Identification of Potential Crucial Immune-related Genes and Key Pathways in Oral Lichen Planus

Lou Geng

Shanghai Jiao Tong University School of Medicine Affiliated Ninth People's Hospital

Xingming Zhang

Shanghai Jiao Tong University School of Medicine Affiliated Ninth People's Hospital

Yi Tang

Shanghai Jiao Tong University School of Medicine Affiliated Ninth People's Hospital

Wenli Gu (✉ lilygu8@126.com)

Shanghai Jiao Tong University School of Medicine Affiliated Ninth People's Hospital

<https://orcid.org/0000-0001-7162-6017>

Research

Keywords: oral lichen planus, bioinformatic analysis, biomarkers, immune-related genes, transcription factors, pathway

Posted Date: August 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-728444/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Oral lichen planus (OLP) is a chronic autoimmune oral mucosal disease which seriously affect the life quality of the patients. But till now, the exact etiology and pathogenesis of OLP still remains unclear. The aim of our study was to find the key molecules and pathway involved in the pathogenesis mechanisms of OLP, therefore, provide reference for identifying more effective therapeutic strategies for OLP. Data from GSE52130 were downloaded from GEO data sets for analysis, and a total of 255 differentially expressed genes were selected out. Then by means of a series of gene enrichment analysis, we illustrated that most of these important genes were enriched in the focal adhesion pathway and metabolism pathways. In addition, we also constructed a protein-protein interaction (PPI) network and identified 8 hub genes which may play the important role in the development of OLP, including THBS2, COL3A1, IGFBP3, COL1A2, COL1A1, PTGS2, IL1B, ARG1. Lastly, we also predicted transcription factors of these hub genes, then constructed the corresponding transcription regulatory network. Taken together, our data indicated that the focal adhesion pathway and metabolism pathways, as well as 8 crucial immune-related genes might be most likely involved in the development of OLP, which would provide important information and objective basis for elucidating the pathogenesis of OLP and providing more effective targeted immunotherapy strategies for OLP.

1. Background

Oral lichen planus (OLP) is a chronic autoimmune oral mucosal disease affecting 1–2% population worldwide(1, 2), and is characterized by its recurrence and chronic protraction course. OLP has been defined as an oral potentially malignant disease (OPMD) by WHO in 2005 with a malignant change rate of 0.4–1.5 percent(2, 3), especially with a high-incidence malignancy in middle-aged and older adults. It can be divided into erosive and non-erosive OLP according to the condition of the base mucosa of the lesion. The disease mainly invades buccal mucosa, tongue and gingiva, so most patients may manifest ulcers, erosion, papules and mucosal exfoliation in the oral cavity, while others obtain rough, numb and painful feelings (4). Pain is one of the OLP common symptoms and may even interfere with the patient's speech, eating and swallowing.

It has been known that infiltration of T lymphocytes in lamina propria, formation of keratinocytes and destruction of basement membrane are typical histopathological features of OLP(5–7). The helper T cells are the main lymphocytes in the OLP, which can activate cytotoxic T cells and aggravate the local immune response, and thus induce apoptosis of keratinocytes(8).

But till now, the exact etiology and pathogenesis of OLP remains unclear. At present, it is recognized that many pathogenic factors, such as autoimmune response, mental stress, infection and hypersensitivity, might be involved in the development of OLP(9).The treatment of this disease is mainly with glucocorticoids, nystatin, surface anesthesia, antibiotic mouthwash, and so on, but the effect is not ideal. Meanwhile these local or systemic immunosuppressors can usually only alleviate clinical symptoms, and

have obvious side effects for long-term use(10). Therefore, elucidating the mechanisms underlying the pathogenesis and identifying more effective therapeutic strategies for OLP is essential.

Moreover, OLP has been regarded as one of the precancerous lesions of oral squamous cell carcinoma (OSCC), which is the most dominant type of oral cancer, accounting for more than 90% of them(11, 12). As a result, it is urgent to clarify the malignant transformation mechanism of OLP to OSCC, that might benefit to improve the ability of early diagnosis of OSCC and provide new and more effective treatment measures.

2. Discussion

OLP is a common chronic mucocutaneous inflammatory disease, and also regarded as a potentially malignant oral disorder by WHO, because 1.63% of lesions initially diagnosed as OLP evolved into oral squamous carcinomas. In this study, we totally screened 255 differentially expressed genes from OLP epithelium samples through the array dataset GSE52130, in which 187 up-regulated and 68 down-regulated genes. Next, the results of gene enrichment analysis suggested that the differentially expressed genes in the OLPE samples were mainly enriched in the focal adhesion pathway and metabolism pathway.

It has been known that metabolism become the focus of etiology of immune diseases(13, 14), which could coordinate the proliferation and differentiation of T cells. And OLP is T cell mediated inflammatory disorder, and metabolic pathway has been regarded as important parts of OLP pathophysiological mechanism. Furthermore, the metabolic changes of oral lichen planus cells are closely related to the development of disease and the malignant transformation. Studies have shown that succinate accumulates in OLP and OSCC at both tissue and cell levels, which activates the hypoxia-inducible factor-1 α (HIF-1 α) pathway and induces apoptosis, so succinate plays key roles in metabolic changes during malignant transformation from OLP to OSCC(15).

Other researchers found that the mTOR pathway was up-regulated in OLP patients, which played an important role in the immune metabolism of T cells. Targeted mTOR glycolysis pathway could significantly inhibit the proliferation of T cells and block its DNA synthesis, thereby induce cell apoptosis and regulate Th17 subsets differentiation^(16, 17). In addition, HIF1 α and phospholipase D2 (PLD2) are highly expressed in local T cells of OLP, and HIF1 α could upregulate the expression of PLD2 and promote T cellular immunity of OLP through glycolysis⁽¹⁸⁾. So, the metabolic changes of OLP are especially worthy of our further exploration, and the intervention to OLP metabolism might be a new therapeutic schedule.

On the other hand, the adhesion pathway is enriched in various diseases including cancer. As we known, the focal adhesion and ECM-receptor interaction pathways has been found to be involved in the development of OSCC(19). In addition, it had been reported that HPV-mediated cervical malignancy might disrupt the process of cell homeostasis because of local immunosuppression, then damage the focal

adhesion and decomposition of extracellular matrix, therefore promote the invasion, diffusion and metastasis of cancer cells(20). Besides, it has been found that integrin $\alpha 3$ could recruit the c-Src/extracellular signal-regulated protein kinase cascade, and induce the phosphorylation of focal adhesion kinase, thus enhance migration and invasion of cervical cancer cells, and promote angiogenesis through matrix metalloproteinase 9(21). Therefore, targeting the focal adhesion pathway is expected to become a new therapeutic strategy to slow down the development of the disease and improve the patient's condition.

Moreover, we have identified THBS2, COL3A1, IGFBP3, COL1A2, COL1A1, PTGS2, IL1B, ARG1 as hub genes in the protein-protein interaction (PPI) network analysis, which might be involved in the development of OLP.

Thrombospondin 2 (THBS2) is a disulfide-linked homotrimeric glycoprotein which belongs to the thrombospondin family, and mediates cell-to-cell and cell-to-matrix interactions. It had been reported that THBS2 was abnormally upregulated in colorectal cancer (CRC), and associated with poor prognosis. It could interact with Toll receptor 4, which upregulated mRNA expression of GLUT1, HK2, ALDOA, PKM2 and LDHA, and enhanced the glycolysis ability in CRC cells, therefore promotes cell proliferation(22). In addition, the gene silence of THBS2 could affect the proliferation, migration, invasion and apoptosis of gastric cancer cells through PI3K-Akt signaling pathway(23). Furthermore, researchers also found that, THBS2 might have a potential predictive effect on the response of immunotherapy in gastric cancer(24).

It has been well known that the molecules of the collagen family play important functions in the progression of many diseases. Collagen families, especially COL1A1, COL1A2, may be potential prognostic biomarkers of gastric cancer and its immune-related therapeutic targets(25). They were found to be positively correlated with macrophage infiltration and M2 macrophage marker expression, meanwhile, the expression of collagen was related to the sensitivity and drug resistance of gastric cancer cells to therapeutic drugs. In addition, COL3A1 can promote the metastasis of triple-negative breast cancer (TNBC) cells, while methyltransferase-like 3 (METTL3) can down-regulate the expression of COL3A1 by increasing the level of m6A, therefore, inhibiting the expression of METTL3 might induce poor prognosis caused by COL3A1 m6A modification in TNBC patients(26). Furthermore, it had been found that COL3A1 and COL1A1 had close relationship with immune infiltration in many malignant tumors, such as gastric cancer, esophageal cancer, and mesothelioma(27, 28).

IGFBP3, insulin-like growth factor binding protein 3, had been found significantly highly expressed in OSCC cell lines, which might directly interact with EGFR and the phosphorylated DNA-dependent protein kinase catalytic subunit in the irradiated cell nucleus and reduce radiosensitivity. In addition, enhanced expression of IGFBP-3 could be associated with the poor prognosis of OSCC(29). Researchers also found that T cell could accumulate within the tumor when absent of IGFBP-3, which indicated that IGFBP-3 might inhibit T cell infiltration. Moreover, IGFBP-3 could increase myeloid-derived suppressor cells ratio, hence inhibit T cell accumulation, which might make targeted IGFBP-3 immunotherapy possible(30).

PTGS2, also known as COX2, had been identified as a key node in the gene regulatory network related to immunophenotype of head and neck cancer(31). It is a key enzyme in prostaglandin biosynthesis and plays an important role in the immune mechanism of many diseases. It has been reported that PTGS2 is a potential intrinsic regulator of pancreatic ductal adenocarcinoma (PDA) cells, which promotes shaping the low immune tumor microenvironment and the resistance to immunotherapy, and the knockdown of PTGS2 in PDA tumor cells could increase tumor antigenicity and reduce tumor immunosuppression. Therefore, the pharmacological inhibition of PTGS2 could induce PDA sensitive to immunotherapy⁽³²⁾. These results showed that targeting immunosuppressive signals arising from the tumor cells may be a sensitization strategy for successful immunotherapy.

IL1 β is a member of the interleukin 1 cytokine family, an important mediator of inflammatory response, which is widely involved in a variety of cell activities, including cell proliferation, differentiation and apoptosis, and is also important in tumor development. It is said that IL1 β produced by tumor cells might promote epithelial-to-mesenchymal transition, invasion, migration and bone grafting(33). In addition, IL1R antagonists or IL1 β antibodies could reduce metastasis in humanized models of spontaneous breast cancer metastasis to bone. Furthermore, IL-1 β could mediate immunosuppression in tumor microenvironment, and accordingly provide a promising immunotherapy for pancreatic cancer(34).

Arginase catalyzes the hydrolysis of arginine to ornithine and urea, including two mammalian isoforms (Arg I and II) with different localization and physiological functions. Among them, Arg I is a cytosolic enzyme expressed primarily in the liver as a component of the urea cycle. On the other hand, tumor-associated macrophages (TAM) represent a group of heterogeneous cells with immunosuppressive and tumorigenic functions in the tumor microenvironment, which can modulate key cancer-associated activities and influence the response to different therapies. According to reports, Arg1 is an immunosuppressive signal to be found on the TAM in growing tumors, and its expression defines an immunosuppressive TAM subset(35), therefore attract increasing attention.

Moreover, we also identified that TBP, TCF12, EP300, RAD21, CREB3L1 might be the regulators of OLP, which would be worthy of further exploration in order to provide clues for improving clinical treatment effects of OLP.

3. Conclusion

Generally speaking, through bioinformatic analysis, we indicated that the focal adhesion pathway and metabolism pathways, as well as 8 crucial immune-related genes might be likely involved in the development of OLP, which would provide important information for elucidating the pathogenesis of OLP, meanwhile threw a new light on exploring more effective targeted therapy strategies for this disease.

4. Methods

4.1 Acquisition of sample information

Firstly, we got the clinical sample's information of OLP from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). Totally, 23 tissue samples were retrieved from GSE52130, we selected the top 14 ones, including 7 OLP epithelium samples (OLPE) and 7 healthy control oral epithelium (COE) samples. OLP patients were diagnosed clinically and histologically according to WHO diagnostic criteria. Normal mucosal tissues with no clinically visible inflammation were collected from retained wisdom teeth extraction operations and were used as healthy control samples.

To do further analysis, we downloaded all RNA information of these 14 selected samples. These data were based on the GPL6244 Affymetrix Human Gene 1. ST Array platform, transcript (gene) version (Affymetrix Inc, Santa Clara, CA, USA). As the information and data of these selected samples was all downloaded from the public database, there is unnecessary to obtain the patient consent or ethics committee approval the consent of the patient or the approval of the ethics committee.

4.2 Data further process

Next, we normalized and processed the original expression matrix by means of R. After that, we also selected out differentially expressed genes (DEGs) by the limma package(36). The P-value of these genes was calculated by t-test method and the adjusted P-value was calculated by Benjamini and Hochberg's method. The DEGs were screened out by the following selection criteria: at least a 1.5-fold change between COE samples and OLPE samples and with adjusted P-value < 0.05.

4.3 Enrichment analysis

It was well known that Gene set enrichment analysis (GSEA) could sequence the genes according to the differential expression degree of the two samples, then detect whether the preset gene set was enriched at the top or bottom of the sequencing table(37). Afterwards, all genetic information of OLPE and COE samples was uploaded to GSEA for further analysis. Next, we also analyze the differential expressed genes in OLPE through Database for annotation, visualization and integrated discovery (DAVID, <https://david.ncifcrf.gov/>) v6.8, which including the information of molecular function (MF), biological process (BP) and cell composition (CC)(38). Pathway analysis was conducted to find out which cell pathways might be involved in the changes in DEGs. Therefore, we could identify crucial pathway which might have relationship with these DEGs.

Moreover, we further performed the pathway analysis by means of Ingenuity Pathway Analysis (IPA) database, which was a graphical interface bioinformatics software based on cloud computing, and could analyze, integrate and understand the omics data from the perspective of biological pathway, and help to reveal the hidden biological significance of experimental data, and show the value of data obtained(39).

Therefore, we uploaded the 255 differentially expressed genes to IPA software to perform canonical pathway and molecule function analysis. The statistical cut-off of the above analysis was determined by Fisher's exact test P-value, which evaluated the correlations between meaningful molecules and known processes from random matches, and the z-score, which assessed the directional effects of one molecule on another or the effects of multiple molecular changes in the data set on biological processes, thus

assessed the match of observed and predicted regulatory patterns with a prediction for the activation state(39). We used the default parameters, both P-value < 0.05 and absolute value of z-score > 2 is considered significant.

4.4 Gene ontology (GO) and pathway enrichment analyses

It has been well recognized that Gene ontology (GO) includes three categories, respectively molecular function, biological processes, and cellular components, which is a tool to define, structure and control vocabulary for gene annotation(40). In addition, Kyoto Encyclopedia of Genes and Genomes (KEGG) is a knowledge base for systematic analysis of gene function in terms of the networks of genes and molecules, which helps researchers to link genomic information with higher order functional information(41).

On the other hand, Reactome is pathway database which provides intuitive bioinformatics tools for the visualization, interpretation and analysis of pathway knowledge. Meanwhile, it is a free open-source access database provided articles written by experts and peer reviewed on reactions and biological pathways in the human body.

In this part of the study, $P < 0.05$ was used as the threshold value, and the number of genes enriched in each pathway was ≥ 2 .

4.5 Gene cluster identification and protein-protein interaction (PPI) network analysis

The DEGs in OLPE samples were uploaded to STRING to obtain the protein network interaction diagram, in which a combination score of > 0.4 was set as a threshold value(42). Next, the result of STRING analysis was imported into Cytoscape v.3.7.2, and protein cluster analysis of differential expressed genes was conducted using Molecular Complex Detection (MCODE) plug-in(43, 44). The genes contained in the gene cluster with the highest scores were imported into the STRING to draw the protein interaction network and further analysis which biological processes this gene cluster was participated in.

Accordingly, we could regard the nodes with higher degrees of interaction as hub nodes.

4.6 Analysis of hub genes and transcription factors associated with OLP

As we all know, Hub genes are key molecules in the signaling pathway, and they are closely related to the development of diseases. Related research on this enables to better understand the disease and provide a solid theoretical basis for treatment. Comparative toxicogenomics database (CTD; <http://ctdbase.org/>), a powerful and open database, includes artificial selection of core information on chemical-gene, chemical-disease and gene-disease interactions, with which we can obtain toxicogenomic-related information(45). In this respect, we input 'oral lichen planus' as a key word into CTD database with the default parameters, and obtained a series of genes related to OLP, which was used to support the acquisition of our hub genes. On the other hand, CytoHubba is a Cytoscape plugin app, which has provided a simple and

convenient method to help us explore key nodes in biological networks, from which we finally selected EcCentricity method to probe the PPI network for hub genes.

Subsequently, the Cytoscape plugin iRegulon, which made use of cis-regulatory sequence analysis to reverse-engineer the transcriptional regulatory network underlying a co-expressed gene set, was used to analyze transcription factors regulating marker genes(46). It could integrate the transcription factor information from databases such as Transfac, Jaspar, Encode, Swissregulon and Homer, and detect enriched transcription factor motifs and optimal sets of their direct targets by means of genome-wide ranking and recovery. Parameter settings were as follows: Minimum identity between orthologous genes = 0.05 and maximum false discovery rate on motif similarity = 0.001. The Normalized Enrichment Score (NES) was the output result. The higher the scores were, the more reliable the results were. Transcription factors and target gene pairs with NES > 5 was selected.

5. Results

5.1 Sample information processing and screening of differentially expressed genes

On the basis of the sample information, 255 differentially expressed genes were picked out from the OLPE samples, in which 187 genes were up-regulated and 68 genes were down-regulated. The screening criteria for differentially expressed genes were as follows: adjust P-value < 0.05 and $|\log_2 \text{Fold Change}| \geq 1.5$. According to the analysis of these gene expressions, the heatmap and the volcano plot was made as shown in Fig. 1.

5.2 GO and KEGG enrichment analysis

Then, GO and KEGG enrichment analysis was conducted on 255 differentially expressed genes (DEGs) in the OLPE samples by using R. Genes ontology and biological functional analysis indicated that the 255 DEGs in OLPE samples were most related to skin development in the biological process, and extracellular matrix structure constituent in the molecular function, in addition, cornified envelope in the cellular component as shown in Fig. 2A-2C.

On the other hand, KEGG pathway analysis showed the 255 DEGs genes were significantly enriched in protein digestion and absorption, ECM-receptor interaction, focal adhesion, and metabolism pathways, such as drug metabolism-cytochrome P450, glycine, serine and threonine metabolism, tryptophan metabolism, and so on, as shown in Fig. 2D.

In addition, through analysis of the pathway enrichment, it indicated that of differentially expressed genes in OLPE samples were mainly enriched in IL-17 signaling pathway. It has been well known that OLP is an autoimmune disease, in which IL-17 mRNA and protein with higher expression, so IL-17 signaling pathway may have a critical role in OLP(47). And recent research found that Renin could significantly

upregulate the expression of IL-17 by promoting STAT4 phosphorylation in oral keratinocytes, which provided promising potential targeted therapies for OLP patients(48).

5.3 The differentially expressed genes in the OLPE samples were mainly enriched in Focal Adhesion Pathway and Metabolism Pathway

Gene set enrichment analysis (GSEA), DAVID software was used for enrichment analysis of the samples' genes. Firstly, all gene expression information in OLPE and COE samples was uploaded to GSEA software, and the hallmark gene set database was used to analyze genes at the overall level of expression profile. The significantly enriched gene sets were set at a default cut-off as $P\text{-value} < 0.05$ and $FDR < 0.25$. The enrichment analysis of gene sets indicated that the focal adhesion pathway, ECM-receptor interaction, primary immunodeficiency, autoimmune thyroid disease, viral myocarditis and cytokine receptor interaction were significantly enriched in OLPE samples as shown in Fig. 3. Accordingly, it was reasonably concluded that OLP might be relate to primary immunodeficiency diseases and autoimmune driven diseases, such as autoimmune thyroid disease. Moreover, the enrichment analysis also showed that there are some common pathways and pathophysiological mechanisms between OLP and immune diseases, which giving us a deeper understanding of OLP.

GO and pathway enrichment analysis was conducted on 255 differentially expressed genes in the OLPE samples through DAVID and Reactome online database. Next, we further performed GO, KEGG pathway analyses respectively for up-regulated and down-regulated DEGs above. The results showed that up-regulated DEGs were mainly enriched in these GO terms, such as keratinization, keratinocyte differentiation, peptide cross-linking, epidermis development and extracellular matrix organization, as shown in Table I. But down-regulated differential genes were not enriched at the threshold of $FDR < 0.05$.

In addition, it was showed that the KEGG and Reactome pathways of upregulated DEGs were significantly enriched in ECM-receptor interaction, protein digestion and absorption, amoebiasis and focal adhesion, but those pathways of downregulated DEGs were enriched in biological oxidations, drug metabolism-cytochrome P450, and metabolic pathways, as shown in Table II.

5.4 Further pathway analysis by using IPA

To further validate our results and identify crucial molecules that involved in the progress of the OLP, a total of 255 differentially expressed genes were uploaded to IPA for core analysis. The advantage of IPA software has updated its literature information every week, quarterly updated its function, at the same time, all data have been passed three rounds of manual quality control, in which only high-impact databases and magazines were selected. Therefore, the analysis results are much more accurate and persuasive. We used IPA software to further validate the results of DEGs. The advantage of IPA software is to update the literature information every week, quarterly update function, all data through three rounds

of manual quality control, only select high-impact databases and magazines, so the results are more persuasive.

A total of 20 canonical pathway analysis results showed that 15 pathways were activated, and 5 pathways were inhibited, as shown in Fig. 4A. Specifically, intrinsic prothrombin activation pathway had the highest activation scores ($-\log(\text{B-H p-value}) = 5.68$, $z\text{-score} = 2.828$). The remaining high activation pathways included neuroprotective role of THOP1 in Alzheimer's disease ($-\log(\text{B-H p-value}) = 2.9$, $z\text{-score} = 2.828$), p38 MAPK signaling ($-\log(\text{B-H p-value}) = 2.72$, $z\text{-score} = 2.121$), dendritic cell maturation ($-\log(\text{B-H p-value}) = 1.9$, $z\text{-score} = 2.646$), GP6 signaling pathway ($-\log(\text{B-H p-value}) = 1.63$, $z\text{-score} = 2.449$), cholecystokinin/gastrin-mediated signaling ($-\log(\text{B-H p-value}) = 1.62$, $z\text{-score} = 2.449$), acute phase response signaling ($-\log(\text{B-H p-value}) = 1.88$, $z\text{-score} = 2.236$), Toll-like receptor signaling ($-\log(\text{B-H p-value}) = 1.74$, $z\text{-score} = 2.236$), and phospholipases ($-\log(\text{B-H p-value}) = 1.42$, $z\text{-score} = 2.000$). On the contrary, the inhibitory pathway was in turn SPINK1 pancreatic cancer pathway ($-\log(\text{B-H p-value}) = 4.74$, $z\text{-score} = -2.828$), PPAR signaling ($-\log(\text{B-H p-value}) = 3.07$, $z\text{-score} = -2.121$), xenobiotic metabolism PXR signaling pathway ($-\log(\text{B-H p-value}) = 1.88$, $z\text{-score} = -2.121$), super pathway of melatonin degradation ($-\log(\text{B-H p-value}) = 1.47$, $z\text{-score} = -2.000$).

On the other hand, the heatmap could indicate the results of disease and function analysis, which showed there were 10 main functional modules about 255 differentially expressed genes in the OLPE samples, namely organismal injury and abnormalities, renal and urological disease, cancer, small molecule biochemistry, molecular transport, lipid metabolism, Cell-To-Cell signaling and interaction, hematological system development and function, cellular movement, hematological system development and function, immune cell trafficking. Among these main modules, the significantly activated relevant functions were release of lipid ($P\text{-value} = 1.2\text{E-}03$, $z\text{-score} = 2.93$), cancer ($P\text{-value} = 2.56\text{E-}05$, $z\text{-score} = 2.77$), release of fatty acid ($P\text{-value} = 1.08\text{E-}03$, $z\text{-score} = 2.764$), hematopoietic neoplasm ($P\text{-value} = 1.48\text{E-}03$, $z\text{-score} = 2.592$), release of eicosanoid ($P\text{-value} = 1.61\text{E-}03$, $z\text{-score} = 2.581$), growth of malignant tumor ($P\text{-value} = 1.98\text{E-}03$, $z\text{-score} = 2.436$), quantity of catecholamine ($P\text{-value} = 1.76\text{E-}03$, $z\text{-score} = 2.376$), release of prostaglandin E2 ($P\text{-value} = 1.86\text{E-}6$, $z\text{-score} = 2.361$), transport of molecule ($P\text{-value} = 7.78\text{E-}4$, $z\text{-score} = 2.265$), cancer of cells ($P\text{-value} = 1.28\text{E-}6$, $z\text{-score} = 2.228$), activation of phagocytes ($P\text{-value} = 8.66\text{E-}4$, $z\text{-score} = 2.211$), concentration of dopamine ($P\text{-value} = 1.81\text{E-}3$, $z\text{-score} = 2.204$), neoplasia of cells ($P\text{-value} = 3.32\text{E-}4$, $z\text{-score} = 2.190$), leukocyte migration ($P\text{-value} = 1.51\text{E-}3$, $z\text{-score} = 2.150$), cell movement of leukocytes ($P\text{-value} = 6.93\text{E-}4$, $z\text{-score} = 2.125$), malignant solid tumor ($P\text{-value} = 1.49\text{E-}5$, $z\text{-score} = 2.093$), cell movement of granulocytes ($P\text{-value} = 4.2\text{E-}3$, $z\text{-score} = 2.078$), neurotransmission ($P\text{-value} = 5.08\text{E-}3$, $z\text{-score} = 2.049$), activation of myeloid cells ($P\text{-value} = 1.19\text{E-}3$, $z\text{-score} = 2.003$), replication of hepatitis C virus ($P\text{-value} = 3.74\text{E-}3$, $z\text{-score} = -2.219$), as shown in Fig. 4B. From the point of view of the analysis, we could draw conclusions that OLP might have a close relationship with tumor, and those upregulated DEGs were mainly enriched in cell migration and movement pathways. In addition, the analysis results also indicated that OLP's occurrence might be related to hepatitis C virus infection.

5.5 Protein-protein interaction (PPI) network analysis

In order to filter out the hub genes from the differentially expressed genes in the OLPE sample, we next uploaded 255 differentially expressed genes to the STRING for further analysis and obtained 245 nodes and 586 edges. Local clustering coefficient was 0.411 and PPI enrichment P-value < 1.0e-16, then the TSV format file was downloaded and processed with Cytoscape as shown in Fig. 5A. MCODE was used to process the network data to identify gene clusters as shown in Fig. 5B-D, and genes in the first three gene cluster with the highest score ranking were selected for BP enrichment analysis. Next, we imported these cluster genes into Metascape database to analyze, and then found that the genes in the gene cluster 1 and cluster 2 were mainly involved in cornification and formation of the cornified envelope as shown in Table III, which was consistent with the pathological results and basic pathological process of oral lichen planus.

But actually, what interested us most was cluster 3, which enriched more related ways and might play more important roles in OLP. The 29 genes in gene cluster 3 mainly participated in post-translational protein phosphorylation, inflammatory response to antigenic stimulus, fever generation, blood vessel development, signaling by PDGF, response to nutrient levels, response to inorganic substance, negative regulation of cytokine-mediated signaling pathway and regulated exocytosis pathway, all of which were with high statistical significance according to P values.

5.6 Hub genes and transcription factors associated with OLP

By means of EcCentricity method, we identified THBS2, COL3A1, IGFBP3, COL1A2, COL1A1, PTGS2, IL1B, ARG1 as hub genes, as shown in Fig. 6A. Moreover, we further verified these results in CTD database, and concluded that these hub genes mainly enriched the focal adhesion and metabolism pathway. From table V, it could be clearly found that these hub genes were enriched in PI3K-Akt signaling pathway, focal adhesion, ECM-receptor interaction, cytokine-cytokine receptor interaction and metabolism.

Furthermore, we downloaded focal adhesion KEGG pathway diagram, as shown in Fig. 6B. Focal adhesion and PI3K-Akt signaling pathway, ECM-receptor interaction, cytokine-cytokine receptor interaction exits a crosslink. It was well known that the focal adhesion pathway plays the major role in the pathogenesis of OLP(49), which was also related to hematopoiesis, tumor metastasis, vascular diseases and malignant transformation(50–52). On the other hand, the metabolism disorder had close relationship with the development of OLP. It had been shown that the HIF1 α /PLD2 axis was associated with glycolysis and induces T cell immunity in oral lichen planus(18). As metabolic changes are significant during malignant transformation of primary OLP cells, it's important to focus on changes in metabolism(15).

In addition, the transcription regulatory network of these hub genes was shown in Fig. 7. Among them, the transcription factors with an NES score > 5 consisted of TBP (TATA-Box Binding Protein, NES = 5.274), TCF12 (Transcription Factor 12, NES = 5.091), EP300 (E1A Binding Protein P300, NES = 5.583), RAD21 (NES = 5.925), CREB3L1 (CAMP Responsive Element Binding Protein 3 Like 1, NES = 5.946).

Declarations

Ethics approval and consent to participate

Not necessary.

Consent for publication

Not applicable.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This work was supported by National Natural Science Foundation of China (No.81101642).

Authors' Contributions

Wenli Gu designed this work. Lou Geng and Xingming Zhang performed bioinformatics analyses and wrote the manuscript. Yi Tang revised the manuscript. All authors have read the final version of this manuscript.

Acknowledgements

The authors gratefully acknowledge the financial supported by the National Natural Science Foundation of China (Grant No. 81101642).

Availability of data and material

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.ncbi.nlm.nih.gov/geo/>.

References

1. Lodi G, Scully C, Carrozzo M, Griffiths M, Sugerman PB, Thongprasom K. Current controversies in oral lichen planus: report of an international consensus meeting. Part 1. Viral infections and

- etiopathogenesis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2005;1001:40–51. doi:10.1016/j.tripleo.2004.06.077.
2. Fitzpatrick SG, Hirsch SA, Gordon SC. The malignant transformation of oral lichen planus and oral lichenoid lesions: a systematic review. *J Am Dent Assoc.* 2014;1451:45–56. doi:10.14219/jada.2013.10.
 3. Warnakulasuriya S, Johnson NW, van der Waal I. Nomenclature and classification of potentially malignant disorders of the oral mucosa. *J Oral Pathol Med.* 2007;3610:575–80. doi:10.1111/j.1600-0714.2007.00582.x.
 4. van der Waal I. Oral lichen planus and oral lichenoid lesions; a critical appraisal with emphasis on the diagnostic aspects. *Med Oral Patol Oral Cir Bucal.* 2009;147:E310-4.
 5. Weber B, Schlapbach C, Stuck M, Simon HU, Borradori L, Beltraminelli H, et al. Distinct interferon-gamma and interleukin-9 expression in cutaneous and oral lichen planus. *J Eur Acad Dermatol Venereol.* 2017;315:880–6. doi:10.1111/jdv.13989.
 6. Sklavounou-Andrikopoulou A, Chrysomali E, Iakovou M, Garinis GA, Karameris A. Elevated serum levels of the apoptosis related molecules TNF-alpha, Fas/Apo-1 and Bcl-2 in oral lichen planus. *J Oral Pathol Med.* 2004;337:386–90. doi:10.1111/j.1600-0714.2004.00221.x.
 7. Wang F, Tan YQ, Zhang J, Zhou G. Insulin-like growth factor 1 exhibits the pro-autophagic and anti-apoptotic activity on T cells of oral lichen planus. *Int J Biol Macromol.* 2019;133:640–6. doi:10.1016/j.ijbiomac.2019.04.158.
 8. Wang H, Zhang D, Han Q, Zhao X, Zeng X, Xu Y, et al. Role of distinct CD4(+) T helper subset in pathogenesis of oral lichen planus. *J Oral Pathol Med.* 2016;456:385–93. doi:10.1111/jop.12405.
 9. Lodi G, Scully C, Carrozzo M, Griffiths M, Sugerman PB, Thongprasom K. Current controversies in oral lichen planus: report of an international consensus meeting. Part 2. Clinical management and malignant transformation. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2005;1002:164–78. doi:10.1016/j.tripleo.2004.06.076.
 10. Yang H, Wu Y, Ma H, Jiang L, Zeng X, Dan H, et al. Possible alternative therapies for oral lichen planus cases refractory to steroid therapies. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2016;1215:496–509. doi:10.1016/j.oooo.2016.02.002.
 11. Choi S, Myers JN. Molecular pathogenesis of oral squamous cell carcinoma: implications for therapy. *J Dent Res.* 2008;871:14–32. doi:10.1177/154405910808700104.
 12. Scully C, Bagan J. Oral squamous cell carcinoma: overview of current understanding of aetiopathogenesis and clinical implications. *Oral Dis.* 2009;156:388–99. doi:10.1111/j.1601-0825.2009.01563.x.
 13. Almeida L, Lochner M, Berod L, Sparwasser T. Metabolic pathways in T cell activation and lineage differentiation. *Semin Immunol.* 2016;285:514–24. doi:10.1016/j.smim.2016.10.009.
 14. Patel CH, Powell JD. Targeting T cell metabolism to regulate T cell activation, differentiation and function in disease. *Curr Opin Immunol.* 2017;46:82–8. doi:10.1016/j.coi.2017.04.006.

15. Yang Q, Sun H, Wang X, Yu X, Zhang J, Guo B, et al. Metabolic changes during malignant transformation in primary cells of oral lichen planus: Succinate accumulation and tumour suppression. *J Cell Mol Med*. 2020;242:1179–88. doi:10.1111/jcmm.14376.
16. Zhang N, Zhang J, Tan YQ, Du GF, Lu R, Zhou G. Activated Akt/mTOR-autophagy in local T cells of oral lichen planus. *Int Immunopharmacol*. 2017;48:84–90. doi:10.1016/j.intimp.2017.04.016.
17. Wang F, Zhang J, Zhou G. The mTOR-glycolytic pathway promotes T-cell immunobiology in oral lichen planus. *Immunobiology*. 2020;2253:151933. doi:10.1016/j.imbio.2020.151933.
18. Wang F, Zhang J, Zhou G. HIF1alpha/PLD2 axis linked to glycolysis induces T-cell immunity in oral lichen planus. *Biochim Biophys Acta Gen Subj*. 2020;18647:129602. doi:10.1016/j.bbagen.2020.129602.
19. Yang S, Wang Y, Ren J, Zhou X, Cai K, Guo L, et al. Identification of diagnostic and prognostic lncRNA biomarkers in oral squamous carcinoma by integrated analysis and machine learning. *Cancer Biomark*. 2020. doi:10.3233/CBM-191215.
20. Xuan WX, Li JJ, Zhang QC, Sun GN, Xu ZW, Sun ZF, et al. Protein expression shift and potential diagnostic markers thorough proteomics profiling of tuberculous pleurisy biopsy tissues. *Int J Infect Dis*. 2020. doi:10.1016/j.ijid.2020.07.051.
21. Du Q, Wang W, Liu T, Shang C, Huang J, Liao Y, et al. High Expression of Integrin alpha3 Predicts Poor Prognosis and Promotes Tumor Metastasis and Angiogenesis by Activating the c-Src/Extracellular Signal-Regulated Protein Kinase/Focal Adhesion Kinase Signaling Pathway in Cervical Cancer. *Front Oncol*. 2020;10:36. doi:10.3389/fonc.2020.00036.
22. Xu C, Gu L, Kuerbanjiang M, Wen S, Xu Q, Xue H. Thrombospondin 2/Toll-Like Receptor 4 Axis Contributes to HIF-1alpha-Derived Glycolysis in Colorectal Cancer. *Front Oncol*. 2020;10:557730. doi:10.3389/fonc.2020.557730.
23. Ao R, Guan L, Wang Y, Wang JN. Silencing of COL1A2, COL6A3, and THBS2 inhibits gastric cancer cell proliferation, migration, and invasion while promoting apoptosis through the PI3k-Akt signaling pathway. *J Cell Biochem*. 2018;1196:4420–34. doi:10.1002/jcb.26524.
24. Wang Y, Zheng K, Chen X, Chen R, Zou Y. (2021). Bioinformatics analysis identifies COL1A1, THBS2 and SPP1 as potential predictors of patient prognosis and immunotherapy response in gastric cancer. *Biosci Rep*. 411 doi:10.1042/BSR20202564.
25. Chen Y, Chen W, Dai X, Zhang C, Zhang Q, Lu J. Identification of the collagen family as prognostic biomarkers and immune-associated targets in gastric cancer. *Int Immunopharmacol*. 2020;87:106798. doi:10.1016/j.intimp.2020.106798.
26. Shi Y, Zheng C, Jin Y, Bao B, Wang D, Hou K, et al. Reduced Expression of METTL3 Promotes Metastasis of Triple-Negative Breast Cancer by m6A Methylation-Mediated COL3A1 Up-Regulation. *Front Oncol*. 2020;10:1126. doi:10.3389/fonc.2020.01126.
27. Zhang T, Wang BF, Wang XY, Xiang L, Zheng P, Li HY, et al. Key Genes Associated with Prognosis and Tumor Infiltrating Immune Cells in Gastric Cancer Patients Identified by Cross-Database Analysis. *Cancer Biother Radiopharm*. 2020;359:696–710. doi:10.1089/cbr.2019.3423.

28. Zhang C, Liu S, Wang X, Liu H, Zhou X, Liu H. (2021). COL1A1 Is a Potential Prognostic Biomarker and Correlated with Immune Infiltration in Mesothelioma. *Biomed Res Int.* 2021, 5320941. doi:10.1155/2021/5320941.
29. Sakata J, Hirose A, Yoshida R, Matsuoka Y, Kawahara K, Arita H, et al. (2020). Enhanced Expression of IGFBP-3 Reduces Radiosensitivity and Is Associated with Poor Prognosis in Oral Squamous Cell Carcinoma. *Cancers (Basel).* 122 doi:10.3390/cancers12020494.
30. Scully T, Scott CD, Firth SM, Sedger LM, Pintar JE, Twigg SM, et al. Enhancement of mammary tumour growth by IGFBP-3 involves impaired T cell accumulation. *Endocr Relat Cancer.* 2018;252:111–22. doi:10.1530/ERC-17-0384.
31. Feng B, Shen Y, Pastor Hostench X, Bieg M, Plath M, Ishaque N, et al. Integrative Analysis of Multi-omics Data Identified EGFR and PTGS2 as Key Nodes in a Gene Regulatory Network Related to Immune Phenotypes in Head and Neck Cancer. *Clin Cancer Res.* 2020;2614:3616–28. doi:10.1158/1078-0432.CCR-19-3997.
32. Markosyan N, Li J, Sun YH, Richman LP, Lin JH, Yan F, et al. Tumor cell-intrinsic EPHA2 suppresses anti-tumor immunity by regulating PTGS2 (COX-2). *J Clin Invest.* 2019;1299:3594–609. doi:10.1172/JCI127755.
33. Tulotta C, Lefley DV, Freeman K, Gregory WM, Hanby AM, Heath PR, et al. Endogenous Production of IL1B by Breast Cancer Cells Drives Metastasis and Colonization of the Bone Microenvironment. *Clin Cancer Res.* 2019;259:2769–82. doi:10.1158/1078-0432.CCR-18-2202.
34. Winograd R, Simeone DM, Bar-Sagi D. A novel target for combination immunotherapy in pancreatic cancer: IL-1beta mediates immunosuppression in the tumour microenvironment. *Br J Cancer.* 2021. doi:10.1038/s41416-021-01303-2.
35. Arlauckas SP, Garren SB, Garriss CS, Kohler RH, Oh J, Pittet MJ, et al. Arg1 expression defines immunosuppressive subsets of tumor-associated macrophages. *Theranostics.* 2018;821:5842–54. doi:10.7150/thno.26888.
36. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;437:e47. doi:10.1093/nar/gkv007.
37. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;10243:15545–50. doi:10.1073/pnas.0506580102.
38. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;41:44–57. doi:10.1038/nprot.2008.211.
39. Kramer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics.* 2014;304:523–30. doi:10.1093/bioinformatics/btt703.
40. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;251:25–9. doi:10.1038/75556.

41. Wixon J, Kell D. The Kyoto encyclopedia of genes and genomes–KEGG. Yeast. 2000;171(200004):48–55. doi:10.1002/(SICI)1097-0061. 8 :AID-YEA2>3.0.CO;2-H (., 17:1<4.
42. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45D1:D362–8. doi:10.1093/nar/gkw937.
43. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;1311:2498–504. doi:10.1101/gr.1239303.
44. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4:2. doi:10.1186/1471-2105-4-2.
45. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, King BL, McMorran R, et al. The Comparative Toxicogenomics Database: update 2017. Nucleic Acids Res. 2017;45D1:D972–8. doi:10.1093/nar/gkw838.
46. Janky R, Verfaillie A, Imrichova H, Van de Sande B, Standaert L, Christiaens V, et al. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput Biol. 2014;107:e1003731. doi:10.1371/journal.pcbi.1003731.
47. Lu R, Zeng X, Han Q, Lin M, Long L, Dan H, et al. (2014). Overexpression and selectively regulatory roles of IL-23/IL-17 axis in the lesions of oral lichen planus. *Mediators Inflamm*. 2014, 701094. doi:10.1155/2014/701094.
48. Ge X, Xie H, Nguyen T, Zhao B, Xu J, Du J. (2020). Renin Promotes STAT4 Phosphorylation to Induce IL-17 Production in Keratinocytes of Oral Lichen Planus. *iScience*. 234, 100983. doi:10.1016/j.isci.2020.100983.
49. Becker J, Schuppan D. Altered expression of extracellular matrix proteins and integrins in oral lichen planus (OLP). J Oral Pathol Med. 1995;244:159–64. doi:10.1111/j.1600-0714.1995.tb01158.x.
50. Khadilkar RJ, Ho KYL, Venkatesh B, Tanentzapf G. Integrins Modulate Extracellular Matrix Organization to Control Cell Signaling during Hematopoiesis. Curr Biol. 2020. doi:10.1016/j.cub.2020.06.027.
51. Murphy JM, Rodriguez YAR, Jeong K, Ahn EE, Lim SS. Targeting focal adhesion kinase in cancer cells and the tumor microenvironment. Exp Mol Med. 2020;526:877–86. doi:10.1038/s12276-020-0447-4.
52. Hwang JS, Eun SY, Ham SA, Yoo T, Lee WJ, Paek KS, et al. PPARdelta modulates oxLDL-induced apoptosis of vascular smooth muscle cells through a TGF-beta/FAK signaling axis. Int J Biochem Cell Biol. 2015;62:54–61. doi:10.1016/j.biocel.2015.02.014.

Tables

Table I

GO analysis of differentially expressed genes (FDR<0.05)

GO ID	Terms	Number	FDR
Upregulated genes			
GO_BP terms			
GO:0031424	keratinization	18	1.04E-19
GO:0030216	keratinocyte differentiation	20	5.53E-19
GO:0018149	peptide cross-linking	17	1.36E-17
GO:0008544	epidermis development	15	1.03E-10
GO:0030198	extracellular matrix organization	14	8.24E-05
GO_CC terms			
GO:0001533	cornified envelope	18	2.53E-20
GO:0005576	extracellular region	45	6.69E-08
GO:0005615	extracellular space	39	7.85E-07
GO:0005578	proteinaceous extracellular matrix	13	0.010895389
GO:0070062	extracellular exosome	49	0.020247725
GO:0031012	extracellular matrix	13	0.029446628
GO_MF terms			
GO:0005198	structural molecule activity	21	5.18E-10
GO:0005201	extracellular matrix structural constituent	10	2.26E-05
GO:0005149	interleukin-1 receptor binding	5	0.007719439
BP, biological process; CC, cellular component; MF, molecular function; GO, gene ontology.			

Table II

KEGG and Reactome pathway enrichment analyses for differentially expressed genes (P<0.05)

ID	Terms	Count	P-Value
up-regulated			
hsa04512	ECM-receptor interaction	9	3.06E-07
hsa04974	Protein digestion and absorption	8	4.79E-06
hsa05146	Amoebiasis	7	1.65E-04
hsa04510	Focal adhesion	9	1.80E-04
hsa04611	Platelet activation	6	0.00336272
hsa04151	PI3K-Akt signaling pathway	9	0.005079306
hsa04145	Phagosome	6	0.006175871
hsa05144	Malaria	4	0.006529778
hsa04726	Serotonergic synapse	5	0.0109323
hsa00590	Arachidonic acid metabolism	4	0.011940849
hsa05323	Rheumatoid arthritis	4	0.031412058
hsa04912	Rheumatoid arthritis	4	0.034212174
hsa04750	Inflammatory mediator regulation of TRP channels	4	0.041234535
hsa04940	Type I diabetes mellitus	3	0.042639686
hsa02010	ABC transporters	3	0.046388375
down-regulated			
R-HSA-211859	Biological oxidations	11	1.950E-7
R-HSA-5579029	Metabolic disorders of biological oxidation enzymes	3	1.380E-4
hsa00982	Drug metabolism-cytochrome P450	4	0.002280447
hsa00340	Histidine metabolism	3	0.003265313
hsa00260	Glycine, serine and threonine metabolism	3	0.010053474
hsa00380	Tryptophan metabolism	3	0.010557146
hsa00330	Arginine and proline metabolism	3	0.016186014
hsa01100	Metabolic pathways	10	0.037458967

Table III

Top3 gene cluster analysis by Metascape database

cluster 1	Terms	Count	P-Value
GO:0070268	cornification	9	-21.14
cluster 2			
R-HSA-6809371	Formation of the cornified envelope	18	-40.17
GO:0070268	cornification	9	-16.2
cluster 3			
R-HSA-8957275	Post-translational protein phosphorylation	6	-8.54
GO:0002437	inflammatory response to antigenic stimulus	5	-7.92
GO:0001660	fever generation	3	-6.74
GO:0001568	blood vessel development	9	-6.74
R-HSA-186797	Signaling by PDGF	4	-6.18
GO:0031667	response to nutrient levels	7	-5.8
GO:0010035	response to inorganic substance	6	-4.34
GO:0001960	negative regulation of cytokine-mediated signaling pathway	3	-4.05
GO:0045055	regulated exocytosis	6	-3.53

Table VI

Hub genes and their KEGG pathway

ID	Gene Name	KEGG_PATHWAY
ARG1	arginase 1	Arginine biosynthesis; Arginine and proline metabolism; Metabolic pathways; Biosynthesis of antibiotics; Biosynthesis of amino acids; Amoebiasis
COL1A1	collagen type I alpha 1 chain	PI3K-Akt signaling pathway; Focal adhesion; ECM-receptor interaction; Platelet activation; Protein digestion and absorption; Amoebiasis
COL1A2	collagen type I alpha 2 chain	PI3K-Akt signaling pathway; Focal adhesion; ECM-receptor interaction; Platelet activation; Protein digestion and absorption; Amoebiasis
COL3A1	collagen type III alpha 1 chain	PI3K-Akt signaling pathway; Focal adhesion; ECM-receptor interaction; Platelet activation; Protein digestion and absorption; Amoebiasis
IGFBP3	insulin like growth factor binding protein 3	p53 signaling pathway; transcriptional misregulation in cancer
IL1B	interleukin 1 beta	MAPK signaling pathway; Cytokine-cytokine receptor interaction; NF-kappa B signaling pathway; Osteoclast differentiation; Toll-like receptor signaling pathway; NOD-like receptor signaling pathway; Cytosolic DNA-sensing pathway; Hematopoietic cell lineage; TNF signaling pathway
PTGS2	prostaglandin-endoperoxide synthase 2	Arachidonic acid metabolism; Metabolic pathways; NF-kappa B signaling pathway; VEGF signaling pathway; TNF signaling pathway; Retrograde endocannabinoid signaling; Serotonergic synapse; Ovarian steroidogenesis; Oxytocin signaling pathway; Regulation of lipolysis in adipocytes; Leishmaniasis; Pathways in cancer; Chemical carcinogenesis; MicroRNAs in cancer; Small cell lung cancer
THBS2	thrombospondin 2	Phagosome; PI3K-Akt signaling pathway; Focal adhesion; ECM-receptor interaction

Figures

Figure 1

The differentially expressed genes between OLPE and COE were shown on the heatmap and the volcano plot. A. Heat map of differentially expressed genes. Blue color represented a lower and red color represented a higher expression level, while grey color demonstrated no differential expression. B. The volcano plot of differentially expressed genes. Blue dots represented significantly down-regulated genes, while red dots represented significantly up-regulated genes. COE, control oral epithelium; OLPE, oral lichen planus epithelium samples.

Figure 2

GO and KEGG enrichment analysis was conducted on 255 DEGs in the OLPE samples. A. Biology process enrichment of the DEGs. B. Molecular function enrichment of the DEGs. C. Cellular component enrichment of the DEGs. D. KEGG pathway enrichment of the DEGs.

Figure 3

h.all. v 6.2.symbols.gmt [Hallmarks] gene set database was used to analyze the whole gene expression value of the OLPE and COE samples. GSEA first filtered the gene set according to the number of genes contained in the gene set, with the minimum number of 15 genes and the maximum number of 500 genes by default. Significant gene sets were cut-off by $FDR < 0.25$ and $P\text{-value} < 0.05$.

Figure 4

Further pathway analysis by using IPA was conducted on 255 DEGs in the OLPE samples. A. The canonical pathway analysis of IPA. The colors depth in the bar chart represented the z-score, and generally, an absolute z-score greater than 2 was considered meaningful. B. Disease and function analysis of IPA. Orange meant z-score>0, blue meant z-score<0, and grey meant no z-score; z-score>2 meant the function was significantly activated, and z-score<-2 means the function was significantly inhibited.

Figure 5

Protein-protein interaction network analysis was processed with Cytoscape and different clusters was analyzed by MCODE. A. The STRING database file was uploaded into Cytoscape v.3.7.2 to draw Protein-protein interaction network, and different clusters analyzed by MCODE were noted with different colors. The significance of P-Value was shown by the size of node. The color of the nodes showed the change of the DEGs, and the smaller the P-Value is, the larger the diameter of node is. The color of the edge represented the value of combined score from 0.4 to 1, grey to dark. B. The Cluster 1 had 19 nodes and 171 edges, and the cluster score is 19.00. C. The Cluster 2 network had 9 nodes and 36 edges with 9.00 cluster score. D. The Cluster 3 network had 29 nodes and 89 edges with 6.357 cluster score, in which the up-regulated nodes were colored in red, but the down-regulated ones were colored in blue.

A

B

Figure 6

Hub genes associated with OLP was showed on the hub genes network and the focal adhesion KEGG pathway diagram. A. Top 8 genes were identified as hub genes by means of EcCentricity method, all of which were up-regulated. B. The focal adhesion pathway diagram was downloaded from KEGG website.

Figure 7

The transcription factors of the hub genes with an NES score > 5 was predicted by iRegulon and visualized by the Cytoscape. These transcription factors included TBP(A), TCF12(B), RAD21(C), EP300(D), CREB3L1(E). Among them, enrichment score threshold=3.0, ROC threshold for AUC calculation=0.03, rank threshold=5000, Maximum false discovery rate (FDR) on motif similarity ≤ 0.001 .