

Bioinformatic Analysis of Plus Gene Expression Related to Progression from Leukoplakia to Oral Squamous Cell Carcinoma

Jaime Guzman de Avila¹, Carlos Silvera Redondo², Antistio Alviz-Amador^{3*}

Abstract

Introduction: Leukoplakia is one of the most frequently found lesions in the oral cavity, with a probability of 17 to 24% of becoming malignant cells in a period of 30 years. Objective: To identify differentially expressed gene profiles of leukoplakia and its progression to oral squamous cell carcinoma, essential for the discovery of new biomarkers to predict and prevent the presence of diseases in the oral cavity. **Methods:** Initially, gene profiles of GSE85514 and GSE160042 from the Gene Expression Omnibus database were used. Differentially expressed genes were identified using GEO2R. The CLUEGO plugin in Cytoscape was used for DEG functionality and enrichment analysis. Finally, a protein-protein interaction (PPI) network was constructed using Cytoscape from data collected online from the STRING server. **Results:** According to the MCC algorithm, the 10 most found gene sequences were HNRNPU, SMC1A, PAFAH1B1, EHMT1, SPTBN4, OLFM1, NCAM1, SF3B3, FGF2, and UBE2I; with HNRNPU, SMC1A, and PAFAH1B1 being the most representative of the modules. **Conclusions:** We were able to describe the gene sequences that promote the progression from leukoplakia to oral squamous cell carcinoma. Within these genes, the HNRNPU, SMC1A, and PAFAH1B1 constitute the main promising therapeutic targets to counteract the progression of oral cancer, they could also be important biomarkers for the diagnosis and classification of the disease.

Keywords: Leukoplakia- squamous cell carcinoma- GEO2R- HNRNPU- SMC1A- PAFAH1B1.

Asian Pac J Cancer Prev, 23 (11), 3833-3842

Introduction

Nowadays, oral health is a topic of crucial importance, because the quality of life and general health of people has an impact on their overall care. Through the oral cavity, human beings carry out some necessary actions for life, such as swallowing food, gaseous exchange to carry out breathing, and emission of sounds for the articulation of language. The world health organization WHO defines oral health as the absence of oral or facial pain, oral infections or injuries, gum disease, cavities, tooth loss, and other pathologies such as disorders that limit the ability of normal functioning (World Health Organization, 2021).

The lesions of the oral cavity most studied by stomatology include benign, premalignant, and malignant lesions (González et al., 2006). The benign ones consist of processes of inflammation, irritation, and oral manifestations of systemic diseases. The premalignant, develop pre-neoplastic lesions (leukoplakia, lichen planus, dyskeratosis congenita, palatal lesions associated with inverted smoking, erythroplasia, among others that include an immunological component) which present a high risk of

progressing to malignant lesions such as cancer (Guerrero Brito et al., 2020). Finally, malignant lesions, often associated with oral cancer such as oral squamous cell carcinoma, usually follow a condition called potentially malignant disorder, which regularly progresses to cancer (Macey et al., 2015).

Leukoplakia is one of the most frequent lesions in the oral cavity, with a 17% to 24% probability of becoming malignant cells in a 30-year period (Sol and Meir, 1984; Warnakulasuriya et al., 2007). These oral malformations are associated with hyperkeratosis, dysplasia, or carcinomas, which have multiple etiological factors, including genetics, with alcoholism and smoking being the most relevant.

Head and neck squamous cell carcinomas are one of the most destructive diseases due to their rapid progression and invasion through the different tissues of the body, using lymphatic and blood dissemination routes. This is characterized by a considerably aggressive biological behavior with a high rate of metastasis and death if not detected in time (Sasahira and Kirita, 2018). Currently, oral cancer is one of the most frequent malignant tumors in

¹Department of Oral Medicine, School of Dentistry, University of Cartagena Cartagena, Colombia. ²Department of Genetics, Universidad del Norte, Barranquilla, Colombia. ³Pharmacology and Therapeutic Group, Faculty of Pharmaceutical Sciences, University of Cartagena, Health Science Campus, Colombia. *For Correspondence: aalviza@unicartagena.edu.co

the world. According to the world cancer statistics carried out by the international agency for research on cancer, there have been 377,773 new cases and 177,757 deaths by 2020, a number that is worrying due to the increase in both adult and young patients. Reports indicate that the incidence of oral cancer will continue to increase by 41.7% by the year 2040 (an estimated 535,157 new cases), which predicts an increase in the mortality rate corresponding to 47% (i.e., 261,000 mortality). In Latin America and the Caribbean, rates of 17888 new cases and a mortality rate of 7,548 were reported. In the particular case of Colombia, the incidence rates of new cases in recent years were 914 and the mortality rate was 1.4 per 100,000 inhabitants in Colombia (Bray et al., 2018).

Regarding the diagnosis of oral cancer, there are different screening methods. Due to the accessibility of the oral cavity, clinical examination (visual inspection) is currently the most frequently used method to detect visible lesions. On the other hand, for clinically undetectable lesions, it is necessary to use an adjuvant method so that the visual examination detects premalignant lesions earlier. The most frequently used adjuvant methods are toluidine blue, brush biopsy, and fluorescence imaging. However, there is insufficient solid evidence to determine whether the use of these techniques would increase the detection of oral malignancies in screening programs.

Likewise, as adjuvants of the visual examination, there are some tests in saliva to find different cellular and tissue markers from a molecular perspective, such as the genomic analysis using Microarray which provides additional information to that obtained in the clinical examination and the histopathological study. The family of markers that can be found through genomic analysis may include those of tumor growth, tumor suppressors, and antitumor response, as well as markers for angiogenesis, tumor invasion and metastatic potential, cell surface, intracellular, derivatives of arachidonic acid, and enzymes.

Biomarkers for the early detection of oral cancer are of great importance because molecular responses to the pathogenesis of the disease can lead to early screening.

Genetic modifications of gene expression can lead to the production of malignant tumors such as oral squamous cell carcinoma. The evolution of gene sequencing and high-throughput DNA Microarray analysis has proven to be a predictive model of numerous genetic alterations from differential gene expression profiles, which are related to the genesis and progression of tumors and malignant cells.

Taking all of the above into account, it should be noted that computational methods such as the search for hub genes in databases have become very useful tools for the identification of predictable biomarkers of many diseases, useful for diagnosis, prediction, and/or behavior analysis of some genes that are related to the progression of leukoplakia to oral squamous cell carcinoma, that is, the carcinogenesis process.

For this reason, the aim of this work was the identification of differential gene expression profiles for oral lesions in the progression of oral squamous cell carcinoma and the elucidation of the interactions between them and the signaling pathways involved that are essential for the detection of new biomarkers that

allow us to predict and prevent the presence of diseases in the oral cavity.

Materials and Methods

Data collection

The data sets were obtained from the GEO database (Barrett et al., 2013) (<https://www.ncbi.nlm.nih.gov/gds>), using the following keywords: “Oral Squamous Cell Carcinoma (OSCC)”, Homo sapiens (Organism), Expression profiling by MPSS (Study Type), tissue (Attribute name). Once the systematic search of the gene expression profiles was carried out, two data sets, GSE85195 (Bhosale et al., 2017) and GSE160042 (Zhuang et al., 2021), were collected for analysis. Differentially expressed RNAs between patients with leukoplakia and oral squamous cell carcinoma were identified by GEO2R from array GSE85195, while differentially expressed long non-coding RNAs between control patients with tongue tissue and oral squamous cell carcinoma of the tongue were identified by GEO2R with the data set GSE160042 by microarray.

Analysis of RNA and differentially expressed non-coding RNAs in GEO2R, were selected with a criterion of an adjusted P-value of <0.05 and a logarithm change of (CF) ≥ 2 .

Dataset selection criteria

- That the studies generated by the datasets used the microarray and RNA-seq hybridization methodology for the identification of genes related to the progression of leukoplakia to squamous cell carcinoma of the oral cavity.
- That the data sets were of human beings.
- Considering that in the latest studies the most accepted etiologies of oral cavity carcinoma are lncRNA and miRNA alterations, it was stipulated that the datasets should have reported both.
- That the records of the data sets at the time of the bioinformatic analysis were from the last 5 years.

Gene ontology functional enrichment analysis

The original gene codes and open reading frames (ORFs) were filtered and analyzed by GEO2R. Gene ontology (GO) data including biological pathways, molecular functions, and cellular components, were analyzed by the functional enrichment tool of the FunRich software (Pathan et al., 2015).

Analysis of protein-protein interaction networks

Differentially expressed gene profiles were evaluated from the protein-protein interaction network using the STRING database (von Mering et al., 2003), which used the selection criteria of $FC > 2$ and $P < 0.05$. Analysis and visualization of the protein-protein interaction network were performed with the Cytoscape software (version 3.9.0) (Paul Shannon et al., 2003), which allowed us to integrate biomolecular interaction networks with differential expression gene data with a unified reading frame. The main genes derived from the nodes with the highest score were obtained from module detection with the MCODE algorithm with selection criteria of

Cut-off = 0.2; K-core = 2; max-Depth = 100.

Identification of Hub genes

For the identification of the Hub genes of the protein-protein interaction network in Cytoscape, the complementary tool CytoHubba (Chin et al., 2014) was used, which employs an algorithm called Maximal Clique Centrality (MCC). This method is used to efficiently calculate the central nodes that identify the main genes involved in the interaction network.

Results

Identification of differentially expressed genes

According to the analysis performed by GEO2R, a total of 14347 differentially expressed genes were identified between patients with oral disorders (leukoplakia) and oral tumoral tissue in the first data set (GSE85195), of which 7,662 genes were upregulated and 6685 were downregulated. Similarly, in the second dataset (GSE160042) 30,606 differentially expressed open reading frame sequences were identified between control patients and patients with squamous cell carcinoma of the tongue, including 13,763 upregulated and 16843 downregulated. The volcano plot shows the differentially expressed genes and sequences (see Figure 1).

A total of 10,932 differentially expressed genes were filtered for the GSE85195 data set, which presented 5,680 up-regulated and 4,819 down-regulated genes, sharing 433 genes with a p value < 0.05 and CF > 2. For the second data set GSE160042, 21,679 gene sequences were filtered, including 8815 up-regulated and 11,052 down-regulated genes, with 1,815 genes shared. According to the above, it was possible to detect 65 co-expressed gene sequences for both groups of data, which can relate the progression of leukoplakia with oral tumoral tissue and squamous cell carcinoma of the tongue, as seen in Figure 2.

Biological processes

The functional enrichment analysis of the 65 differentially co-expressed genes by the FunRich software is represented in Figure 3 of panel A, where it was shown that the genes are mainly enriched, in 31.7%, by unknown biological processes, while the metabolic regulation of nucleobases, nucleosides, nucleotides, and nucleic acids represents 25.4%; Energy pathways and metabolism are also involved with 11.1%, as well as signal transduction and cellular communication with a value of 12.7%.

Cellular components and molecular functions

On the other hand, in Figure 3, panel B shows the cellular components of co-expressed genes with 54.7% for the nucleus and 43.4% for the cytoplasm. While panel C shows the molecular functions for differentially co-expressed genes in the progression of leukoplakia with oral tumoral tissue and squamous cell carcinoma of the tongue, finding 30.2% for unknown molecular function, 9.5% for transcriptional regulation activity, 6.3% for DNA binding, and 4.8% for transcription factor activity and ubiquitin-specific protease activity.

GO enrichment pathways

The biological pathways for differentially co-expressed gene sequences in the progression of leukoplakia with oral tumoral tissue and squamous cell carcinoma of the tongue showed that 22.7% of genes would be expressed in signaling events mediated by proteoglycans, while 18.2% would do so in any of the following processes: mRNA transcript formation and maturation, gene expression, signal transduction, AKT-mediated PI3K intracellular signaling events, mTOR signaling pathway, and VEGFR and VEGF receptor signaling pathways. It should be noted that the pathways related to splicing and mRNA processing reached 13.6% each, as can be seen in Figure 4.

Clinical phenotype

Figure 5, shows the main clinical effects that were

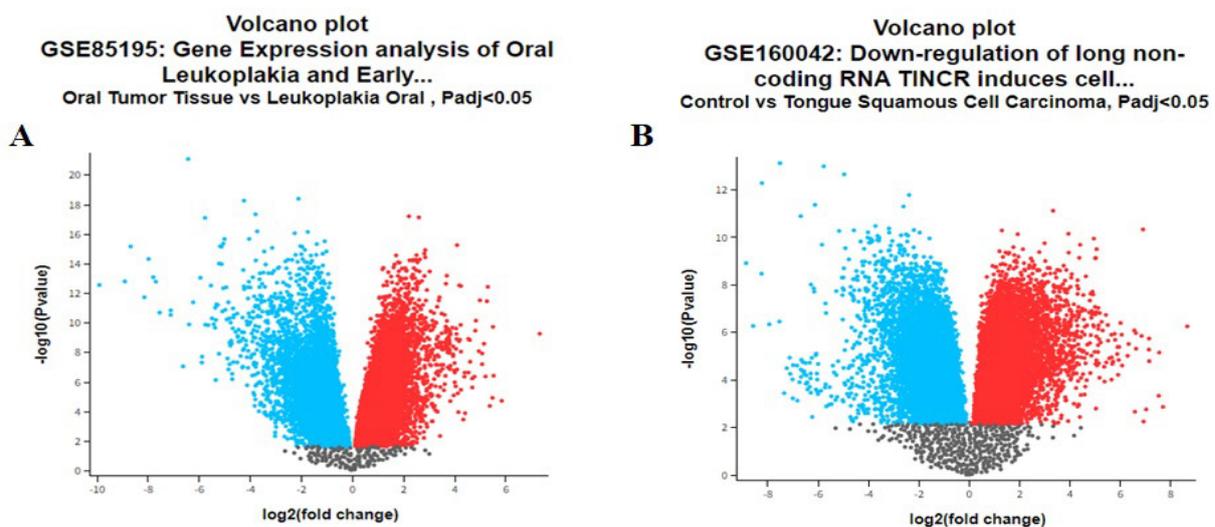


Figure 1. Volcano Plot. A, GSE85195 dataset of differentially expressed genes in patients with oral tumor and oral leukoplakia; B, GSE160042 dataset of differentially expressed ORF sequences in control patients and patients with squamous cell carcinoma of the tongue.

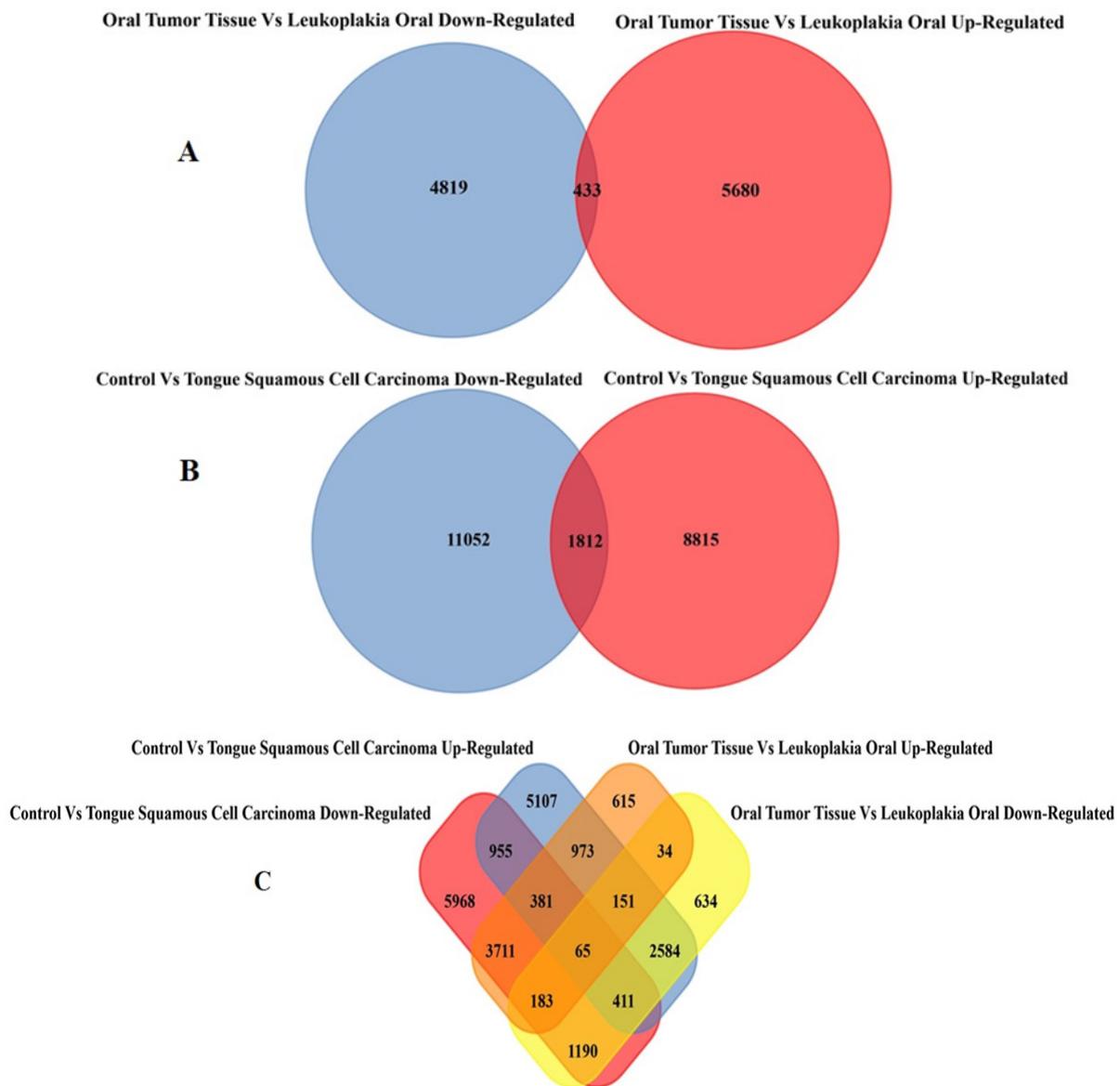


Figure 2. Venn Diagram. A. Venn diagram of the GSE85195 data set of differentially expressed genes in patients with oral tumor and oral leukoplakia. B. Venn diagram of the GSE160042 dataset of differentially expressed ORF sequences in control patients and tongue squamous cell carcinoma patients. C. Venn Diagram of Co-Expressed Gene Sequences.

associated with head and neck cancer with a value of 60%, and a significant percentage of 40% related to head cancer. However, several clinical phenotypes are linked to these co-expressed genes, such as symptoms precipitated by alcohol, caffeine, fatigue, stress, exertion, ovulation, and menstruation with a value of 20%. Additionally, damage to the teeth, lips, skin, nails, and hair.

Protein-protein interaction network

The analysis of the protein-protein interaction network between differentially co-expressed genes was established through the STRING database tool, generating 29 nodes and 42 edges. A significant module that includes 4 nodes and 8 edges was obtained using the MCODE algorithm (selected genes with yellow color), which are found to have a higher score and are related to the HNRNPU, UBE2I, SMC1A, and SF3B3 gene sequences as observed in Figure 6 (specifically in panel A).

Identification of hub genes

Selected hub genes of the protein-protein interaction network were obtained using the CytoHubba plugin MCC algorithm, as shown in Figure 6 (panel B, C and D). According to the MCC algorithm, the top 10 genes in the network include the gene sequence from HNRNPU, SMC1A, PAFAH1B1, EHMT1, SPTBN4, OLFM1, NCAM1, SF3B3, FGF2, and UBE2I; being HNRNPU, SMC1A and PAFAH1B1 the most representative of the different modules.

Discussion

In the present study, bioinformatics analysis was carried out based on the expression of the data sets GSE85195 and GSE160042, of which it was possible to detect 65 co-expressed gene sequences in each group, which can relate the progression of potentially malignant

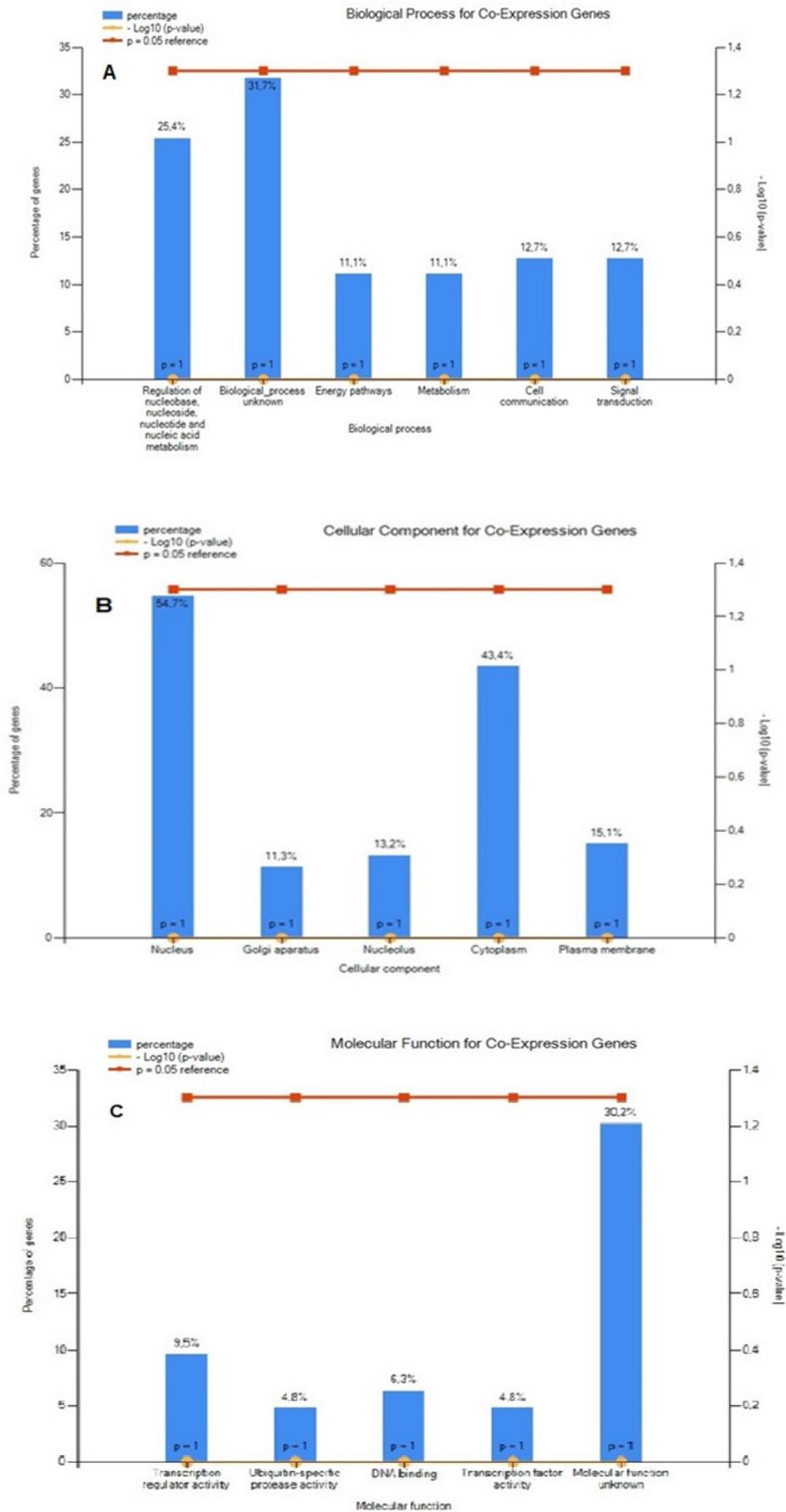


Figure 3. Functional Enrichment of GO. A, Biological processes for differentially co-expressed genes; B, Cellular component for co-expressed genes. C. Molecular functions for differentially co-expressed genes in the progression of Leukoplakia with oral tumor tissue and Squamous cell carcinoma of the tongue.

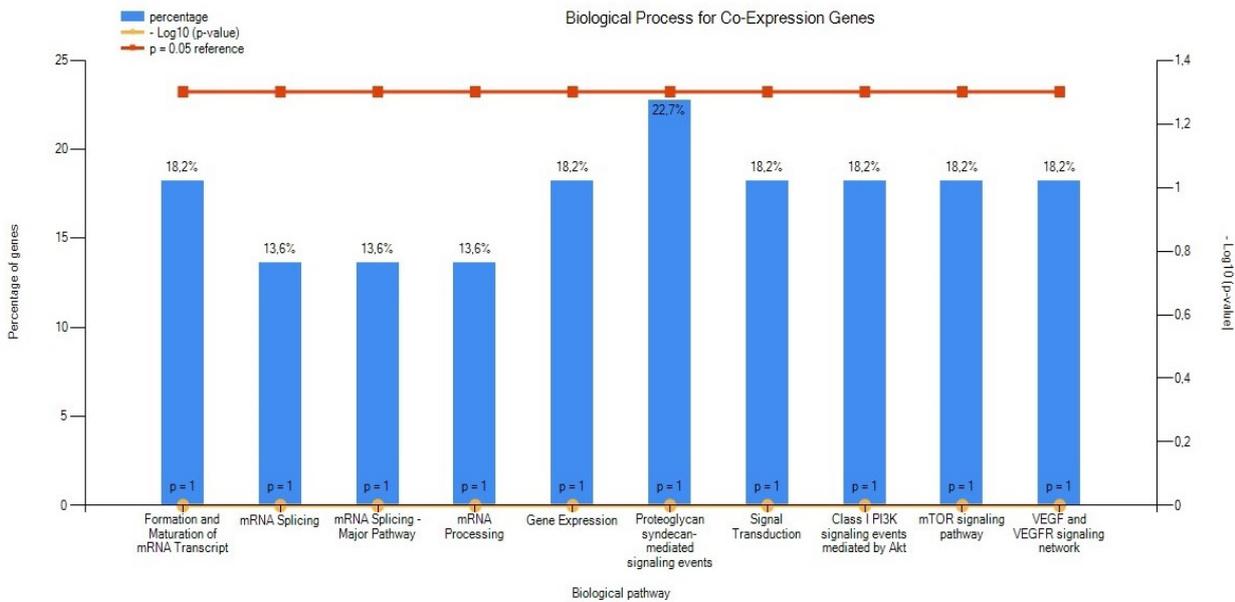


Figure 4. Enrichment of GO Pathways. Major biological pathways for differentially co-expressed gene sequences in the progression of leukoplakia with oral tumor tissue and squamous cell carcinoma of the tongue.

disorders of the oral cavity with the progression of oral squamous cell carcinoma that has been reported by Bhosale et al., (2017); Zhuang et al., (2021); Bhosale et al., (2017) and Zhuang et al., (2021).

Initially, differentially expressed genes were identified using the omnibus GEO2R tool, of all the filtered genes for both datasets, 973 were upregulated and 1,190 were downregulated, these genes could be involved in tongue carcinoma and progression from leukoplakia to oral tumor, these results highlight the complexity of squamous cell carcinoma of the oral cavity. Taking into account that to date multiple studies attempt to elucidate the molecular

mechanism of oral tumorigenesis, which ranges from endogenous genes related to metabolic pathways and different biological processes such as the Wnt5B activator pathway to already reported non-coding genes capable of expressing miRNA and mRNA (Chen et al., 2021), but without a doubt, the present work offers results that could specifically describe the progression from leukoplakia to oral cancer.

Regarding functional enrichment, it is important to highlight that the unknown biological processes were the ones that presented the highest percentage for the expressed genes, which may be because currently,

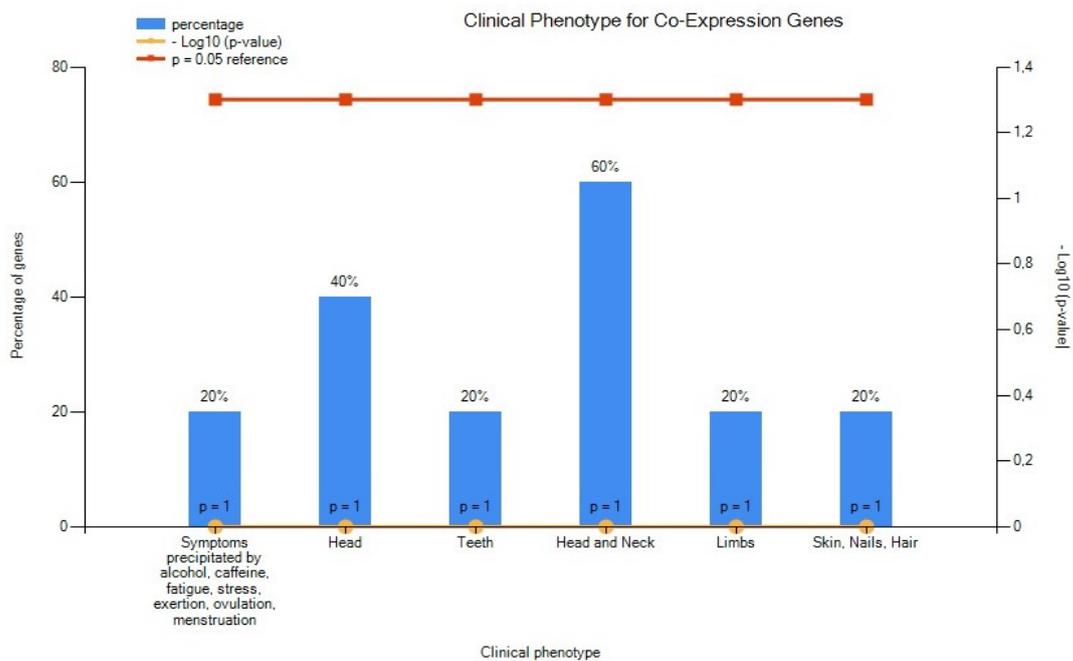


Figure 5. Clinical Phenotype. Main clinical effects for differentially co-expressed gene sequences on the progression of leukoplakia with oral tumor tissue and squamous cell carcinoma of the tongue

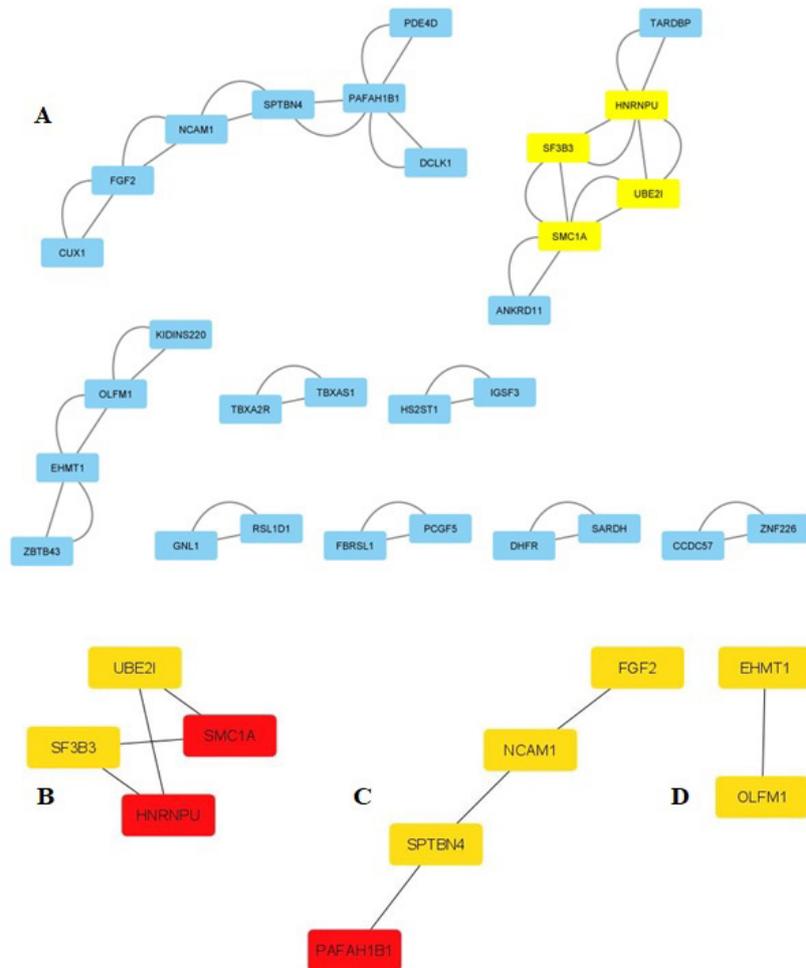


Figure 6. Protein-Protein Interaction Network and Identification of Hub Genes. A, Gene sequences that are co-expressed in the interaction network (blue color). Main gene sequences that are co-expressed in the interaction network (yellow color); B, Differentially co-expressed gene sequences of 4 nodes and 4 edges; C, Differentially co-expressed gene sequences of 4 nodes and 3 edges. D. Differentially co-expressed gene sequences of 2 nodes and 1 edge

the most studied promising markers for oral cancer are non-protein-coding genes such as those related to microRNA and mRNA, of which there is no specific understanding of their relationship with oral cancer, although there is a role that should be studied in depth. Therefore, it is of vital interest to continue investigating each of these genes to explain a mechanism related to tumorigenesis and progression from leukoplakia to oral cancer. It should be taken into account that since the regulation of nucleobases, nucleosides, nucleotides, and nucleic acid metabolism were the processes that reached the second place for the co-expressed genes identified in the present study, the hypothesis of the roles played by microRNAs, snRNAs and mRNAs could be confirmed. Similarly, signal transduction and cell communication were the processes that were mostly found in functional enrichment (see Figure 3).

Within the enrichment of cellular components, it was revealed that these genes were mainly involved in the nucleus and cytoplasm; and in small percentages related to the plasmatic membrane, nucleolus, and the Golgi apparatus. In addition, the analysis of molecular functions shows that many of their main activities are unknown and that, moderately, these genes are involved in the regulation

of transcription, DNA binding, transcription factor activity, and activity of ubiquitin-specific proteases. These results are very consistent with what was reported by Chen et al., (2021), who found genes involved in nuclear and cytosolic processes, highlighting that ubiquitin-like proteins, specifically ISG15, induce lymphovascular progression and non-coding genes for microRNA and mRNA. The latter could be linked to those previously mentioned unknown biological processes found in the present study.

Similarly, Chen et al., (2016) found that an increased ubiquitin-specific protease 14 (USP14) was associated with in vitro squamous cell carcinoma cell lines and with clinical samples associated with a poor survival rate (Chen et al., 2016). Which demonstrates a ratio of 4.8% ubiquitin-specific protease activity reported in the current bioinformatics analysis.

Regarding the result of GO enrichment pathways, several biological pathways could be classified, such as the proteoglycan-syndecan pathway. Numerous publications show that differentially expressed proteoglycans have been found during extracellular matrix remodeling in oral squamous cell cancer, such as agrin and perlecan (Kawahara et al., 2014; Mishra et al., 2013; Tilakaratne

et al., 2009). In addition, the expression of cell surface proteoglycans is frequently altered in head and neck squamous cell carcinomas primarily expressed in the oral cavity, as reported by Anna Farnedi et al 2015 (Farnedi et al., 2015), who identified NG2/CSPG4 and syndecan-2 as the only predictors of relapse and overall survival.

On the other hand, many pathways involving the formation and maturation of mRNA were found in the enrichment of pathways such as the spliceosome process, this is explained by Zhou et al., (2010) who confirmed the role of lncRNA-miRNA- mRNA ceRNA in the initiation and progression of oral cancer (Zhou et al., 2019). Similarly, Rishabh et al., (2020) reported that the family of MicroRNAs (miRNAs) have become master modulators of gene expression in various cellular and biological processes, therefore, the aberrant expression of these molecules has been associated with oral cancers.

The signaling networks of VEGFR/VEGF receptors and the PI3K/akt/Mtor signaling pathway that continue downstream in the cell cytosol are shown in Figure 4 as a result of pathway enrichment, their role in oral squamous cell cancer has also been widely reported.

Hani Al-Shareef et al., (2012) evaluated the expression of VEGF/VEGFR in lymph node metastasis in squamous cell carcinoma of the tongue and concluded that these vascular growth factors are expressed together with CCR7 in a non-independent way and constitute a predictive factor of the progression of the illness (Sasahira et al., 2012). Likewise, Li et al., (2016) have reported the central role of the VEGF/VEGFR axis in the progression of squamous cell carcinoma, they have even highlighted the role of this axis in the modulation by microRNA such as miR-126, which is inhibited in squamous cell carcinoma, demonstrating its role in VEGF/VEGFR modulation, this is important since the role of miRNAs has also previously been discussed in the present bioinformatic analysis (Al-Shareef et al., 2016).

As part of a complementary analysis to the GO functional enrichment for differentially co-expressed genes, the main clinical phenotypic effects related to these genes were identified, and clearly, head and neck cancer was the most frequent phenotype, which is very logical considering that oral squamous cell carcinoma is the most prevalent within the group of head and neck cancers, as reported in several studies by The Global Cancer Observatory (Sung et al., 2021). In addition, the rest of the phenotypes related to expressed genes linked to alcohol, tobacco, and stress constitute one of the main risk factors related to oral carcinoma (Chen et al., 2021).

It is important to emphasize that the protein-protein interaction network and the identification of Hub genes carried out in the present bioinformatic analysis revealed that the key genes identified with the highest score were HNRNPU, SMC1A, PFAH1B1, EHMT1, SPTBN4, OLFM1, NCAM1, SF3B3, FGF2, and UBE2I. The HNRNPU gene sequence codes for a heterogeneous nuclear ribonucleoprotein U related to DNA and RNA binding proteins, which are involved in several cellular processes, such as nuclear chromatin remodeling, DNA replication and repair, and RNA splicing (Ye et al., 2015; Zietzer et al., 2020). The SMC1A gene sequence encodes

protein 1A, which regulates chromosome maintenance and is involved in chromosome cohesion during the cell cycle and DNA repair (Musio, 2020). PFAH1B1 encodes the beta subunit of platelet-activating factor acetylhydrolase IB, which has the main function of catalyzing the hydrolysis of acetyl groups at the sn-2 position of PAF and its analogs and participates in the inactivation of PAF (Dinday et al., 2017).

The HNRNPU gene sequence has been reported by proteomic analysis to be overexpressed in dysplasia with oral cancer compared to normal tissue using western blot and mass spectrometry (Kumar et al., 2015). It should be noted that the role of HNRNPU is involved in cell proliferation, RNA splicing, and cell survival in oral squamous cell carcinoma. This association with the RNA splicing axis confirms the results presented in the functional enrichment where the majority of co-expressed genes affected the RNA maturation process.

Another of the genes that presented the highest score in the networks of interaction and identification of hub genes was SMC1A, whose mutations and deregulation have been reported to be highly relevant in different malignant carcinomas. For example, Zhang et al., (2013) demonstrated in vitro in A549 and H1299 lung cancer cell lines that were infected with a lentivirus, that SMC1A was expressed through shRNA, promoting the proliferation of cells to their tumor form, with which SMC1A is a regulator of tumor proliferation in lung carcinoma, but this could also be related to different squamous cell carcinomas such as oral cavity.

The other gene with the highest score in the bioinformatic analysis was PFAH1B1, whose overexpression according to Lo et al., (2012) increases migration and invasion in lung cancer cells, while suppression of PFAH1B1 decreases cell migration and invasion, and alters the organization of peri-cellular polyfibronectin assemblies and cell microtubules. In vivo tumor metastasis assay confirmed that knockdown of PFAH1B1 in lung cancer cells markedly reduced their ability to metastasize in animals.

Overall, the present study describes the gene networks that promote the progression of leukoplakia to oral squamous cell carcinoma. Within these networks, the HNRNPU, SMC1A, and PFAH1B1 genes are promising therapeutic targets for a treatment that counteracts the progression of oral cancer. They could also be important biomarkers for the diagnosis and classification of the disease.

In conclusion, This study proposes the identification of 10 hub genes as determinants in the progression of potentially malignant disorders to oral squamous cell carcinoma, where up and downregulated expressed genes can be linked, being the genes of the protein-protein interaction network with similar tendencies of expression those who obtained the highest score in the databases. The HNRNPU, SMC1A, and PFAH1B genes must be analyzed together at the same time, thus being able to obtain a specific profile of the patient with lesions suggestive of malignancy in the early stages of the disease and thus make timely diagnoses and prognoses. These genes also become therapeutic targets as important

alternatives to prevent tumorigenesis. We highlight the importance of the use of computational tools in the development of new drugs, allowing treatments to be adjusted with precision by analyzing the behavior of molecules with the best pharmacokinetic and toxicological characteristics.

Author Contribution Statement

Antistio Alvíz-Amador conceived and designed the study. Antistio Alvíz-Amador, Jaime Guzman de Avila perform all calculation and data collection. Antistio Alvíz-Amador, Jaime Guzman de Avila wrote the manuscript. All the authors read and approved the manuscript.

Acknowledgements

We thank the computer time provided by the University of Cartagena and the Professors Antistio Alvíz-Amador and Rafael Pineda-Aleman

Ethical considerations

Not applicable

Conflict of interest

None.

References

Al-Shareef H, Hiraoka SI, Tanaka N, et al (2016). Use of NRP1, a novel biomarker, along with VEGF-C, VEGFR-3, CCR7 and SEMA3E, to predict lymph node metastasis in squamous cell carcinoma of the tongue. *Oncol Rep*, **36**, 2444–54.

Barrett T, Wilhite SE, Ledoux P, et al (2013). NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res*, **41**, 991–5.

Bhosale PG, Cristea S, Ambatipudi S, et al (2017). Chromosomal alterations and gene expression changes associated with the progression of leukoplakia to advanced gingivobuccal cancer. *Transl Oncol*, **10**, 396–409.

Bray F, Ferlay J, Soerjomataram I, et al (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, **68**, 394–424.

Chen SH, Hsiao SY, Chang KY, Chang JY (2021). New insights into oral squamous cell carcinoma: From clinical aspects to molecular tumorigenesis. *Int J Mol Sci*, **22**, 1–21.

Chen X, Wu J, Chen Y, et al (2016). Ubiquitin-specific protease 14 regulates cell proliferation and apoptosis in oral squamous cell carcinoma. *Int J Biochem Cell B*, **79**, 350–9.

Chin CH, Chen SH, Wu HH, et al (2014). cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*, **8**, 1–7.

Dinday MT, Girskis KM, Lee S, Baraban SC, Hunt RF (2017). PAFAH1B1 haploinsufficiency disrupts GABA neurons and synaptic E/I balance in the dentate gyrus. *Sci Rep-Uk*, **7**, 1–10.

Farnedi A, Rossi S, Bertani N, et al (2015). Proteoglycan-based diversification of disease outcome in head and neck cancer patients identifies NG2/CSPG4 and syndecan-2 as unique relapse and overall survival predicting factors. *BMC Cancer*, **15**, 1–19.

González AP, López MA, Martínez LV (2006). Clinical and epidemiologic behavior of oral cancer. *Revista Cubana de*

Estomatologia, **43**.

Guerrero Brito M, Pérez Cabrera D, Hernández Abreu NM (2020). Lesiones bucales premalignas en pacientes con hábito de fumar. *Med Electrónica*, **24**, 159–64.

Kawahara R, Granato DC, Carnielli CM, et al (2014). Agrin and perlecan mediate tumorigenic processes in oral squamous cell carcinoma. *PLoS One*, **9**, 1–16.

Kumar M, Matta A, Masui O, et al (2015). Nuclear heterogeneous nuclear ribonucleoprotein D is associated with poor prognosis and interactome analysis reveals its novel binding partners in oral cancer. *J Transl Med*, **13**, 1–15.

Li YL, Zhao H, Ren XB (2016). Relationship of VEGF/VEGFR with immune and cancer cells: staggering or forward?. *Cancer Biol Med*, **13**, 206–14.

Lo FY, Chen HT, Cheng HC, Hsu HS, Wang YC (2012). Overexpression of PAFAH1B1 is associated with tumor metastasis and poor survival in non-small cell lung cancer. *Lung Cancer*, **77**, 585–92.

Macey R, Walsh T, Brocklehurst P, et al (2015). Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions (Review). *Cochrane Database of Sys Rev Diagn*, **5**.

Mishra M, Chandavarkar V, Naik VV, Kale AD (2013). An immunohistochemical study of basement membrane heparan sulfate proteoglycan (perlecan) in oral epithelial dysplasia and squamous cell carcinoma. *J Oral Maxillofac Pathol*, **17**, 31–5.

Musio A (2020). The multiple facets of the SMC1A gene. *Gene*, **743**, 144612.

Pathan M, Keerthikumar S, Ang CS, et al (2015). FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics*, **15**, 2597–601.

Paul S, Andrew M, Owen O, et al (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res*, **13**, 426.

Rishabh K, Khadilkar S, Kumar A, et al (2021). Review micromas as modulators of oral tumorigenesis—a focused review. *Int J Mol Sci*, **22**, 1–32.

Sasahira T, Kirita T (2018). Hallmarks of cancer-related newly prognostic factors of oral squamous cell carcinoma. *Int J Mol Sci*, **19**, 2413.

Sasahira T, Kurihara M, Bhawal UK, et al (2012). Downregulation of miR-126 induces angiogenesis and lymphangiogenesis by activation of VEGF-A in oral cancer. *Br J Cancer*, **107**, 700–6.

Sol S, Meir G, FL (1984). Oral leukoplakia and malignant transformation. *J Cancer*, **53**, 563–8.

Sung H, Ferlay J, Siegel RL, et al (2021). Global cancer statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*, **71**, 209–49.

Tilakaratne WM, Kobayashi T, Ida-Yonemochi H, et al (2009). Matrix metalloproteinase 7 and perlecan in oral epithelial dysplasia and carcinoma in situ: An aid for histopathologic recognition of their cell proliferation centers. *J Oral Pathol Med*, **38**, 348–55.

Von Mering C, Huynen M, Jaeggi D, et al (2003). STRING: A database of predicted functional associations between proteins. *Nucleic Acids Res*, **31**, 258–61.

Warnakulasuriya S, Johnson NW, Van Der Waal I (2007). Nomenclature and classification of potentially malignant disorders of the oral mucosa. *J Oral Pathol Med*, **36**, 575–80.

World Health Organization (2021). Oral health. https://www.who.int/health-topics/oral-health/#tab=tab_2.

Ye J, Beetz N, O'Keefe S, et al (2015). HnRNP U protein is required for normal pre-mRNA splicing and postnatal heart development and function. *Proc Nat Acad Sci U S A*, **112**,

3020–9.

- Zhang YF, Jiang R, Li JD, et al (2013). SMC1A knockdown induces growth suppression of human lung adenocarcinoma cells through G1/S cell cycle phase arrest and apoptosis pathways in vitro. *Oncol Lett*, **5**, 749–55.
- Zhou RS, Zhang EX, Sun QF, et al (2019). Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of tongue. *BMC Cancer*, **19**, 1–10.
- Zhuang Z, Huang J, Wang W, et al (2021). Down-regulation of long non-coding RNA TINCR induces cell dedifferentiation and predicts progression in oral squamous cell carcinoma. *Front Oncol*, **10**, 1–14.
- Zietzer A, Hosen MR, Wang H, et al (2020). The RNA-binding protein hnRNPU regulates the sorting of microRNA-30c-5p into large extracellular vesicles. *J Extracell Vesicles*, **9**, 1–21.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.