Genetics and Molecular Research

Screening of crucial long non-coding RNAs in oral epithelial dysplasia by serial analysis of gene expression

X. Han, Y.B. Wei, G. Tian, Z. Tang, J.Y. Gao and X.G. Xu

Department of Stomatology, Changhai Hospital, Second Military Medical University, Shanghai, China

Corresponding author: X.G. Xu
E-mail: xxgxxgdd@163.com

Genet. Mol. Res. 14 (4): 11729-11738 (2015)
Received January 30, 2015
Accepted August 18, 2015
Published October 2, 2015
DOI http://dx.doi.org/10.4238/2015.October.2.6

ABSTRACT. Oral epithelial dysplasia (OED) is a premalignant lesion of the oral mucosa. Considering the poor 5-year survival rate of oral cancer, further investigation is needed in order to determine the pathogenesis of OED. In the present study, serial analysis of gene expression (SAGE) data from patients with OED were compared to normal controls to identify differentially expressed genes (DEGs). SAGE data were obtained from the Gene Expression Omnibus, and included samples from patients with mild, moderate, or severe dysplasia. The DEGs were identified using the edgeR software package and functional-enrichment analysis was performed with the DAVID (https://david.ncifcrf.gov/) software program. The co-expression network was constructed using the CoExpress software and target genes of long non-coding RNAs (lncRNAs) were predicted according to the proximity between the lncRNAs and mRNAs in the genome. A total of 517 DEGs were identified, including 409 mRNAs and 108 lncRNAs. Functional-enrichment analysis showed that mRNAs and lncRNAs involved in epithelial cell differentiation, epithelium development, and epidermal cell differentiation were significantly enriched in the DEGs. Thirty-eight potential regulatory relationships were unveiled between lncRNAs and mRNAs, and

two subnetworks were discovered by analyzing the topological properties of the co-expression network. In conclusion, we have identified key mRNAs and lncRNAs in OED, and these findings may aid in understanding the pathogenesis of OED and advance potential future treatments

Key words: Oral epithelial dysplasia; Long non-coding RNAs; SAGE; Differentially expressed genes

INTRODUCTION

Oral epithelial dysplasia (OED) is a premalignant lesion of the oral mucosa and can be graded as mild, moderate, or severe according to the cellular histology. Clinically, OED may be defined as leukoplakia (white lesion), erythroplakia (red lesion), or leukoerythroplakia (mixed lesion), depending on the clinical phenotype. The rate of malignant transformation of oral leukoplakia into oral mucosal squamous cell carcinoma (OMSCC) varies (Gupta et al., 1980; Schepman et al., 1998). Despite advances in surgery, radiotherapy and chemotherapy, the 5-year survival rate for oral cancer has not improved significantly over the past decade, and remains at approximately 50% (Silverman, 2001).

Many techniques have been used to study the complicated molecular mechanisms underlying OED. Banerjee et al. (2005) examined the transcriptome of OED and identified nearly 1700 differentially expressed genes (DEGs). Ohkura et al. (2005) reported the differential expression of several members of the keratin family by reverse transcription-polymerase chain reaction and immunohistochemical analyses. These reports have achieved some insight into the mechanism of OED and other publications have identified several potential biomarkers, such as heat shock protein 27 (Leonardi et al., 2002), p63 (Takeda et al., 2006), Ki-67 (Takeda et al., 2006) and p53 (Chiang et al., 2000).

Long non-coding RNAs (lncRNAs) are a class of mRNA-like transcripts that do not have the capacity to code proteins; they have a variety of functions, including roles in epigenetics and gene regulation (Mercer et al., 2009). The aberrant expression of lncRNAs has been associated with human cancers, suggesting a role in tumorigenesis (Gibb et al., 2011a; Yang et al., 2011).

In this study, we compared serial analysis of gene expression (SAGE) data from patients with OED to normal controls, to identify differentially expressed mRNAs and lncRNAs and the regulatory relationships between them. The results of this study provide an interesting perspective on OED and its underlying mechanisms.

MATERIAL AND METHODS

Data source

The oral premalignant lesion lncRNA profile dataset was obtained from the Gene Expression Omnibus (GEO) database under GSE31021 accession number (Gibb et al., 2011b), which included two mild dysplasia, four moderate dysplasia, and four severe dysplasia samples. In addition, six normal oral samples previously deposited as GSE8127 were used as normal controls (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8127). These two sets of data were based on the SAGE–Seq platform (GPL4 SAGE:10:NlaIII:*Homo sapiens*).

Annotation of SAGE tags

An expression matrix was constructed for unique SAGE tags of all samples using a Perl script. SAGE tags were then mapped to Unigene IDs (ftp://ftp1.nci.nih.gov/pub/SAGE/HUMAN/Hs_short.best_gene.gz). If more than one tag was mapped to a single Unigene ID, average tag counts were calculated and then assigned to the Unigene ID. Finally, two matrices were obtained: the unmapped tag matrix and the Unigene ID mapped matrix. Each Unigene ID with tag counts <2 and unmapped tags were removed.

In the Unigene ID mapped matrix, annotation tags were regarded as candidate lncRNAs if they did not have a gene name or contained any of the following: non-coding, non-protein, cDNA, transcribed locus, clone IMAGE, chr (#), orf (#), hypothetical, family with sequence similarity, FLJ (#), or KIAA (#). All remaining annotation tags were classified as mRNAs. In the unmapped tag matrix, unmapped tags were matched to lncRNA reference sequences using the Bowtie software (Langmead et al., 2009). The forward tag matches were retained and reverse tag matches were removed. Finally, candidate lncRNAs in the Unigene ID mapped matrix and matched lncRNAs in the unmapped tag matrix were combined as the lncRNA expression matrix. Therefore, the mRNA expression matrix and the candidate lncRNA expression matrix were acquired.

Screening of DEGs

The lncRNA and mRNA expression matrices were combined into a gene expression matrix. The DEGs were identified using the edgeR Bioconductor package (Robinson et al., 2010) based on the exact negative binomial test (Robinson and Smyth, 2008). Benjamini and Hochberg (1995) published an algorithm that was used to adjust the P values and to obtain false-discovery rate (FDR). The genes with a log |(Fold-Change)| > 1 and FDR < 0.05 were considered as DEGs.

Functional-enrichment analysis

The Gene Ontology project (Ashburner et al., 2000) is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases. The DAVID (https://david.ncifcrf.gov/) software program (Huang da et al., 2009) was used to identify over-represented gene ontology (GO) terms in biological process and pathways. A P value <0.05 was considered as the threshold for analysis using the hypergeometric distribution.

Gene co-expression analysis

The selected DEGs were analyzed by the CoExpress software (http://www.bioinformatics.lu/CoExpress/), which predicts the interactive relationship between genes by calculating the co-expression coefficient. The interactive pairs with a |Pearson correlation| > 0.6 were obtained and all other parameters were set at default. All co-expression pairs were visualized using Cytoscape (Shannon et al., 2003).

Identification of potential target genes of lncRNAs

The potential target genes of lncRNAs were predicted according to their proximity to lncRNAs in the genome. Overlap or close distance in chromosomal location between lncRNAs and mRNA allowed us to identify potential gene targets.

Topology analysis of the co-expresssion network

The topological features (e.g., degree distribution and clustering coefficient) of the co-expression network between the lncRNAs and mRNAs were analyzed to identify the critical lncRNAs. Molecular complex detection (MCODE, Cytoscape plug-in) detects dense and connected regions based on local network density - a modified measure of the clustering coefficient to screen gene cluster (Bader and Hogue, 2003). If the gene cluster included related lncRNAs then they may be important candidates. The interactive nodes with a degree >5 and a cluster score >10 were selected in this study.

RESULTS

Annotation of the SAGE tags

Annotation results of SAGE tags are shown in Figure 1. Sample GSM194652 had the highest number of genes and lncRNAs.

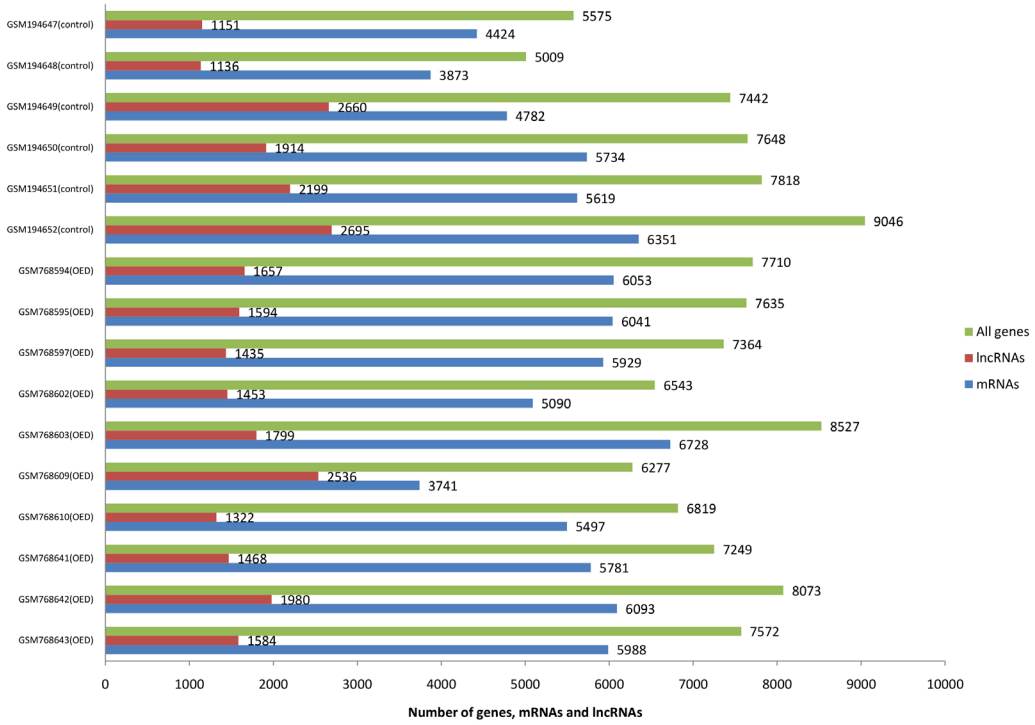


Figure 1. SAGE tag annotation results of the 16 samples (OED and control).

Differentially expressed mRNAs and lncRNAs

A total of 517 DEGs were identified, including 409 mRNAs and 108 lncRNAs (Figure 2). Among the differentially expressed mRNAs, 275 were upregulated (the top five were PSCA,

AMY2B, TJP3, VSIG2, and SOCS5) and 134 were downregulated (the top 5 were CA2, IGHA1, MMP1, HST2, and FABP4). Among the differentially expressed lncRNAs, 87 were upregulated (e.g., LINC00675, HCG22, MIR17HG, FAM3D, LINC00152, FAM45A, FAM129B, and NEAT1) and 21 were downregulated (e.g., SNHG6, HCG11, and LINC00116).

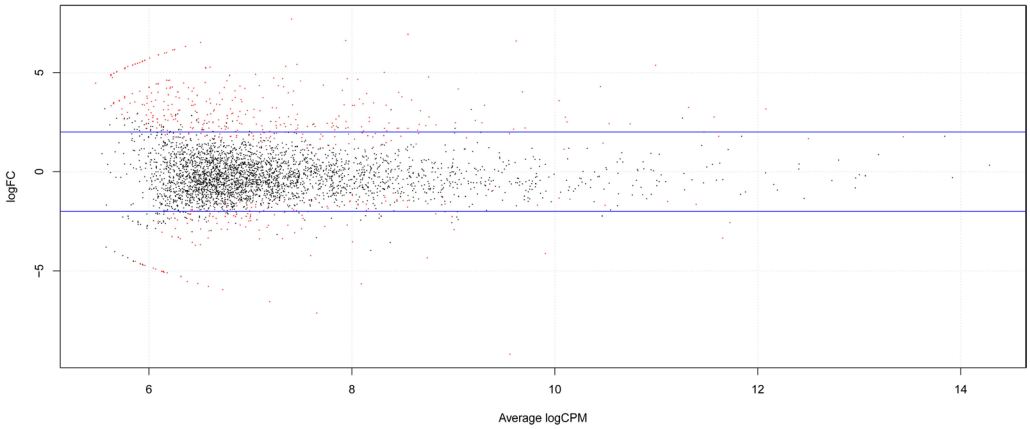


Figure 2. Distribution of logFC and logCPM of the DEGs. Red dots represent DEGs, black dots represent un-DEGs and blue lines indicate the range of |logFC| = 2.

Functional enrichment for differentially expressed mRNAs

Using the DAVID software, we found that the differentially expressed mRNAs were enriched into 47 GO terms, among which the most significant were epithelial cell differentiation, epithelium development, and epidermal cell differentiation (Table 1).

Table 1. The top Gene Ontology (GO) terms enriched for the differentially expressed genes.

GO Term	Description	P value	Genes
GO:0030855	epithelial cell differentiation	7.54E-05	B4GALT1, ELF3, S100A7, DHRS9, EHF, SPINK5, SCEL, LAMA3, RHCG, SPRR2A, CNFN, TGM3, EMP1
GO:0006091	generation of precursor metabolites and energy	1.17E-04	NDUFA4, NDUFB4, LDHA, ACO2, FDXR, IDH3B, BPGM, PPP1CC, ATP5G3, NDUFA12, NDUFB2, NDUFS6, ATP6V1C2, PPP1R3C, UQCRH, ATPIF1, ERO1L, PDHA1, COX17, ATP5J
GO:0060429	epithelium development	2.47E-04	B4GALT1, ELF3, S100A7, DHRS9, EHF, SPINK5, SCEL, PFN1, LAMA3, RHCG, SPRR2A, ALDH1A3, CNFN, TGM3, CA2, EMP1
GO:0006955	immune response	6.06E-04	IGHG1, ITGAL, KYNU, S100A7, IL18, NLRX1, VTN, DEFB4A, CCL5, FTH1, B2M, SQSTM1, RNASE7, TICAM1, TAP1, IGHA1, SEMA3C, DEFB1, APLN, IL1A, CD27, CRISP3, IL1RN, IGJ, CCL18, APOL1, ANXA11, AICDA, CTSC, HSPD1, HLA-DRA
GO:0009913	epidermal cell differentiation	0.006	LAMA3, S100A7, SPRR2A, CNFN, TGM3, SPINK5, SCEL
GO:0006119	oxidative phosphorylation	0.007	NDUFA4, NDUFS6, NDUFB4, ATP6V1C2, UQCRH, ATP5G3, ATP5J, NDUFB2
GO:0006766	vitamin metabolic process	0.008	KYNU, ACADM, RFK, ALDH1A3, CRABP2, DHRS9, NAPRT1
GO:0006916	anti-apoptosis	0.008	PGAP2, SQSTM1, F3, EEF1A2, SERPINB2, NFKBIA, PIM2, PRNP, MYC, CD27, TAX1BP1, IL1A
GO:0008544	epidermis development	0.010	LAMA3, ELF3, S100A7, SPRR2A, CRABP2, CNFN, TGM3, SPINK5, SCEL, EMP1, FABP5
GO:0022900	electron transport chain	0.016	NDUFA4, NDUFS6, NDUFB4, UQCRH, FDXR, ERO1L, NDUFA12, NDUFB2

Co-expression between differentially expressed mRNAs and lncRNAs

Using the threshold value of |Pearson correlation| > 0.6, we obtained 3449 relationship pairs, which were then visualized by the Cytoscape software (Figure 3).

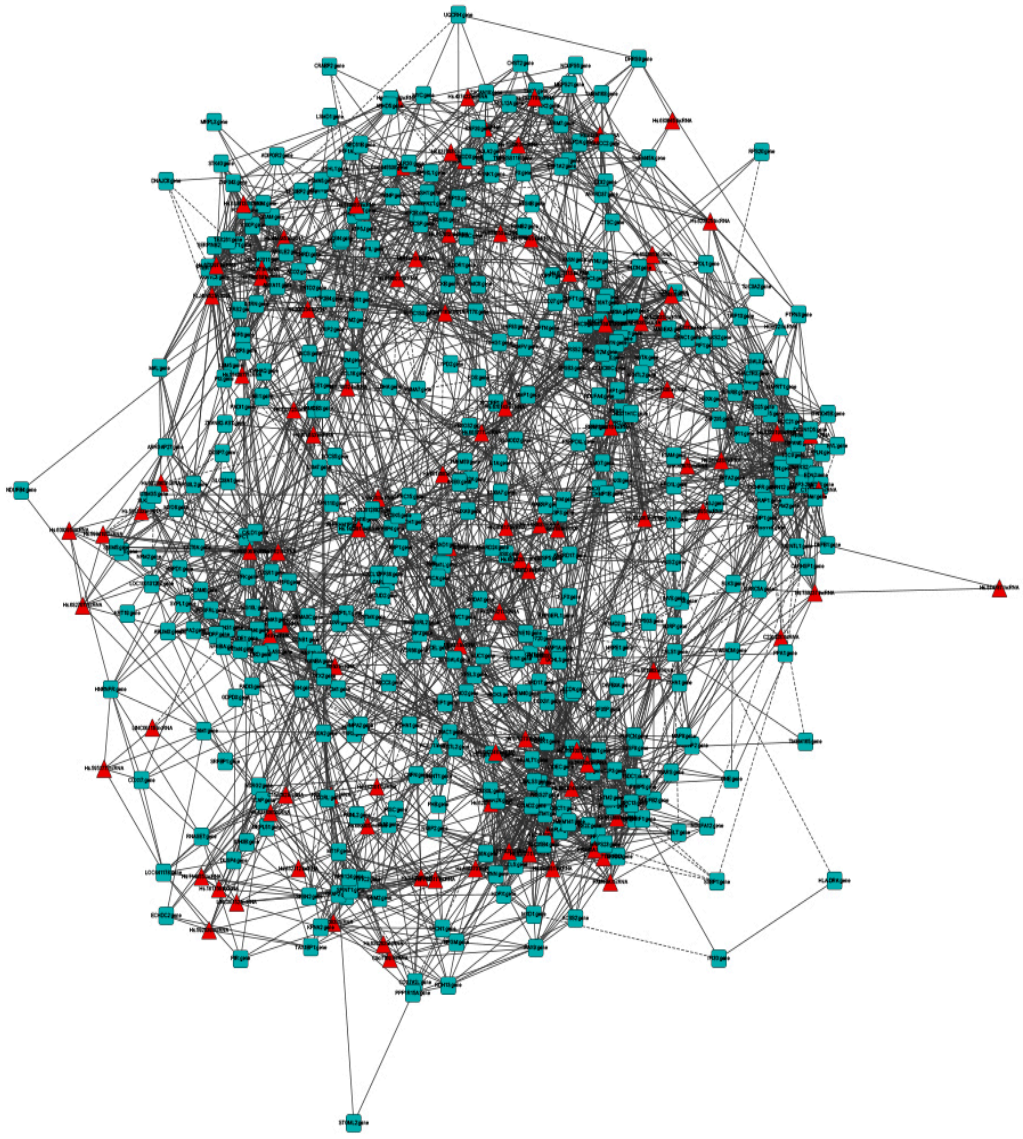


Figure 3. Co-expresssion network of differentially expressed mRNAs and lncRNAs. The green squares represent differentially expressed mRNAs and the red triangles represent differentially expressed dlncRNAs. The solid lines indicate the positive correlation between genes and the dashed lines indicate the negative correlation between genes.

Predicted target genes of lncRNAs

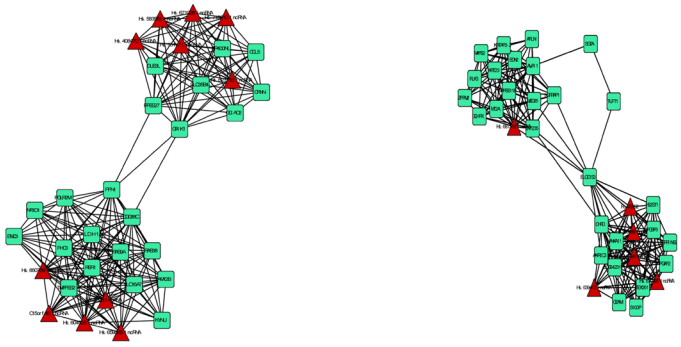
According to the genomic distances between lncRNAs and mRNAs, we obtained 38 potential regulatory relationships between lncRNAs and mRNAs (Table 2). Among them, there was overlap between two relationship gene pairs in genome location, including Hs.402083(+) and CKAP2(-); LINC00675(+) and KRTAP3-2(+); Hs.585133(+) and FDXR(-); Hs.659665(+) and DDI2(+); and Hs.713907(+) and EEF1A2(+).

Table 2. Regulatory relationship pairs between lncRNAs and mRNAs.

Gene	Location of gene	lncRNA	Location of lncRNA
KIF2B	17q22	LOC645638	17q23.1
AVPI1	10q24.2	Hs.97536	10q26
EHF	11p12	Hs.710610	11p15-11q25
MYEOV	11q13	Hs.710610	11p15-11q25
ARNTL2	12p12.2-p11.2	Hs.604538	12p13-12q24.3
SLC16A7	12q13	Hs.604538	12p13-12q24.3
SLC16A7	12q13	MIR17HG	13q31.3
TBC1D15	12q21.1	Hs.693806	12q13
CKAP2	13q14	Hs.402083	13q14
ALDH1A3	15q26.3	C15orf48	15q21.1
KRTAP3-2	17q12-q21	LINC00675	17p13.1-p12
ARHGAP27	17q21.31	Hs.585133	17q22-17q23
FDXR	17q24-q25	Hs.585133	17q22-17q23
RER1	1p36	Hs.659665	1p36.2-1q44
DDI2	1p36.21	Hs.659665	1p36.2-1q44
GRHL1	2p25.1	Hs.658800	2p32
FLNB	3p14.3	Hs.594960	3q29
MAP9	4q32.1	Hs.676371	4p16-4q35
SPINK5	5q32	Hs.671928	5q31
UBD	6p21.3	C6orf162	6q15-q16.1
MYO6	6q13	C6orf162	6q15-q16.1
LETM2	8p11.23	Hs.614103	8p23-8q24.3
FABP5	8q21.13	Hs.614103	8p23-8q24.3
PIM2	Xp11.23	Hs.600254	xp22,xp11,xq12-28
SLC38A1	12q13.11	Hs.693806	12q13
TSPAN31	12q13.3	Hs.6061	12q24.2-.3
AMY2B	1p21	Hs.659665	1p36.2-1q44
TP53BP2	1q41	C1orf85	1q22
EEF1A2	20q13.3	Hs.713907	20p13-20q13.3
SOCS5	2p21	Hs.560908	2p25-2q37
DHRS9	2q31.1	Hs.680646	2q31
ANKRD17	4q13.3	Hs.480068	4q21
AREG	4q13.3	Hs.480068	4q21
PABPC4L	4q28.3	Hs.480068	4q21
ELMOD2	4q31.1	Hs.480068	4q21
FDCSP	4q13	Hs.570860	4q33
LYRM7	5q23.3	Hs.670140	5q32-5q33
INVS	9q31	Hs.651983	9q21

Topological properties of the co-expression network

Topological characteristics of the co-expression network were investigated to screen for important lncRNAs based upon the degree of relationship. The entire co-expression network was analyzed with MCODE and a score >10 was set as the threshold. Two subnetworks were revealed and visualized with Cytoscape (Shannon et al., 2003) (Figure 4A and B). In Cluster A, 11 lncRNAs were included and the seed was Hs.604538. In Cluster B, 5 lncRNAs were included and the seed was heparin sulfate 2-O-sulfotransferase 1 (HS2ST1).



A B

Figure 4. Two subnetworks extracted from the co-expression network. The green squares represent differentially expressed mRNAs and the red triangles represent differentially expressed lncRNAs. Subnetwork (**A**) included 11 lncRNAs and Hs.604538 was used as the seed. Subnetwork (**B**) included 5 lncRNAs and HS2ST1 was used as the seed.

DISCUSSION

A total of 517 DEGs were identified in our study of samples from normal and OED patients, including 409 mRNAs and 108 lncRNAs. Pathways closely related to OED, including epithelial cell differentiation, epithelium development, and epidermal cell differentiation, were significantly enriched in differentially expressed mRNAs. Pathways involved in immune response were also enriched, which correlated with the progression of oral epithelium from hyperkeratosis to dysplasia and carcinoma (Gannot et al., 2002). In addition, pathways in apoptosis and programmed cell death were significantly enriched in DEGs and it is well known that abnormal apoptosis plays an important role in tumorigenesis (Evan and Vousden, 2001). In fact, Macluskey et al. (2000) found that disease progression in the oral mucosa is accompanied by increases in both epithelial proliferation and apoptosis. Piattelli et al. (2002) reported that an inverse relationship is found between bcl-2 expression, cell proliferation (MIB-1, an E3 ubiquitin-protein ligase involved in regulating apoptosis) and the apoptotic index. In our study, several genes associated with keratinocyte proliferation were identified. It has been reported that induction of Kruppel-like factor 4 (KLF4) in basal keratinocytes blocks the proliferation-differentiation switch and initiates squamous epithelial dysplasia (Foster et al., 2005). Therefore, DEGs involved in these pathways may be good candidates for future therapeutic targets.

To further decipher the role of lncRNAs in the pathogenesis of OED, the target genes of lncRNAs were predicted. Thirty-eight potential regulatory relationships were identified between lncRNAs and mRNAs. Solute carrier family 16 member 7 (SLC16A7) is a predicted gene target of MIR17HG. SLC16A7 is a member of the monocarboxylate transporter family and is responsible for metabolite transportation; its upregulation has been observed in renal cell carcinoma (Riss et al., 2006). MIR17HG is an miRNA cluster comprised of at least six different miRNAs. Pais et al. (2013) reported that deletion of MIR17HG decreases the activity of the mTOR (mammalian target of rapamycin) pathway in Burkitt lymphoma cells. Previous studies have shown that mTOR plays a critical role in several pathways that are involved in human cancer (Morgensztern and McLeod, 2005; Sabatini, 2006). In our study, four genes were predicted as targets of Hs.480068, one of which is amphiregulin (AREG), a member of the epidermal growth factor family. It has been reported that AREG is upregulated in oral lichen planus (Kumagai et al., 2010) suggesting that it may play a role in the development of OED.

Topological characteristics of the co-expression network were analyzed to discover two subnetworks, consisting of lncRNAs and mRNAs with close interactions that may play collective roles in the development of OED. Hs.604538 was the seed in one subnetwork while HS2ST1 was the seed in the other subnetwork. Cornulin (CRNN) was one of the gene targets in the subnetwork with Hs.604538 as the seed, and it has been reported to play a role in the mucosal/epithelial immune response and epidermal differentiation (Contzler et al., 2005). Schaaij-Visser et al. (2010) found that decreased expression of CRNN in oral leukoplakia is significantly associated with the presence of hyperkeratosis but not malignant transformation. However, Hsu et al. (2013) found that loss of CRNN expression is related to advanced tumor stage and poor survival in patients with esophageal squamous cell carcinoma. Future studies are needed to fully elucidate the role of CRNN in OED.

In this study, DEGs were identified in OED by comparative analysis of SAGE data of patients with OED to normal controls. Significant biological pathways were revealed by functional-enrichment analysis, and key lncRNAs and mRNAs were identified through co-expression network analysis. This information will potentially guide future research into novel therapies for OED and advance our understanding of its pathogenesis.

Conflicts of interest

The authors declare no conflict of interest.

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet*. 25: 25-29.

Bader GD and Hogue CW (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4: 2.

Banerjee AG, Bhattacharyya I and Vishwanatha JK (2005). Identification of genes and molecular pathways involved in the progression of premalignant oral epithelia. *Mol. Cancer Ther*. 4: 865-875.

Benjamini Y and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57: 289-300.

Chiang CP, Lang MJ, Liu BY, Wang JT, et al. (2000). Expression of p53 protein in oral submucous fibrosis, oral epithelial hyperkeratosis, and oral epithelial dysplasia. *J. Formosan Med. Assoc.* 99: 229-234.

Contzler R, Favre B, Huber M and Hohl D (2005). Cornulin, a new member of the “fused gene” family, is expressed during epidermal differentiation. *J. Invest. Dermatol.* 124: 990-997.

Evan GI and Vousden KH (2001). Proliferation, cell cycle and apoptosis in cancer. *Nature* 411: 342-348.

Foster KW, Liu Z, Nail CD, Li X, et al. (2005). Induction of KLF4 in basal keratinocytes blocks the proliferation-differentiation switch and initiates squamous epithelial dysplasia. *Oncogene* 24: 1491-1500.

Gannot G, Gannot I, Vered H, Buchner A, et al. (2002). Increase in immune cell infiltration with progression of oral epithelium from hyperkeratosis to dysplasia and carcinoma. *Br. J. Cancer* 86: 1444-1448.

Gibb EA, Brown CJ and Lam WL (2011a). The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* 10: 1-17.

Gibb EA, Enfield KS, Stewart GL, Lonergan KM, et al. (2011b). Long non-coding RNAs are expressed in oral mucosa and altered in oral premalignant lesions. *Oral Oncol*. 47: 1055-1061.

Gupta PC, Mehta FS, Daftary DK, Pindborg JJ, et al. (1980). Incidence rates of oral cancer and natural history of oral precancerous lesions in a 10-year follow-up study of Indian villagers. *Community Dent. Oral Epidemiol*. 8: 287-333.

Hsu PK, Kao HL, Chen HY, Yen CC, et al. (2013). Loss of CRNN expression is associated with advanced tumor stage and poor survival in patients with esophageal squamous cell carcinoma. *J. Thorac. Cardiovasc. Surg.* 147: 1612-1618.

Huang da W, Sherman BT and Lempicki RA (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc*. 4: 44-57.

Kumagai K, Horikawa T, Gotoh A, Yamane S, et al. (2010). Up-regulation of EGF receptor and its ligands, AREG, EREG, and HB-EGF in oral lichen planus. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod*. 110: 748-754.

Langmead B, Trapnell C, Pop M and Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to

the human genome. *Genome Biol*. 10: R25.
Leonardi R, Pannone G, Magro G, Kudo Y, et al. (2002). Differential expression of heat shock protein 27 in normal oral mucosa, oral epithelial dysplasia and squamous cell carcinoma. *Oncol. Rep*. 9: 261-266.
Macluskey M, Chandrachud LM, Pazouki S, Green M, et al. (2000). Apoptosis, proliferation, and angiogenesis in oral tissues. Possible relevance to tumour progression. *J. Pathol*. 191: 368-375.
Mercer TR, Dinger ME and Mattick JS (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet*. 10: 155-159.
Morgensztern D and McLeod HL (2005). PI3K/Akt/mTOR pathway as a target for cancer therapy. *Anticancer Drugs* 16: 797-803.
Ohkura S, Kondoh N, Hada A, Arai M, 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.
Pais F, Lee S, Rodic V, Barth MJ, et al. (2013). Transcription activator-like effector nucleases (TALENs)-mediated deletion of MIR17HG in Burkitt lymphoma cells decreases mTOR pathway activity and increases chemosensitivity. *Blood* 122: 243-243.
Piattelli A, Rubini C, Fioroni M, Iezzi G, et al. (2002). Prevalence of p53, bcl-2, and Ki-67 immunoreactivity and of apoptosis in normal oral epithelium and in premalignant and malignant lesions of the oral cavity. *J. Oral Maxillofac. Surg*. 60: 532-540.
Riss J, Khanna C, Koo S, Chandramouli GV, et al. (2006). Cancers as wounds that do not heal: differences and similarities between renal regeneration/repair and renal cell carcinoma. *Cancer Res*. 66: 7216-7224.
Robinson MD and Smyth GK (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* 9: 321-332.
Robinson MD, Mccarthy DJ and Smyth GK (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139-140.
Sabatini DM (2006). mTOR and cancer: insights into a complex relationship. *Nat. Rev. Cancer* 6: 729-734.
Schaaij-Visser TB, Bremmer JF, Braakhuis BJ, Heck AJ, et al. (2010). Evaluation of cornulin, keratin 4, keratin 13 expression and grade of dysplasia for predicting malignant progression of oral leukoplakia. *Oral Oncol*. 46: 123-127.
Schepman KP, Van Der Meij EH, Smeele LE, and Van Der Waal I (1998). Malignant transformation of oral leukoplakia: a follow-up study of a hospital-based population of 166 patients with oral leukoplakia from The Netherlands. *Oral Oncol*. 34: 270-275.
Shannon P, Markiel A, Ozier O, Baliga NS, et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 13: 2498-2504.
Silverman S (2001). Demographics and occurrence of oral and pharyngeal cancers. The outcomes, the trends, the challenge. *J. Am. Dent. Assoc*. 132: 7S-11S.
Takeda T, Sugihara K, Hirayama Y, Hirano M, et al. (2006). Immunohistological evaluation of Ki-67, p63, CK19 and p53 expression in oral epithelial dysplasias. *J. Oral Pathol. Med*. 35: 369-375.
Yang Z, Zhou L, Wu LM, Lai MC, et al. (2011). Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann. Surg. Oncol*. 18: 1243-1250.