Journal of Cancer Science & Therapy
ISSN: 1948-5956

Kondoh et al., J Cancer Sci Ther 2012, S18
DOI: 10.4172/1948-5956.S18-001

Research Article Open Access

Comprehensive Study of Oral Squamous Cell Carcinoma Patients Using Blood Samples and Gene Expression Profiles

Nobuo Kondoh[1]*, Eiji Takayama[1], Masako Kamiya[1], Harumi Kawaki[1], Masayuki Motohashi[2], Yasunori Muramatsu[2], Michio Shikimori[2], Kenji Mitsudo[3] and Iwai Tohnai[3]

[1]Department of Oral Biochemistry, Asahi University School of Dentistry, Mizuho-shi, Hozumi 1851Gifu 501-0296, Japan
[2]Department of Oral Surgery, Asahi University School of Dentistry, Mizuho-shi, Hozumi 1851Gifu 501-0296, Japan
[3]Department of Oral and Maxillofacial Surgery, Yokohama City University Graduate School of Medicine, Yokohama-shi, Kanagawa 236-0004, Japan

Abstract

Oral squamous cell carcinoma (OSCC) is an aggressive malignancy which shows a variable degree of malignant behavior. To identify molecular signatures and establish a new diagnostic model for oral malignancies, we have identified marker genes representing pre-malignant and malignant phenotypes of oral mucosal lesions. The expression of marker genes was examined by quantitative reverse transcription-PCR. Then, we created discriminatory predictor models using Fisher's linear discriminant analysis and leave-one-out cross validation. These models were applicable for the diagnoses of pre-malignant leukoplakias (LPs), and of invasion status for advanced OSCCs.

The clinical course of various cancers is also influenced by host immune response. Our preliminary data using flow cytometric analysis demonstrate that the percentage of $CD4^+CD57^+$ T cells in peripheral blood lymphocyte was higher in the high grade OSCCs than that in the low grade ones. Furthermore, lipopolysaccharides (LPS)-induced *ex-vivo* production of Interferon (IFN)-γ from peripheral blood cells was highest in stage I patients and gradually decreased during the course of OSCC progression up to stage III. These decreased levels in the early stages were inversely correlated with tumor size.

In this review, we propose that the usage of the immunological status of OSCC patients combined with the molecular signatures of tumor tissues could provide valuable indices for diagnosis of oral malignancies.

Keywords: Oral squamous cell carcinoma (OSCC); Leukoplakia (LP); Interferon (IFN)-γ $CD4^+CD57^+$ T cells

Introduction

Leukoplakias (LPs) are white lesions that include hyperplasias and dysplasias of the oral mucosa, and often undergo malignant transformation to oral squamous cell carcinoma (OSCC) [1]. The histological features associated with LP, and with OSCC, have been described previously [2]. The dysplasias are classified as mild, moderate or severe, based on histopathological findings, and these designations are thought to be the sequential phases of oral carcinogenesis. However, we sometimes experience discordance between the pathological diagnosis of these lesions and the corresponding prognosis in an individual patient. It has been suggested that the detection of clonal genetic changes, such as loss of heterozygosity or microsatellite instability, in both primary OSCC and pre-malignant lesions could be a more informative method of monitoring these cancer patients [3]. The staging and grading of OSCCs are still, for the most part, dependent on traditional clinicopathological observations [3]. In line with our approach [4], there are several attempts for the identification of molecular biomarkers associated with specific phenotypes of head and neck squamous cell carcinomas [5-7]. Although, these approaches are effective to observe the wide variety of genetic alterations in oral malignancy, in order to accurately diagnose cancer subtypes, supervised learning methods are the most suitable [8-11]. Thus, we attempted to generate a multigene classifier for the diagnosis of pre-malignant-to-malignant transition. Using cDNA microarray and QRT-PCR techniques, a comprehensive gene expression profile was generated and compared among OSCC and LP tissues. We subsequently defined a list of 27 marker genes that are either significantly elevated or downregulated in OSCCs, compared with LPs. Among these genes, predictor gene sets for OSCC-LP classification were determined by Fisher's linear discriminant analysis (LDA) and validated by the leave-one-out cross validation [12].

On the other hand, OSCC is an aggressive malignancy which shows a variable degree of malignant behavior. The biological characteristics of this cancer are not yet well understood. We also attempt to generate diagnostic classifiers for OSCCs of an advanced stage [13]. Metastasis is a major cause of local relapse and death after definitive therapy in patients with these tumors. There are several reported approaches for the identification of molecular biomarkers and for predicting patients that are at high risk of recurrence [9,14-18]. Among the conventional staging and grading systems used for OSCC tissues, we focused upon the Yamamoto-Kohama's (YK's) mode of invasion [19], since the criteria involved can be largely correlated with prognosis, particularly in the case of lymph node metastases [20,21]. However, pre-operative clinicopathological estimations of the invasion status of these lesions are often inaccurate, because of the difficulty in biopsy sampling of the deepest portion of the invasive front, which is closely related to lymph node metastasis. It has been suggested that the genetic backgrounds

***Corresponding author:** Nobuo Kondoh, Department of Oral Biochemistry, Asahi University School of Dentistry, Mizuho-shi, Hozumi 1851, Gifu 501-0296, Japan, Tel: 81 58 1416; Fax: 81 58 329 1417; E-mail: nkondoh@dent.asahi-u.ac.jp

Received November 01, 2012; **Accepted** November 24, 2012; **Published** November 26, 2012

Citation: Kondoh N, Takayama E, Kamiya M, Kawaki H, Motohashi M, et al. (2012) Comprehensive Study of Oral Squamous Cell Carcinoma Patients Using Blood Samples and Gene Expression Profiles. J Cancer Sci Ther S18: 001. doi:10.4172/1948-5956.S18-001

associated with malignant phenotypes are equally responsible for the bulk of the resulting primary tumor [22]. Hence, using primary OSCCs, we have attempted to generate molecular classifiers that can predict the YK's mode of invasion using Fisher's LDA, and have evaluated the diagnostic significance of our findings. Our present results suggest that these differentially expressed genes can provide valuable prognostic tools for LP or OSCC patients.

Although there are several attempts to molecularly diagnose advanced tumors, the predictive values for metastases based on the gene expression profiles of various cancer cell types can vary from 75% to 100% [9,15,17]. It has been postulated that several discrete steps comprise the biological cascade leading to tumor metastasis, including the evasion of growth suppression, invasiveness, motility, detachment, angiogenesis, the contribution of tumor-associated macrophages, lymphangiogenesis, vascular adhesion, homing, and resistance to the innate immune response [23]. These should be defined not only by the phenotype of the cancer cells themselves, but also by the conditions of the host microenvironment [23]. We had previously reported that a specific subset of T-lymphocyte, blood cytokine levels and cytokine-producing capability of peripheral blood cells (PBCs) are well correlated with the progression of gastric cancer [24] and hepatocellular carcinoma [25]. Therefore, we attempt to diagnose the malignancy by evaluating systemic sign based on immunological status of OSCC patients. Hence, we attempt to evaluate general conditions of OSCC patients.

Corroborating information from both local tissues and patient's conditions could be meaningful to accomplish quite accurate diagnosis of oral malignancy.

Discrimination of Oral LP Subtypes and OSCC Using Gene Expression Signatures

The identification of potential marker genes between LP and OSCC

To identify potential marker genes that are differentially expressed between LP and OSCC, cDNA microarray analyses were performed using RNA mixtures of 5 OSCC and of 5 LPs [12]. Among the 16,600 target cDNAs on the chip arrays, 63 genes were highly expressed (3-fold or more) in the OSCC mixture, compared with the LP mixture. In addition, 55 genes were preferentially expressed (3-fold or more) in the LP mixture [12]. To validate the identified marker gene candidates, QRT-PCR analysis using several OSCC and LP tissues was performed. We focused finally on 27 differentially expressed marker genes, among which 15 (denoted as LP) were overexpressed in LPs compared to OSCCs, whilst 12 genes were upregulated in OSCCs and in some moderate- to severe-dysplasias (denoted as SC) (Table 1).

Supervised classification based on Fisher's LDA

The goal of this study is to establish a clear distinction between LP and OSCC, based upon a molecular classification. In order to identify marker gene sets that could discriminate between OSCCs and LPs, a supervised classification approach using LDA was performed [12]. The expression of these 27 marker genes was analyzed among 27 OSCCs

[#]Gene ID	Accession No.	Gene name (symbol)	Category	[$]Classifier
LP1	NM_000900	Matrix Gla protein (MGP)	Extra Cellular Matrix	A
LP8	NM_021005	Nuclear receptor subfamily 2, group F, member 2 (NR2F2)	Nuclear Receptor	A
LP12	NM_003062	Slit homolog 3 (Drosophila) (SLIT3)	Secretory Protein	A
LP21	NM_006121	Keratin 1 (KRT1)	Epithelial/Cytoskeretal	A
LP29	X14640	Keratin 13 (KRT13)	Epithelial/Cytoskeretal	A
LP17	NM_002725	Proline arginine-rich end leucine-rich repeat protein (PRELP)	Extra Cellular Matrix	
LP19	NM_016190	Chromosome 1 open reading frame 10 (C1orf10)	Epithelial/Protein modification	
LP27	NM_006732	FBJ murine osteosarcoma viral oncogene homolog B (FOSB)	Trans cription factor	
LP5	NM_001937	Dermatopontin (DPT)	Extra Cellular Matrix	
LP28	NM_003245	Transglutaminase 3 (TG3)	Epithelial/Protein modification	A
LP4	NM_022003	FXYD domain containing ion transport regulator 6 (FXYD6)	Membrane Protein/Receptor	A
LP15	NM_002404	Microfibrillar-associated protein 4 (MFAP4)	Extra Cellular Matrix	
LP22	NM_001387	Dihydropyrimidinase-like 3 (DPYSL3)	Metabolism	
LP2	NM_001831	Clusterin (CLU)	Anti Apoptotic	
LP16	NM_001311	Cysteine-rich protein 1 (intestinal) (CRIP1)	Immune response/Development	
SC1	NM_001565	Chemokine (C-X-C motif) ligand 10 (CXCL10)	Interferon-induced	A
SC5	NM_005562	Laminin, gamma 2 (LAMC2), transcript variant 1	Extra Cellular Matrix	A
SC13	NM_006350	Follistatin (FST), transcript variant FST317	Secretory protein	A
SC43	NM_144646	Immunoglobulin J polypeptide, linker protein (IGJ)	Secretory protein	A
SC3	NM_002899	Retinol binding protein 1, cellular (RBP1)	Nuclear Receptor	
SC27	NM_198966	Parathyroid hormone-like hormone (PTHLH)	Secretory protein	
SC41	NM_000358	Transforming growth factor, beta-induced, 68kDa (TGFBI)	Secretory protein	
SC9	NM_033255	Epithelial stromal interaction 1 (EPSTI1)	Epidermal specific	
SC6	NM_006417	Interferon-induced protein 44 (IFI44)	Interferon-induced	
SC44	NM_001549	Interferon-induced protein with tetratricopeptide repeats 3 (IFIT3)	Interferon-induced	
SC10	NM_017414	Ubiquitin specific protease 18 (USP18)	Ubiquitine proteasome	
SC7	NM_000067	Carbonic anhydrase II (CA2)	Metabolism	

[#]Genes upregulated in leukoplakia (LP) or oral sqamous cell carcinoma (SC)
[$]A, 11 predictor genes that can discriminatte between LPs and OSCCs

Table 1: Leukoplakia-(LP) and OSCC-(SC) dominant marker genes.

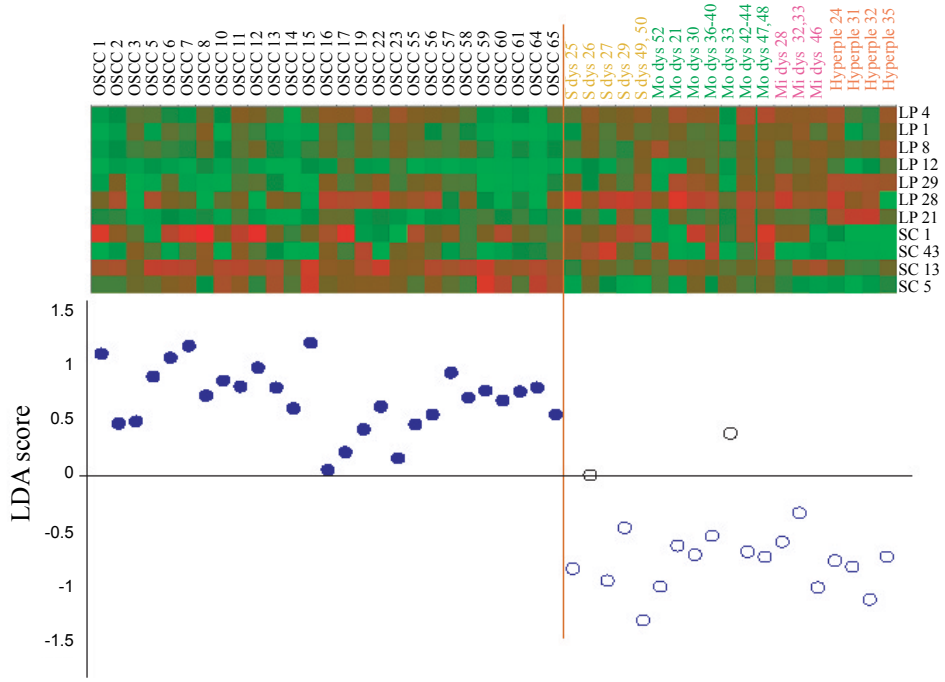

Figure 1: Expression profiles of 11 predictor genes and discrimination by LDA scores. The relative gene expression levels are represented as color intensity on a logarithmic scale (high, red; low, green; black, medium). The LDA score for each sample is given by the linear discrimination function. Filled circles, OSCCs; open circles, LPs. This figure was quoted from [12].

and 19 LPs, including hyperplasias and dysplasias. This approach involves parameter (gene) selection by the use of a stepwise increment and genetic algorithm. When the Fisher's ratio was employed as a score, a model with 11 parameters was selected as the best model. The stability of this model was examined by the leave-one-out cross validation (loo) method. The LDA score for each sample is given as the following linear discrimination function:

Score= - 0.231 (LP1) + 0.223 (LP4) - 0.0537 (LP28) - 0.0734 (LP21) - 0 .892 (LP12) - 0.0617 (LP29) - 0.282 (LP8) + 0.0122 (SC1) + 0.0669 (SC13) - 0.0684 (SC43) - 0.0366 (SC5).

The score was obtained when the levels of 11 genes were substituted into the equation. As shown in figure 1, the scores for OSCCs become plus, while that for LPs became minus. However, the absolute values of scores have no meaning. With an exception of moderately-differentiated dysplasia, *Mo dys 33*, all sample sets were correctly discriminated by the 11 marker genes selected by the Fisher's ratio. The optimal prediction accuracy with this set of 11 genes was 97.8% (loo).

Discrimination of Non-aggressive and Aggressive Primary OSCC Using Gene Expression Signatures

The identification of marker genes for non-aggressive and aggressive OSCCs

To identify potential marker genes that are differentially expressed between non-aggressive and aggressive OSCCs, cDNA microarray analysis was performed using RNA mixtures from metastasis-negative and less invasive SCCs and also from relatively aggressive OSCCs that have a developing lymph node and/or local metastasis [13]. Among the 16,600 target cDNAs on the chip arrays used, 46 genes were found to be highly expressed in the aggressive OSCCs compared with the non-aggressive ones, and a further 37 genes showed the opposite pattern.

The expression levels of all these marker gene candidates were verified by QRT-PCR analysis using 64 OSCC tissues. Among the 83 marker gene candidates, we selected, 53 genes showed markedly different expression levels that could be associated with an YK's mode of invasion transition, T classification and/or lymphnode metastasis ($p<0.06$). Of these, 29 were found to be markedly down-regulated (Table 2) and 24 were observed to be up-regulated (Table 3), concomitantly with the acquisition of an invasive phenotype. Some of the marker genes demonstrated differential expression along with T grades and/or metastasis, however, the number was restrictive [13].

Supervised classifications to establish predictor models for YK's mode of invasion

It has been reported that YK's mode of invasion is largely correlated with the incidence of lymph node metastasis [20,21]. Histological criteria of the YK's mode of invasion are defined as the following 5 grades: grade 1, well defined borderline; 2, cords, less marked borderline; 3, groups of cells, no distinct borderline; 4C, diffuse invasion with cord-like type; and 4D, with diffuse type invasion [19]. Since, we had isolated a large number of candidate marker genes for the invasion status of an OSCC; we attempted to establish a molecular classification along with the YK's mode of invasion using biopsied OSCC samples. Using the same strategies as already mentioned [13], total RNAs were isolated from primary OSCCs of both node-positive and –negative patients, and marker gene selection was performed between them. We isolated 53 marker genes characteristic for YK's mode of invasion.

As we have reported [12,13], a supervised classification approach based on LDA fitted with a step-wise increment method was performed on this same panel of marker genes in the 64 patient samples. Then, we created four discriminatory predictor models (from LDA-1 to -4) based on from 16 to 25 gene signatures (Tables 2 and 3), which could best distinguish the five established grades of YK's mode of invasion.

Symbol	Accession no.	Gene category and name	Classifier#			
		Cyto skeleton-associated				
TGM3	NM_003245	Transglutaminase 3	A		C	D
CLSP	NM_017422	Calmodulin-like skin protein. Associated with TGM3.			C	
KRT17	NM_000422	Keratin 17.				
LOC144501	XM_096612	Similar to cytokeratin (AA 1-513)		B		D
SPRR1B	NM_003125	Small proline-rich protein 1B (cornifin).				
ZNF185	NM_007150	Zinc finger protein 185 (LIM domain).		B		
KRT1	NM_006121	Homo sapiens keratin 1 (epidermolytic hyperkeratosis).				
		Membrane				
TM7SF2	NM_003273	Transmembrane 7 superfamily member 2.		B		
AQP3	NM_004925	Aquaporin 3, Channering, water transport.			C	
C4.4A	NM_014400	GPI-anchored metastasis-associated protein homolog.	A			
FLJ11036	NM_018306	Hypothetical protein FLJ11036, transmembrane protein 40 (TMEM40),		B	C	D
		Cell adhesion				
CEACAM5	NM_004363	Carcinoembryonic antigen-related cell adhesion molecule 5.				
HSPC159	NM_014181	Homo sapiens galectin-related protein, HSPC159 protein.		B		
		Protease modification				
CSTB	NM_000100	Cystatin B (stefin B) (CSTB), cysteine protease inhibitors			C	
KLK7	NM_005046	Kallikrein 7 (chymotryptic, stratum corneum) (KLK7), transcript variant 1	A		C	
MPN	NM_031948	Pancreasin (MPN), novel tryptic serine peptidase expressed primarily by the pancreas			C	D
WFDC12	NM_080869	WAP four-disulfide core domain 12 (WFDC12), functions as a protease inhibitor			C	D
SLPI	NM_003064	Secretory leukocyte protease inhibitor (antileukoproteinase), secreted inhibitor of serine proteases.		B	C	
		Signal transduction				
RAB25	NM_020387	RAB25, member RAS oncogene family.				
S100A12	NM_005621	S100 calcium binding protein A12 (calgranulin C),involved in specific calcium-dependent signal transduction	A			
		Cell growth/differentiation				
HOP	NM_032495	Homeodomain-only protein, transcript variant 1	A			
CSRP2	NM_001321	Cysteine and glycine-rich protein 2, development and cellular differentiation.	A		C	
CRABP2	NM_001878	Cellular retinoic acid binding protein 2.				
		Proliferation differentiation transformation				
CDA	NM_001785	Cytidine deaminase.	A	B		
CBR3	NM_001236	Carbonyl reductase 3.		B	D	
FOSB	NM_006732	Homo sapiens FBJ murine osteosarcoma viral oncogene homolog B.				
		Others				
ODC1	NM_002539	Ornithine decarboxylase 1.	A		C	
SULT2B1	NM_004605	Sulfotransferase family, cytosolic, 2B, member 1.	A		C	D
RNASE7	NM_032572	Ribonuclease, RNase A family, 7.		B		
D4S234E	NM_014392	DNA segment on chromosome 4 (unique) 234 expressed sequence.		B	C	D
UBD	NM_006398	Homo sapiens ubiquitin D (UBD),				
APOBEC3A	NM_145699	Homo sapiens apolipoprotein B mRNA editing enzyme, catalyticpolypeptide-like 3A.				
APM2	NM_006829	Adipose specific 2.			C	

#A, classifyer for LDA-1; B, for LDA-2; C, for LDA-3; D, for LDA-4, respectively

Table 2: Marker genes down-regulated in advanced OSCCs.

The stability of this model was examined using the leave-one-out cross validation (loo) method and then compared with that of the stepwise increment method. The LDA score for each sample is given by the following linear discrimination functions:

LDA-1 (for YK-1 vs. -2, -3, -4C and -4D); Score = 0.394866 + 0.229884 (HOP) + 0.211169 (CKM) + 0.324503 (CDA) - 0.620754 (CSRP2) - 0.338919 (C1S) + 0.004993 (ODC1) + 0.136036 (TNFSF10) +0.05849 (TAGLN) - 0.104157 (NK4) + 0.15035 (HLA-DBP1) - 0.164207 (HLA-DMB) + 0.359343 (GRCC10) + 0.159684 (GLG1) - 0.097888 (C4.4A) + 0.078056 (KLK7) - 0.161743 (S100A12) + 0.031256 (SULT2B1) - 0.131336 (TGM3).

LDA-2 (for YK-1, -2 vs. -3, -4C and -4D); Score = 0.986068 - 0.010085 (CLSP) - 0.03203 (RNASE7) + 0.088897 (LOC14450) - 0.250882 (SLPI) + 0.03128 (ZNF185) + 0.182597 (CKM) + 0.224274 (CDA) - 0.360236 (D4S234E) + 0.398142 (TNFSF10) - 0.136753 (CBR3) + 0.213823 (MYL9) + 0.045013 (MMP11) - 0.222319 (HLA-DMB) + 0.116148 (ACTN1) - 0.124201 (HSPC159) - 0.267171 (FLJ11036) - 0.538498 (AFG3L2) + 0.221296 (TM7SF2).

LDA-3 (for YK-1, -2, -3 vs. -4C and -4D); Score = - 0.328603 + 0.128613 (CCL19) - 0.00548 (CLSP) + 0.345872 (FHL1) - 0.39984 (SLPI) - 0.023614 (APM2) - 0.052443 (AQP3) - 0.040878 (CSTB) + 0.24299 (CSRP2) - 0.141004 (D4S234E) - 0.00906 (ODC1) + 0.211764 (TNFSF10) + 0.046316 (TAGLN) + 0.157812 (PLAU) - 0.038082 (NK4) - 0.165494 (HLA-DMB) + 0.337967 (LGALSI) - 0.369404 (GLG1) -

Symbol	Accession no.	Gene category and name	Classifier#			
		Cell recognition				
HLA-DPB1	NM_002121	major histocompatibility complex, class II, DP beta 1 (HLA-DPB1)	A			
HLA-DMB	NM_002118	major histocompatibility complex, class II, DM beta (HLA-DMB)	A	B	C	
GLG1	NM_012201	golgi apparatus protein 1 (GLG1), E-selectin ligand-1, MG-160, cysteine-rich fibroblast growth factor receptor	A		C	D
		Complements				
C1S	NM_001734	complement component 1, s subcomponent (C1S)	A			
C1QG	NM_172369	complement component 1, q subcomponent, gamma polypeptide (C1QG)				D
C1QA	NM_015991	complement component 1, q subcomponent, alpha polypeptide (C1QA)				
		ECM remodeling				
PLAU	NM_002658	plasminogen activator, urokinase (PLAU)			C	D
MMP11	NM_005940	matrix metalloproteinase 11 (stromelysin 3) (MMP11)		B		D
MMP2	NM_004530	matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase) (MMP2)				D
COL1A1	NM_000088	collagen, type I, alpha 1 (COL1A1)			C	
COL6A1	NM_001848	collagen, type VI, alpha 1 (COL6A1)				D
COL3A1	NM_000090	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant) (COL3A1)				
LGALSI	NM_002305	lectin, galactoside-binding, soluble, 1 (galectin 1) (LGALS1)			C	D
		Cyto skeleton-associated				
TAGLN	NM_003186	transgelin (TAGLN)	A		C	
ACTN1	NM_001102	actinin, alpha 1 (ACTN1)		B		
CKM	NM_001824	creatine kinase, muscle (CKM)	A	B		
MYL9	NM_006097	myosin, light polypeptide 9, regulatory (MYL9)		B		
FHL1	NM_001449	four and a half LIM domains 1 (FHL1)			C	
		Cytokines				
CCL19	NM_006274	chemokine (C-C motif) ligand 19 (CCL19)			C	
NK4	NM_004221	natural killer cell transcript 4 (NK4),interleukin 32 (IL32), transcript variant 2	A		C	
TNFSF10	NM_003810	tumor necrosis factor (ligand) superfamily, member 10 (TNFSF10)	A	B	C	
		Others				
GRCC10	NM_138425	hypothetical protein BC009925 (LOC113246), (GRCC10)	A			
AFG3L2	NM_006796	AFG3 ATPase family gene 3-like 2 (yeast) (AFG3L2), nuclear gene encoding mitochondrial protein		B	C	
HIST1H2BK	NM_080593	histone 1, H2bk (HIST1H2BK)				D

#A, classifyer for LDA-1; B, for LDA-2; C, for LDA-3; D, for LDA-4, respectively

Table 3: Marker genes up-regulated in advanced OSCCs.

0.090304 (COL1A1) - 0.372974 (FLJ11036) + 0.217867 (AFG3L2) + 0.19053 (KLK7) + 0.010609 (MPN) + 0.092418 (SULT2B) - 0.138387 (WFDC12) + 0.098447 (TGM3).

LDA-4 (for YK-1, -2, -3, -4C vs. -4D); Score = -0.771245 + 0.071189 (LOC144501) - 0.401477 (D4S234E) - 0.066541 (CBR3) + 0.080463 (PLAU) - 0.14224 (MMP2) - 0.045855 (MMP11) + 0.481743 (LGALSI) - 0.106972 (HIST1H2BK) - 0.412974 (GLG1) + 0.087268 (C1QG) + 0.026049 (COL6A1) - 0.513483 (FLJ11036) + 0.019158 (MPN) + 0.0567 (SULT2B1) - 0.259138 (WFDC12) + 0.209169 (TGM3).

The gene sets for each LDA equation are summarized in tables 2 and 3. The optimal prediction accuracies examined by the leave-one-out cross validation of the predictor gene sets were 93.8% (LDA-1), 95.3% (LDA-2), 92.2% (LDA-3) and 93.8% (LDA-4), respectively. As a validation test, these 4 LDA models were then applied to data from 13 independent primary OSCCs. The prediction fidelity to the pathological observations of these tumors was 77% (LDA-1), 85% (LDA-2), 77% (LDA-3) and 100% (LDA-4) (Table 4). Among these 13 samples, five demonstrated discrepancies between the pathological and molecular diagnoses. However, four cases remained within one rank (SCC126, 134, 125 and 132), whereas only one case (SCC129) demonstrated fluctuation on three ranks. Interestingly, most of the discrepancies (80%) could be attributed to an over-estimation by molecular diagnosis (Table 4).

Diagnostic Significance of Cytokine Production and T cell Subsets in the Peripheral Blood Cells from OSCC Patients

As our next approach, we attempt to evaluate general conditions of OSCC patients. In the previous study, we have reported that $CD4^+CD57^+$ T, a subset of $CD4^+$ (conventional helper) T, cells are increased in PBCs of tumor patients including hepatocellular carcinoma and gastric cancer [24,25]. Thus, we performed fluorescence activated cell sorting (FACS) analysis to see $CD4^+CD57^+$ T cells in PBCs of OSCC patients. We also focused on immunological status of OSCC patients. To this end, we examined LPS-Induced cytokine production and/or T lymphocyte subsets in peripheral blood cells from them. In this assay, blood sample was stimulated by lipopolysaccharides (LPS), resulting in the activation of Th1 cells which could release IFN-γ. After 48 h of the stimulation, the production of IFN-γ was assayed by ELISA. As shown in figure 2A, our preliminary data demonstrated that LPS-induced IFN-γ-producing capability of PBCs from OSCC patients was higher in stage I and decreased in a step-wise manner up to stage III during the course of tumor progression. In contrast, the IFN-γ production seemed to be increased in stage IV, compared to that in

Patients	YK grade(I) (Pathological diag.)	Classification by LDA models[a] LDA1	LDA2	LDA3	LDA4	YK grade(II) (Molecular diag.)
SCC129	1	+	+	+	-	4C (↑↑↑)[b]
SCC126	2	+	+	-	-	3 (↑)
SCC132	2	-	-	-	-	1(↓)
SCC131	3	+	+	-	-	3
SCC133	3	+	+	-	-	3
SCC127	3	+	+	-	-	3
SCC139	3	+	+	-	-	3
SCC137	3	+	+	-	-	3
SCC134	3	+	+	+	-	4C (↑)
SCC138	3	+	+	-	-	3
SCC125	3	+	+	+	-	4C (↑)
SCC130	4C#	+	+	+	-	4C
SCC124	4C	+	+	+	-	4C
Fidelity (%)[c]		77	85	77	100	

[a]Gray backgrownd denotes discrepancy between pathological and molecular diagnoses.
[b]One rank over (↑)/Under (↓)estimation compared with the pathologycal diagnosis.
[c]Fidelity to the pathological diagnosis.
#After intensive examination, the pathological diagnosis has been corrected from grade 2 to 4C.

Table 4: Prediction of invasion status by LDA models in 13 test OSCC samples.

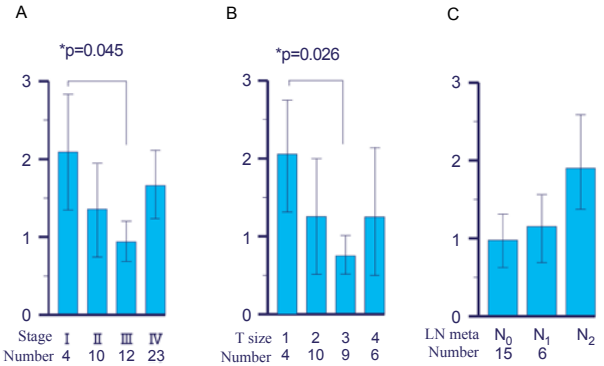

Figure 2: LPS-induced IFN-γ production in PBC from OSCC patients; OSCCs were divided by clinical stages (A), T grades (tumor sizes) (B) or lymphnode status (C). Peripheral blood samples from OSCC patients were stimulated for 24h. Induced IFN-γ was detected by ELISA. The amount of IFN-γ (ng/ml) is in vertical axis. Stages (I, II, III, IV), T grades (1, <2 cm; 2, 2-4 cm; 3, 4-6 cm; 4, 6 cm<), lymphnode status (N_0, no regional lymph node metastasis; N_1, metastasis in a single ipsilateral lymph node, 3 cm or less in greatest dimension; N_2, metastasis in a single ipsilateral lymph node 3 cm ~ 6 cm in dimension) and patient numbers are in abscissa axis, respectively. Evaluation of T grades was performed using blood samples of OSCC patients harboring no metastasis. Evaluation of limphnode status was performed using blood samples of OSCC patients classified as stages III and IV.

stage III patients. This decreased levels in the early stages were inversely correlated with tumor size (Figure 2B), while the levels regain in the last stage seemed to be associated with lymph node metastasis (Figure 2C). However, to reach this conclusion, we have to examine additional samples (manuscript in preparation). As shown in figure 3, the ratio of $CD4^+$ $CD57^+$ T cells against $CD4^+$ (conventional helper) T cells was gradually increased in a step-wise manner up to stage III during the course of tumor progression.

Discussion

In order to identify marker gene candidates, we first screened differential gene expression between OSCCs and LPs. We identified 27 marker genes representing the differential expression between LPs and OSCCs. To further identify marker gene sets and establish appropriate algorism that can sufficiently discriminate between OSCCs and LPs, a supervised classification approach based on LDA was performed. After intensive parameter selection and cross validation, we reached an optimal prediction with a set of 11 genes. According to our classification, however, a moderately differentiated displasia sample, *Mo dys 33*, was insistently classified as an OSCC. Interestingly, *Mo dys 33* had been clinically diagnosed and treated as an OSCC, because of its cancerous macroscopic appearances and history of multiple OSCC [12]. Therefore, a molecular diagnosis based on these 11 genes may, in part, predict clinical features or genetical background of the patient rather than histological grades.

After intensive parameter selection and cross validation, we reached optimal predictions for the YK's mode of invasion with four sets of marker genes [13]. As a validation test, a data set of 13 independent primary OSCCs were applied to the 4 LDA models and our results demonstrated that the prediction fidelity of these models with the pathological observations was higher than 77%. Among the inconsistencies found between the molecular and pathological diagnoses, four out of five remained within one rank up or down and could be attributed to an over-estimation, rather than under-estimation, by molecular diagnosis. In general, the pathological diagnosis complies with the highest grade among limited numbers of tissue sections. In contrast, however, OSCC tissues are molecularly diagnosed using homogenously extracted RNAs from a certain volume of tissue sample. Thus, there may be an oversight during the pathological inspection, but not in the molecular-based inspection. In fact, case OSCC130 listed in Table 4 had first been diagnosed as grade YK-2 tumors, which are two ranks below its molecular diagnosis. However, this case was pathologically re-evaluated and agreement was found with the molecular diagnosis (i.e. an YK-4C grading was assigned). Gene expression may thus obediently reflect the cellular potency of a tumor whereas the pathological appearance of a lesion may demonstrate cell behavior which is more or less modified by its tissue microenvironment.

There are several attempts to molecularly diagnose advanced tumors; However, in line with our results, the predictive values for metastases based on the gene expression profiles of various cancer cell types can vary from 75% to 100% [9,15,17]. There are several discrete steps that comprise the biological cascade leading to metastasis, including, for example, invasion, cell homing and evasion from the

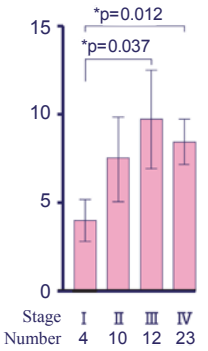

Figure 3: Representative FACS analysis of $CD4^+CD57^+$ T cells. The ratio (%) of $CD4^+CD57^+$ T cells against CD4+PBC is in vertical axis. Stages and Patient numbers are in abscissa axis, respectively.

innate immune system. Thus, the clinical prognosis may be affected not only by the phenotype of cancer cells, the so called "seed", but also by the conditions of the host microenvironment, the "soil" [23]. There should be an interaction between cancer cells and host immunity during epithelial-mesenchymal transition (EMT). During EMT, Snail is a major transcription factor involved in cancer metastasis partly by inducing multiple immunosuppression and immunoresistant mechanisms [26]. Our preliminary data demonstrated that LPS-induced *ex-vivo* production of IFN-γ from peripheral blood cells (PBC) of OSCC patients is inversely correlated with the tumor progression. In addition, the ratio of $CD4^{+}CD57^{+}$ T cells was gradually increased along with the tumor progression.

Our results strongly suggest that corroborating information from both local tissues and patient's immunological indicators should be essential to accomplish quite accurate diagnosis of oral malignancy.

References

1. López M, Aguirre JM, Cuevas N, Anzola M, Videgain J, et al. (2003) Gene promoter hypermethylation in oral rinses of leukoplakia patients--a diagnostic and/or prognostic tool? Eur J Cancer 39: 2306-2309.
2. Bloor BK, Seddon SV, Morgan PR (2001) Gene expression of differentiation-specific keratins in oral epithelial dysplasia and squamous cell carcinoma. Oral Oncol 37: 251-261.
3. Forastiere A, Koch W, Trotti A, Sidransky D (2001) Head and neck cancer. N Engl J Med 345: 1890-1900.
4. Ohkura S, Kondoh N, Hada A, Arai M, Yamazaki Y, et al. (2005) Differential expression of the keratin-4, -13, -14, -17 and transglutaminase 3 genes during the development of oral squamous cell carcinoma from leukoplakia. Oral Oncol 41: 607-613.
5. Schmalbach CE, Chepeha DB, Giordano TJ, Rubin MA, Teknos TN, et al. (2004) Molecular profiling and the identification of genes associated with metastatic oral cavity/pharynx squamous cell carcinoma. Arch Otolaryngol Head Neck Surg 130: 295-302.
6. Carinci F, Lo Muzio L, Piattelli A, Rubini C, Palmieri A, et al. (2005) Genetic portrait of mild and severe lingual dysplasia. Oral Oncol 41: 365-374.
7. Ginos MA, Page GP, Michalowicz BS, Patel KJ, Volker SE, et al. (2004) Identification of a gene expression signature associated with recurrent disease in squamous cell carcinoma of the head and neck. Cancer Res 64: 55-63.
8. Whipple ME, Mendez E, Farwell DG, Agoff SN, Chen C (2004) A genomic predictor of oral squamous cell carcinoma. Laryngoscope 114: 1364-1354.
9. Roepman P, Wessels LF, Kettelarij N, Kemmeren P, Miles AJ, et al. (2005) An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. Nat Genet 37: 182-186.
10. Iizuka N, Oka M, Yamada-Okabe H, Mori N, Tamesa T, et al. (2002) Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method. Cancer Res 62:3939-3944.
11. Ramaswamy S, Tamayo P, Rifkin R, Mukherjee S, Yeang CH, et al. (2001) Multiclass cancer diagnosis using tumor gene expression signatures. Proc Natl Acad Sci USA 98: 15149-15154.
12. Kondoh N, Ohkura S, Arai M, Hada A, Ishikawa T, et al. (2007) Gene expression signatures that can discriminate oral leukoplakia subtypes and squamous cell carcinoma. Oral Oncol 43: 455-462.
13. Kondoh N, Ishikawa T, Ohkura S, Arai M, Hada A, et al. (2008) Gene expression signatures that classify the mode of invasion of primary oral squamous cell carcinomas. Mol Carcinog 47: 744-756.
14. Marchet A, Mocellin S, Belluco C, Ambrosi A, DeMarchi F, et al. (2007) Gene expression profile of primary gastric cancer: towards the prediction of lymph node status. Ann Surg Oncol 14: 1058-1064.
15. Kato Y, Uzawa K, Saito K, Nakashima D, Kato M, et al. (2006) Gene expression pattern in oral cancer cervical lymph node metastasis. Oncol Rep 16: 1009-1014.
16. O'Donnell RK, Kupferman M, Wei SJ, Singhal S, Weber R, et al. (2005) Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity. Oncogene 24: 1244-1251.
17. Chung CH, Parker JS, Karaca G, Wu J, Funkhouser WK, et al. (2004) Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. Cancer Cell 5: 489-500.
18. Chung CH, Parker JS, Ely K, Carter J, Yi Y, et al. (2006) Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor-kappaB signaling as characteristics of a high-risk head and neck squamous cell carcinoma. Cancer Res 66: 8210-8218.
19. Yamamoto E, Kohama G, Sunakawa H, Iwai M, Hiratsuka H (1983) Mode of invasion, bleomycin sensitivity, and clinical course in squamous cell carcinoma of the oral cavity. Cancer 51: 2175-2180.
20. Kaihara T, Kusaka T, Kawamata H, Oda Y, Fujii S, et al. (2001) Decreased expression of E-cadherin and Yamamoto-Kohama's mode of invasion highly correlates with lymph node metastasis in esophageal squamous cell carcinoma. Pathobiology 69: 172-178.
21. Nakayama A, Ogawa A, Fukuta Y, Kudo K (1999) Relation between lymphatic vessel diameter and clinicopathologic parameters in squamous cell carcinomas of the oral region. Cancer 86: 200-206.
22. Ramaswamy S, Ross KN, Lander ES, Golub TR (2003) A molecular signature of metastasis in primary solid tumors. Nat Genet 33: 49-54.
23. Gupta GP, Massagué J (2006) Cancer metastasis: building a framework. Cell 127: 679-695.
24. Chochi K, Ichikura T, Majima T, Kawabata T, Matsumoto A, et al. (2003) The increase of CD57+ T cells in the peripheral blood and their impaired immune functions in patients with advanced gastric cancer. Oncol Rep 10: 1443-1448.
25. Shiraki T, Takayama E, Magari H, Nakata T, Maekita T, et al. (2011) Altered cytokine levels and increased CD4+CD57+ T cells in the peripheral blood of hepatitis C virus-related hepatocellular carcinoma patients. Oncol Rep 26: 201-208.
26. Kudo-Saito C, Shirako H, Takeuchi T, Kawakami Y (2009) Cancer metastasis is accelerated through immunosuppression during Snail-induced EMT of cancer cells. Cancer Cell 15: 195-206.