CANCER GENOMICS & PROTEOMICS *11*: 127-140 (2014)

Profiling of Chromosomal Changes in Potentially Malignant and Malignant Oral Mucosal Lesions from South and South-East Asia Using Array-comparative Genomic Hybridization

MAI LILL SUHR LUNDE[1], ERIC ROMAN[1], SAMAN WARNAKULASURIYA[2], RAVI MEHROTRA[3], JUSSI LARANNE[4], ENDRE N. VASSTRAND[5] and SALAH O. IBRAHIM[5]

[1]*Department of Biomedicine, University of Bergen, Bergen, Norway;*
[2]*Department of Oral Medicine and Pathology, Guy's, King's and St Thomas' NH5 Trust Hospitals, WHO Collaborating Centre for Oral Cancer and Precancer, King's College London, London, U.K.;*
[3]*Department of Pathology, Moti Lal Nehru Medical College, Allahabad, India;*
[4]*Department of Otorhinolaryngology, Tampere University Hospital, Tampere, Finland;*
[5]*Department of Clinical Dentistry, Periodontology, University of Bergen, Bergen, Norway*

Abstract. *Background/Aim: Using array-CGH, the present study aimed to explore genome-wide profiles of chromosomal aberrations in samples of oral cancer (OC), oral submucous fibrosis (OSF) and their corresponding normal oral mucosa from Indian (n=18) and OC from Sri Lankan (n=12) patients with history of BQ use, and correlate the findings to other clinicopathological parameters. A second aim was to verify the results from the array-CGH by selecting a candidate gene, S100A14, and examine its expression and genetic polymorphisms by immunohistochemistry (IHC) and restriction fragment length polymorphism (RFLP) using samples from both populations and from multi-national archival DNA and paraffin-embedded samples of OC. Results: In OC and OSF samples, 80 chromosomal regions (harboring 349 genes) were found as deleted or amplified. Out of the 349 genes, 34 (including several S100 gene family members) were found to be deleted and 30 (containing NOTCH4, TP53 and ERBB2) were found as amplified in OSF and OC cases. 285 genes (including TP53, ERBB2 and BRCA1) were found either as deleted in one population or amplified in the other. Few chromosomal alterations were found to be exclusive to either OC or OSF samples alone. IHC demonstrated down-regulation and transfer of S100A14 protein expression from membrane to cytoplasmic. RFLP showed differential distribution between Asian samples compared to African and Western samples at 461 G>A SNP. Conclusion: The present study provides findings on chromosomal aberrations likely to be involved in pathogenesis of OC and OSF. Findings of chromosomal changes harboring genes previously found in OC examined from Western, African and Asian populations demonstrate the importance of these changes in development of OC, and the existence of common gene-specific amplifications/deletions, regardless of source of samples or attributed risk factors. We report a down-regulation of S100A14 expression to be a significant marker in association with loss of 1q21 in 70% of OC samples.*

Correspondence to: Dr. Salah O. Ibrahim, Department of Clinical Dentistry, Faculty of Medicine and Dentistry, University of Bergen, Årstadveien 19, 5009 Bergen, Norway. Tel: +47 55586656, Fax: +47 55588561, e-mail: salah.ibrahim@iko.uib.no

Key Words: Oral squamous cell carcinoma, oral submucous fibrosis, betel quid, array-CGH, S100A14 SNPs.

Worldwide, oral squamous cell carcinoma (OSCC) is one of the 10 most common causes of cancer deaths and the sixth most common malignancy (1). In India and Sri Lanka, besides changes in betel-quid (BQ) chewing habits, trends of alcohol consumption show an increasing intake in the last decades, a sign of westernization of the cultural habits that influence the incidence of OSCC (2, 3). OSCCs account for up to 30% of all cancers in India, with lymph node metastasis being the most important prognostic factor for a patient with OSCC. The 5-year survival rate can decrease to below 20% when cervical metastasis is present, because of the rich lymphatic network of the majority of intra-oral anatomic sites (4). In India, 50% of OSCC were reported to involve the buccal mucosa, whereas in many western countries the same was reported to be only 5% (5). OSCC is the most common cancer in Sri Lanka, accounting for approximately 20%-30% of all cancers in males. Although tobacco and alcohol use are involved in the etiology of OSCC, in Sri Lanka, BQ chewing is closely related to the high incidence of this type of cancer

(6). This habit also induces precancerous lesions, including epithelial dysplasia and oral submucous fibrosis (OSF) (6). For BQ chewing, a mixture of fresh areca nut, slaked lime from seashells, fresh betel leaf, and slightly dried tobacco is introduced into the mouth (6).

Most of the OSCC cases in South and South East Asia arise in the presence of clinical premalignant or potentially malignant conditions including oral leukoplakias, oral lichen planus, and OSF (7). Among these, OSF, a chronic progressive disorder, described first in the early 1950s is predominantly seen in India, Sri Lanka, China (e.g., Hunan and Hainan), Taiwan, Bangladesh, and Pakistan. BQ is the most important risk factor for development of OSF. Malignant transformation rate of OSF was found to be in the range of 7%-13% depending on the study population (8). It has been suggested that the identification of a potential predictor for the transformation of OSF to OSCC might be favorable for early diagnosis, early therapy and prevention of OSCC (9).

Regardless the source, and like all solid tumors, OSCCs are characterized by multiple chromosomal alterations and are genetically complex (10). Dependencies between the numerous genetic alterations lead to observed karyotypic complexity, which results in the distinct biological behavior of OSCC (11). For example, node-positive OSCCs are biologically-aggressive and have poor prognosis when compared to node-negative OSCC (12). This indicates that different genetic pathways of progression exist in OSCCs, leading to the molecular subtypes with distinct clinical outcomes. Hence, it is necessary to identify the genetic alterations and the interactions between OSCC that form multiple progression pathways. This approach may aid in the better understanding of the biology of OSCC.

Comparative genomic hybridization (CGH), a known genome-wide profiling technique, has revealed non-random pattern of genetic alterations in OSCC (13-18). An early study has suggested that genomic alterations in OSCC may be more uniform than those of other solid tumors, but a more recent study demonstrated that the initiation and progression of OSCC involves divergent biological pathways (16). Frequently reported chromosomal alterations in OSCC are gains at 1q, 3q, 5p, 7q, 8q, 9q, 11q, 12p, 14q and 15q and deletions at 1q, 3q, 1p, 3p, 4p, 5p, 7q, 8p, 10p, 11q, 13q, and 18q (13, 16-21). In contrast to conventional CGH, microarray-based CGH (array-CGH) enables high-resolution analysis of chromosomal alterations, allowing the detection of chromosomal deletions and amplification down to DNA segments as small as 5-10 kb.

To date, the pathogenetic pathways followed by OSCC associated with BQ use from Indian subcontinent have not been thoroughly investigated. Elucidating the divergent routes in these OSCCs could provide information about molecular subtypes, which might support treatment decisions. With this long-term goal in mind, in the current study, we used array-CGH to explore genome-wide profiles of chromosomal aberrations in samples of OSCC and OSF from Indian (n=18) and OSCC from Sri Lankan (n=12) patients with history of BQ use, and the findings were correlated to other clinicopathological parameters. To verify the findings in this study, we further selected one candidate gene, *S100A14*, which was found deleted in the samples from both populations and in previous cases of OSCC examined from Sudan and Norway, and examined its expression and genetic polymorphisms by IHC and RFLP using multi-national archival DNA and paraffin-embedded samples of OSCC.

Materials and Methods

Patients. Primary samples of OSCC (n=24) and OSFs (n=6) with their corresponding pair-wise normal controls were obtained from consecutive patients with previously untreated OSCC/OSF from Sri Lanka (n=12, average age 58.8 years; range 43 to 70; SD±7.57) and India (n=18, average age 46.9 years; range 18 to 70; SD±18.4). After surgery, tissue samples (malignant, pre-malignant and normal controls) were submerged in the tissue storage and RNA stabilization solution RNA*later*™ (Ambion, Inc., Woodlands, TX, USA), and were transported to the Department of Biomedicine at the University of Bergen, where they were stored at –20˚C until DNA purification and array-CGH experiments. In Sri Lanka, all tissue samples were collected from the Department of Maxillofacial Surgery, University Dental School at Peradeniya, Sri Lanka, and in India, all samples were collected from the Department of Maxillofacial Surgery, Moti Lal Nehru Medical College, Allahabad, India. The study was approved by corresponding Committees for Medical Ethics at the University of Peradeniya and at Moti Lal Nehru Medical College, Allahabad, India. All tumors were staged following the 1987 UICC staging system, and had their histopathological diagnosis confirmed by three of the authors (SW/RM/SOI) using either fresh frozen/or 10% formalin-fixed, paraffin embedded tissue sections stained with haematoxylin and eosin (H&E). The tumors were histologically graded into highly-, moderate- or poorly-differentiated carcinomas (22). To rule-out gene expression alterations because of stromal cell contamination, we confirmed pathologically that each tumor specimen contained ≥70% tumor tissue by analysis of the corresponding H&E-stained sections. Data on clinicopathological parameters for all 30 patients are represented in Table I. Histopathological diagnosis of the OSF cases was confirmed by two of the authors (RM/SOI) using formalin-fixed, paraffin-embedded tissue sections stained with H&E. Diagnosis was performed according to Pindborg *et al*. (23).

Array CGH. Total genomic DNA was extracted from all samples using DNeasy Purification Kit (QIAGEN Inc., Valencia, CA, USA), following the manufacturer's protocol. 500 ng of total DNA were digested overnight at 37˚C with DpnII (New England Biolabs, Ipswich, MA, USA), and tumor DNA was labeled with Cy3-dCTP and control DNA (Human Universal Reference DNA) with Cy5-cCTP (NEN Life Science Products Inc., Boston, MA, USA) by random priming. The labeled DNA samples were combined and mixed with 135 μg human Cot-1 DNA (Roche Diagnostics Corp.,

Table I. *Patient clinicopathological data*.

Sample	Age (years)	Diagnosis	Tumor site	Smoking	ST	Alcohol
Sri Lanka 1	67	OSCC	Alveolus	+ +	+ +	+ +
Sri Lanka 2	56	OSCC	Buccal	–	+ +	–
Sri Lanka 3	64	OSCC	Alveolus	–	+ +	–
Sri Lanka 4	43	OSCC	Floor	+ +	+ +	+ +
Sri Lanka 5	50	OSCC	Buccal	–	+ +	–
Sri Lanka 6	55	OSCC	Alveolus	–	+ +	–
Sri Lanka 7	68	OSCC	Alveolus	+ +	+ +	+ +
Sri Lanka 8	58	OSCC	Buccal	+ +	+ +	–
Sri Lanka 9	63	OSCC	Lateral tongue	+ +	–	+ +
Sri Lanka 10	59	OSCC	Buccal	+	–	+ +
Sri Lanka 11	70	OSCC	Buccal	+ +	+ +	+ +
Sri Lanka 12	64	OSCC	Alveolus	–	+ +	–
India 1	50	OSCC	Alveolus	–	++	–
India 2	65	OSCC	Buccal mucosa	++	++	–
India 3	50	OSCC	Floor of mouth	–	++	–
India 4	50	OSCC	Tounge	–	++	–
India 5	60	OSCC	Lip	–	++	–
India 6	45	OSCC	Buccal mucosa	–	++	–
India 8	65	OSCC	Lip	++	–	–
India 11	70	OSCC	Tongue	–	++	–
India 13	40	OSCC	Tongue	++	++	–
India 14	70	OSCC	Buccal mucosa	ND	ND	ND
India 15	65	OSCC	Alveolus	–	++	–
India 16	27	OSCC	Buccal mucosa	–	++	+
India 22	35	OSF	Buccal mucosa	–	++	–
India 23	22	OSF	Buccal mucosa	–	++	–
India 24	25	OSF	Buccal mucosa	–	++	–
India 25	66	OSF	Buccal mucosa	–	ND	–
India 26	18	OSF	Buccal mucosa	–	++	–
India 27	22	OSF	Buccal mucosa	–	++	–

OSCC, Oral squamous cell carcinoma; OSF, oral submucous fibrosis; ST, smokeless, tobacco; ++, regular (daily); +, occasionally; –, non-user; ND, no data.

Indianapolis, IN, USA). The microarrays used contained 4,549 BAC (bacterial artificial chromosome) and PAC (P1 artificial chromosome) clones representing the human genome at ~1 Mb resolution, as well as the minimal tiling- path between 1q12-q25. The genomic 1 Mb arrays used were produced at the Norwegian Microarray Consortium (NMC; www.microarray.no) based on the clone set provided by Dr. Nigel Carter at the Wellcome Trust Sanger Institute, United Kingdom, to the Norwegian Microarray Consortium at the RR-HF/UiO Microarray Core Facility in Oslo, Norway as described in detal by Meza-Zepeda *et al*. (24). Arrays were produced where each clone is represented by four DNA spots on the array. Hybridization was performed using an automated hybridization station GeneTAC/HybArray (Genomic Solutions, Ann Arbor, MI, USA) with agitation for 42 h at 37˚C, and arrays were scanned using an Agilent G2565B scanner (Agilent Technologies Inc, Santa Clara, CA, USA). The method has been described in detail by Meza-Zepeda *et al*. (24).

Data analysis. The array CGH data have been documented according to the MIAME guidelines and deposited in the ArrayExpress repository under accession no E-MTAB-819. Images acquired after scanning were analyzed using GenePix Pro 3.1 software (Molecular Devices Corp., Sunnyvale, CA, USA). Necessary adjustments of spots were performed, and intensity ratio for the two dyes were calculated for each spot before further data processing, as previously described (25) by using M-CGH, a MATLAB toolbox designed for the analysis of array CGH experiments. Clones belonging to chromosomes 1 to 22 with a known unique chromosomal position in the Ensembl database (http://www.ensembl.org/index.html) were analyzed. Detailed clone information, chromosomal regions, gene content, gene function and possible cancer association was obtained from the Ensembl and Genecards® databases (http://www.genecards.org/). Alterations of specific chromosomal regions >0.3 and <-0.3 on log2 scale found in a minimum of 20% of the Sri Lankan and Indian patients were considered for further investigations.

Chromosomal alterations were further examined separately for each population. Chromosomal regions that were found to be either amplificated or deleted in a minimum of 18% of one population were defined as unique only if the other population showed no alterations in the same region. Hierarchical clustering of the Indian samples was performed.

Tissue specimens and immunohistochemistry (IHC) for the selected candidate gene S100A14. We selected the S100A14 gene, located in chromosome region 1q21 and examined its expression and

genetic variants by IHC and RFLP. This gene was selected because of our previous findings on its possible role as a tumor suppressor gene associated with p53 and G_1-arrest, and invasion in OSCC. Immunohistochemical analysis of the S100A14 protein was performed on corresponding archival formalin-fixed, paraffin-embedded tissue specimens of the OSCC cases examined from Sri Lanka using Autostainer universal staining system (DAKO-USA, Carpinteria, CA) as described previously (26). Grading and staging of the lesions were performed as previously described (27).

For IHC, antigen retrieval was performed by microwave treatment in Tris-EDTA buffer, pH 9.0 (DAKO). After blocking with 3% BSA in TBST, rabbit polyclonal anti-human S100A14 primary antibody (10489-1-AP, Proteintech, Chicago, IL, USA, 1:500 dilutions) was applied. After washing, anti-rabbit secondary antibody conjugated with horseradish peroxidase labeled polymer (EnVision System, DAKO) was applied. Presence of antigen was visualized by staining with 3, 30 - diaminobenzidine (DAKO), counterstained with hematoxylin (DAKO) and mounted with EuKit mounting medium. Sections incubated with 3% BSA instead of primary antibody served as negative controls (28).

Evaluation of IHC. Tissue sections were examined using a light microscope for the S100A14 expression. IHC evaluation was performed on areas showing normal, dysplastic and malignant changes with main focus on the invading islands of the OSCC specimens. S100A14 staining was semi-quantitatively evaluated by manually counting the immunostained cells (at least 500 cells were counted in 3 representative areas, at 400× magnification) expressing either membranous, mixed membranous/cytoplasmic or only cytoplasmic S100A14. Based on the number of positive cells with respect to sub-cellular localization of the S100A14, OSCCs were categorized into 3 groups with low (0-9% positive cells), moderate (10-49% positive cells) and high (50-100% positive cells) scores (28).

Restriction fragment length polymorphism (RFLP). To search for single nucleotide polymorphisms (SNPs) in the *S100A14* gene, we performed polymerase chain reaction (PCR) with two different sets of primers within the *S100A14* gene (Table II) and searched for two SNPs, 461 G>A (rs11548103) and 1545 A>T (rs11548102) using DNA samples from the cases used for array-CGH. For comparison and further verification of the role of *S100A14*, we also included archival OSCC DNA samples from India (n=60), Sudan (n=65), Norway (n=24) and Finland (n=47). In addition, archival DNA samples from patients diagnosed with OSF (n=43) from India were also included. PCR-RFLP analysis was performed using a 25 μl reaction mixture containing 100 ng DNA template, 2 pmol of each primer, 0.2 mM dNTPs, 2,5-3,5 mM $MgCl_2$, 1X Gold® PCR Buffer (Applied Biosystems) and 0.5 U AmpliTaq Gold® DNA polymerase (Applied Biosystems). The reaction was carried out with an initial denaturation at 95˚C for 2 min, following 35 cycles of 94˚C at 30 sec, 62˚C (461 G>A) for 1 min or 64˚C (1545 A>T) for 30 sec, and 72˚C for 45 sec. A final elongation step of 7 min at 72˚C was included. Forward primer used for the 461 A>G PCR reaction was 5'-TCTCCTCAGCTCAGGCCTACCC-3', and reverse: 5'-TGAC CGACACTGTCCCATGGTAC-3', and primers for the 1545A>T PCR reaction was 5'- TGTACCCTAGCCTGCAC CTGTCC-3' (forward) and 5'- TCCCTCTGGTGGAGACTCCTCC-3' (reverse). The PCR products were digested with restriction enzymes *Kpn* I (461 G>A), *Mnl* I (1545 A>T) and separated by agarose gel electrophoresis.

Sequencing. Following analysis of the PCR-RFLP results, we selected 10 PCR-products of the 461 A>G SNP for further verification by DNA sequencing. The PCR reaction, performed as described earlier, was run on a 3% agarose gel, excised from the gel and purified with Quiagen gel extraction kit (Qiagen Inc., Valencia, CA, USA) and ligated into the TOPO pCR®2.1® vector (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. After ligation, the vector with the insert was transformed into XL10-Gold Ultracompetent *E. coli* cells (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's protocol, plated on LB^{amp} plates and grown overnight. Clones were controlled for insert by PCR method. Further, positive clones were picked and grown overnight in 10 mL LB^{amp}, and plasmids extracted using the Promega plasmid prep kit (Promega, Madison, WI, USA). The sequencing reaction was performed at the DNA sequencing facility at the University of Bergen (www.seqlab.uib.no) using the manufacturer's protocol for the ABI PRISM BigDye Cycle Sequencing Ready Reaction Kit (Applied Biosystems).

Results

Patients. Among the 12 Sri Lankan patients with OSCC studied [3 (25%) females and 9 (75 %) were males], 10 (83%) were regular betel quid chewers, 6 (50%) were both cigarette smokers and alcohol users, while one (8%) patient was a regular smoker without other habits. The two (16%) non-betel quid chewers smoked and drank alcohol regularly. One of the 12 cancers had the verrucous variant of carcinoma while all the others were of squamous cell type. 16 (89%) of the Indian patients were males, and 2 (11%) were females. 12 (67%) were diagnosed with OSCC and 6 (33%) with OSF. One patient (5.5%) was a non-ST user, and 15 (83%) were chewing tobacco or betel quid. Two patients (11%) were smokers and ST-users, while the non-ST user was a smoker. One (5.5%) of the Indian patients was an alcohol user. For two of the patients, data on tobacco/alcohol habits were missing. The clinicopathological data are listed in Table I.

Chromosomal gains and losses common for the two populations. We searched for genomic chromosomal aberrations in 24 cases of OSCC and 6 cases of OSF using array-CGH, resulting in a total of 349 candidate genes found to be common for the two populations. These genes were located in 80 chromosomal regions. 34 of the 349 genes were found as deleted in both populations (Table II), 30 were amplificated (Table III), while 282 genes represented differentially altered regions in the tumors from Sri Lanka versus India (Data not shown but available on request). Of interest, 72 of the 349 candidate genes have been reported in our previous OSCC gene expression profiling/array-CGH studies from Sudan, Norway, Sri Lanka, Sweden and UK (Table IV). 14 chromosomal regions harbored the 34 candidate genes representing losses common for both populations. Among these was the chromosomal region 1q21, found to be deleted in a total of 70% of the patients,

Table II. *Chromosomal losses found in both populations*

Clone ID	Chromosome region	Basepair position	Gene sympol/Gene name	Sri Lanka no. (%)	India no. (%)	Total no. (%)
rp1-19k8	1q21	152,881,021-152,884,362	*IVL*/Involucrin	4 (30)	5 (28)	9 (30)
	1q21	153,330,330-153,604,513	*S100A1-9*/S100 calcium binding protein A1-A9	8 (67)	13 (72)	21 (70)
	1q21	153,346,184-153,348,125	*S100A12*/S100 calcium binding protein 12			
	1q21	153,591,263-153,606,873	*S100A13*/S100 calcium binding protein A13			
		153,586,731-153,589,462	*S100A14*/S100 calcium binding protein A14			
	1q21	153,579,362-153,585,644	*S100A16*/S100 calcium binding protein A16			
rp11-48o20	1q23.2	159,881,683-159,962,080	*PEA15*/phosphoprotein enriched in astrocytes 1	3 (25)	4 (22)	7 (23)
rp11-440p5	2p16.1	60,699,590-60,834,598	*REL*/V-rel reticuloendotheliosis viral oncogene homolog (avian)	12 (100)	4 (22)	16 (5)
rp11-59l22	2q32.3	193,207,674-193,365,513	*TMEFF2*/Transmembrane protein with EGF-like and two follistatin-like domains	8 (67)	16 (89)	24 (80)
rp1-137d17	6q27	169,746,291-169,834,193	*THBS2*/Thrombospondin 2	9 (75)	5 (27)	14 (47)
rp11-22c8	8q11.21	49,816,050-49,935,508	*SNAI2*/Snail homolog 2 (Drosophila)	6 (50)	4 (22)	10 (33)
rp11-150p21	8q22.3	103,734,091-103,889,486	*KLF10*/Kruppel-like factor 10	10 (83)	4 (22)	14 (47)
rp11-79f7	8q23.1	107,752,315-107,903,070	*ANGPT1*/Angiopioetin 1	7 (58)	3 (17)	10 (33)
rp11-16g11	8q24.13	122,724,076-122,806,813	*HAS2*/Hyaluronan synthase 2	10 (83)	9 (50)	19 (63)
rp6-98a24	8q24.22	134,057,314-134,184,861	*TG*/Thyroglobulin	4 (33)	10 (56)	14 (47)
	8q24.22	135,490,030-135,725,291	*WISP1*/WNT1 inducible signaling pathway protein 1			
	8q24.22	135,490,030-135,725,291	*NDRG1*/*N*-myc downstream regulated 1			
rp11-438p9	9q31.2	110,279,707-110,477,307	*KLF4*/Kruppel-like factor (gut)	6 (50)	4 (22)	10 (33)
rp3-416j11	11p15.5	2,105,785-2,182,749	*TSPAN32*/Tetraspanin32	10 (83)	5 (28)	15 (50)
rp11-179b7	11q22.3	104,298,333-104,459,820	*CASP4*/caspase 4, apoptosis-related cysteine peptidase *CASP5*/caspase 5, apoptosis-related cysteine peptidase	3 (25)	9 (50)	12 (40)
rp11-59h1	12p13.1	12,880,858-13,000,779	*CDKN1B*/Cyclin-dependent kinase inhibitor 1B (p27, Kip1)	9 (75)	10 (56)	19 (63)
rp11-13c13	12p13.2	10,232,160-10,398,790	*KLRK1*/Killer cell lectin-like receptor subfamily K, member 1	3 (25)	6 (33)	9 (30)
rp11-95g6	13q12.2	28,970,944-29,072,962	*CDX2*/Caudal type homeobox	11 (92)	5 (28)	16 (53)
	13q12.3	28,874,489-29,069,232	*FLT1*/Fms-related tyrosine kinase 1			
rp5-1025a1	20p11.21	25,008,412-25,102,971	*CST7*/cystatin F (leukocystatin)	10 (83)	4 (22)	14 (47)

harboring several of the *S100* gene family members including *S100A1, S100A2* and *S100A14*, which all have been reported in our earlier multi-national studies (20, 29, 30). We also found *REL* (V-rel reticuloendotheliosis viral oncogene homolog), an oncogene located in 2p16, and 8q23 which harbors *ANGPT1* (Angiopoietin 1), involved in angiogenesis, *WISP1* (WNT1 inducible signaling pathway protein 1), related to cell survival, found to be deleted by Roman *et al*. (20), and *NDRG1* (N-myc downstream regulated 1) involved in cell growth and differentiation. Two apoptosis-related genes, caspase 4 (*CASP4*) and caspase 5 (CASP5), located in chromosome 11q22 were also found to be lost. Gains found in both populations were located in 13 regions, hosting 33 candidate genes. Three genes were located in region 1p36, *EPHA2* (EPH receptor A2), *MFAP2* (Microfibrillar-associated protein 2) and *SDHB* (Succinate dehydrogenase complex, subunit B). Epha2 is related to angiogenesis and tumor neovascularization, while mfap2 is an ECM constituent. Sdhb is involved in oxidative phosphorylation and the citrate cycle. *NOTCH 4* (Notch 4 homolog, Drosophila), located in 6p21 is a member of the notch signaling network, reported in our earlier work (27), while *IL2RA* (Interleukin 2 receptor, alpha) encodes a cytokine located in region 10p15. We found three more cytokines, *TNFSF9, TFNSF14* and *CD70*, located in region 19p3. Matrix metallopeptidase 15, *MMP15*, is harbored by region 16q21. Protein kinase C alpha, *PRKCA* (17q24), is involved in a number of pathways, as with the apoptosis gene *BCL2* (B-cell CLL/lymphoma 2, located in region 18q21) and *VAV1* (vav 1 guanine nucleotide exchange factor) located in 19p3.

We also identified a large group of 66 chromosomal domains that were either deleted or amplified in both populations. To a large extent, the domains were deleted in the Sri Lankan tumors and amplified in the Indian tumors. Domain 1p36 harbors the genes *PIK3CD* (phosphoinositide-3-kinase) *MAD2L2* (MAD2 mitotic arrest deficient-like 2 (yeast), both involved in several biological pathways, *CDC42* (cell division cycle 42), and SFN (stratifin), both are cell cycle genes. *CSF3R*, colony stimulating factor 3 receptor, *RRAGC*, (ras-related GTP binding C) are both located in 1p34. Chromosome 2q22-24 hosts *ACVR2A* (activin A receptor, type IIA), *CXCR1* and *CXCR2*, all involved in cytokine-cytokine receptor interaction. *CDC25A* is a cell-

Table III. *Gains found in both populations*.

Clone ID	Chromosome region	Basepair position	Gene sympol/Gene name	Sri Lanka no. (%)	India no. (%)	Total no. (%)
rp11-45i3	1p36.13	16,450,832-16,482,582	*EPHA2*/EPH receptor A2	7 (58)	7 (39)	14 (47)
	1p36.13	17,307,997-17,307,330	*MFAP2*/Microfibrillar-associated protein 2			
	1p36.13	17,345,217-17,380,665	*SDHB*/Succinate dehydrogenase complex, subunit B			
rp1-93n13	6p21.32	32,162,620-32,191,844	*NOTCH4*/Notch homolog 4 (Drosophila)	5 (42)	5 (28)	10 (33)
	6p21.32	32,808,494-32,812,480	*PSMB8*/Proteasome (prosome, macropain) subunit, beta type, 8			
	6p21.32	32,781,544-32,806,599	*TAP2*/Transporter 2, ATP-binding cassette, sub-family B			
rp11-208g20	7q36.1	151,832,007-152,133,628	*MLL3*/Myeloid/lymphoid or mixed-lineage leukemia protein 3	8 (67)	5 (28)	13 (43)
	7q36.1	152,341,864-152,373,250	*XRCC2*/X-ray repair complementing defective repair in Chinese hamster cells 2			
	7q36.1	152,456,834-152,552,463	*ACTR3B*/ARP3 actin-related protein 3 homolog B (yeast)			
rp11-298k24	10p15.1	6,052,652-6,104,333	*IL2RA*/interleukin 2 receptor, alpha	3 (25)	4 (22)	7 (23)
		6,186,843-6,277,508	*PFKFB3*/6-phosphofructo-2-kinase/ fructose-2,6-biphosphatase 3			
rp11-138n3	11q13.2	67,250,512-67,258,574	*AIP*/Aryl hydrocarbon receptor interacting protein	4 (33)	6 (33)	10 (33)
	11q13.2	67,351,066-67,354,131	*GSTP1*/Glutathione S-transferase pi 1			
	11q13.2	67,820,326-67,888,736	*CHKA*/Choline kinase alpha			
rp11-110j12	12q24.23	118,103,333-118,262,962	*PEBP1*/Phosphatidylethanolamine binding protein 1	5 (42)	13 (72)	18 (60)
rp11-355d13	15q15.3	43,637,067-43,804,629	*TP53BP1*/Tumor protein p53 binding protein 1	4 (33)	4 (22)	8 (27)
rp11-282m16	15q23	67,835,021-68,099,461	*MAP2K5*/Mitogen-activated protein kinase kinase 5	5 (42)	8 (44)	13 (43)
	15q23	68,346,572-68,480,402	*PIAS1*/Protein inhibitor of activated STAT, 1			
rp11-481j2	16q21	58,408,000-58,582,474	*MMP15*/Matrix metallopeptidase 15 (membrane-inserted)	6 (50)	8 (44)	14 (47)
rp11-332h18	17q23.2	59,304,937-59,492,292	*BRIP1*/BRCA1-interacting protein C-terminal helicase 1	4 (33)	6 (33)	10 (33)
rp11-4f22	17q24.2	64,258,294-64,433,375	*PRKCA*/Protein kinase C, alpha	5 (42)	7 (39)	12 (40)
rp11-267c19	18p11.32	657,604-673,578	*TYMS*/Thymidylate synthetase	8 (68)	6 (33)	14 (47)
	18p11.32	721,588-812,327	*YES1*/v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1			
rp11-13I22	18q21.33	136,754,542-136,920,723	*BCL2*/B-cell CLL/lymphoma 2	4 (33)	8 (44)	12 (40)
ctd-3113p16	19p13.3	571,325-583,492	*BSG*/Basigin (Ok blood group)	4 (33)	6 (33)	10 (33)
	19p13.3	917,342-921,014	*KISS1R*/KISS1 receptor			
rp11-30f17	19p13.3	6,531,010-6,535,939	*TNFSF9*/Tumor necrosis factor (ligand) superfamily, member 9	9 (75)	4 (22)	13 (43)
	19p13.3	6,664,568-6,670,599	*TNFSF14*/Tumor necrosis factor (ligand) superfamily, member 14			
	19p13.3	6,583,311-6,591,163	*CD70*/CD70 molecule			
	19p13.3	6,772,722-6,857,371	*VAV1*/vav 1 guanine nucleotide exchange factor			

cycle gene located in chromosomal domain 3p21. Region 6p21 contains *DAXX* (death-domain associated protein), involved in the MAPK signaling pathway, *BAK1* (BCL2-antagonist/killer 1), apoptosis-related and *ITPR3* (inositol 1,4,5-triphosphate receptor, type 3), involved in many different pathways.

Further, *RAC1* (Ras-related C3 botulinum toxin substrate) and *NUDT1* (nudix-type motif 1) are hosted by chromosomal domain 7p22.3. *AZGP1* (alpha-2-glycoprotein) and *CUX1* (cut-like homeobox 1) are located in 7q22, while 7q31 harbors caveolin 1 and 2 (*CAV1* and *CAV2*), as well as *SPAM1* (sperm adhesion molecule 1). The caveolins 1 and 2 are associated with several malignancies, and spam1 is involved in tumor invasion and metastasis. A fibroblast growth factor receptor, *FGFR1*, is oriented in 8p11.22. Chromosome 9q21-22 contains *DAPK1* (death-associated protein kinase 1), *COL15A1* (collagen, type XV, alpha 1) and *TGFBR1* (transforming growth factor, beta receptor 1). Tenascin C (*TNC*), an ECM constituent, and *DEC1* (deleted in esophageal cancer 1) are hosted by 9q32. *COL13A1* (collagen, type VIII, alpha 1) and *PLAU* (plasminogen activator, urokinase) are related to cell proliferation and migration.

Alterations in chromosomal domain 11q13 represent a number of genes, among them *VEGFB* (vascular endothelial growth factor B), *BAD* (BCL2-associated agonist of cell death), *ESRRA* (estrogen-related receptor alpha), *BRMS1* (breast cancer metastasis suppressor 1), RHOD (ras homolog gene family, member 1), *CCND1* (cyclin D1), *FGF3* and

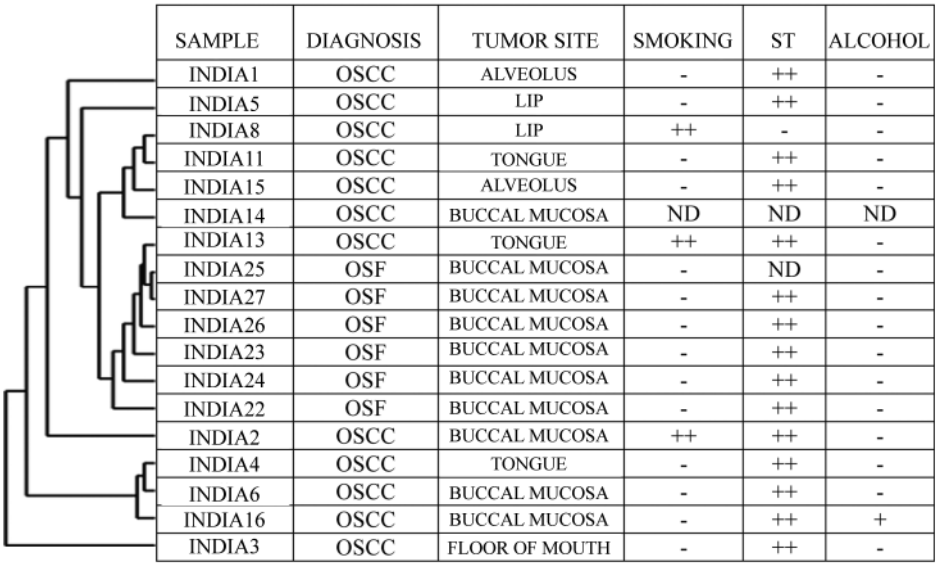
Table IV. *Genes reported both in our earlier work and in the present study.*

Gene symbol/Gene name	Reference
AZGP1/Alpha-2-glycoprotein, zinc binding	(31)
BAG4/BCL2-assoicated athanogene 4	(21)
BAK1/BCL2-antagonist/killer 1	(21)
BAX/BCL2-associated X protein	(37)
BCL2L1/BCL2-like 1	(21)
BPIL1/bactericidal/permeability-increasing protein-like 1	(31)
BRMS1/breast cancer metastasis suppressor 1	(31)
CAV1/Caveolin 1	(31, 54)
CAV2/Caveolin 2	(31)
CCND1/Cyclin D1	(21)
COL11A2/Collagen, type IX, alpha II	(21)
CTTN/cortacin	(21)
CD70/CD70 molecule	(37)
CDX2/Caudal type homeobox	(21)
CRK/v-crk sarcoma virus CT10 oncogene homolog (avian)	(21)
DAPK1/Death-associated protein kinase 1	(37)
DAXX/Death-domain associated protein	(21)
ECM1/Extracellular matrix protein 1	(21)
EIF4A/eukaryotic translation initiation factor 4A	(31)
EP300/E1A binding protein p300	(21)
FASLG/Fas ligand (TNF superfamily, member 6)	(21)
FASN/Fatty acid synthase	(30)
FAU/Finkel-Biskis-Reilly murine sarcoma virus (FBR-MuSV)	(21, 37)
FGF3/Fibroblast growth factor 3	(21)
FGF4/Fibroblast growth factor 4	(21)
FGFR1/Fibroblast growth factor receptor 1	(21)
FLT1/Fms-related tyrosine kinase 1	(21)
HAS3/Hyaluronan synthase 3	(31)
IVL/Involucrin	(21)
JUNB/Jun B proto-oncogene	(37)
LIF/Leukemia inhibitory factor	(21)
KLK11/Kallikrein 11	(31)
LYN/v-yes-1 Yamaguchi sarcoma viral related oncogene homolog	(21)
MCL1/Myeloid cell leukemia sequence 1	(21)
MTA2/metastasis associated 1 family, member 2	(21)
MAPK3/mitogen-activated protein kinase 3	(21)
NME1/Non-metastatic cells 1, protein (NM23A) expressed in	(31)
NME2/non-metastatic cells 2, protein (NM23B) expressed in	(37)
NOSIP/Nitric oxide synthase interacting protein	(31)
NOTCH4/Notch homolog 4 (Drosophila)	(37)
ORAOV1/Oral cancer overexpressed 1	(21)
OSM/Oncostatin M	(21)
PAX7/Paired box 7	(21)
PLAU/Plasminogen activator, urokinase	(31)
PRMT1/Protein arginine methyltransferase	(30, 31)
RHOD/Ras homolog gene family, member D	(31)
RRAGC/Rag C protein	(31)
S100A1/S100 calcium binding protein A1	(21, 31)
S100A2/S100 calcium binding protein A2	(21, 36)
S100A3-A9/S100 calcium binding protein A3-A9	(21)
S100A12/S100 calcium binding protein A12	(21)
S100A13-14/S100 calcium binding protein A13-14	(21)
SFN/Stratifin	(31)
STAR/steroidogenic acute regulatory protein	(21)
TAPBP/tapasin	(21)
TGFB1/transforming growth factor, beta 1	(37)
THBS2/Thrombospondin 2	(37)
TNFSF9/tumor necrosis factor (ligand) superfamily, member 9	(21)
TNC/Tenascin C	(29, 31)
TP53/Tumor protein 53	(21)
THY1/Thy-1 cell surface antigen1	(31)
VAV1/Vav 1 guanine nucleotide exchange factor	(21)
VEGFB/Vascular endothelial growth factor B	(21)
WISP1/WNT1 inducible signaling pathway protein 1	(21)
WNT10B/Wingless-type MMTV integration site family, member 10B	(36)

FGF4 (fibroblast growth factor 3 and 4). *CDK4*, cyclin-dependent kinase 4, is located in region 12q14, *CDK2AP1* (cyclin-dependent kinase 2 associated protein 1) in 12q24. *MMP14* (matrix metallopeptidase-14) and *BCL2L2* (BCL2-like 2) are oriented in region 14q11.2. Region 16p11 harbors *ITGAL* (integrin, alpha L) and *MAPK3* (mitogen-activated protein kinase 3), while *HAS3* (hyaluronan synthase 3) and *TERF2* (telomeric repeat binding factor 2).

Chromomosal domain 17p13 represents another altered region which harbors many genes, among others *HIC1* (hypermethylated in cancer 1), *CLDN7* (claudin 7), *EIF4A1* (eukaryotic translation initiation factor 4A, *CD68* and *TP53* (tumor protein 53). The deletion/amplification of larger parts of chromosome 17q affected several genes, *ERBB2* (v-erb-b2 erythroblastic leukemia viral oncogene homolog 2), *STAT3* (signal transducer and activator of transcription 3), *BRCA1* (breast cancer 1, early onset) *MAP3K14, NME1* (non-metastatic cells 1), and *FASN* (fatty acid synthase). Further, 19p13 was deleted or copied affecting genes like CDKN2D (cyclin-dependent kinase inhibitor 2D), JUNB (jun B proto-oncogene) and *JAK3* (janus kinase 3). 19q13 harbors *PLAUR* (plasminogen activator, urokinase receptor), BBC3 (BCL2 binding component), *BAX* (BCL2-associated X protein), NOSIP (Nitric oxide synthase interacting protein) and *PRMT1* (Protein arginine methyltransferase 1). Several of these genes were reported in our previous work, including *SFN, RRAGC, DAXX, BAK1, AZGP1, CAV1, CAV2, DAPK1, TNC, VEGF, FAU, CCND1, FGF3, FGF4, MMP14, EIF4A1, TP53, FASN, JUNB* and *BAX* (Table IV).

Chromosomal alterations in OSCC compared to OSF. In total, there were only a few chromosomal changes in the



SAMPLE	DIAGNOSIS	TUMOR SITE	SMOKING	ST	ALCOHOL
INDIA1	OSCC	ALVEOLUS	-	++	-
INDIA5	OSCC	LIP	-	++	-
INDIA8	OSCC	LIP	++	-	-
INDIA11	OSCC	TONGUE	-	++	-
INDIA15	OSCC	ALVEOLUS	-	++	-
INDIA14	OSCC	BUCCAL MUCOSA	ND	ND	ND
INDIA13	OSCC	TONGUE	++	++	-
INDIA25	OSF	BUCCAL MUCOSA	-	ND	-
INDIA27	OSF	BUCCAL MUCOSA	-	++	-
INDIA26	OSF	BUCCAL MUCOSA	-	++	-
INDIA23	OSF	BUCCAL MUCOSA	-	++	-
INDIA24	OSF	BUCCAL MUCOSA	-	++	-
INDIA22	OSF	BUCCAL MUCOSA	-	++	-
INDIA2	OSCC	BUCCAL MUCOSA	++	++	-
INDIA4	OSCC	TONGUE	-	++	-
INDIA6	OSCC	BUCCAL MUCOSA	-	++	-
INDIA16	OSCC	BUCCAL MUCOSA	-	++	+
INDIA3	OSCC	FLOOR OF MOUTH	-	++	-

Figure 1. *Hierarchical clustering of the Indian samples in relation to patients' clinocopathological data. The OSF samples are clustered in the same subgroup, including only one OSCC sample (India 13). There was no other obvious relationship between sample clustering and clinicopathological data.*

OSCC samples that were not found in the OSF samples, and none affecting only OSFs. There were more deletions than amplifications exclusive to OSCC, predominantly in chromosome 8q21-24. Chromosomes 7p11, 9p24 and 20p11 were also among the altered regions. Genes of interest located in these regions included among others *EGFR*, shown to be over-expressed in oral cancer and other malignancies, leading to cell proliferation and survival (31-35). This gene has frequently been approached as a therapeutic target in a number of cancers, including OSCC (33, 34). Hierarchical clustering showed that the OSF samples clustered together in one subgroup, only including one of the OSCC samples (Figure 1).

IHC results. IHC was performed and analyzed in tissue samples showing normal, dysplastic and malignant changes. In the adjacent normal mucosa, S100A14 protein was found to be strongly expressed in the cell membrane of the epithelial cells. Some of the epithelial cells showed cytoplasmic staining, but none showed nuclear staining. To a large extent, the dysplastic lesions showed the same expression as the adjacent normal tissues, but the expression was more heterogeneous in its strength in the dysplastic areas. Across the OSCC samples, a variable pattern of staining was seen in each tumor. The S100A14 protein expression was found to be clearly down-regulated and sometimes lost in the OSCC samples, particularly in the invading tumor islands. The sub-cellular S100A14 expression pattern was also seen to be changed from plasma membrane to the cytoplasm, as shown in Figure 2.

RFLP results. PCR products including the SNPs 461 G>A and 1545 A>T were digested with *Kpn* I and *Mnl* I, respectively. The results of the RFLP are presented in Table V and Figure 3. For the 461 G>A SNP, there was a tendency for the Asian (India/Sri Lanka) samples and Scandinavian (Norway, Finland) to be dominated by the heterozygote genotype GA and homozygote GG. 46% of the Scandinavian samples had the GA genotype, while 74% of the Asian samples (both OSCC and OSFs) represented the heterozygote. For the African (Sudanese) samples, there was an even distribution of the heterozygotic and the homozygotic GG genotypes, but no cases of the AA genotype were found. When separating OSF from the OSCC, there was a significantly higher number of the A allele represented among the OSF samples compared to OSCCs.

Discussion

In the present study, we employed array-CGH to profile chromosomal imbalances in cases of OSCC and OSF from South-East Asian countries, in particular, India and Sri Lanka, and identified significantly altered 80 chromosomal regions including 1p36, 1q21, 2p, 7q, 9q, 10q, 11q13, 16p11, 17p13, 17q12-22, 19p13, 19q13 and 20q11. The chromosomal regions identified were found to harbor 349 candidate genes. Out of these genes, 34 (Table II) were found to be located in deleted regions in the samples from both populations (including several S100 gene family members, *REL, ANGPT1, WISP1, CASP4* and *CASP5*), 30

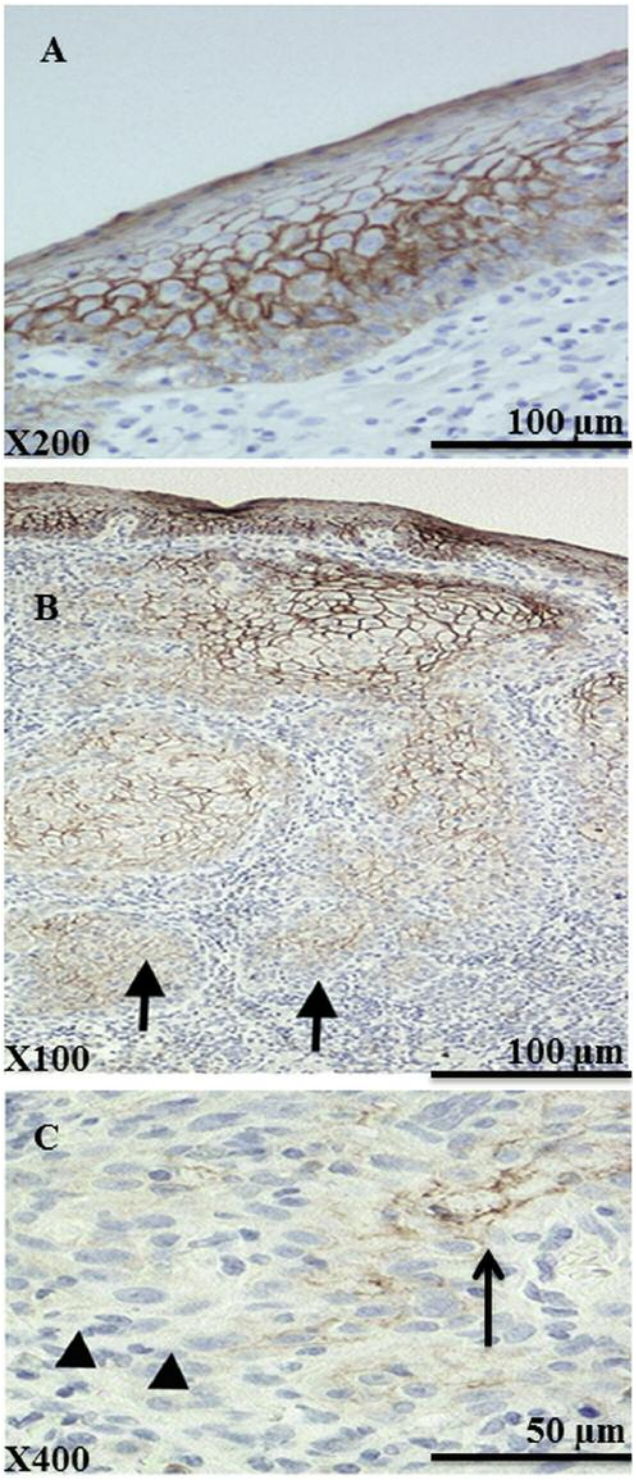


Figure 2. *Expression of the S100A14 protein in NHOM and OSCC. A. NHOM, showing strong epithelial membranous expression of the S100A14. B. Invasive OSCC showing strong expression of the S100A14 in the centre of the tumor and weak expression in the invasive front. C. Invasive OSCC showing reduced to lost expression of the S100A14.*

(Table III) were found to be located in amplified chromosomal regions (including *NOTCH4, XRCC2, MMP15* and *BCL2*) while 285 (Data not shown but available on request) others were located in chromosomal regions as either deleted in one population or amplified in the other. Examples of these genes included *MTOR, SFN, RRAGC, EMC1, CDC25C, LYN, TNC, PLAU, BAD, FGF3, FGF4, CDK4, MAPK3, EIF4A1, TP53, ERBB2, BRCA1, JUNB, JAK3, BAX* and *PRMT1*. Further, we found similar results for 72 of the 349 candidate genes identified either as up- or down-regulated or showing similar chromosomal changes in our previous studies (Table IV) of OSCCs from Western, Asian and African populations (20, 29, 30, 36, 37). These included among others *S100A14*, several of the other S100 gene family members, *WISP1, NOTHC4, TNFSF9, VAV1, RRAGC, ECM1, BAK1, AZGP1, CAV1, CAV2, LYN, DAPK1, VEGFB, FAU, CCND1, FGF3, FGF4, EIF4A1, TP53, FASN, JUNB, PRMT1* and *EP300*. These genes represent both genes with a known relation to OSCC development, and genes novel to this malignancy, and differ from those found in other epithelial tumors. Our current findings strengthen the results of our earlier studies and open the doot for further studies of several of these genes.

Located in chromosome 1q21 and deleted in 70% of the total number of the cases examined, we found a group of *S100* calcium-binding genes. These genes are known to be involved in the development of different human cancers, and have been suggested to play various roles in association with *p53* functions, cell-cycle arrest and tumor invasion (38). Altered gene expression/DNA aberrations were reported in *S100* genes in our previous studies (21, 31, 37, 38). Based on this, we further extended our focus on the S100A14 protein, a calcium-binding protein of the S100 protein family located in chromosome 1q21 found to be differentially expressed in a number of tumors (39). In this study, the chromosomal 1q21 region harboring *S100A14* was found to be deleted, while in a previous CGH study performed on OSCCs from Sudan and Norway, an amplification of the 1q21 region has been reported (21). An in-depth study on gene expression profile of *S100A14* and 18 other *S100* gene family members in OSCCs from Sudanese patients showed down-regulation of *S100A14* (3836). This result is in agreement with others using esophageal SCC (40). Furthermore, a recent work by our group showed that over-expression of S100A14 in OSCC-derived cells harboring wild-type p53 induced G1-arrest and hereby suppressing cell proliferation (39). In the current work, we selected *S100A14* gene that has been found as deleted in samples from both populations and examined its expression and genetic variants by immunohistochemistry and RFLP in multi-national archival samples. By IHC, we demonstrated a decreased expression of the S100A14 protein in OSCC tissues compared to normal oral mucosa, and a change in sub-

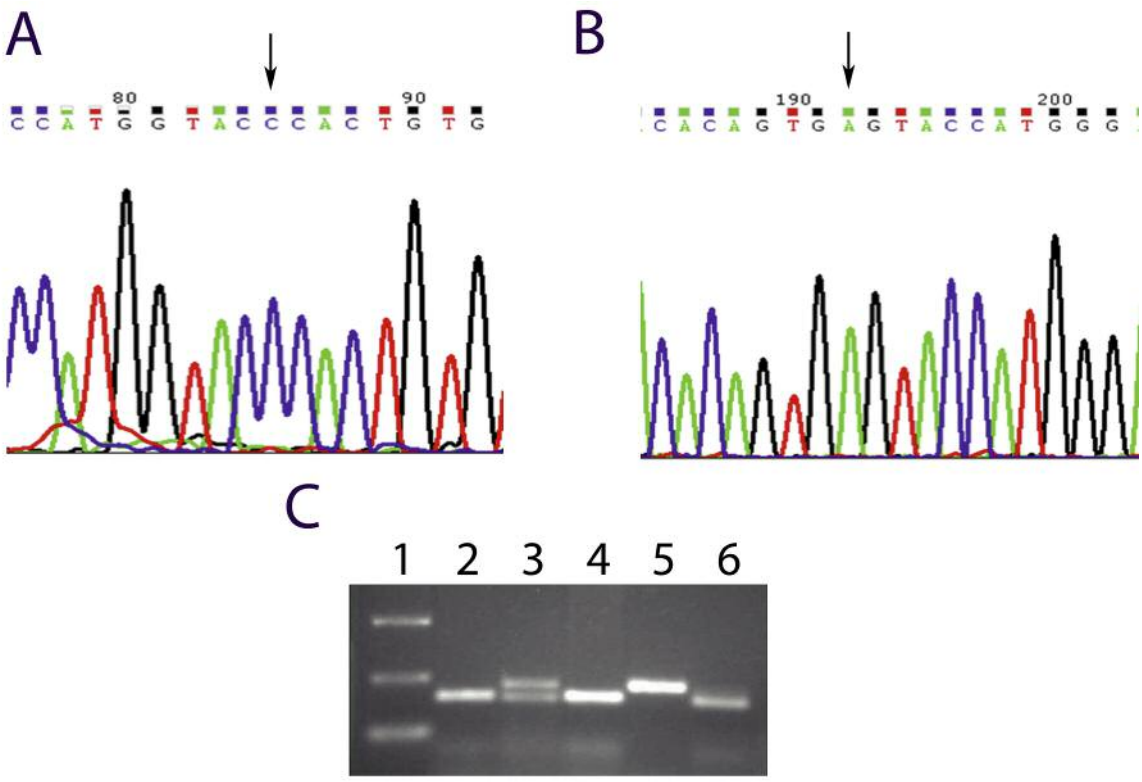


Figure 3. *DNA sequencing and RFLP of the 461 G>A S100A14 SNP. A. 461 AA SNP, and 461 GG SNP (B). Arrows show the nucleotide position where base change has occurred. C. RFLP of the genetic variants of S100A14. Lane 1: DNA ladder; Lane 2: GG genotype; Lane 3: GA genotype; Lane 4: GG genotype; Lane 5: AA genotype; Lane 6: GG genotype.*

cellular localization from the plasma membrane to the cytoplasm. The decreased expression found confirms our findings of deletions in chromosomal region 1q21, where *S100A14* is located. RFLP resulted in a differential distribution between premalignant/African samples and Western/Asian samples in one SNP. These results emphasize the essential role of *S100A14* in OSCC tumorigenesis.

Chen *et al*. (40) identified four different SNPs in the *S100A14* gene, one of these being related to a p53 binding site, reducing the expression of S100A14. Bearers of this SNP, 461 G>A, had an increased risk of developing esophageal cancer (40). Also, S100A14 was found to regulate the invasive potential of two OSCC cell lines, involving down- or up-regulation of matrix metalloproteinases *MMP1* and *MMP9* (28). In this study, we have chosen to focus on S100A14, and have performed immunohistochemistry on the Sri Lankan samples, and furthermore SNP analysis using samples from different populations. We performed the RFLP study examining two SNPs in the *S100A14* gene, mainly 461 G>A and 1545 A>T (Table V). Interestingly, the OSF samples from India were dominated by the heterozygotic genotype in the case of SNP 461 G>A, with very few samples harboring the homozygotic alleles. The European samples harbored the GG (38%) and GA (46.5%) genotype in a larger proportion than the AA genotype (15.5%). Interestingly, none of the Sudan samples harbored the AA genotype, but were found to harbor only the GA or GG genotypwa, distributed evenly. Chen *et al*. demonstrated that patients bearing the 461A allele have a higher risk of developing esophageal cancer, particularly smokers bearing the AA genotype, and that this allele is associated with reduced expression of S100A14. Our results show that there is a larger proportion of patients from Asia and Europe harboring the A allele in total (89% and 62%, respectively) than in the Sudan (47%). This might be related to the different tobacco habits practiced in these regions. Our results also indicate that there is also a high risk of developing OSCC in patients with OSF harboring the 461A genotype, as 94% of the OSF cases were showed to be bearers of the A allele. This observation warrants further in-depth studies with a focus on large number of OSF cases to be studied.

The tumor suppressor gene *TP53* is involved in many cellular processes, preventing abnormal cell division and

Table V. *Results of SNP investigation.*

SNP	461 G>A			
Population	GG	GA	AA	Total
Finland/Sweden/Norway	27	33	11	71
Sudan	36	32	0	65
India/Sri Lanka	5	11	3	19
OSF India	3	43	8	54
Total no of samples	71	119	22	212
SNP	1545 A>T			
Population	AA	AT	TT	Total
Finland/Sweden/Norway	10	26	13	49
Sudan	11	12	2	23
India/Sri Lanka	10	11	3	24
OSFM	1	10	13	24
Total no of samples	32	59	31	120

proliferation by inducing apoptosis, DNA repair, cell-cycle arrest and inhibition of angiogenesis (41). Chen *et al*. (40) showed that p53 regulates transcription of S100A14, and we suggested that the S100A14 protein may have a role in regulation of p53 activity by favoring nuclear accumulation of p53 (39). *TP53* mutations occur in a great number of human cancers, including more than 50% of OSCCs, leading to a dysfunctional protein. More than 70% of the mutations affect the DNA-binding site (42, 43). *p53* inactivation also occurs through alterations in genes encoding proteins involved in regulation of p53, like S100A14, or viral protein binding. Tobacco products contain several carcinogens responsible for mutations in *TP53*. For smokeless tobacco, the main tobacco habit among patients in this study, tobacco-specific nitrosamines (TNSAs) like NNK and NNN are the strongest carcinogens (43, 44). Previously, we showed a relationship between use of oral snuff (toombak) in Sudan and *p53* mutations (43). Also, betel quid or smokeless tobacco chewing may be responsible for *p53* mutations (45-47). Alterations in *p53*-hosting chromosome part 17p13.1 were found in this study, deleted in the Sri Lankan tumors and amplified in the Indian tumors. Our study involving OSCC cases from Sudan and Norway reported this region as amplified (20).

Further, S100A14 was demonstrated to up-regulate p21 and p27 protein expression in CalH3 cells, leading to G1-arrest (39, 41). p21, encoded by the gene *CDKN1A*, and regulates cell cycle progression through inhibition of Cyclin E/CDK2 and CyclinD1/CDK4 complexes, preventing cell cycle G_1/S transition. *CDKN1B*, encoding protein p27, also regulates G_1/S transition to binding of the cyclin/CDK complexes. In this study, the chromosomal region 12p13.1, harboring the CDKN1B gene, was found to be deleted in 75% of the Sri Lankan and 56% of the Indian samples, leading to decreased activity of p27. CyclinD1, binding to CDK4 and CDK6 to regulate G_1/S-transition, were also among the genes located in altered regions in this study, deleted in Sri Lankan samples, but amplified in Indian samples. In our previous array-CGH study with Norway/Sudan OSCCs, the chromosomal region of this gene was amplified (20), and in another study including pre-malignant and malignant oral lesions cyclin D1 was found as over-expressed (49-47). Our findings of alterations in these genes, important for cell cycle progression, may indicate their important role in tumorigenesis of OSCC.

Among the other genes of specific interest found in this and our earlier OSCC microarray/array-CGH work, are *CAV1* and *CAV2*, encoding caveolin 1 and 2, structural components of the caveolae, 50-100 nm flask shaped invaginations of the plasma membrane in a variety of cells, abundant in adipocytes, vascular endothelial cells, fibroblasts and epithelial cells (48-50). The caveolae are involved in vesicular transport and cellular signaling, and the caveolins are important for the caveolae functions as they are scaffolding proteins interacting with signaling molecules and their binding partners, bringing them in close proximity with one another (48-50). *Caveolin 1*, and to a certain extent *caveolin 2*, are implicated in several malignancies (51-53), and have been reported as up-regulated in oral cancer in our earlier work (29, 53). *WISP1* (WNT1 inducible signaling pathway protein 1), a member of the cysteine-rich, glycosylated signaling protein family located in chromosome 8q24, associated with cell proliferation, attenuation of p53-mediated apoptosis in response to DNA damage. Wisp1 is highly expressed in fibroblasts, and has been reported to be over-expressed in colon cancer, and may promote tumor growth (54-56). *VAV1* (Vav 1 guanine nucleotide exchange factor) is a hematopoietic cell specific oncogene and a guanine nucleotide exchange factor (GEF) involved in regulation of Rho GTPases, again organizing the actin cytoskeleton. *Vav 1* has been suggested to play a role in T cell development and activation, and has also been suggested to play a role in tumor progression (55). Both *WISP1* and *VAV1* were found to be altered in tumor samples previously examined from Sudan and Norway (20). *VEGFB*, vascular endothelial growth factor B, also reported by Roman *et al*, encodes a protein playing a role in angiogenesis and stimulation of endothelial-cell growth and migration. Vegbf is found to be over-expressed in several malignancies, and is a ligand for *FLT1* (Fms-related tyrosine kinase 1, also named *VEGFR1*). Flt1 is, not surprisingly, also involved in angiogenesis, and may be a possible anti-angiogenic therapeutic target (56). TNC (Tenascin C) encodes an ECM protein expressed during the wound healing process, and also

Table VI. *Altered chromosomal regions exclusive to OSCC compared to OSF samples.*

Clone ID	Chromosome region	Basepair position	Gene sympol/Gene name	OSCC no. (%)
Losses				
rp5-1091e12	7p11.2	55,086,714-55,324,313	*EGFR*/epidermal growth factor receptor	10 (83)
rp11-446h10	8q13.1	67,955,314-67,996,018	*COPS5*/COP9 constitutive photomorphogenic homolog subunit 5 (Arabidopsis)	4 (33)
rp11-21c17	8q13.2	70,378,859-70,573,150	*SULF1*/Sulfatase 1	6 (59)
rp11-27n21	8q21.13	80,947,749-81,143,467	*TPD52*/Tumor proteinD52	4 (33)
rp11-24h3	8q22.1	98,656,407-98,740,998	*MTDH*/Methaderin	4 (33)
rp11-150p21	8q22.3	103,264,501-103,425,069 103,661,007-103,668,130	*UBR5*/ubiquitin protein ligase E3 component n-recognin 5 KLF10/Kruppel-like factor 10	4 (33)
rp11-536k17	8q24.11	117,654,369-117,779,164	*EIF3H*/Eukaryotic translation initiation factor 3, subunit H	5 (42)
rp11-494n20	8q24.11	118,806,729-119,124,092	*EXT1*/Extostosin 1	5 (42)
rp11-532m24	8q24.13	125,486,979-125,500,155 125,563,031-125,740,730	*RNF139*/Ring finger protein 139 *MTSS1*/Metastasis suppressor 1	4 (33)
rp11-218i7	9p24.1	5,890,802-5,910,606	MLANA/Melan-A	4 (33)
rp11-283f6	9p24.1	6,757,641-7,175,648	*KDM4C*/lysine (K)-specific demethylase 4	4 (33)
rp11-153m9	16q22.2	71,392,619-71,424,341	*CALB2*/Calbindin 2	4 (33)
rp4-788l2	20p11.21	23,016,057-23,017,314	*SSTR4*/somatostatin receptor 4	5 (42)
rp5-1025a1	20p11.21	24,929,866-24,940,564	*CST7*/cystatin F	4 (33)
Gains				
rp11-241p12	8p23.1	9,413,424-9,639,856	*TNSK*/Tankyrase TRF1-interacting ankyrin-related ADP-ribose polymerase	5 (42)
rp1-127l4	22q12.3	32,340,447-32,353,590	*YWHAH*/tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	4 (33)

related to tumor growth, invasion and angiogenesis (57). This gene was reported to have an altered expression in OSCCs from Sri Lanka and India, and was down-regulated in an OSCC cell line over-expressing S100A14 protein. This study indicates a role for TNC in metastasis and invasion, related to *MMP9*, among others (30). A relation between MMP9 expression and tenascin C has been demonstrated in breast cancer (58). MMP9 also has a prognostic significance in tongue cancer (59).

When searching for differences in DNA copy number between OSCC and OSF samples, we found only a few clones with no alterations in the OSF cases compared to OSCC cases, and none that harbored chromosomal changes only in the OSF cases and none of the OSCCs. Predominantly, the alterations of chromosomal parts were deletions (Table VI). Interestingly, one of the genes localized in a deleted chromosomal region is epidermal growth factor receptor (*EGFR*), an oncogene involved in a number of malignancies, including oral cancer (31, 33, 60). Over-expression of EGFR can be related to tumorigenesis and metastasis, and is thought to be an early event in OSCC development, as over-expression also has been demonstrated in premalignant lesions (31, 33, 34, 60). Chen *et al* associated EGFR over-expression with poor survival in OSCC patients with a history of betel-quid chewing (61). The *EGFR* gene is an important therapeutic target for OSCC and other malignancies where different therapeutic strategies for *EGFR* in OSCC have been approached, with promising results (33, 60). Our results however contradict earlier findings, as *EGFR* is localized in a chromosomal region found to be deleted, and not found in OSF. However, the OSF sample size was too small to draw a certain conclusion from this result.

Conclusion

This study showed that array-CGH is a powerful tool in investigating genomic alterations in OSCC and OSF samples. The present work strengthened our previous findings of genes being common for OSCC regardless of ethnicity and life-style. This study and our previous work have demonstrated the down-regulation of expression of S100A14 in OSCC, suggesting that this protein might be used as a molecular biomarker in OSCC and OSF. Findings of differential distribution of the A allele in the SNP 461 G>A among OSCC and OSF from Asia, Africa and Europe can be related to tobacco habits and cancer susceptibility, particularly for OSF transition into OSCC.

Conflicts of Interest

The Authors declare that they have no competing interests.

Acknowledgements

We thank Inger Ottesen, Marte I. Flydal, Siren Hammer Østvold, Gunvor Øijordsbakken, Mamta Singh and Dipak Sapkota for their skilled technical assistance. We also thank Kjell Petersen and Pawel Sztromwasser at the Norwegian Microarray Consortium for data analysis assistance.

Refrences

1 Cancela MDC, Voti L, Guerra-Yi M, Chapuis F, Mazuir M and Curado MP: Oral cavity cancer in developed and in developing countries: Population-based incidence. Head & Neck *32(3)*: 357-367, 2009.
2 Ariyawardana A and Warnakulasuriya S: Declining oral cancer in Sri Lanka: Are we winning the war after heading the cancer league table? Oral Dis *17(7)*: 636-641. doi: 10.1111/j.1601-0825.2011.01809.x. Epub 2011 Jul 18, 2011.
3 Rogers SN, Brown JS, Woolgar JA *et al*: Survival following primary surgery for oral cancer. Oral Oncol *45(3)*: 201-211, 2009.
4 Warnakulasuriya S: Global epidemiology of oral and oropharyngeal cancer. Oral Oncol *45(4-5)*: 309-316, 2009.
5 IARC: Betel-quid and areca nut chewing and some areca-nut derived nitrosamines. IARC Monographs, International Agency for Research on Cancer, 2004.
6 Warnakulasuriya S, Johnson NW and Van Der Waal I: Nomenclature and classification of potentially malignant disorders of the oral mucosa. J Oral Pathol Med *36(10)*: 575-580, 2007.
7 Murti PR, Bhonsle RB, Pindborg JJ, Daftary DK, Gupta PC and Mehta FS: Malignant transformation rate in oral submucous fibrosis over a 17-year period. Community Dent Oral Epidemiol *13(6)*: 340-341, 1985.
8 Kerr AR, Warnakulasuriya S, Mighell AJ *et al*:. A systematic review of medical interventions for oral submucous fibrosis and future research opportunities. Oral Dis *17(Suppl 1)*: 42-57, 2011
9 Baudis M: Genomic imbalances in 5918 malignant epithelial tumors: an explorative meta analysis of chromosomal CGH data. BMC Cancer 7: 226, 2007.
10 Höglund M, Gisselsson D, Mandahl N *et al*: Multivariate analyses of genomic imbalances in solid tumors reveal distinct and converging pathways of karyotypic evolution. Genes Chromosomes Cancer *31*: 156-171, 2001.
11 Shingaki S, Takada M, Sasai K *et al*:. Impact of lymph node metastasis on the pattern of failure and survival in oral carcinomas. . Am J Surg *185*: 278-284, 2003.
12 Hermsen M, Joenje H, Arwert F *et al*: Assessment of chromosomal gains and losses in oral squamous cell carcinoma by comparative genomic hybridisation. Oral Oncol *33*: 414-418, 1997.
13 Ita M, Oga A, Hayatsu Y *et al*: The relationship of genetic aberrations detected by comparative genomic hybridization to DNA ploidy and tumor size in human oral squamous cell carcinomas. . J Oral Pathol Med *29*: 226-231, 2000.
14 Lin S-C, Chen Y-J, Kao S-Y, Hsu M-T, Lin C-H and Yang S-C: Chromosomal changes in betel-associated oral squamous cell carcinomas and their relationship to clinical parameters. Oral Oncol *38*: 266-273, 2002.
15 Noutomi Y, Oga A, Uchida K *et al*: Comparative genomic hybridization reveals genetic progression of oral squamous cell carcinoma from dysplasia via two different tumorigenic pathways. J Pathol *210*: 67-74, 2006.
16 Weber R, Scheer M, Born I *et al*: Recurrent chromosomal imbalances detected in biopsy material from oral premalignant and malignant lesions by combined tissue microdissection, universal DNA amplification and comparative genomic hybridization. . Am J Pathol *153*: 295-303, 1998.
17 Wolff E, Girod S, Liehr T *et al*: Oral squamous cell carcinomas are characterized by a rather uniform pattern of genomic imbalances detected by comparative genomic hybridisation. Oral Oncol *34*: 186-190, 1998.
18 Martin CL, Reshmi SC, Ried T *et al*:. Chromosomal imbalances in oral squamous cell carcinoma: examination of 31 cell lines and review of the literature. Oral Oncol *44(4)*: 369-382, 2008.
19 Patmore H, Cawkwell L, Stafford N and Greenman J: Unraveling the Chromosomal Aberrations of Head and Neck Squamous Cell Carcinoma: A Review. Annals of Surgical Oncology *12(10)*: 831-842, 2005.
20 Roman E, Meza-Zepeda L, Kresse S, Myklebost O, Vasstrand EN and Ibrahim SO: Chromosomal abberations in head and neck squamous cell carcinomas in Norwegian and Sudanese populations by array comparative genomic hybridization. Oncol Rep *20*: 825-843, 2008.
21 Scully C and Bagan J: Oral squamous cell carcinoma overview. Oral Oncology *45(4-5)*: 301-308, 2009.
22 Cawson RA and Eveson JW: Oral pathology and diagnosis. William Heinemann Medical Books: London, United Kingdom, 1987.
23 Pindborg JJ and Sirsat SM: Oral submucous fibrosis. Oral Surg Oral Med Oral Pathol *22(6)*: 764-779, 1966.
24 Meza-Zepeda LA, Kresse SH, Barragan-Polania AH *et al*: Array comparative genomic hybridization reveals distinct DNA copy number differences between gastrointestinal stromal tumors and leiomyosarcomas. Cancer Res *66(18)*: 8984-8993, 2006.
25 Wang J, Meza-Zepeda LA, Kresse SH and Myklebost O. M-CGH: analysing microarray-based CGH experiments. BMC Bioinformatics *5*: 74, 2004.
26 Costea DE, Loro LL, Dimba EAO, Vintermyr OK and Johannessen AC: Crucial effects of fibroblasts and keratinocyte growth factor on morphogenesis of reconstituted human oral epithelium. J Invest Dermatol *121(6)*: 1479-1486, 2003.
27 Dysvik B, Vasstrand EN, Løvlie R *et al*: Gene expression profiles of head and neck carcinomas from Sudanese and Norwegian patients reveal common biological pathways regardless of race and lifestyle. Clin Cancer Res *12(4)*: 1109-1120, 2006.
28 Sapkota D, Bruland O, Costea DE, Haugen H, Vasstrand EN and Ibrahim SO: S100A14 regulates the invasive potential of oral squamous cell carcinoma derived cell-lines *in vitro* by modulating expression of matrix metalloproteinases, MMP1 and MMP9. European Journal of Cancer *47(4)*: 600-610, 2011. doi: 10.1016/j.ejca.2010.10.012. Epub 2010 Nov 10, 2011
29 Lunde MLS, Warnakulasuriya S, Sand L, Hirsch JM, Vasstrand EN and Ibrahim SO: Gene expression analysis by cDNA microarray in oral cancers from two Western populations. Anticancer Res *30(4)*: 1083-1091, 2010.
30 Suhr ML, Dysvik B, Bruland O *et al*: Gene expression profile of oral squamous cell carcinomas from Sri Lankan betel quid users. Oncol Rep *18(5)*: 1061-1075, 2007.

31 Scully C, Field JK and Tanzawa H: Genetic aberrations in oral or head and neck squamous cell carcinoma 2: chromosomal aberrations. Oral Oncology *36(4)*: 311-327, 2000.
32 Tsantoulis PK, Kastrinakis NG, Tourvas AD, Laskaris G and Gorgoulis VG: Advances in the biology of oral cancer. Oral Oncol *43(6)*: 523-534, 2007.
33 Normanno N, DE luca A, Bianco C *et al*: Epidermal growth factor receptor (EGFR) signaling in cancer. Gene *366(1)*: 2-16, 2006.
34 Caponigro F, Milano A, Ottaiano A and Iaffaioli RV: Epidermal growth factor receptor as a major anticancer drug target. Expert Opin Ther Targets *10(6)*: 877-888, 2006.
35 Ibrahim SO, Aarsaether N, Holsve MK *et al*: Gene expression profile in oral squamous cell carcinomas and matching normal oral mucosal tissues from black Africans and white Caucasians: the case of the Sudan vs. Norway. Oral Oncol *39(1)*: 37-48, 2003.
36 Sapkota D, Bruland O, Bøe OE *et al*: Expression profile of the S100 gene family members in oral squamous cell carcinomas. Journal of Oral Pathology & Medicine *37(10)*: 607-615, 2008.
37 Pietas A, Schlüns K, Marenholz I, Schäfer BW, Heizmann CW and Petersen I: Molecular Cloning and Characterization of the Human S100A14 Gene Encoding a Novel Member of the S100 Family. Genomics *79(4)*: 513-522, 2002.
38 JI J, Zhao L, Wang X *et al*: Differential expression of S100 gene family in human esophageal squamous cell carcinoma. Journal of Cancer Research and Clinical Oncology *130(8)*: 480-486, 2004.
39 Sapkota D, Costea D, Blø M *et al*: S100A14 induces G1-arrest through nuclear accumulation of p53 and p53-dependent up-regulation of p21 in oral carcinoma derived cells. Oral Oncol *48(3)*: 219-225, 2012. doi: 10.1016/j.oraloncology.2011. 10.001. Epub 2011 Oct 26.42, 2012.
40 Chen H, Yu D, Luo A *et al*: Functional role of S100A14 genetic variants and their association with esophageal squamous cell carcinoma. Cancer Res *69(8)*: 3451-3457, 2009.
41 Goh AM, Coffill CR and Lane DP: The role of mutant p53 in human cancer. J pathol *223(2)*: 116-26. doi: 10.1002/path.2784. Epub 2010 Oct 25, 2011.
42 Hecht SS: Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nature Reviews 3(October), 2003.
43 Ibrahim SO, Vasstrand EN *et al*: Mutations of the p53 gene in oral squamous-cell carcinomas from Sudanese dippers of nitrosamine-rich toombak and non-snuff-dippers from the Sudan and Scandinavia. Int J Cancer *81(4)*: 527-534, 1999.
44 Chiba I, Muthumala M, Yamazaki Y *et al*: Characteristics of mutations in the p53 gene of oral squamous-cell carcinomas associated with betel-quid chewing in Sri Lanka. Int J Cancer *77(6)*: 839-842, 1998.
45 Ralhan R, Agarwal S, Nath N, Mathur M, Wasylyk B and Srivastava A: Correlation between p53 gene mutations and circulating antibodies in betel- and tobacco-consuming North Indian population. Oral Oncol *37(3)*: 243-250, 2001.
46 Warnakulasuriya KA and Ralhan R: Clinical, pathological, cellular and molecular lesions caused by oral smokeless tobacco--a review. J Oral Pathol Med *36(2)*: 63-77, 2007.
47 Raju B, Mehrotra R, Oijordsbakken G, Al-sharabi AK, Vasstrand EN and Ibrahim SO: Expression of p53, cyclin D1 and Ki-67 in pre-malignant and malignant oral lesions: association with clinicopathological parameters. Anticancer Res *25(6C)*: 4699-4706, 2007.
48 Chidlow JR J and Sessa W: Caveolae, caveolins and cavins: complex control of cellular signalling and inflammation. Cardiovascular Res *86*: 219-225, 2010.
49 Cohen AW, Hnasko R, Schubert W and LisantI MP: Role of caveolae and caveolins in health and disease. Physiol Rev *84(4)*: 1341-1379, 2004.
50 Engelman JA, Zhang X, Galbiati F *et al*: Molecular Genetics of the Caveolin Gene Family: Implications for Human Cancers, Diabetes, Alzheimer Disease, and Muscular Dystrophy. AmJHumGenet *63*: 1578-1587, 1998.
51 Ando T, Ishiguro H, Kimura M *et al*: The overexpression of caveolin-1 and caveolin-2 correlates with a poor prognosis and tumor progression in esophageal squamous cell carcinoma. Oncol Rep *18(3)*: 601-609, 2007.
52 Bau DT, Tsai MH, Tsou YA *et al*: The association of Caveolin-1 genotypes with oral cancer susceptibility in Taiwan. Ann Surg Oncol *18(5)*: 1431-8, 2011.
53 Ibrahim S, Miron T, Krohn M, Amaratunga A, Warnakulasuriya S and Vasstrand E: Tumour-associated Proteins in Oral Squamous Cell Carcinomas by Proteomics. Cancer Genomics Proteomics *2*: 353-364, 2005.
54 Pennica D, Swanson TA, Welsh JW *et al*: WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors. Proc Natl Acad Sci USA *95(25)*: 14717-14722, 1998.
55 Hornstein I, Alcover A and Katzav S: Vav proteins, masters of the world of cytoskeleton organization. Cell Signal *16(1)*: 1-11, 2004.
56 Fischer C, Mazzone M, Jonckx B and Carmeliet P: FLT1 and its ligands VEGFB and PlGF: drug targets for anti-angiogenic therapy? Nat Rev Cancer *8(12)*: 942-956, 2008.
57 Lyons AJ and Jones J: Cell adhesion molecules, the extracellular matrix and oral squamous carcinoma. Int J Oral Maxillofac Surg *36(8)*: 671-679, 2007.
58 Kalembeyi I, Inada H, Nishiura R, Imanaka-Yoshida K, Sakakura T and Yoshida T: Tenascin-C upregulates matrix metalloproteinase-9 in breast cancer cells: direct and synergistic effects with transforming growth factor beta1. Int J Cancer *105(1)*: 53-60, 2003.
59 Bello IO, Soini Y and Salo T: Prognostic evaluation of oral tongue cancer: means, markers and perspectives (II). Oral Oncol *46(9)*: 636-643, 2010.
60 Moon C, Chae YK and LEE J: Targeting epidermal growth factor receptor in head and neck cancer: lessons learned from cetuximab. Exp Biol Med (Maywood) *235(8)*: 907-920, 2010.
61 Chen IH, Chang JT, Liao CT, Wang HM, Hsieh LL and Cheng AJ: Prognostic significance of EGFR and Her-2 in oral cavity cancer in betel quid prevalent area cancer prognosis. Br J Cancer *89(4)*: 681-686, 2003.

Received May 1, 2014
Revised June 9, 2014
Accepted June 10, 2014