displayed an apparent change in their expression level in normal controls compared to tumours included ferritin light subunit, fast skeletal myosin, fibrinogen and a hypothetical protein (Table IV).

Discussion

Buccal mucosa represents the primary site for OSCC development in Sri Lanka (3, 8), but in Sudan, the gum, lip, buccal or floor of the mouth represent the primary sites (4, 5), contrary to the tongue and other oropharyngeal cancers which represent the primary site of cancers in Western countries, where cigarette smoking and heavy alcohol consumption are the main causative factors (1-3). Smokeless tobacco and BQ are the major oral cancer causative factors for the two populations studied here. This renders the squamous epithelium at these sites in patients from Sudan (*toombak* users) and Sri Lanka (BQ users) to be susceptible to genotoxic damage by these products, hence the development of OSCC. In this work, changes of the proteome in pools of OSCCs from these two populations were examined, and significant changes in 6 to 8 differentially-expressed proteins were found by antibody microarrays in the cases examined from Sudan and Sri Lanka, respectively. The findings demonstrate a relative measure of protein abundance by comparing protein levels in one sample (tumour) to those of a second sample (normal), which compete for binding to the same antibody target. As far as we know, for 8 of the proteins found (NEK2, p56dok2, Gap1m, neuroglycan C, pp120 src

substrate, CAF-1 p150, EBP50, PMF-1), this is the first report of altered expression of these proteins in oral cancers (OCs). These proteins are involved in critical cellular processes including, among others, chromatin assembly during DNA replication, transcription, transport, GTPase-activating proteins and cellular structure and adhesion. Two proteins, CAF-1 p150 and Adenovirus 5E1A, showed the highest fold changes in the OSCCs from Sudan. CAF-1 p150 has been found to interact directly with PCNA, and its use as a proliferation marker and prognostic indicator in malignant and benign breast lesions has been reported (22). The adenovirus E1A (early region 1A) gene has been found to activate the transcription of c-Fos, c-Jun, JunB and c-Myc, repress the transcription of *c-erbB2/neu* and EGFR, interact and alter the function of pRB, p27, cyclin A-CDK2, cyclin E-CDK2, CtBP and p300/CBP, among others (23). However, although evidence for E1A involvement in human cancers is lacking based on *in vivo* studies (23), its expression has recently been shown to induce apoptosis in a panel of head and neck cancer cell lines independently of their p53 status (24). Our observation of high change folds of the two proteins in the cancers from Sudan might suggest an important role of the two proteins in carcinogenesis of these tumours and warrants further investigations. For the OSCCs from Sri Lanka, three proteins $p56^{dok-2}$, Nek2 and TGF-β1 showed high change folds. $p56^{dok-2}$, a member of the Dok-related proteins family that binds to RasGAP, plays a novel role as a negative regulator of signal transduction and cell proliferation mediated by cytokines (25), and is implicated in regulation of centrosome separation and spindle formation (26). TGF-β1, inhibits the growth of cells of epithelial origin by down-regulating components of the cell cycle and up-regulating cell cycle inhibitors, and it acts in late G1-phase preventing further progression to the G1/S-phase transition (27). Expression of Nek2 has been reported in human cancer cell lines and primary tumours, being elevated 2- to 5-fold in cell lines derived from a range of human tumours, and its role as a novel potent target for chemotherapeutic intervention in breast cancer has been suggested (26). Although no studies in oral or head and neck SCCs have reported changes in expression of these proteins, except for TGF-β1, the fact that they were only detected in cancers from Sri Lanka might suggest their importance in carcinogenesis of tumours associated with BQ use. Although the results from the antibody microarrays need to be confirmed by other methods like immunoblotting and immunohistochemistry, we were unable to perform these analyses due to lack of sufficient tissue samples and have, therefore, restricted our findings to protein overexpression higher than 1.7-fold or less than 0.6-fold changes. Nevertheless, and for proper interpretation of our findings from this array experiment, we plan to test these proteins by immunohistochemistry in a larger sample size of OCs in an ongoing independent study involving tissues of Sudanese and Sri Lankan patients with these tobacco habits to address their exact roles.

In proteomics, the most widely used techniques are still the 2-DE maps (12, 28). Using 2-DE-MALDI-TOF-MS, we identified up-regulation of S100A7 and S100A9, which belong to the S100 protein family. Findings of S100A9 are in agreement with other studies in OSCCs (29). The S100 gene family, composed of at least 20 members, share a common structure defined in part by the Ca^{2+}-binding EF-hand motif, and are implicated in the immune response, differentiation, cytoskeleton dynamics, enzyme activity, Ca^{2+} homeostasis and growth (30). It has been suggested that many S100 proteins may promote cancer progression through specific roles in cell survival and apoptosis pathways (30). Our findings on S100A7 and S100A9, which are also in agreement with cDNA microarray results conducted in cases of OCs from Sudan, highlight the role of these proteins in promoting the development of this cancer. In addition, other S100 proteins (S100A2, S100A4, S100A6) were also among the genes found to be differentially-expressed using cDNA microarrays (data not shown). Therefore, it will be of interest to investigate the role of these proteins during oral carcinogenesis, in particular in cases of OCs from the two populations. We have also found elevation of epithelial cell marker protein 1 (stratifin) in the OCs examined, which is in agreement with previous studies in OSCCs (31). Stratifin, known as tumour suppressor protein 14-3-3 sigma, plays significant roles in a number of cellular activities, such as cell cycle progression, differentiation and apoptosis, by regulating various cytoplasm signalling pathways, and it reduces cell growth by G2 arrest (32). Although decreased expression of stratifin has been reported in neuroendocrine lung tumours (33) and in bladder SCC (34), high expression of stratifin has been found in several skin diseases including SCC (35) and in OCs (31), which are in agreement with our findings. Nevertheless, this discrepancy in the expression in different cancers demands further work to interpret these findings.

Underexpression of glutathione transferase and lactoglobulin was found. The former, a member of the two main cytosolic enzyme families involved in the detoxification process, is involved in catalyzing the glutathione conjugation reaction with electrophilic compounds and carcinogens (36). Our findings of underexpression of this protein in the OSCCs examined might indicate failure of the self-protective mechanism in cells exposed to *toombak* and BQ, therefore rendering them more susceptible to these carcinogens. Although the findings of underexpression of this protein by our study are different from studies showing its overexpression in OCs from the buccal mucosa (31), in other oral SCCs (37-39) and in carcinomas (40-42), the fact that there are several reports related to the expression of

this protein in OCs, might suggest its use as a potential tumour marker and as a drug target in OCs. Moreover, we found underexpression of lactoglobulin, a member of the whey proteins; to our knowledge, this is the first report in OCs. Whey proteins possess antioxidant and antitumour effects in addition to several other functions, and the primary mechanism by which whey is thought to exert its effects is by intracellular conversion of the amino acid cysteine to glutathione (reviewed in 43). Findings of underexpression of glutathione transferase and lactoglobulin in OCs might suggest the role of the failure of the antitumour and antioxidant effects of these proteins in carcinogenesis of these tumours.

On the other hand, significant underexpression of fast skeletal myosin, ferritin light subunit and the blood-clotting protein fibrinogen was found in the NCs when compared to the OSCCs examined from Sudan and Sri Lanka. The dramatic alterations in the expression of skeletal myosin found in the tumours examined, reflects muscle destruction during cancer invasion and migration process, which are in agreement with other reports in OCs (29). The similarity between these results supports the importance of myosin expression as a marker for tumour invasion and the migration process (29), and its role in these processes warrants further studies. Fibrin(ogen) deposition is known to be a histological feature of a variety of pathological processes including inflammation and wound healing, but its role in cancer is not as well appreciated or understood (44, 45). Using a number of basic and clinical studies (reviewed in 44), the role of fibrin in the progression of a variety of tumours has been demonstrated, and its association with tumours has been shown to result from extravasation and subsequent clotting of fibrinogen from plasma. Most fibrinogen that is extravasated from blood vessels gets clotted and cross-linked within minutes, and the resulting fibrin forms a three-dimensional gel that is present in association with individual tumour cells or clumps of tumour cells (reviewed in 44). It remains, however, to be investigated further in OCs. With respect to ferritin, a known tumour marker that belongs to the acute-phase protein family synthesized by the liver, it has been used as a biomarker in different cancers, and some investigators have reported an increase in ferritin levels in patients with a variety of tumours including head and neck SCCs, where it has been reported as a useful prognostic marker (46). The dramatic significant underexpression of this protein observed in our study strongly suggests that this protein may therefore function as a potential tumour marker in cases from Sudan and Sri Lanka.

In this study, pools of proteins were used in order to establish the protein expression profile(s) of OSCCs from two populations. This was done to balance for the complexity and individual difference(s) of the tumours used, exposure of different anatomical sites to the same carcinogen, and the fact that this is a population-based comparative study. However, the results obtained for some of the proteins identified are in agreement with other studies that had selected tumours confined to one site (29, 31). Our results are, therefore, as consistent as can be expected, although the total number of measurements from the patient population has been reduced. One of the main problems with protein profiles of cancer tissues is the heterogeneous nature of the tumour samples. Sources of heterogeneity include a variable proportion of cancer cells relative to non-cancer cells, and a variable degree of tissue damage, apoptosis and necrosis (47). Tumour samples were used that were confirmed to contain more that 70% tumour cells. In addition, primary tumour biopsies and normal tissues from the same patients were analyzed. The primary tumour biopsies may contain several different cell types other than carcinoma cells, so the exact definition of major cellular changes during the conversion of normal to malignant oral cancer is limited due to defects in the multiplicity of 2-DE features.

In conclusion, the observed over- and underexpressed proteins reported in this study might be of value in correlation with clinicopathological data to define clinically useful biomarkers in OCs from the two populations. Since identification of functioning protein(s) is essentially needed to provide more effective therapy for patients suffering from OCs, the findings of this study portray global protein changes that appear to be specific to OCs from Sudan and Sri Lanka. Although only a few specific proteins were found to be either over- or underexpressed in both population groups and with lack of concordance, this may reflect differences in patterns of tobacco use in the two populations, in particular *toombak* use in Sudan and BQ mixed with tobacco among Sri Lankans. Results for some proteins, however, agree with other studies in OCs, which may provide the potential for testing these proteins as diagnostic, prognostic biomarkers and/or targets for therapy in larger studies involving cases of OCs from several populations. In particular, glutathione transferase, ferritin, fast skeletal myosin, fibrinogen and a battery of S100 proteins warrant further investigations in these cancers.

Acknowledgements

The Norwegian Research Council, Felleslegatet til Fordel for Biologisk Forskning fond, Meltzer Høyskolefond and the UICC International Cancer Technology Transfer Fellowship (author S.O.I.) supported this study. We greatly appreciate the skilled technical assistance of Ms. Inger Ottesen. The authors would like to thank Mr. Olav Majaavatten and Ms. Jill Opsahl at the Proteomic Unit of the University of Bergen, Norway (http://www.probe.uib.no/) for their excellent technical help. We thank Professor Ali M. Idris and the personnel of the Toombak

and Smoking Research Centre, Khartoum, Sudan, for their logistic support. We thank Dr. Osman A. Algindi and Dr. Abdelaal Saeed of the Department of Oral and Maxillofacial Surgery, University of Khartoum, Sudan, and Dr. Toru Nagao, former JICA Technical Coordinator in Sri Lanka, for their logistic support. Finally, we thank Professor Johan R. Lillehaug and Professor Harald B. Jensen of the Department of Molecular Biology, University of Bergen, Norway, for critically reading and commenting on the manuscript.

References

1 Moore SR, Johnson NW, Pierce AM and Wilson DF: Review: The epidemiology of mouth cancer: a review of global incidence. Oral Dis *6*: 65-74, 2000.

2 Moore SR, Johnson NW, Pierce AM and Wilson DF: Review: oral oncology: the epidemiology of tongue cancer: a review of global incidence. Oral Dis *6*: 75-84, 2000.

3 Franceschi S, Bidoli E, Herrero R and Munoz N: Comparison of cancers of the oral cavity and pharynx worldwide: etiological clues. Eur J Cancer (Oral Oncol) *36B*: 106-115, 2000.

4 Ferlay J, Parkin DM, Pissani P and Globacan I: Cancer incidence and mortality worldwide. IARC Press CD Rom, Lyon, IARC, 2004.

5 Idris AM, Ahmed HM and Malik MOA: Toombak dipping and cancer of the oral cavity in the Sudan: case-control study. Int J Cancer *63*: 477-480, 1995.

6 Idris AM, Ibrahim SO, Vasstrand EN *et al*: The Swedish snus and the Sudanese toombak: are they different? Oral Oncol *34*: 558-566, 1998.

7 Idris AM, Nair J, Ohshima H *et al*: Unusually high levels of carcinogenic tobacco-specific nitrosamines in Sudan snuff (toombak). Carcinogenesis *12*: 1115-1118, 1991.

8 IARC (2004): Tobacco habits other than smoking; betel-quid and areca nut chewing; and related nitrosamines. *In*: IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. IARC Scientific Publications, Lyon *37*: 141-202, 2004.

9 Hoffmann D, Brunnemann KD, Prokopczyk B and Djordjevic MV: Tobacco-specific N-nitrosamines and Areca-derived N-nitrosamines: chemistry, biochemistry, carcinogenicity, and relevance to humans. J Toxicol Environ Health *41*: 1-52, 1994.

10 Schramm A, Apostolov O, Sitek B, Pfeiffer K *et al*: Proteomics: techniques and applications in cancer research. Klin Padiatr *215*: 293-297, 2003.

11 Marvin LF, Roberts MA and Fay LB: Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in clinical chemistry. Clin Chim Acta *337*: 11-21, 2003.

12 Glokler J and Angenendt P: Protein and antibody microarray technology. J Chromatogr B Analyt Technol Biomed Life Sci *797*: 229-240, 2003.

13 Schliephake H: Prognostic relevance of molecular markers of oral cancer–a review. Int J Oral Maxillofac Surg *32*: 233-245, 2003.

14 He Q-Y, Chen J, Kung H-F, Yuen A P-W and Chiu J-F: Identification of tumor-associated proteins in oral tongue squamous cell carcinoma by proteomics. Proteomics *4*: 271-278, 2004.

15 Ghafouri B, Tagesson C and Lindahl M: Mapping of proteins in human saliva using two-dimensional gel electrophoresis and peptide mass fingerprinting. Proteomics *3*: 1003-1015, 2003.

16 Wu W, Tang X, Hu W, Lotan R *et al*: Identification and validation of metastasis-associated proteins in head and neck cancer cell lines by two-dimensional electrophoresis and mass spectrometry. Clin Exp Metastasis *19*: 319-326, 2002.

17 Rodrigo MC, Martin DS, Redetzke RA and Eyster KM: A method for the extraction of high-quality RNA and protein from single small samples of arteries and veins preserved in RNA*later*. J Pharmacol Toxicol Methods *47*: 87-92, 2002.

18 Cawson RA and Eveson J: Oral Pathology and Diagnosis 13: pp. 10-13, Gower, London, 1987.

19 Rabilloud T, Adessi C, Giraudel A and Lunardi J: Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis *18*: 307-316, 1997.

20 Berven FS, Karlsen OA, Murrell JC and Jensen HB: Multiple polypeptide forms observed in two-dimensional gels of *Methylococcus capsulatus* (Bath) polypeptides are generated during the separation procedure. Electrophoresis *24*: 757-761, 2003.

21 Karlsen OA, Ramsevik L, Bruseth LJ *et al*: Characterization of a prokaryotic haemerythrin from the methanotrophic bacterium *Methylococcus capsulatus* (Bath). FEBS J *272*: 2428-2440, 2005.

22 Polo SE, Theocharis SE, Klijanienko J, Savignoni A *et al*: Chromatin assembly factor-1, a marker of clinical value to distinguish quiescent from proliferating cells. Cancer Res *64*: 2371-2381, 2004.

23 Frisch SM and Mymryk JS: Adenovirus-5 E1A: paradox and paradigm. Nat Rev Mol Cell Biol *3*: 441-452, 2002.

24 Flinterman M, Gäken J, Farzaneh F and Tavassoli M: E1A-mediated suppression of EGFR expression and induction of apoptosis in head and neck squamous carcinoma cell lines. Oncogene *22*: 1965-1977, 2003.

25 Gugasyan R, Quilici C, Grail D *et al*: Dok-related protein negatively regulates T cell development *via* its RasGTPase-activating protein and Nck docking sites. J Cell Biol *158*: 115-125, 2002.

26 Hayward DG, Clarke RB, Faragher AJ, Pillai MR, Hagan IM and Fry AM: The centrosomal kinase Nek2 displays elevated levels of protein expression in human breast cancer. Cancer Res *64*: 7370-7376, 2004.

27 Derynck R: TGF-beta-receptor-mediated signaling. Trends Biochem Sci *19*: 548-553, 1994.

28 Hamdan M and Righetti PG: Modern strategies for protein quantification in proteome analysis: advantages and limitations. Mass Spectrom Rev *21*: 287-302, 2002.

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

30 Donato R: S100: a multigenic family of calcium-modulated proteins of the EF-hand type with intracellular and extracellular functional roles. Int J Biochem Cell Biol *33*: 637-668, 2001.

31 Chen J, He QY, Yuen AP and Chiu JF: Proteomics of buccal squamous cell carcinoma: the involvement of multiple pathways in tumorigenesis. Proteomics *4*: 2465-2475, 2004.

32 Hermeking H, Lengauer C, Polyak K, He TC *et al*: 14-3-3 Sigma is a p53-regulated inhibitor of G2/M progression. Mol Cell *1*: 3-11, 1997.

33 Yatabe Y, Osada H, Tatematsu Y, Mitsudomi T *et al*: Decreased expression of 14-3-3 sigma in neuroendocrine tumors is independent of origin and malignant potential. Oncogene *21*: 8310-8319, 2002.

34 Ostergaard M, Rasmussen HH, Nielsen HV, Vorum H *et al*: Proteome profiling of bladder squamous cell carcinomas: identification of markers that define their degree of differentiation. Cancer Res *57*: 4111-4117, 1997.
35 Lodygin D, Yazdi AS, Sander CA, Herzinger T *et al*: Analysis of 14-3-3 sigma expression in hyperproliferative skin diseases reveals selective loss associated with CpG-methylation in basal cell carcinoma. Oncogene *22*: 5519-5524, 2003.
36 Tsuchida S and Sato K: Glutathione transferases and cancer. Crit Rev Biochem Mol Biol *27*: 337-384, 1992.
37 Lafuente A, Maristany M, Arias C, Cuchi A *et al*: Glutathione and glutathione S-transferases in human squamous cell carcinomas of the larynx and GSTM1 dependent risk. Anticancer Res *18*: 107-111, 1998.
38 Li TJ, Hirayama Y and Kitano M: Glutathione S-transferase pi-class as a tumour marker in lingual preneoplastic and neoplastic lesions of rats and humans. Virchows Arch *431*: 37-43, 1997.
39 Zhang L, Xiao Y and Priddy R: Increase in placental glutathione S-transferase in human oral epithelial dysplastic lesions and squamous cell carcinomas. J Oral Pathol Med *23*: 75-79, 1994.
40 Palackal NT, Lee SH, Harvey RG, Blair IA *et al*: Activation of polycyclic aromatic hydrocarbon trans-dihydrodiol proximate carcinogens by human aldo-keto reductase (AKR1C) enzymes and their functional overexpression in human lung carcinoma (A549) cells. J Biol Chem *277*: 24799-24808, 2002.
41 Scuric Z, Stain SC, AndersonWF and Hwang JJ: New member of aldose reductase family proteins overexpressed in human hepatocellular carcinoma. Hepatology *27*: 943-950, 1998.
42 Takahashi M, Fujii J, Miyoshi E, Hoshi A *et al*: Elevation of aldose reductase gene expression in rat primary hepatoma and hepatoma cell lines: implication in detoxification of cytotoxic aldehydes. Int J Cancer *62*: 749-754, 1995.
43 Marshall K: Therapeutic applications of whey protein. Altern Med Rev *9*: 136-156, 2004.
44 Costantini V and Zacharski LR: The role of fibrin in tumor metastasis. Cancer Metastasis Rev *11*: 283-290, 1992.
45 Costantini V and Zacharski LR: Fibrin and cancer. Thromb Haemost *69*: 406-414, 1993.
46 Maxim PE and Velti RW: Serum ferritin as a tumor marker in patients with squamous cell carcinoma of the head and neck. Cancer *57*: 305-311, 1986.
47 Ou K, Seow TK, Liang RC, Ong SE and Chung MC: Proteome analysis of a human heptocellular carcinoma cell line, HCC-M: an update. Electrophoresis *22*: 2804-2811, 2001.

Received September 23, 2005
Accepted October 27, 2005