Clinical Medicine Insights: Oncology

R E V I E W

Gene Expression Profiling for the Purposes of Biomarker Discovery in Oral Potentially Malignant Lesions: A Systematic Review

Ahmad A. AbdulMajeed[1,2] and Camile S. Farah[1,2,*]

[1]The University of Queensland, UQ Centre for Clinical Research, Herston, Queensland, Australia. [2]The University of Queensland, School of Dentistry, Brisbane, Queensland, Australia. Corresponding author email: c.farah@uq.edu.au

Abstract: Early and accurate diagnosis of oral potentially malignant lesions (OPML) is of critical importance in preventing malignant transformation. Although histopathological interpretation of the degree of epithelial dysplasia is considered the gold standard for diagnosis, this method is subjective and lacks sensitivity. Therefore, many attempts have been made to identify objective molecular biomarkers to improve diagnosis. Microarray technology has the advantage of screening the expression of the whole genome making it one of the best tools for searching for novel biomarkers. However, microarray studies of OPMLs are limited, and no review has been published to highlight and compare their findings. In this paper, we systematically review all studies that have incorporated microarray analyses in the investigation of gene profile alterations in OPMLs and suggest a set of commonly dysregulated genes across multiple gene expression profile studies. This list of common genes may help focus selection of markers for further analysis regarding their importance in the diagnosis and prognosis of OPMLs.

Keywords: oral cancer, oral potentially malignant lesion, biomarker, microarray, systematic review

This article is available from http://www.la-press.com.

Introduction

Oral squamous cell carcinoma (OSCC) usually develops from oral potentially malignant lesions (OPMLs).[1] Early and accurate diagnosis of OPMLs is of critical importance in preventing malignant transformation.[2] The current gold standard for diagnosis is histopathological interpretation of the degree of epithelial dysplasia on a biopsy specimen.[1] However, histopathologic diagnosis is subjective and lacks sensitivity, namely that there is no agreement on which of the features of dysplasia are important in predicting progression, and there is both inter- and intra-observer variation in interpreting the degree of epithelial dysplasia.[3–5] Therefore, many attempts have been made to identify objective molecular biomarkers for diagnosis using different types of approaches such as loss of heterozygosity, DNA ploidy, telomerase activity, methylation, and gene expression analysis. Nonetheless, these efforts have failed to characterize or predict the behavior of OPMLs since studies have been based on analyzing one or a few markers, despite the well known fact that carcinogenesis is dictated by the expression of thousands of genes along complex molecular pathways. Therefore, a new strategy for discovering useful molecular biomarkers through analyzing the expression of the entire genome at different stages of oral carcinogenesis is required.

Microarray technology (cDNA- and oligonucleotide-based microarrays) allows rapid screening of the whole genome.[6,7] This technique has helped elucidate many significant genetic events that may lead to cancer, and has revealed new pathways in the pathophysiology of tumorigenesis. In addition, it is one method currently used in the search for novel biomarkers which have allowed the successful molecular classification of several cancers regarding their stage, metastasis, recurrence potential, prognostic outcome and response to therapy.[8–10] The strength of microarrays lies in their ability to perform simultaneous analysis of tens of thousands of genes at a time, raising the probability of discovering novel markers. However, questions have been raised regarding the reproducibility and reliability of microarray experiments. While microarrays can be used to determine mRNA levels, it is impossible to predict protein concentration or activity.[11] Regardless of these limitations, if appropriate candidate markers are applied, purpose-designed arrays can be used one day to obtain expression fingerprints in routine diagnostic protocols of OPMLs, similar to commercial multigene assays (20–70 signature genes) available for breast cancer prognosis and prediction.[12]

Methodology

To identify all studies that have incorporated microarray analyses in the investigation of gene profile alterations in OPMLs, we searched the PubMed medical literature database for the following keywords: "(oral dysplastic or oral dysplasia) OR (potentially malignant) AND (microarray or gene expression profile)". Supplemental PubMed searches for references cited by review articles were undertaken to identify any additional manuscripts not included in the primary queries. After exclusion of non-related articles, 15 studies were included in this review.

In order to define a set of commonly dysregulated genes in OPMLs across multiple gene expression profile studies, we prepared a universal datasheet containing all differentially expressed genes extracted from microarray studies on OPMLs. We attempted to obtain all dysregulated gene sets but we were only able to extract published tables and supplementary data from 9 out of 15 published articles.[13–21]

Direct matching for repeated genes was not feasible because authors published their results using various forms of gene identification (eg, gene name, gene symbol, Genbank accession number, Affymetrix probe set ID, or Unigene cluster ID). Therefore, we used standardized gene identification by converting all these forms into Genbank accession number utilizing Clone/Gene ID converter tool (http://idconverter. bioinfo.cnio.es/IDconverter.php).[22] We searched for duplicate genes in the spreadsheet and constructed a set of commonly dysregulated genes.

Results

The microarray studies which are included in this study are shown in Table 1. A list of commonly dysregulated genes in OPMLs across multiple gene expression profile studies was prepared. The concordance between studies was low because of differences in sample number, clinical diagnosis, histologic grading, microarray platforms, experimental design, and analysis methods. This lack of agreement between studies was not surprising as this constraint is a common criticism of expression profiling studies.

Table 1. Studies using microarray analysis of OPMLs.[a]

Authors	Tissue samples	Genes number in the array	Alteration in gene expression	Further analysis	Validation	Post microarray investigation
Mendez et al[32]	18 Normal 2 OPML 26 invasive OSCC	7000	OSCC vs. normal ↑ 239 ↓ 75	Hierarchical clustering	Quantitative reverse transcriptase-PCR for 4 ↑ genes 1 ↓ gene	NA
Ha et al[33]	6 normal 8 leukoplakia 7 invasive carcinoma	>12,000	Leukoplakia vs. normal ↑ 108 ↓ 226 Malignant vs. leukoplakia ↑ 5 ↓ 13	Hierarchical clustering	Quantitative reverse transcriptase-PCR for 1 ↑ gene	NA
Banerjee et al[14]	6 paired OPMLs and normal adjacent tissue 2 unpaired OMPL tissue 2 paired OMPLs and verrucous carcinoma	>23,000	↑ 1,300 ↓ 400	Hierarchical clustering	Real time-PCR for 4 genes IHC for 2 genes	In vitro oral cancer progression model
Carinci et al[13]	11 normal 5 mild dysplasia 4 severe dysplasia	19,200	Normal vs. mild dysplasia ↑ 161 ↓ 109 Mild vs. severe dysplasia ↑ 63 ↓ 118	Average Linkage Analysis	NA	NA
Carinci et al[34]	9 dysplasia 8 OSCC without metastasis (TWM) 11 OSCC with metastasis (MT)	19,200	TWM vs. dysplasia ↑ 24 ↓ 9 MT vs. TWM ↑ 89 ↓ 66	Hierarchical clustering	NA	NA
Odani et al[16]	6 Normal tissues adjacent to lesion 4 Leukoplakia 2 OSCC	8,800	Leukoplakia vs. normal ↑ 8 ↓ 10	NA	Quantitative reverse transcriptase-PCR for loricrin	NA
Vigneswaran et al[38]	2 Normal cell line 2 premalignant (leukoplakia) cell line 2 primary OSCC cell line 2 metastatic OSCC cell line	22,500	NA	NA	Real time-PCR for EMMPRIN	IHC Western blot for EMMPRIN
Hunter et al[84]	4 normal cultures 19 dysplasia cultures 16 OSCC cultures	NA	NA	Spectral clustering	Quantitative reverse transcriptase-PCR	Western blot IHC
Kondoh et al[15]	19 leukoplakia 27 OSCC	16,600	Leukoplakia vs. OSCC ↑ 53 ↓ 63	Hierarchical clustering	Quantitative reverse transcriptase-PCR	

(*Continued*)

Table 1. (*Continued*)

Authors	Tissue samples	Genes number in the array	Alteration in gene expression	Further analysis	Validation	Post microarray investigation
Hu et al[40]	4 normal 4 OSF	14,500	OSF vs. normal ↑ 716 ↓ 149	Hierarchical clustering	Quantitative reverse transcriptase-PCR 3 ↑ genes 2 ↓ genes	Western blot IHC
Li et al[17]	4 normal 4 OSF	14,500	OSF vs. normal ↑ 661 ↓ 129	Hierarchical clustering	Quantitative reverse transcriptase-PCR for 5 genes	NA
Kuribayashi et al[18]	10 leukoplakia OSCC cell line	>29,000	Mild vs. severe dysplasia ↑ 16	Hierarchical clustering	Quantitative reverse transcriptase-PCR	NA
Watanabe et al[19]	3 leukoplakia 3 early tongue cancer Normal tissue	NA	Dysplasia vs. normal ↑ 5 ↓ 10	NA	IHC for PI3-AKT	NA
Saintigny et al[20]	35 leukoplakia progressed to OSCC 51 leukoplakia did not progress to OSCC	NA	NA	CoxBoost algorithm	NA	NA
Sumino et al[21]	11 Normal tissue 11 Oral dysplasia 11 OSCC	35,035	Dysplasia vs. normal tissue ↑ 890 ↓ 17 OSCC vs. dysplastic tissues ↑ 461 ↓ 1,388	Hierarchical clustering	Quantitative reverse transcriptase-PCR	IHC for ISG15

Note: [a](↑) over-expression, (↓) under-expression, (NA) not applicable, (IHC) Immunohistochemistry, (PCR) Polymerase chain reaction, (OPML) oral potentially malignant lesion.

Nevertheless, we identified 31 genes with common expression changes in at least two independent studies (Table 2). Some of these genes have roles in human carcinogenesis supporting their use as potential diagnostic markers for OPML. However, literature mining for these genes showed that the majority of these had not been validated by alternative experimental methods.

Discussion

Several studies have incorporated microarray analyses in the investigation of the genetic profile of oral cancer,[6,7,20,21,23–31] but the literature on microarray analysis of OPMLs is limited. Microarray studies of OPMLs have had different objectives. Some have aimed at identification of molecular subtypes and drawing predictive molecular signatures, while others have utilized microarray analysis to perform global gene expression and then chose one or more genes for further investigation. Only two studies found a limited number of genes that could be used to discriminate mild dysplasia from severe dysplasia[15,18] and one study suggested a prediction model for cancer progression.[20] In this review we highlight some of the major findings from microarray studies undertaken on OPMLs.

In 2002, Mendez et al compared the gene expression profiles of 26 OSCC and 2 OPMLs with 18 normal oral tissue samples. Hierarchical clustering analysis showed that patients with OPMLs displayed a higher degree of relatedness to OSCC compared to normal samples.[32] Although the sample size of OPMLs in this study was too small to draw general conclusions, a similar result was found in a later larger sample.[33] Similarly, Ha et al found a small difference in gene expression between OPMLs and malignant lesions, compared to differences between normal and OPMLs. In addition, unsupervised hierarchical clustering revealed that both the malignant and OPML groups tended to form related but distinct clusters.[33] Data from these two studies demonstrate that a greater proportion of transcriptional alterations occur during the transition from normal to potentially malignant mucosa, than in the transition from potentially malignant to malignant tissue.[32,33]

In 2005, Banerjee et al showed a large number of significantly altered genes in OPMLs. They identified 1300 genes as significantly activated and 400 genes as significantly repressed.[14] Among the genes found

to be dysregulated in OPMLs, were genes functionally known to be involved in angiogenic stimulation, apoptosis, cell mitosis, cell cycle control, cell signaling, DNA repair, epithelial to mesenchymal differentiation, oncogenesis, invasiveness, immunoregulation, and protein translation. The expression of several genes was validated by real-time-PCR and immunohistochemistry (IHC).[14] Although this study was not designed to identify a gene signature classifier, it provided a molecular basis for persistent proinflammatory conditions in oral premalignant tissues.

In an attempt to define the genetic signature of dysplastic progression, Carinci et al performed a comparison between lingual dysplasia (5 mild and 4 severe) and 11 normal oral tissue specimens. Microarray analysis showed remarkable differences in 270 expressed genes (161 up-regulated and 109 down-regulated) between normal tissue and mild dysplasia and 181 genes (63 up- and 118 down-regulated) in mild versus severe dysplasia. The authors suggested that those genes could be used for classification of dysplasia at the molecular level.[13] In another study, Carinci et al compared 9 dysplastic lesions with 8 tumors without metastasis and found only 33 genes that were differentially expressed.[34] However, they did not confirm their results by additional methods, despite the well known importance of validating gene expression results by other techniques.[15,35] It should be noted that the true test of a classifier is whether it predicts accurately in independent data. A final validation on an independent cohort is vital before a gene signature can be used in clinical practice.

In 2006, Odani et al compared 4 cases of oral leukoplakia and 2 cases of OSCC with their normal matched tissue obtained from normal sites adjacent to these lesions. All cases of leukoplakia showed overexpression of 8 genes as well as under-expression of 10 genes that belonged to diverse functional groups.[16] Most of the genes found to be over-expressed were genes included in the keratinocyte cytoskeleton network, while the set of down-regulated genes included genes associated with the cornified cell envelope of epithelial cells, epithelial cell adhesion, cancer antigen, keratinocyte activation, transcription of epithelial cells, and epithelial cell embryonicity. The number of altered genes was very small compared with the results of other studies. This may be because normal matched tissue in this study was obtained

Table 2. Common gene alterations seen in OPMLs identified by microarray studies.[a]

Gene symbol	Gene name	Results
LOR	Loricrin	↑ in oral leukoplakia[16] ↑ in OSF[17]
CLSP	Calmodulin-like skin protein	↑ in oral leukoplakia[16] ↑ in OPML[14]
KRT1	Keratin 1	↑ in oral leukoplakia vs. OSCC[15] ↑ in OPML[14]
KRT10	Keratin 10	↑ in Oral leukoplakia[16] ↑ in OPML[14]
KRT19	Keratin 19	↑ in oral dysplasia[19] ↓ in OSF[17]
CXCL9	Chemokine (C-X-C motif) ligand 9	↑ in OSF[17] ↑ in OPML[14]
CXCL10	Chemokine (C-X-C motif) ligand 10	↑ in OSCC vs. oral leukoplakia[15] ↑ in OPML[14]
CXCL13	Chemokine (C-X-C motif) ligand 13	↑ in OSF[17] ↑ in OPML[14]
EVA1	Epithelial V-like antigen 1 (also known as MPZL2)	↓ in mild dysplasia vs. normal[13] ↑ in severe vs. mild dysplasia[13] ↑ in OPML[14]
HSPA4L	Heat shock 70 kDa protein 4-like	↑ in mild dysplasia vs. normal[13] ↑ in OPML[14]
HSPB3	Heat shock 27 kDa protein 3	↑ in OSF[17] ↓ in OPML[14]
HPSE	Heparanase	↑ in mild dysplasia vs. normal[13] ↓ in severe vs. mild dysplasia[13] ↑ in OPML[14]
PTPRZ1	Protein tyrosine phosphatase receptor-type Z polypeptide 1	↑ in Oral leukoplakia[16] ↑ in OPML[14]
COMP	Cartilage oligomeric matrix protein	↑ in OSF[17] ↑ in OPML[14]
LRRC15	Leucine-rich repeat containing 15	↑ in OSF[17] ↑ in OPML[14]
DPYSL3	Dihydropyrimidinase-like 3	↑ in oral leukoplakia vs. OSCC[15] ↑ in OPML[14]
ELF3	E74-like factor 3	↓ in Oral leukoplakia[16] ↓ in OSF[17]
BHLHB2	Basic helix-loop-helix domain containing class B 2.	↓ in mild dysplasia vs. normal[13] ↑ in OPML[14]
POSTN	Periostin osteoblast specific factor	↓ in mild dysplasia vs. normal[13] ↑ in OPML[14]
SERPINB1	Serpin peptidase inhibitor clade B (ovalbumin) member 1	↑ in OPML[14] ↓ in OSF[17]
CH25H	Cholesterol 25-hydroxylase	↑ in mild dysplasia vs. normal[13] ↑ in OPML[14]
ARL1	ADP-ribosylation factor-like 1	↑ in mild dysplasia vs. normal[13] ↓ in severe vs. mild dysplasia[13] ↑ in OPML[14]
GTF2E2	General transcription factor IIE polypeptide 2 beta 34 kDa	↓ in mild dysplasia vs. normal[13] ↑ in severe vs. mild dysplasia[13] ↑ in OPML[14]
AGA	Aspartylglucosaminidase	↓ in mild dysplasia vs. normal[13] ↑ in OPML[14]
SEC23A	Sec23 homolog A (S. cerevisiae)	↓ in mild dysplasia vs. normal[13] ↑ in OPML[14]

(*Continued*)

Table 2. (*Continued*)

Gene symbol	Gene name	Results
ACTA1	Actin alpha 1 skeletal muscle	↓ in OPML[14] ↑ in OSF[17]
CSRP3	Cysteine and glycine-rich protein 3 (cardiac LIM protein)	↑ in OSF[17] ↓ in OPML[14]
IFI44	Interferon-induced protein 44	↑ in mild dysplasia vs. normal[13] ↑ in OSCC vs. oral leukoplakia[15]
USP18	Ubiquitin specific peptidase 18	↓ in mild dysplasia vs. normal[13] ↑ in severe vs. mild dysplasia[13] ↑ in OSCC vs. oral leukoplakia[15]
MYH2	Myosin heavy polypeptide 2	↑ in OSF[17] ↓ in OPML[14]
DPYD	Dihydropyrimidine dehydrogenase	↓ in oral dysplasia[19] ↑ in OPML[14]

Note: [a](↑) over-expressed, (↓) under-expressed, (OPML) oral potentially malignant lesion.

from normal sites close to the lesions and according to the concept of field cancerization, adjacent areas of tissue share common genetic changes.[36,37]

Other microarray studies were performed using dysplastic cell lines. Vigneswaran et al used microarrays to conduct a global analysis of gene expression in a set of cell lines to identify genes that were differentially expressed during malignant transformation of oral epithelial cells. They found that the expression of extracellular matrix metalloproteinase inducer (EMMPRIN) was markedly up-regulated in potentially malignant cells compared to normal oral and epidermal keratinocytes. To confirm this result, they used real time-PCR, western blot and IHC for validation. They concluded that EMMPRIN over-expression was an important early event in oral carcinogenesis.[38] Hunter et al tested the hypothesis that mortal and immortal OSCCs involve distinct transcriptional changes, by performing microarray analysis of primary cultures of 4 normal oral mucosa biopsies, 19 dysplasias, and 16 OSCCs. They concluded that there were divergent mortal and immortal pathways for OSCC development via intermediate dysplasias.[39] Nevertheless, there are potential limitations of using cell lines as opposed to using actual premalignant samples for microarray analysis. Although cell lines can produce relatively large amounts of good-quality RNA for comparative expression analyses, established cell lines do not completely resemble the original source tissue.

Two studies have investigated gene expression profiles of oral submucous fibrosis (OSF). Li et al demonstrated increased expression of 661 genes and decreased expression of 129 genes in OSF samples with respect to normal tissue.[17] The study of Hu et al showed that 716 genes were up-regulated and 149 genes were down-regulated in OSF. The altered genes were related to different ontology groups such as immune response, inflammatory response and epithelial-mesenchymal transition.[40] Hierarchical clustering analysis revealed two clearly distinct groups of normal and OSF according to their gene expression profile.[17,40] In microarray studies, due to large inter-individual variability, an adequate sample size is a necessity. In both studies only 4 OSF and 4 normal samples were used for the development of gene signatures. Significantly larger patient groups are needed before these classifier can be applied clinically.

In order to generate classifiers for OSCC and leukoplakia, Kondoh et al analyzed 27 OSCC and 19 leukoplakias using oligonucleotide microarrays. Using pooled samples, they identified an 11-gene predictor set that could best distinguish OSCC from leukoplakia. Furthermore, they found that seven of these gene predictors could be used to differentiate mild dysplasia from severe dysplasia.[15] Kuribayashi et al revealed that 16 selected genes could be used to differentiate mild dysplasia from severe dysplasia.[18] However, both studies did not use histologically normal tissues as controls.

Watanabe et al compared the gene expression profile of 3 oral leukoplakias and 3 early tongue cancers to normal epithelium using microarray analysis. They used laser micro-dissection to procure only the

dysplastic cells. They found that 5 genes were significantly over-expressed and 10 genes under-expressed in oral epithelial dysplasia.[19] More recently, Sumino et al also used laser micro-dissection to analyze the gene expression profile in normal tissues compared to oral dysplasia and in oral dysplasia compared to OSCC. They identified 15 candidate genes with continuously increasing or decreasing expression during oral carcinogenesis.[21] The strength of these studies lies in the advantage of using laser micro-dissection which allows for the precise isolation of individual cells of interest. However, owing to the small size of the arrayed samples, the true representation of the disease may be questioned.

In an interesting study, Saintigny et al investigated gene expression alterations in 35 oral dysplastic samples that progressed to OSCC compared with 51 samples that did not develop OSCC over a median follow-up of 6.08 years.[20] The authors were able to develop a 29-transcript prediction model that had a prediction error rate of 8%. However, a limitation of this study lies in the fact that all patients were enrolled in a clinical chemoprevention trial which would change the progression outcome.[20]

Among the genes consistently reported at a highly significant rate in OPMLs are components of the keratinocyte cytoskeleton network such as loricrin (LOR), calmodulin-like skin protein (CLSP), keratin 1 (KRT1), keratin 10 (KRT10) and keratin 19 (KRT19). LOR is a major constituent of cornified cell envelopes which is strongly expressed at the later stage of epithelial differentiation.[41] In two microarray studies, LOR expression was significantly up-regulated with the largest fold change in oral leukoplakia[16] and OSF.[17] These results were validated by semi-quantitative reverse transcription-PCR. IHC staining of paraffin embedded tissue of normal samples showed no detectable expression of LOR, while 63.6% of OSF cases exhibited intensive staining. Nevertheless, expression of LOR and keratins is dependent on whether the dysplasia is keratinized or not, reducing the potential value of these genes in the diagnosis and prognosis of OPMLs. A statistically significant association with histologic grade of OSF was also found.[17] CLSP was found to be upregulated in oral leukoplakia[16] and OPMLs.[14] A role for CLSP in late keratinocyte differentiation has been suggested.[42] Changes in the type and distribution of keratins have been observed during oral carcinogenesis.[43–48] E74-like factor 3 (ELF3), another gene related to keratinocyte differentiation, was the only gene found to be down-regulated in two microarray studies,[16,17] however it has been found to be over-expressed in other tumors such as lung adenocarcinoma and synovial sarcoma.[49,50]

Other genes of note include those related to immunologic response such as C-X-C motif ligand chemokines CXCL9, CXCL10, CXCL13, USP18, IFI44, and epithelial V-like antigen 1 (EVA1). Members of the CXC chemokine family are known for their ability to regulate angiogenesis, an important hallmark of carcinogenesis.[51,52] Dysregulation of three subtypes of chemokine was noticed in OPMLs in different microarray studies[14,15,17] suggesting a role for CXC chemokine expression in the progression to oral cancer. This is supported by the fact that CXCL9 was found to be up-regulated in OSCC[25] and has been proposed as a useful candidate biomarker for breast cancer screening.[53] In addition, a functional role for CXCL13 in OSCC has been implicated in other studies.[54,55] USP18 is a member of the ubiquitin-specific protease family which is known to be an interferon-stimulated gene 15 (ISG15) specific isopeptidase that removes ISG15 from its conjugated proteins.[56] It has been found to control IFN-b stimulated genes which regulate a group of immune response related genes.[57] Dysregulation of USP18 was found in OPMLs in two microarray studies.[13,15] Another study found that USP18 was significantly over-expressed in OSCC compared to normal controls.[25] Duex et al suggested that USB18 may have oncogenic properties through controlling microRNAs and cancer cell growth.[58]

IFI44 is an interferon stimulating gene. It has been demonstrated that IFI44 over-expression leads to cell cycle arrest in vitro.[59] Carinci et al and Kondoh et al both found that IFI44 was dysregulated in OPMLs.[13,15] In another two studies, IFI44 was found to be over-expressed in OSCC suggesting a role for this gene in oral carcinogenesis.[60,61] EVA1, a member of the immunoglobulin superfamily[62] was found to be down-regulated in mild dysplasia but up-regulated in severe dysplasia[13] and OPMLs.[14] It has been shown that EVA1 is one of the significantly down-regulated genes in the transition from prostatic intraepithelial neoplasia to prostate cancer.[63]

The list of commonly dysregulated genes also contains two types of heat shock proteins, namely, heat shock 70 kDa protein 4-like (HSPA4 L) and heat shock 27 kDa protein 3 (HSPB3). Heat shock proteins are synthesized by cells in response to a variety of stress conditions, including carcinogenesis.[64] Experimental evidence suggests that these proteins may be associated with tumor progression by inhibiting apoptosis.[65] Microarray studies found that HSPA4 L was up-regulated in OPMLs.[14,17] Other studies found significant correlation between heat shock 70 kDa protein expression and the severity of oral dysplasia.[65,66] A five year follow up study showed that the median transition time (premalignancy to malignancy) was significantly shorter in cases showing over-expression of heat shock 70 kDa protein.[66] Contradictory results have been found for HSPB3 in microarray studies. One study found that HSPB3 was up-regulated in OSF,[17] while another found that it was down-regulated in OPMLs.[14] IHC found no significant difference for cytoplasmic expression of heat shock 27 kDa protein between oral leukoplakia with or without epithelial dysplasia.[65]

Other sets of genes commonly seen to be up-regulated in OPMLs which are also found to be over-expressed in other cancer types include heparanase (HPSE), protein tyrosine phosphatase receptor-type Z polypeptide 1 (PTPRZ1), cartilage oligomeric matrix protein (COMP), leucine-rich repeat containing 15 (LRRC15), and dihydropyrimidinase-like 3 (DPYSL3). HPSE is an endoglycosidase that cleaves heparan sulphate complex (glycosaminoglycan consisting of polysaccharide expressed on the cell surface and in the extracellular matrices).[67] It has been found that cancer metastasis correlated with high levels of HPSE activity,[68] and that HPSE expression could serve as an indicator of aggressive potential and poor prognosis in cervical cancer.[69] COMP is a tissue specific non-collagenous matrix protein,[70] highly up-regulated in OSF with the largest fold change. Immunohistochemical analysis has shown that the expression of COMP is significantly associated with histologic grade of OSF.[17] Its high expression has also been observed in hepatocellular carcinoma.[71] LRRC15 is a cell surface glycoprotein normally expressed only in the invasive cytotrophoblast layer of the placenta. Various studies indicate that LRRC15 is frequently over-expressed in

different types of cancer such as prostate and breast cancer.[72,73] DPYSL3 is a developmentally regulated protein, strongly expressed in early embryonic post-mitotic neural cells and in adult brain in regions that retain neurogenesis, such as granular neurons in dentate gyrus. It is believed to play a role in neuronal differentiation, axonal outgrowth, and possibly, neuronal regeneration.[74] Significantly different DPYSL3 expression was found between metastasized and non-metastasized head and neck squamous cell carcinoma.[75]

Contradictory results have been found for some genes such as basic helix-loop-helix domain containing class B 2 (BHLHB2), periostin osteoblast specific factor (POSTN), and serpin peptidase inhibitor clade B (ovalbumin) member 1 (SERPINB1), as one study reported significant up-regulation, whereas another reported down-regulation.[13,14,16,17] BHLHB2 is a basic helix-loop-helix (bHLH) domain-containing protein that acts as a transcriptional repressor. Falvella et al found BHLHB3 transcript levels were low in three human lung cancer cell lines and down-regulated in human lung adenocarcinomas as compared to normal lung tissue suggesting a potential role for BHLHB2 protein as a tumor suppressor of lung cancer.[76] POSTN is a mesenchyme-specific gene, known to be over-expressed in human breast cancer.[77] SERPINB1 was found to be down-regulated in prostatic intraepithelial neoplasia.[63]

Other identified genes such as cholesterol 25-hydroxylase (CH25H) and aspartylglucosaminidase (AGA) have no well documented known role in carcinogenesis but their reproducible expression in multiple studies, and their ability to pass noise and error which usually occur in microarray experiments may indicate a true biologic role for these novel genes in oral carcinogenesis, however further studies are needed to validate this.

Examination of the common gene list (Table 2) shows it to be devoid of traditional markers for OPMLs. This could be explained by the possible absence of probes for these genes in the diverse array platforms employed in previous microarray studies. A wide range of probe numbers (7000–29000) was used in different studies. Furthermore, carcinogenesis is controlled by multiple molecular pathways and opposing modulation of different genes can produce a similar final effect, thus studies on

limited samples can reach discordant findings.[13] For example, the list is lacking cancer stem cell markers such as ALDH1 and CD44.[78] Cancer stem cells, by definition, have the ability to propagate, and to differentiate into mature, functional cells. Premalignant conditions may serve as a model for the cancer stem cell concept.[79] Additionally, the list is lacking cancer metabolism markers such as LDH5 and TKTL1.[80] As in most solid tumors, oral cancer displays dramatically altered glucose metabolism.[81] Various markers which are associated with aerobic glycolysis, have been found to be correlated with malignant transformation of oral epithelial dysplasia.[80,81]

The applicability of the common gene list is highlighted in a recent immunohistochemistry study undertaken in our group, where we investigated the usefulness of 5 genes (CLSP, ELF3, USP18, IFI44, CXCL13) for further analysis. Our selection was based on the fact that these genes were associated with human cancer but had not been examined in oral dysplasia or OSCC by other methods. We found that significant alterations in the expression of CLSP, ELF3, and IFI44 which were initially identified by microarray studies were associated with similar changes in protein expression in epithelial cells based on immunohistochemical analyses.[82]

Thus, the common gene list elucidated in this review may help focus selection of markers for further analysis regarding their importance in diagnosis and prognosis of OPMLs. A limitation of the true value of this list stems from a lack of statistical power due to low sample numbers in some studies, unavailability of raw data in others (since some authors published only part of their gene list), variation of clinical data such as age, sex, smoking/ alcohol habits, and site of lesion, and heterogeneity of OPML diagnostic criteria as most studies did not diagnose lesions according to the degree of epithelial dysplasia proposed by the WHO.[83] Further validation of the usefulness of these markers for OPML diagnosis is required.

Author Contributions

Conceived and designed the experiments: CSF. Analyzed the data: AAA. Wrote the first draft of the manuscript: AAA. Contributed to the writing of the manuscript: AAA, CSF. Agree with manuscript results and conclusions: AAA, CSF. Jointly developed the structure and arguments for the paper: AAA, CSF. Made critical revisions and approved final version: CSF. All authors reviewed and approved of the final manuscript.

Competing Interests
Author(s) disclose no potential conflicts of interest.

Disclosures and Ethics
As a requirement of publication the authors have provided signed confirmation of their compliance with ethical and legal obligations including but not limited to compliance with ICMJE authorship and competing interests guidelines, that the article is neither under consideration for publication nor published elsewhere, of their compliance with legal and ethical guidelines concerning human and animal research participants (if applicable), and that permission has been obtained for reproduction of any copyrighted material. This article was subject to blind, independent, expert peer review. The reviewers reported no competing interests.

References

1. van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. *Oral Oncol*. 2009;45(4–5):317–23.
2. Sciubba JJ. Oral cancer. The importance of early diagnosis and treatment. *Am J Clin Dermatol*. 2001;2(4):239–51.
3. Abbey LM, Kaugars GE, Gunsolley JC, et al. Intraexaminer and interexaminer reliability in the diagnosis of oral epithelial dysplasia. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 1995;80(2):188–91.
4. Fischer DJ, Epstein JB, Morton TH, Schwartz SM. Interobserver reliability in the histopathologic diagnosis of oral pre-malignant and malignant lesions. *J Oral Pathol Med*. 2004;33(2):65–70.
5. Karabulut A, Reibel J, Therkildsen MH, Praetorius F, Nielsen HW, Dabelsteen E. Observer variability in the histologic assessment of oral premalignant lesions. *J Oral Pathol Med*. 1995;24(5):198–200.
6. Tomioka H, Morita K, Hasegawa S, Omura K. Gene expression analysis by cDNA microarray in oral squamous cell carcinoma. *J Oral Pathol Med*. 2006;35(4):206–11.
7. Tsai WC, Tsai ST, Ko JY, et al. The mRNA profile of genes in betel quid chewing oral cancer patients. *Oral Oncol*. 2004;40(4):418–26.
8. Nagata M, Fujita H, Ida H, et al. Identification of potential biomarkers of lymph node metastasis in oral squamous cell carcinoma by cDNA microarray analysis. *Int J Cancer*. 2003;106(5):683–9.
9. Nagaraja GM, Othman M, Fox BP, et al. Gene expression signatures and biomarkers of noninvasive and invasive breast cancer cells: comprehensive profiles by representational difference analysis, microarrays and proteomics. *Oncogene*. 2006;25(16):2328–38.
10. Adib TR, Henderson S, Perrett C, et al. Predicting biomarkers for ovarian cancer using gene-expression microarrays. *Br J Cancer*. 2004;90(3):686–92.
11. Baginsky S, Hennig L, Zimmermann P, Gruissem W. Gene expression analysis, proteomics, and network discovery. *Plant Physiol*. 2010;152(2):402–10.

12. de Snoo F, Bender R, Glas A, Rutgers E. Gene expression profiling: decoding breast cancer. *Surg Oncol*. 2009;18(4):366–78.

13. Carinci F, Lo Muzio L, Piattelli A, et al. Genetic portrait of mild and severe lingual dysplasia. *Oral Oncol*. 2005a;41(4):365–74.

14. Banerjee AG, Bhattacharyya I, Vishwanatha JK. Identification of genes and molecular pathways involved in the progression of premalignant oral epithelia. *Mol Cancer Ther*. 2005;4(6):865–75.

15. Kondoh N, Ohkura S, Arai M, et al. Gene expression signatures that can discriminate oral leukoplakia subtypes and squamous cell carcinoma. *Oral Oncol*. 2007;43(5):455–62.

16. Odani T, Ito D, Li MH, et al. Gene expression profiles of oral leukoplakia and carcinoma: genome-wide comparison analysis using oligonucleotide microarray technology. *Int J Oncol*. 2006;28(3):619–24.

17. Li N, Jian X, Hu Y, Xu C, Yao Z, Zhong X. Discovery of novel biomarkers in oral submucous fibrosis by microarray analysis. *Cancer Epidemiol Biomarkers Prev*. 2008;17(9):2249–59.

18. Kuribayashi Y, Morita KI, Tomioka H, Uekusa M, Ito D, Omura K. Gene expression analysis by oligonucleotide microarray in oral leukoplakia. *J Oral Pathol Med*. 2009;38(4):356–61.

19. Watanabe S, Sato K, Okazaki Y, Tonogi M, Tanaka Y, Yamane GY. Activation of PI3K-AKT pathway in oral epithelial dysplasia and early cancer of tongue. *Bull Tokyo Dent Coll*. 2009;50(3):125–33.

20. Saintigny P, Zhang L, Fan YH, et al. Gene expression profiling predicts the development of oral cancer. *Cancer Prev Res (Phila)*. 2011;4(2):218–29.

21. Sumino J, Uzawa N, Okada N, et al. Gene expression changes in initiation and progression of oral squamous cell carcinomas revealed by laser microdissection and oligonucleotide microarray analysis. *International Journal of Cancer. Journal International du Cancer*. 2013;132(3):540–8.

22. Alibes A, Yankilevich P, Canada A, Diaz-Uriarte R. IDconverter and IDClight: conversion and annotation of gene and protein IDs. *BMC Bioinformatics*. 2007;8:9.

23. Belbin TJ, Singh B, Barber I, et al. Molecular classification of head and neck squamous cell carcinoma using cDNA microarrays. *Cancer Res*. 2002;62(4):1184–90.

24. Leethanakul C, Knezevic V, Patel V, et al. Gene discovery in oral squamous cell carcinoma through the Head and Neck Cancer Genome Anatomy Project: confirmation by microarray analysis. *Oral Oncol*. 2003;39(3):248–58.

25. Chen C, Mendez E, Houck J, et al. Gene Expression Profiling Identifies Genes Predictive of Oral Squamous Cell Carcinoma. *Cancer Epidemiol Biomarkers Prev*. 2008;17(8):2152–62.

26. Sok JC, Kuriakose MA, Mahajan VB, Pearlman AN, DeLacure MD, Chen FA. Tissue-specific gene expression of head and neck squamous cell carcinoma in vivo by complementary DNA microarray analysis. *Arch Otolaryngol Head Neck Surg*. 2003;129(7):760–70.

27. 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*. 2003;39(1):37–48.

28. Nguyen ST, Hasegawa S, Tsuda H, et al. Identification of a predictive gene expression signature of cervical lymph node metastasis in oral squamous cell carcinoma. *Cancer Sci*. 2007;98(5):740–6.

29. Erdem NF, Carlson ER, Gerard DA. Characterization of gene expression profiles of 3 different human oral squamous cell carcinoma cell lines with different invasion and metastatic capacities. *J Oral Maxillofac Surg*. 2008; 66(5):918–27.

30. Leethanakul C, Patel V, Gillespie J, et al. Gene expression profiles in squamous cell carcinomas of the oral cavity: use of laser capture microdissection for the construction and analysis of stage-specific cDNA libraries. *Oral Oncol*. 2000;36(5):474–83.

31. Irie T, Aida T, Tachikawa T. Gene expression profiling of oral squamous cell carcinoma using laser microdissection and cDNA microarray. *Med Electron Microsc*. 2004;37(2):89–96.

32. Mendez E, Cheng C, Farwell DG, et al. Transcriptional expression profiles of oral squamous cell carcinomas. *Cancer*. 2002;95(7):1482–94.

33. Ha PK, Benoit NE, Yochem R, et al. A transcriptional progression model for head and neck cancer. *Clin Cancer Res*. 2003;9(8):3058–64.

34. Carinci F, Lo Muzio L, Piattelli A, et al. Potential markers of tongue tumor progression selected by cDNA microarray. *Int J Immunopathol Pharmacol*. 2005b;18(3):513–24.

35. Ding Y, Xu L, Jovanovic BD, et al. The methodology used to measure differential gene expression affects the outcome. *J Biomol Tech*. 2007;18(5):321–30.

36. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer*. 1953;6(5):963–8.

37. Papadimitrakopoulou VA, Shin DM, Hong WK. Molecular and cellular biomarkers for field cancerization and multistep process in head and neck tumorigenesis. *Cancer Metastasis Rev*. 1996;15(1):53–76.

38. Vigneswaran N, Beckers S, Waigel S, et al. Increased EMMPRIN (CD 147) expression during oral carcinogenesis. *Exp Mol Pathol*. 2006;80(2):147–59.

39. Hunter KD, Thurlow JK, Fleming J, et al. Divergent routes to oral cancer. *Cancer Res*. 2006;66(15):7405–13.

40. Hu Y, Jian X, Peng J, Jiang X, Li N, Zhou S. Gene expression profiling of oral submucous fibrosis using oligonucleotide microarray. *Oncol Rep*. 2008;20(2):287–94.

41. Mehrel T, Hohl D, Rothnagel JA, et al. Identification of a major keratinocyte cell envelope protein, loricrin. *Cell*. 1990;61(6):1103–12.

42. Mehul B, Bernard D, Schmidt R. Calmodulin-like skin protein: a new marker of keratinocyte differentiation. *J Invest Dermatol*. 2001;116(6):905–9.

43. Bloor BK, Seddon SV, Morgan PR. Gene expression of differentiation-specific keratins in oral epithelial dysplasia and squamous cell carcinoma. *Oral Oncol*. 2001;37(3):251–61.

44. Lindberg K, Rheinwald JG. Suprabasal 40 kd keratin (K19) expression as an immunohistologic marker of premalignancy in oral epithelium. *Am J Pathol*. 1989;134(1):89–98.

45. Coltrera MD, Zarbo RJ, Sakr WA, Gown AM. Markers for dysplasia of the upper aerodigestive tract. Suprabasal expression of PCNA, p53, and CK19 in alcohol-fixed, embedded tissue. *Am J Pathol*. 1992;141(4):817–25.

46. Takeda T, Sugihara K, Hirayama Y, Hirano M, Tanuma JI, Semba I. Immunohistological evaluation of Ki-67, p63, CK19 and p53 expression in oral epithelial dysplasias. *J Oral Pathol Med*. 2006;35(6):369–75.

47. Ohkura S, Kondoh N, Hada A, et al. 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*. 2005;41(6):607–13.

48. Fillies T, Jogschies M, Kleinheinz J, Brandt B, Joos U, Buerger H. Cytokeratin alteration in oral leukoplakia and oral squamous cell carcinoma. *Oncol Rep*. 2007;18(3):639–43.

49. Erdogan E, Klee EW, Thompson EA, Fields AP. Meta-analysis of oncogenic protein kinase Ciota signaling in lung adenocarcinoma. *Clin Cancer Res*. 2009;15(5):1527–33.

50. Allander SV, Illei PB, Chen Y, et al. Expression profiling of synovial sarcoma by cDNA microarrays: association of ERBB2, IGFBP2, and ELF3 with epithelial differentiation. *Am J Pathol*. 2002;161(5):1587–95.

51. Strieter RM, Belperio JA, Phillips RJ, Keane MP. CXC chemokines in angiogenesis of cancer. *Semin Cancer Biol*. 2004;14(3):195–200.

52. Strieter RM, Burdick MD, Mestas J, Gomperts B, Keane MP, Belperio JA. Cancer CXC chemokine networks and tumour angiogenesis. *Eur J Cancer*. 2006;42(6):768–78.

53. Ruiz-Garcia E, Scott V, Machavoine C, et al. Gene expression profiling identifies Fibronectin 1 and CXCL9 as candidate biomarkers for breast cancer screening. *Br J Cancer*. 2010.

54. Yuvaraj S, Griffin AC, Sundaram K, Kirkwood KL, Norris JS, Reddy SV. A novel function of CXCL13 to stimulate RANK ligand expression in oral squamous cell carcinoma cells. *Mol Cancer Res*. 2009;7(8):1399–407.

55. Pandruvada SN, Yuvaraj S, Liu X, et al. Role of CXC chemokine ligand 13 in oral squamous cell carcinoma associated osteolysis in athymic mice. *Int J Cancer*. 2010.

56. Malakhov MP, Malakhova OA, Kim KI, Ritchie KJ, Zhang DE. UBP43 (USP18) specifically removes ISG15 from conjugated proteins. *J Biol Chem*. 2002;277(12):9976–81.

57. Zou W, Kim JH, Handidu A, et al. Microarray analysis reveals that Type I interferon strongly increases the expression of immune-response related genes in Ubp43 (Usp18) deficient macrophages. *Biochem Biophys Res Commun*. 2007;356(1):193–9.

58. Duex JE, Comeau L, Sorkin A, Purow B, Kefas B. Usp18 regulates epidermal growth factor (EGF) receptor expression and cancer cell survival via microRNA-7. *J Biol Chem*. 2011;286(28):25377–86.

59. Hallen LC, Burki Y, Ebeling M, et al. Antiproliferative activity of the human IFN-alpha-inducible protein IFI44. *J Interferon Cytokine Res*. 2007; 27(8):675–80.

60. Ye H, Yu T, Temam S, et al. Transcriptomic dissection of tongue squamous cell carcinoma. *BMC Genomics*. 2008;9:69.

61. Toruner GA, Ulger C, Alkan M, et al. Association between gene expression profile and tumor invasion in oral squamous cell carcinoma. *Cancer Genet Cytogenet*. 2004;154(1):27–35.

62. Guttinger M, Sutti F, Panigada M, et al. Epithelial V-like antigen (EVA), a novel member of the immunoglobulin superfamily, expressed in embryonic epithelia with a potential role as homotypic adhesion molecule in thymus histogenesis. *J Cell Biol*. 1998;141(4):1061–71.

63. Ashida S, Nakagawa H, Katagiri T, et al. Molecular features of the transition from prostatic intraepithelial neoplasia (PIN) to prostate cancer: genome-wide gene-expression profiles of prostate cancers and PINs. *Cancer Res*. 2004;64(17):5963–72.

64. Ha PK, Califano JA. Promoter methylation and inactivation of tumour-suppressor genes in oral squamous-cell carcinoma. *Lancet Oncol*. 2006;7(1): 77–82.

65. Seoane JM, Varela-Centelles PI, Ramirez JR, Cameselle-Teijeiro J, Romero MA, Aguirre JM. Heat shock proteins (HSP70 and HSP27) as markers of epithelial dysplasia in oral leukoplakia. *Am J Dermatopathol*. 2006;28(5):417–22.

66. Kaur J, Srivastava A, Ralhan R. Expression of 70-kDa heat shock protein in oral lesions: marker of biological stress or pathogenicity. *Oral Oncol*. 1998;34(6):496–501.

67. Hulett MD, Freeman C, Hamdorf BJ, Baker RT, Harris MJ, Parish CR. Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis. *Nat Med*. 1999;5(7):803–9.

68. Ikuta M, Podyma KA, Maruyama K, Enomoto S, Yanagishita M. Expression of heparanase in oral cancer cell lines and oral cancer tissues. *Oral Oncol*. 2001;37(2):177–84.

69. Shinyo Y, Kodama J, Hongo A, Yoshinouchi M, Hiramatsu Y. Heparanase expression is an independent prognostic factor in patients with invasive cervical cancer. *Ann Oncol*. 2003;14(10):1505–10.

70. Saxne T, Heinegard D. Cartilage oligomeric matrix protein: a novel marker of cartilage turnover detectable in synovial fluid and blood. *Br J Rheumatol*. 1992;31(9):583–91.

71. Xiao Y, Kleeff J, Guo J, et al. Cartilage oligomeric matrix protein expression in hepatocellular carcinoma and the cirrhotic liver. *J Gastroenterol Hepatol*. 2004;19(3):296–302.

72. Stanbrough M, Bubley GJ, Ross K, et al. Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. *Cancer Res*. 2006;66(5):2815–25.

73. Schuetz CS, Bonin M, Clare SE, et al. Progression-specific genes identified by expression profiling of matched ductal carcinomas in situ and invasive breast tumors, combining laser capture microdissection and oligonucleotide microarray analysis. *Cancer Res*. 2006;66(10):5278–86.

74. Kowara R, Chen Q, Milliken M, Chakravarthy B. Calpain-mediated truncation of dihydropyrimidinase-like 3 protein (DPYSL3) in response to NMDA and H2O2 toxicity. *J Neurochem*. 2005;95(2):466–74.

75. Braakhuis BJ, Senft A, de Bree R, et al. Expression profiling and prediction of distant metastases in head and neck squamous cell carcinoma. *J Clin Pathol*. 2006;59(12):1254–60.

76. Falvella FS, Colombo F, Spinola M, Campiglio M, Pastorino U, Dragani TA. BHLHB3: a candidate tumor suppressor in lung cancer. *Oncogene*. 2008;27(26):3761–4.

77. Shao R, Bao S, Bai X, et al. Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression. *Mol Cell Biol*. 2004;24(9):3992–4003.

78. AbdulMajeed AA, Dalley AJ, Farah CS. Putative cancer stem cell marker expression in oral epithelial dysplasia and squamous cell carcinoma. *J Oral Pathol Med*. 2013:doi: 10.1111/jop.12073.

79. Dalley AJ, AbdulMajeed AA, Upton Z, Farah CS. Organotypic culture of normal, dysplastic and squamous cell carcinoma-derived oral cell lines reveals loss of spatial regulation of CD44 and p75 NTR in malignancy. *J Oral Pathol Med*. 2013;42(1):37–46.

80. Sun W, Liu Y, Glazer CA, et al. TKTL1 is activated by promoter hypomethylation and contributes to head and neck squamous cell carcinoma carcinogenesis through increased aerobic glycolysis and HIF1alpha stabilization. *Clinical Cancer Research*. 2010;16(3):857–66.

81. Zhang X, Han S, Han HY, et al. Risk prediction for malignant conversion of oral epithelial dysplasia by hypoxia related protein expression. *Pathology*. 2013;45(5):478–83.

82. AbdulMajeed AA, Dalley AJ, Farah CS. Loss of ELF3 immunoexpression is useful for detecting oral squamous cell carcinoma but not for distinguishing between grades of epithelial dysplasia. *Ann Diagn Pathol*. 2013;17(4):331–40.

83. Warnakulasuriya S, Reibel J, Bouquot J, Dabelsteen E. Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. *J Oral Pathol Med*. 2008;37(3):127–33.