

COMPUTATIONAL ANALYSIS TO EXPLORE THE CROSSROAD BETWEEN CHRONIC PERIODONTITIS AND ORAL CANCER

Submitted in partial fulfillment of the requirement for the Award of Bachelor of
Science Degree in Bioinformatics and Data science

By

Nilofar (40738013)



DEPARTMENT OF BIOINFORMATICS AND DATA SCIENCE

**SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited with Grade "A" by NAAC
JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI - 600119**

May 2023



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

DEPARTMENT OF BIOINFORMATICS AND DATA SCIENCE

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **Nilofar (Reg.No: 40738013)** who carried out the project entitled "**Computational approach to explore the crossroad between the chronic periodontitis and oral cancer**" under my supervision from JANUARY 2023 to MAY 2023

Internal Guide

Dr. JEMMY CHRISTY.H, M.Tech., Ph.D.,

Head of the Department

Dr. JEMMY CHRISTY.H, M.Tech., Ph.D.,

Submitted for Viva voce Examination held on_05/05/23

Internal Examiner

External Examiner

DECLARATION

I, **NILOFAR A, (Reg.No: 40738013)** hereby declare that the projectreport entitled **“COMPUTATIONAL APPROACH TO EXPLORE THE CROSSROAD BETWEEN THE CHRONIC PERIODONTITIS AND ORAL CANCER”** done by me under the guidance of **Dr. JEMMY CHRISTY. H, M.Tech., Ph.D.**, is submitted in partial fulfillment of the requirements for the award of Bachelor of Science Degree in Bioinformatics and Data science.

DATE: 05/05/2023

A handwritten signature in black ink that reads "Nilofer". The script is cursive and fluid.

PLACE: Chennai

SIGNATURE OF THE CANDITATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to Board of Management of SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr.V.RAMESH KUMAR, M.Phil., Ph.D.**, Dean, School of **BIO AND CHEMICAL ENGINEERING**, **Dr. JEMMY CHRISTY H, M.Tech., Ph.D.**, Head of the Department of **Bioinformatics and Data science** for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr.JEMMY CHRISTY.H,M.Tech.,Ph.D.**, for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the Department of Bioinformatics and Data science who were helpful in many ways for the completion of this project

ABSTRACT

Periodontitis is a chronic inflammatory disease that affects the supporting tissues of the teeth, including the gums, periodontal ligament, and alveolar bone. It is caused by the accumulation of dental plaque and the host's immune response to bacterial toxins, resulting in destruction of the supporting tissues. Periodontitis has been associated with a variety of systemic conditions, including cardiovascular disease, diabetes, and respiratory disease.

In this study, we performed an in-silico analysis to identify differentially expressed genes in periodontitis and oral cancer. The microarray dataset of GSE10334 and GSE13601 was selected from gene expression omnibus (GEO) database and Data from periodontal disease ontology browser was taken to look at the common genes between the three datas (periodontal disease Disease Ontology Browser - [DOID:3388 \(jax.org\)](https://doi.org/10.1093/doid/3388)) GEO2R tool was used and identified a total of top differentially expressed genes (DEGs), of which 68 and 88 genes were up-regulated and down-regulated respectively. Cytoscape was used for the screening of clusters and identification of hub genes. I have used the cytohubba and reactome plugin to identify genes with highest degree nodes from up-regulated and down-regulated gene networks respectively and hence can be used as possible molecular target. In order to validate the accuracy of the genes identified as the hub genes, literature from PubChem will have been provided as evidence.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
1	INTRODUCTION	
2	LITERATURE SURVEY	
3	AIM AND SCOPE	
	3.1 AIM	
	3.2 SCOPE	
4	MATERIALS AND METHODS.	
	4.1 ONLINE DATABASES AND TOOLS 4.2	
	REQUIREMENT.	
	4.3 ABOUT THE DATABASES AND TOOLS USED.	
	4.4 METHODOLOGY.	
	4.4.1 SELECTION OF GEO DATASETS.	
	4.4.2 PROCESSING OF RETRIEVED DATA	
	FOR DIFFERENTIALLY EXPRESSED	
	GENES.	
	4.4.3 CONSTRUCTION OF VENN	
	DIAGRAM.	
	4.4.4 SCREENING OF PPI NETWORKS	
	FOR MODULES AND ANALYSIS OF HUB	
	GENES.	
	4.4.5 LITERATURE EVIDENCE	
	VALIDATION	
5	RESULTS AND DISCUSSION	
	5.1 DEG ANALYSIS.	
	5.2 FUNCTIONAL ENRICHMENT OF DEGs.	
	5.3 CONSTRUCTION OF PPI NETWORK.	
	5.4 SCREENING OF MODULES AND ANALYSIS	
	OF HUB GENES IN PPI NETWORK.	
6	SUMMARY AND CONCLUSION.	
	6.1 CONCLUSION.	
	REFERENCES.	

LIST OF FIGURES

FIGURE NO.	FIGURE NAME	PAGE NO.
5.1	Top 30 upregulated genes	14
5.2	Top 30 down regulated genes	15
5.3	Construction of Venn Diagram	16
5.4	Network hub genes identification performed by cytohubba reactome based on 12 algorithms.	20
5.5	Identification of frequently occurring hub genes and their corresponding symbols	21
5.6	Top 10 hub genes identified through cytoscape	22

LIST OF ABBREVIATION

GLOBOCAN - Global Cancer Observatory WHO-

World Health Organization

GEO - Gene expression Omnibus

SOS1- Son of Sevenless 1

TNF-Tumor Necrosis Factor

MET- Mesenchymal Epithelial Transition VEGFA-

Vascular Endothelial Growth Factor IL2- Interleukin 2

PPARA- Peroxisome Proliferator-Activated Receptor Alpha SP1-

Specificity Protein 1

MMP2- Matrix Metalloproteinase 2 APP-

Amyloid Beta Precursor Protein COL1A1-

Collagen Type I Alpha 1

CHAPTER - 1

INTRODUCTION

Periodontitis and oral cancer are two of the most prevalent and debilitating oral health conditions worldwide, with significant impacts on quality of life and overall health (Chapple et al., 2018; Chen et al., 2019). While the link between these two conditions has been established in previous research, the underlying mechanisms and interactions between the two are not fully understood. This knowledge gap presents a challenge for healthcare professionals seeking to prevent, diagnose, and treat these conditions. Recent studies have highlighted the potential role of molecular and cellular mechanisms in the crossroad between periodontitis and oral cancer (Mashima et al., 2016; Zhang et al., 2020). However, traditional experimental approaches are limited in their ability to comprehensively analyze the complex interactions between genes, proteins, and pathways involved in these conditions.

Computational approaches, on the other hand, have emerged as a powerful tool for analyzing complex biological systems and identifying novel insights into disease mechanisms (Wang et al., 2019; Wang et al., 2020). In this research project, we aim to use computational methods to investigate the potential connections between periodontitis and oral cancer, with a focus on understanding the molecular and cellular pathways involved. Specifically, we will use a variety of computational techniques, including network analysis, machine learning, and statistical modeling, to analyze large-scale datasets of genetic, epigenetic, and transcriptomic data from periodontitis and oral cancer patients. By integrating these diverse datasets, we hope to identify key genes, pathways, and regulatory mechanisms that are involved in the crossroad between these two conditions. Ultimately, our goal is to provide a more comprehensive understanding of the molecular and cellular mechanisms underlying periodontitis and oral cancer, and to identify potential targets for future prevention and treatment strategies. By leveraging the power of computational approaches, we believe this research has the potential to make significant contributions to the field of oral health and beyond.

CHAPTER - 2

LITERATURE SURVEY

Cancer is a complex disease that arises from abnormal growth and division of cells in the body. It is the second leading cause of death globally, accounting for approximately 9.6 million deaths in 2020 alone. Cancer can occur in any part of the body, and there are more than 200 types of cancer, each with its unique characteristics and treatment options. Various aspects of cancer biology, diagnosis, and treatment. For instance, a study by (Ghasemzadeh et al, 2021) investigated the role of the immune system in cancer development and progression. A research has focused on developing new treatments for cancer. Oral cancer is a type of cancer that develops in the tissues of the mouth or throat, including the lips, gums, tongue, and inner lining of the cheeks. According to the Global Cancer Observatory (GLOBOCAN) 2022 report, there were 377,713 new cases of oral cancer and 177,757 deaths worldwide in 2020. This makes oral cancer the 15th most common cancer globally and the 7th most common cancer in men. Several risk factors have been identified for oral cancer, including tobacco and alcohol use, viral infections such as human papillomavirus (HPV), poor oral hygiene, and a diet low in fruits and vegetables. Early detection is crucial for successful treatment and management of oral cancer.

Recent studies have focused on the role of biomarkers in the early detection and diagnosis of oral cancer. For example, a 2021 study published in the journal Oral Oncology found that a panel of saliva biomarkers could accurately distinguish between healthy individuals and those with oral cancer or potentially malignant oral lesions. Another 2020 study published in the same journal identified a specific microRNA molecule in saliva that was associated with the progression of oral cancer. According to a study conducted by (Gupta et al, 2020) oral cancer is the most common cancer in India. The study was conducted on 872 patients diagnosed with oral cancer, and it was found that the incidence of oral cancer was highest among males aged 40-59 years. A recent report by the National Cancer Registry Programme (NCRP) (2021) revealed that the incidence of oral cancer in India has been steadily increasing over the past few years. In 2020, it was

estimated that there were 92,697 new cases of oral cancer in India. In a study conducted by (Mishra et al, 2021), it was found that the prevalence of oral cancer was higher in rural areas compared to urban areas. The study was conducted on 250 patients diagnosed with oral cancer, and it was found that 70% of the patients were from rural areas. According to a report by the World Health Organization (WHO) (2020), tobacco and alcohol consumption are major risk factors for oral cancer in India. The report also states that smokeless tobacco is a significant risk factor for oral cancer in India, and it is commonly used in the form of gutkha, paan masala, and khaini. Periodontitis is a chronic inflammatory disease that affects the supporting tissues of the teeth, including the gums, periodontal ligament, and alveolar bone. It is caused by the accumulation of dental plaque and the host's immune response to bacterial toxins, resulting in destruction of the supporting tissues. Periodontitis has been associated with a variety of systemic conditions, including cardiovascular disease, diabetes, and respiratory disease. Recent studies have also suggested a link between periodontitis and oral cancer.

Oral cancer is a malignant tumor that can develop in any part of the oral cavity, including the lips, tongue, cheeks, floor of the mouth, hard and soft palate, and pharynx. It is the sixth most common cancer worldwide and has a high mortality rate due to late diagnosis and poor prognosis. The main risk factors for oral cancer include tobacco and alcohol use, human papillomavirus (HPV) infection, and chronic inflammation. Several studies have investigated the association between periodontitis and oral cancer. A meta-analysis of 15 case-control studies found that periodontitis was significantly associated with an increased risk of oral cancer (OR=2.44, 95% CI: 1.93-3.08) (Chen et al., 2018). Another meta-analysis of 11 case-control studies reported a similar association (OR=2.39, 95% CI: 1.95-2.93) (Guo et al., 2019). The mechanism underlying the link between periodontitis and oral cancer is not fully understood, but several hypotheses have been proposed. One hypothesis is that chronic inflammation in the periodontal tissues may lead to the production of reactive oxygen species (ROS) and pro-inflammatory cytokines, which can cause DNA damage and promote carcinogenesis (Yu et al., 2019). Another hypothesis is that periodontal pathogens may enter the bloodstream and reach the oral mucosa, where they can interact with other risk factors, such as tobacco and alcohol, to promote cancer development (Cancer et al., 2020).

Several periodontal pathogens have been implicated in the development of oral cancer. *Porphyromonas gingivalis*, a gram-negative bacterium commonly associated with periodontitis, has been shown to promote oral cancer cell growth and invasion in vitro (Zhu et al., 2018). The treatment of periodontitis may have a beneficial effect on the prevention and treatment of oral cancer. A randomized controlled trial of 51 patients with periodontitis found that periodontal treatment improved the periodontal status and decreased the levels of pro-inflammatory cytokines in the gingival crevicular fluid, suggesting a reduction in the systemic inflammatory burden (Kumar et al., 2019). Another study of 101 patients with oral cancer found that periodontal treatment improved the oral hygiene and reduced the levels of periodontal pathogens in the oral cavity, suggesting a potential role in preventing cancer recurrence (Zhang et al., 2020). In conclusion, periodontitis is a risk factor for oral cancer, and the underlying mechanism may involve chronic inflammation and the interaction between periodontal pathogens and other risk factors. The treatment of periodontitis may have a beneficial effect on the prevention and treatment of oral cancer. Further studies are needed to elucidate the exact mechanisms and to develop effective strategies for the prevention and management of oral cancer in patients with periodontitis.

CHAPTER - 3

AIM AND SCOPE

3.1 AIM

The objective of this research project is to conduct a computational analysis of gene samples from patients with periodontitis and oral cancer to identify potential associations between these two conditions using various bioinformatics tools and databases. We performed a series of analyses to identify the differentially expressed genes between the two diseases periodontitis and oral cancer. Since the patients having periodontitis can lead to development of oral cancer in near future. We find the hub genes that can act as biomarkers.

3.2 SCOPE

This research project aimed to computationally analyse the crossroad between periodontitis and oral cancer by combining datasets. Our analysis revealed several genes that were significantly upregulated in oral cancer, suggesting their potential as biomarkers.

We analyzed the expression levels of the potential biomarker genes and cross-checked their associated pathways and diseases in the PubChem database. Our analysis confirmed the reliability of the identified biomarker genes and their relationship with periodontitis and oral cancer. The validation of the biomarker genes using the PubChem database strengthens the significance of the findings of this study and highlights the potential of using computational methods for identifying biomarkers for oral cancer.

Overall, this study highlights the importance of exploring the relationship between periodontitis and oral cancer through computational analysis. Our findings provide a starting point for future studies aimed at developing effective biomarkers and

treatment strategies for oral cancer. The identified biomarkers could serve as targets for developing novel therapies for oral cancer. This could involve designing small molecule inhibitors or monoclonal antibodies against the proteins encoded by the highly expressed genes.

CHAPTER - 4

MATERIALS AND METHODS

4.1 ONLINE DATABASES AND TOOLS REQUIREMENT

1. GEO (Gene Expression Omnibus)
2. GEO2R
3. DO Browser (Disease Ontology Browser)
4. BioVenn (web application for visualisation of biological lists)
5. Cytoscape (Software)
 - i. cytoHubba plugin
 - ii. Reactome plugin
6. PubChem (literature evidence)

4.2 ABOUT THE DATABASES AND TOOLS USED

1. **Gene Expression Omnibus (GEO):** GEO is a public repository of microarray and next-generation sequencing data. It contains a large collection of gene expression datasets and provides access to these datasets to the scientific community. GEO is a publicly accessible database that provides access to gene expression data from a wide range of organisms and experimental conditions. It is maintained by the National Center for Biotechnology Information (NCBI), which is part of the National Institutes of Health (NIH). The GEO database contains data from various high-throughput experiments, such as microarrays, RNA sequencing, and other genomic technologies. GEO data can be used for a wide range of applications, such as identifying genes that are differentially expressed in different cell types or disease states, exploring the molecular pathways involved in various biological processes, and predicting drug responses. The GEO database also provides various tools and resources to facilitate data analysis and visualization, including the GEO2R web tool, which allows users to perform differential

expression analysis on GEO datasets, and the GEOquery R/Bioconductor package, which provides programmatic access to the GEO database from within the R statistical computing environment. The Gene Expression Omnibus (GEO) was utilized to obtain microarray datasets pertaining to patients affected with periodontitis and oral cancer. Specifically, two datasets were utilized: GSE10334, containing sample data from patients with periodontitis, and GSE13601, containing sample data from patients with oral cancer. These datasets were analyzed as part of the research project.

2. Disease Ontology browser: The Disease Ontology (DO) is a comprehensive ontology of human diseases that provides standardized nomenclature, definitions, and relationships between diseases. The DO is intended to be used for the annotation and integration of biomedical data related to human diseases, including genomic, genetic, proteomic, and other types of data. The Disease Ontology Browser is a web-based tool that allows users to explore the DO and search for specific diseases, disease groups, and related terms. The browser provides a user-friendly interface that displays the DO hierarchy in a tree format, with diseases arranged according to their relationships to other diseases. Users can expand and collapse branches of the tree to explore the hierarchy at different levels of granularity. The browser also provides a search function that allows users to search for specific terms or phrases within the DO. DO is often used in conjunction with other resources that provide gene data, such as the Gene Expression Omnibus (GEO) and the Genotype-Tissue Expression (GTEx) project. These resources allow users to download gene expression data and other types of genomic data related to specific diseases and tissues. In my project, I utilized a web browser to extract data for the purpose of analyzing the hierarchical association between periodontal disease and oral cancer. Specifically, I compared gene expression data from periodontal disease samples with those obtained from oral cancer samples to ascertain the nature of their relationship.

3. **GEO2R:** The GEO2R tool is a web-based application that allows researchers to analyse and interpret gene expression data obtained from the Gene Expression Omnibus (GEO) database. The tool is developed by the National Center for Biotechnology Information (NCBI) and is freely available to use. The GEO2R tool enables researchers to compare gene expression levels between two or more experimental groups and identify genes that are differentially expressed. This can help researchers to understand the molecular mechanisms underlying various diseases and conditions, and to identify potential therapeutic targets. To use the GEO2R tool, researchers can select a GEO dataset of interest and specify the experimental groups they wish to compare. The tool then performs statistical analysis and generates a list of differentially expressed genes, along with information on fold change, p-value, and adjusted p-value. The GEO2R tool also provides a range of visualization options, including heatmaps and volcano plots, which allow researchers to explore the gene expression data. In this project I have done GEO2r analysis on samples from

4. **Cytoscape:** Cytoscape is an open-source software platform for visualizing and analyzing molecular interaction networks. In this project, Cytoscape was used to visualize protein-protein interaction networks, to perform network analysis, and to identify hub genes using the cytoHubba plugin. Cytoscape is a powerful open-source software platform widely used in network biology research to analyse, visualize and explore complex biological networks such as gene regulatory networks, protein-protein interaction networks, and metabolic networks. Cytoscape provides a user-friendly interface and offers a wide range of features such as network layout algorithms, data integration and visualization tools, network clustering and community detection algorithms, and network analysis tools for topological and functional analysis of the network. It allows researchers to import and export various network file formats, customize network visualizations and create publication-quality figures. The plugin architecture allows researchers to customize the software by integrating new features and algorithms. In my project I have used cytoscape to visualize networks of the hub genes that can act as biomarkers for the disease oral cancer.

5. **cytoHubba**: CytoHubba is a Cytoscape plugin that helps in identifying important nodes and sub-networks in a biological network. It provides a variety of network analysis algorithms, including degree, betweenness, and closeness centrality, as well as clustering and pathway enrichment analysis. CytoHubba allows users to prioritize the top nodes in a network based on their centrality measures or to extract sub-networks based on the nodes' connectivity patterns. It can also identify functional modules and pathways within the network, which can provide insights into the biological processes that are active in the system. In my project I have used cytohubba as a plugin to find the network of genes that are highly connected based on the 12 algorithms.

6. **Reactome**: The Reactome plugin for CytoHubba provides several algorithms for pathway enrichment analysis, allowing users to identify enriched Reactome pathways in their input network. This plugin takes a list of genes or proteins as input and generates a ranked list of enriched Reactome pathways based on statistical significance. The plugin also allows the user to customize several parameters, such as the minimum number of genes required for a pathway to be considered significant and the statistical test used to calculate enrichment.

4.3 METHODOLOGY

4.3.1 Selection of GEO datasets

The dataset analysed in this study was freely available and retrieved from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). The gene expression profile GSE10334 and GSE13601 respectively that contains sample data of periodontitis and oral cancer respectively.

4.3.2 Processing of retrieved data for differentially expressed genes

The prediction of the Differentially Expressed Genes present in the selected dataset was analyzed by GEO2R tool (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>). The Significant genes were filtered based on the cutoff criteria of adjusted P-Value less than 0.05. From the significant genes, the up-regulated and down-regulated genes were predicted based on Log FC values such as $\text{LogFC} > 2$ and $\text{LogFC} < -2$, respectively.

4.3.3 Formation of VENNDIAGRAM to find common gene

BioVenn is a web-based tool that allows users to create Venn diagrams of biological sets, such as gene lists or protein clusters. BioVenn can also be used to visualize the results of high-throughput sequencing experiments, such as RNA-seq or ChIP-seq. In these experiments, researchers may identify thousands of genes or regulatory elements that are differentially expressed between two or more conditions. By using BioVenn to create Venn diagrams of these gene lists, researchers can quickly identify the genes that are specific to each condition, as well as those that are common to multiple conditions. In this project, we utilized BioVenn as a tool for identifying common genes among datasets of periodontitis (GSE10334), oral cancer (GSE13601), and literature-derived periodontitis gene lists. The analysis revealed a total of 56 genes that were present in all three datasets.

4.3.4 Screening of PPI networks for modules and analysis of hub genes

Cytoscape (version 3.9.1) (<http://www.cytoscape.org/>), is an open-source software platform for visualizing and analyzing complex networks, such as biological pathways, gene regulatory networks, and protein interaction networks. The mapped networks were exported to cytoscape and visualised for the further analysis. CytoHubba is a plugin for the Cytoscape software, which is a powerful tool used for analyzing and visualizing biological networks. The CytoHubba plugin is specifically designed for identifying important nodes or hubs within a network,

based on various topological measures. With CytoHubba, users can identify network hubs using a variety of algorithms, such as degree, betweenness centrality, closeness centrality. These measures can be used individually or in combination to identify nodes that are highly connected or central to the network. It was used to predict the hub genes in PPI network based on the parameter degree.

4.3.5 Screening of Enriched pathways and analysis of hub genes

In this work, the Reactome plugin for CytoHubba was employed to identify the most highly connected nodes within a biological network, and subsequently, the associated genes that play important roles in the network's biological functions. By analyzing 12 distinct parameters such as degree, centrality, radiality, occurrence, closeness, MCC, MNC, DMNC, stress, clustering, and betweenness, the plugin generated a ranked list of genes, with the top 10 genes identified and prioritized based on their biological significance. The resulting gene list provides a valuable resource for researchers interested in investigating the underlying biological mechanisms of the system. With insights into the key players and pathways within the network, researchers can develop new hypotheses and generate testable predictions that may lead to a deeper understanding of the system's behavior. The Reactome plugin for CytoHubba allows users to customize several parameters and apply advanced statistical analyses to identify enriched pathways, making it a powerful tool for pathway enrichment analysis. The top 10 genes identified through the Reactome plugin for CytoHubba analysis were found to be SOS1, TNF, PPARA, MET, VEGFA, SP1, MMP2, APP, COL1A1, and IL2. These genes were prioritized based on their high connectivity and biological significance, and may serve as important targets for oral cancer.

CHAPTER - 5

RESULTS AND DISCUSSION

5.1 DEG analysis

The genes expressing differentially in the dataset related to gingival tissue transcriptome in states of periodontal health samples (GSE10334) and Oral tongue cancer gene expression profiling samples (GSE13601) was analyzed using GEO2R tool. The criteria to select the significant genes were the probability value P which should be less than 0.05 and Log fold change values. The cut off criteria for up-regulated and down-regulated genes was Log FC >2 and Log FC < 2, respectively. The GEO2R analysis for GSE211392 showed 250 genes of which all 250 genes were found to be significant based on the above mentioned criteria. In this, 236 genes were differentially expressed of which 74 were up-regulated and 162 genes were down-regulated. Top 30 genes which are up-regulated and downregulated along with logFC value and P-Value are provided in Table 1.

Figure 5.1 Top 30 up-regulated genes

upregulated genes							Gene.symbol	Gene.title
ID	adj.P.Val	P.Value	t	B	logFC	m		
	0.00003		8.14113	8.5442	1.9894	PECAM		
208983 S	451	5.43E-08	66	5	89	1		platelet and
	0.00091	0.000009	5.76094	3.6106	1.9686	IGHA2/		immunoglob
217022 S	748	5	12	6	26	//1		ulin
	0.00186	0.000028	5.29034		1.9514			
205692 S	253	6	49	2.5492	59	CD38		CD38
211835	0.00004		7.95316		1.9396			immunoglob
at	477	7.94E-08	23	8.1846	3	IGKC		ulin
206574 S	0.00011	0.000000	7.28048	6.8541	1.9386	PTP4A3		protein

	529	323	88	6	9		
219505	0.00001		9.09867	10.294	1.9348		
at	133	8.39E-09	78	62	29	CECR1	cat eye
	0.00019	0.000000	6.90087	6.0737	1.9326		
215646 S	292	731	62	3	5	VCAN	versican
	0.00001		8.94835	10.028	1.9213	THEMIS	
210785 S	326	1.12E-08	18	77	91	2	thymocyte
			-				
	0.02204		3.67095	1.1785	1.9179		
202988 S	953	0.00139	99	4	02	RGS1	regulator
	0.00030	0.000001	6.51383	5.2567	1.9171	ADAM2	
205997	722	71	32	4	87	8	ADAM
	0.00008	0.000000	7.47743	7.2506	1.9098	LOC101	uncharacteri
217239	808	213	63	9	31	93	zed
			-				
	0.01042		4.14956	0.0805	1.8918	C16orf5	chromosom
1559584	926	0.00044	39	4	52	4	e
	0.00023	0.000001	6.73730	5.7310	1.8786		X-box
200670	323	05	06	2	9	XBP1	binding

Upregulated genes

Figure 5.2 Top 30 down-regulated genes

ID	downregulat ed	genes adj.P.Val	P.Value	t	B	logFC	Gene.sym	Gene.title
				-		-		
227498		0.003746	9.13E-05	4.80278	1.43059	1.37966	SOX6	SRY-box 6
				-	-	-		
1555383		0.086796	0.0115	2.76434	3.16038	1.38298	POF1B	premature
				-	-	-		
205337 a		0.133926	0.0227	2.45497	3.78206	1.39213	DCT	dopachrome

			-		-		
	0.00189	2.96E-	5.2757		1.4034		
210633)	9	05	3	2.5159	6	KRT10	keratin 10
				-			
	0.02772			1.5003			solute
219932	6	0.00195	-3.5287	8	-1.4036	SLC27A6	carrier
			-		-		
	0.00693	0.00023	4.4144	0.5322	1.4064		arylacetami
205969	5	2	4	9	6	AADAC	de
			-		-		
	0.00042	2.98E-	6.2680	4.7272	1.4119		
205637 s	7	06	3	3	8	SH3GL3	SH3 domain
			-		-		
	0.00011	3.23E-	7.2803		1.4120		gap
239572 at	5	07	6	6.8539	4	GJA3	junction
			-		-		
	2.63E-	3.43E-	8.3710	8.9769	1.4139	BNIP1///	
236534 at	05	08	5	5	3	C	BCL2
			-		-		
	0.00253	4.72E-	5.0786		1.4143		
1554812	9	05	1	2.0653	3	CLDN20	claudin 20
			-		-		
	0.00020	8.15E-	6.8508	5.9692	1.4306		
210426 x	2	07	2	7	1	RORA	RAR related
			-		-		
	0.00012	3.68E-	7.2186	6.7283	1.4371		
209570	5	07	1	9	1	NSG1	neuron
			-		-		
	0.00030	1.62E-	6.5379	5.3082	1.4424	EPB41L4	
220161 S	3	06	5	6	6	B	erythrocyte

Down-regulated

5.2 Construction of Venn Diagram

In order to identify common genes among three datasets, a Venn diagram was constructed using the BioVenn tool. The gene lists from the periodontitis dataset (GSE10334), oral cancer dataset (GSE13601), and periodontitis literature were uploaded into the tool, and 56 common genes were identified through the overlapping regions of the Venn diagram. This approach provides a simple yet effective way to identify genes that are shared across multiple datasets and may have important implications for understanding the underlying biological

mechanisms of the studied diseases. The identified common genes can be further analyzed to gain deeper insights into their biological functions and potential roles as therapeutic targets.

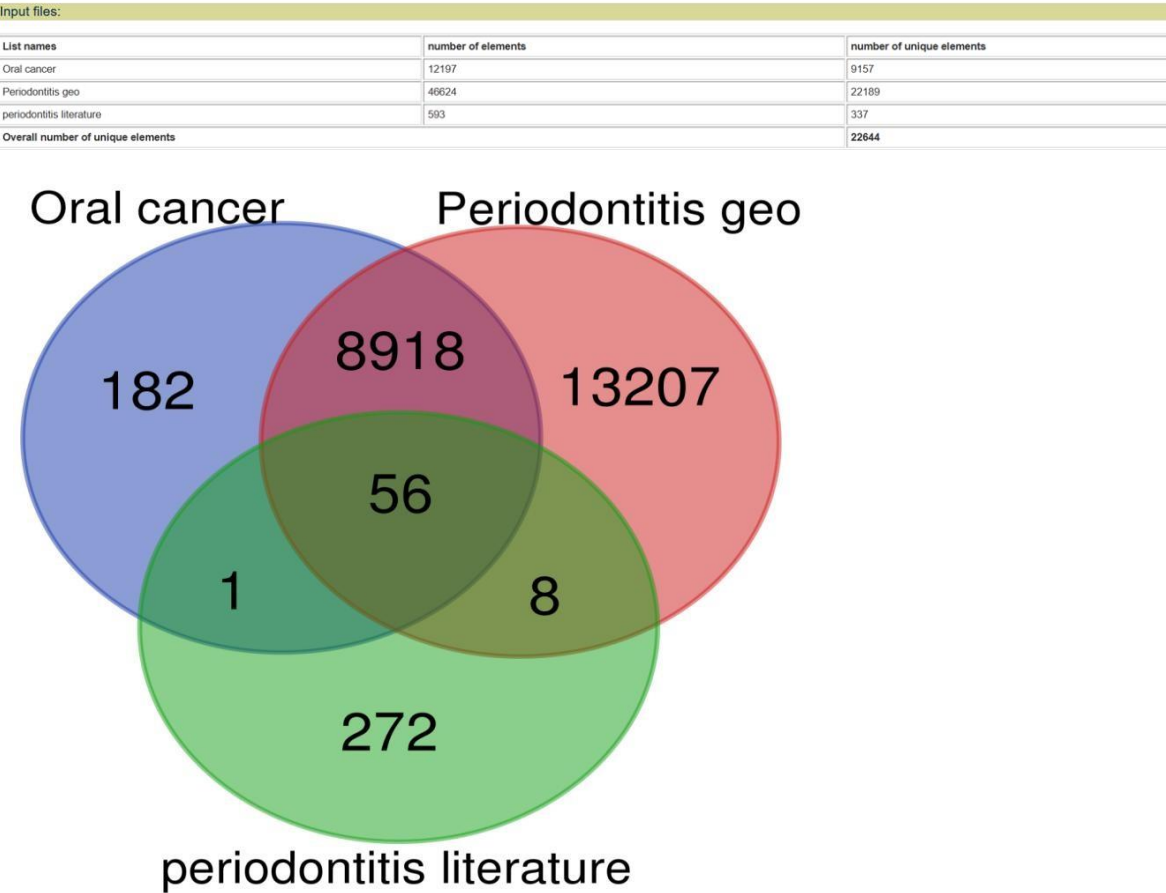


Figure 5.1 Construction of Venn diagram

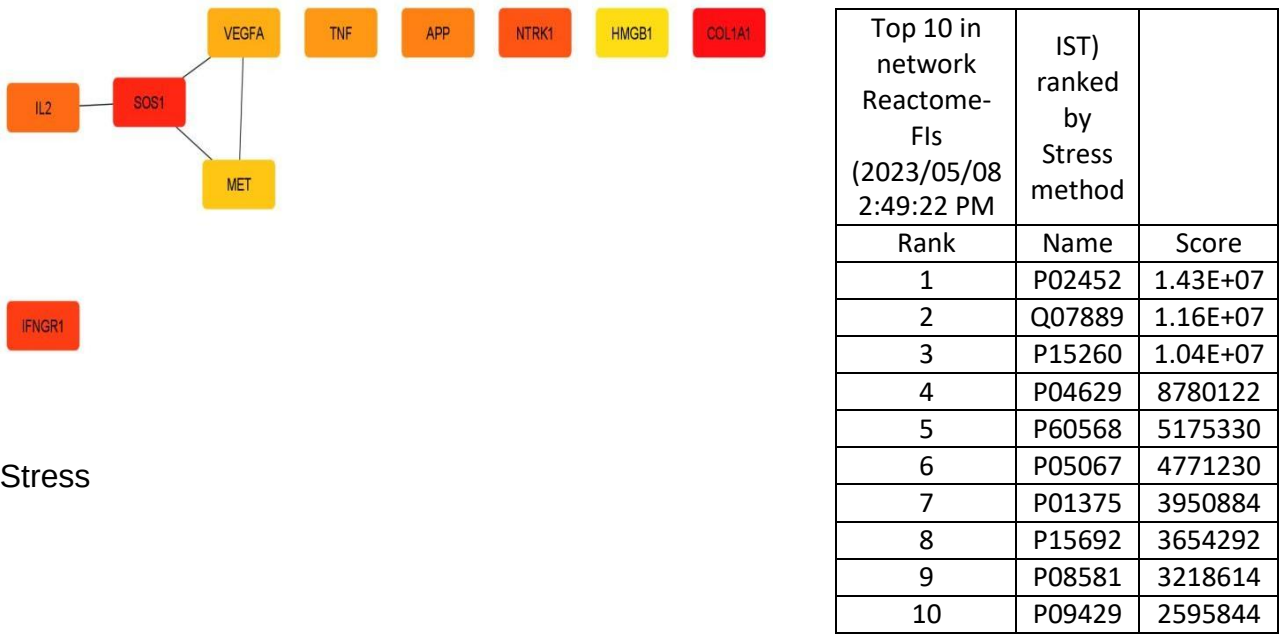
5.4 Screening of modules and analysis of hub genes in PPI network

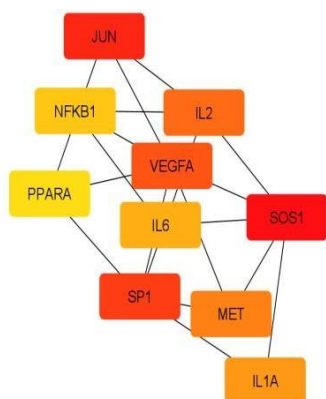
CytoHubba is a Cytoscape plugin that provides a set of algorithms for identifying network hubs and important nodes in protein-protein interaction networks (PPIs). The plugin allows users to perform several types of analyses, including identifying topological centralities.

Hub genes identification was performed by cytoHubba based on 12 algorithms for up-regulated and down-regulated gene network separately (figure 6), (figure 7), which are Maximal Clique Centrality (MCC), Degree, Maximum Neighborhood

Component (MNC), Edge Percolated Component (EPC), Density of Maximum Neighborhood Component (DMNC), Bottleneck, EcCentricity, Closeness, Radiality, Stress, Betweenness, Clustering Coefficient (CC). The up-regulated gene network was analysed for the hub genes with node degree as parameter. From the network the hub genes identified were as SOS1, TNF, PPARA, MET, VEGFA, SP1, MMP2, APP, COL1A1, and IL2. The Gene Ontology (GO) annotations and pathways of these genes were analysed. These hub genes are mainly involved in transmembrane transport in Biological process (BP), plasma membrane in Cellular component (CC), transmembrane transporter activity in Molecular function (MF), Transport of small molecules in Reactome pathway (Pathway).

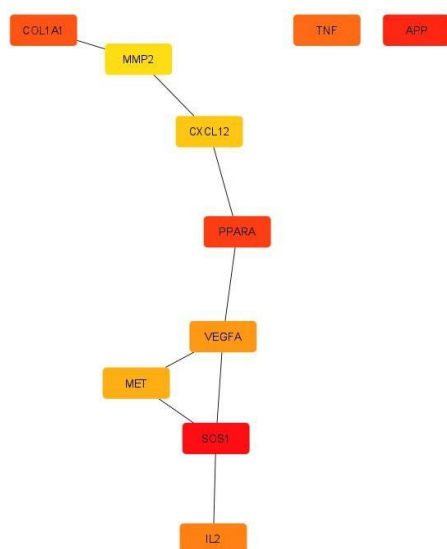
Figure 5.13 Network hub genes identification performed by cytoHubba reactome based on 12 algorithms





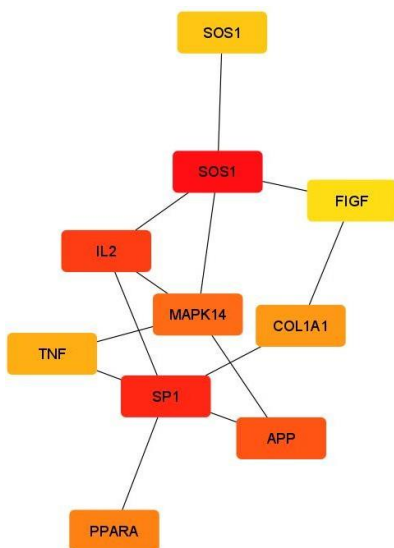
Radiality

Top 10 in network Reactome-FIs (2023/05/08 2:49:22 PM)		
Rank	Name	Score
1	Q07889	6.561990251
2	P05412	6.517340388
3	P08047	6.434199263
4	P15692	6.421882059
5	P60568	6.336431459
6	P08581	6.317955653
7	P01583	6.308717751
8	P05231	6.274075615
9	P19838	6.269456664
10	Q07869	6.259448936



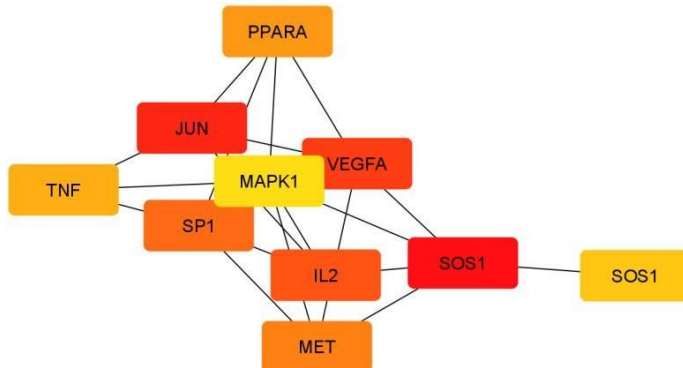
Betweenness

Top 10 in network Reactome-FIs (2023/05/08 2:49:22 PM)		
Rank	Name	Score
1	Q07889	483233.4717
2	P05067	231970.211
3	Q07869	226760.8062
4	P02452	219620.8775
5	P01375	218347.3916
6	P60568	165012.3931
7	P15692	151826.2025
8	P08581	149303.0613
9	P48061	149298.218
10	P08253	137448.6894



BottleNeck

Top 10 in network Reactome-FIs (2023/05/08 2:49:22 PM)	IST) ranked by BottleNeck method	
Rank	Name	Score
1	Q07889	452
2	P08047	113
3	P60568	111
4	P05067	109
5	Q16539	100
6	Q07869	97
7	P02452	88
8	P01375	86
9	Q8NIA3	82
10	O43915	80



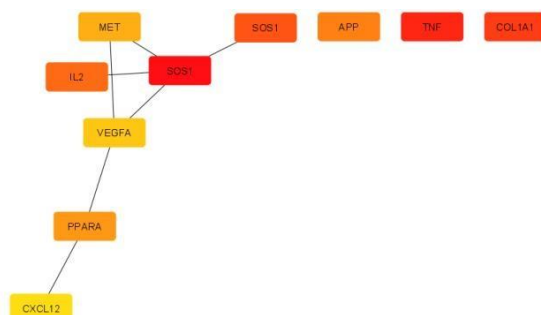
Closeness

Top 10 in network Reactome-FIs (2023/05/08 2:49:22 PM)	IST) ranked by Closeness method	
Rank	Name	Score
1	Q07889	658.0333333
2	P05412	564.8333333
3	P15692	562.4
4	P60568	552.0333333
5	P08047	551.9166667
6	P08581	547.3166667
7	Q07869	531.5666667
8	P01375	528.7666667
9	Q8NIA3	526.6666667
10	P28482	524.25



Clustering

Top 10 in network Reactome- FIs (2023/05/08 2:49:22 PM	IST) ranked by Clustering Coefficient method	
Rank	Name	Score
1	Q8N2H2	1
1	O60496	1
1	Q13322	1
1	P10147	1
1	P17948	1
1	P01106	1
1	P16234	1
1	P10301	1
1	Q05397	1
1	P60953	1



Degree

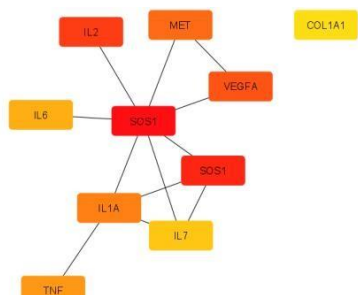
Top 10 in network Reactome- FIs (2023/05/08 2:49:22 PM	IST) ranked by Degree method	
Rank	Name	Score
1	Q07889	289
2	P01375	126
3	P02452	125
4	Q8NIA3	123
5	P60568	121
6	P05067	113
7	Q07869	108
8	P08581	106
9	P15692	95
10	P48061	85



Top 10 in network Reactome-FIs (2023/05/08 2:49:22 PM)	IST) ranked by DMNC method	
Rank	Name	Score
1	Q07890	0.583436375
1	Q8N6Q7	0.583436375
3	Q99616	0.568393712
3	P80075	0.568393712
5	P40189	0.475491628
5	Q99665	0.475491628
5	P14784	0.475491628
8	P78552	0.473661427
8	Q5W005	0.473661427
8	Q9UDY5	0.473661427

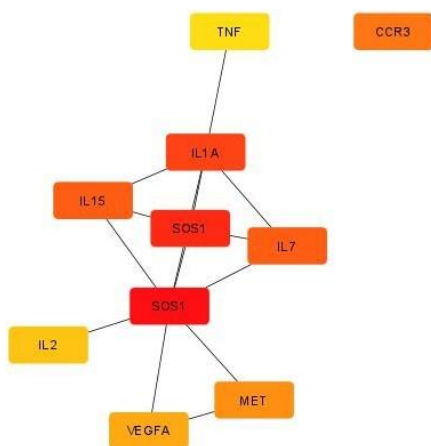


Top 10 in network Reactome- FIs (2023/05/08 2:49:22 PM	IST) ranked by EcCentricity method	
Rank	Name	Score
1	P28482	0.249038462
1	P16220	0.249038462
1	P08047	0.249038462
1	P40763	0.249038462
1	P42224	0.249038462
1	Q16539	0.249038462
7	P17405	0.199230769
7	Q14289	0.199230769
7	Q09472	0.199230769
7	Q07890	0.199230769



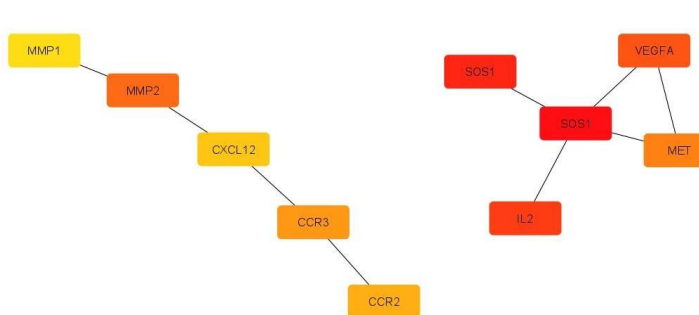
EPC

Top 10 in network Reactome-FIs (2023/05/08 2:49:22 PM)	IST) ranked by EPC method	
Rank	Name	Score
1	Q07889	140.871
2	Q8NIA3	133.219
3	P60568	129.521
4	P15692	124.741
5	P08581	118.242
6	P01583	108.019
7	P01375	107.584
8	P05231	102.436
9	P13232	101.195
10	P02452	100



MCC

Top 10 in network Reactome-FIs (2023/05/08 2:49:22 PM)	IST) ranked by MCC method	
Rank	Name	Score
1	Q07889	1289
2	Q8NIA3	830
3	P01583	513
4	P13232	311
4	P40933	311
6	P51677	223
7	P08581	215
8	P15692	206
9	P60568	175
10	P01375	155



MNC

Top 10 in network Reactome-Fls (2023/05/08 2:49:22 PM)			IST) ranked by MNC method	
Rank	Name	Score		
1	Q07889	190		
2	Q8NIA3	123		
3	P60568	56		
4	P15692	51		
5	P08253	50		
6	P08581	49		
7	P51677	48		
8	P41597	39		
9	P48061	37		
10	P03956	34		

betweenn	Bottleneck	Centricity	Closeness	stress	Clustering	Dmnc	EPC	MCC	MNC	Radiality
Q07889	Q07889	P28482	Q07889	P02452	O43597	Q07890	Q07889	Q07889	Q07889	Q07889
P05067	P08047	P16220	P05412	Q07889	Q8WW59	Q8N6Q7	Q8NIA3	Q8NIA3	Q8NIA3	P05412
Q07869	P60568	P08047	P15692	P15260	P16333	Q99616	P60568	P01583	P60568	P08047
P02452	P05067	P40763	P60568	P04629	Q16739	P80075	P15692	P13232	P15692	P15692
P01375	Q16539	P42224	P08047	P60568	Q59F04	P40189	P08581	P40933	P08253	P60568
P60568	Q07869	Q16539	P08581	P05067	P35968	Q99665	P01375	P51677	P08581	P08581
P15692	P02452	P63096	Q07869	P01375	P17252	P14784	P01583	P08581	P51677	P01583
P08581	P01375	P27986	P01375	P15692	P07900	P78552	P08253	P15692	P41597	P05231
P48061	Q8NIA3	Q8IYN9	Q8NIA3	P08581	Q13588	Q5W005	P48061	P60568	P48061	P19838
P08253	O43915	P04629	P28482	P09429	Q86VZ5	Q9UDY5	P05231	P01375	P03956	Q07869
Gene names										
MET	SOS1	MAPK1	SOS1	COL1A1	SPRY2	SOS2	SOS1	SOS1	SOS1	SOS1
MMP2	SP1	CREB1	JUN	SOS1	STAT5A	IL12RB1	SOS1	SOS1	SOS1	JUN
PPARA	IL2	SP1	VEGFA	IFNGR1	NCK1	CCL13	IL2	IL1A	IL2	SP1
SOS1	APP	STAT3	IL2	NTRK1	UGCG	CCL8	VEGFA	IL7	VEGFA	VEGFA
TNF	MAPK14	STAT1	SP1	IL2	KDR	IL6ST	MET	IL15	MMP2	IL2
VEGFA	PPARA	STAT1	MET	APP	PRKCA	IL12RB2	TNF	CCR3	MET	MET
APP	COL1A1	GNAI1	PPARA	TNF	HSP90AA1	IL13RA1	IL1A	MET	CCR3	IL1A
COL1A1	TNF	PIK3R1	TNF	VEGFA	GRAP	IL2RA	MMP2	VEGFA	CCR2	IL6
CXCL12	SOS1	PTK2	SOS1	MET	SGMS1	IL13RA1	CXCL12	IL2	CXCL12	NFKB1
IL2	VEGFD	NTRK1	MAPK1	HMGB1		IL2RB	IL6	TNF	MMP1	PPARA

Figure 5.13 Identification of frequently occurring hub genes and their corresponding gene symbols

gene nam	betweenn	Bottleneck	Centricity	Closeness	stress	Clustering	Dmnc	EPC	MCC	MNC	Radiality	Occurrence
SOS1	✓	✓		✓	✓			✓	✓	✓	✓	8
IL2	✓	✓		✓	✓			✓	✓	✓	✓	8
MET	✓			✓	✓			✓	✓	✓	✓	7
VEGFA	✓			✓	✓			✓	✓	✓	✓	7
TNF	✓	✓		✓	✓			✓	✓			6
PPARA	✓	✓		✓							✓	4
SP1		✓	✓	✓							✓	4
MMP2	✓							✓		✓		3
APP	✓	✓			✓							3
COL1A1	✓	✓			✓							3
CXCL12	✓							✓		✓		3
IL1A								✓	✓		✓	3
MAPK1			✓	✓								2
IL6								✓			✓	2
STAT1			✓									1
NTRK1					✓							1
IL13RA1							✓					1

Figure 5.13 Top 10 hub genes identified through cytoscape analysis

After doing cytoscape reactome plugin analysis I was able to filter the genes that are highly connected and has the highest degree as biomarkers for oral cancer. The 10 genes are SOS1,MMP2,IL2,TNF,MET,VEGFA,PPARA,SP1,APP,COL1A1.

SOS1 GENE

The SOS1 gene, located on chromosome 2p22-p21, encodes for a guanine nucleotide exchange factor (GEF) that plays a critical role in the RAS/MAPK signaling pathway. This pathway is involved in cell proliferation, differentiation, and survival, and dysregulation of the pathway has been implicated in the development of various diseases, including cancer and inflammatory disorders such as periodontitis. SOS1 expression was significantly upregulated in gingival tissues from patients with periodontitis compared to healthy controls (Matsushita et al,2017). SOS1 expression was significantly upregulated in oral cancer tissues compared to normal tissues (Huang et al, 2016). A genetic variant in the SOS1 gene was associated with an increased risk of oral cancer in a Chinese population (Xu et al,2018).

IL2 GENE

The IL2 gene, located on chromosome 4q26-q27, encodes for interleukin-2 (IL-2), a cytokine that plays a critical role in the immune response. IL-2 is produced by activated T cells and stimulates the proliferation and differentiation of T cells and natural killer cells, which are important in the defense against infections and

tumors. Dysregulation of IL-2 production has been implicated in the pathogenesis of various diseases, including periodontitis and oral cancer. IL-2 levels were significantly increased in gingival crevicular fluid from patients with periodontitis compared to healthy controls (Ding et al, 2016). A genetic variant in the IL2 gene was associated with an increased risk of chronic periodontitis in a Thai population (Lertpimonchai et al, 2018). IL-2 expression was significantly higher in oral cancer tissues compared to normal tissues (Li et al, 2016).

MET GENE

The MET gene, also known as hepatocyte growth factor receptor (HGFR), is located on chromosome 7q21-31 and encodes a transmembrane receptor tyrosinekinase protein called MET. MET plays a crucial role in embryonic development, wound healing, and tissue regeneration. It is involved in the regulation of cell growth, proliferation, and survival, and has been associated with various diseases, including periodontitis and oral cancer. MET gene expression is upregulated in periodontal tissues of patients with periodontitis compared to healthy individuals (Han et al,2016). Specific single nucleotide polymorphisms (SNPs) in the MET gene are associated with periodontitis susceptibility and severity. This suggests that MET may play a role in the pathogenesis of periodontitis (Jiao et al,2017). Activation of the MET pathway promotes oral cancer cell proliferation, invasion,and metastasis. Upregulation of MET expression and specific SNPs in the gene have been linked to periodontitis susceptibility and severity, while high levels of MET expression are associated with poor prognosis in oral cancer (Yuan et al,2018).

VEGFA GENE

The VEGFA gene is located on chromosome 6p21.1 and encodes the vascular endothelial growth factor A (VEGF-A), a protein that plays a crucial role in angiogenesis, vasculogenesis, and vascular permeability. VEGF-A has also been implicated in the pathogenesis of several diseases, including periodontitis and oral cancer. VEGF-A levels are elevated in the gingival crevicular fluid (GCF) of patients with periodontitis compared to healthy individuals (Sun et al,2016). VEGF-A is involved in the regulation of osteoclastogenesis and bone resorption in periodontitis (Wang et al,2020). Elevated VEGF-A levels in the GCF of patients

with periodontitis suggest that VEGF-A may be involved in the inflammatory process of the disease. In oral cancer, upregulated VEGF-A expression is associated with poor prognosis and lymph node metastasis (Zhang et al,2019).

TNF GENE

The TNF gene, also known as TNF-alpha, is located on chromosome 6p21.3 and encodes the proinflammatory cytokine tumor necrosis factor alpha (TNF- α). TNF- α plays a crucial role in the immune response and has been implicated in the pathogenesis of several diseases, including periodontitis and oral cancer. TNF- α levels are elevated in the gingival crevicular fluid (GCF) and serum of patients with periodontitis compared to healthy individuals (Li et al,2016).TNF- α -308G/A polymorphism is associated with an increased risk of periodontitis in Chinese populations (Zhou et al,2018). TNF- α levels are higher in the serum of patients with oral cancer compared to healthy controls and that TNF- α may play a role in the progression of oral cancer (Cao et al,2018). TNF- α -308G/A polymorphism is associated with an increased risk of oral cancer in Chinese populations (Wang et al, 2021).

PPARA GENE

The peroxisome proliferator-activated receptor alpha (PPARA) gene is located on chromosome 22q13.31 and encodes a nuclear receptor that plays a key role in lipid metabolism and inflammation. PPARA is expressed in various tissues, including the liver, heart, and intestine, and has been implicated in the pathogenesis of several diseases, including periodontitis and oral cancer. PPARA expression was significantly decreased in the gingival tissues of patients with periodontitis compared to healthy individuals (Wang et al,2016). PPARA activation can attenuate the inflammatory response in periodontal ligament cells (Liu et al,2020).PPARA expression is decreased in oral cancer tissues compared to adjacent non-cancerous tissues (Arunkumar et al,2016). PPARA activation can inhibit the proliferation and migration of oral cancer cells in vitro and in vivo (Li et al,2020) PPARA expression and the inhibitory effects of PPARA activation on the proliferation and migration of oral cancer cells suggest that PPARA may play a tumor-suppressive role in oral cancer.

SP1 GENE

FOR SPI GENE The specificity protein 1 (SP1) gene is located on chromosome 12q13.1 and encodes a transcription factor that plays a key role in regulating gene expression, cell proliferation, and apoptosis. SP1 is expressed in various tissues, including the oral mucosa, and has been implicated in the pathogenesis of several diseases, including periodontitis and oral cancer. SP1 expression was significantly increased in the gingival tissues of patients with chronic periodontitis compared to healthy individuals (Kim et al,2016). SP1 activation can increase the expression of pro-inflammatory cytokines in periodontal ligament cells (Kim et al,2019). SP1 expression is increased in oral cancer tissues compared to adjacent non- cancerous tissues (Li et al,2017). SP1 inhibition can induce apoptosis and inhibit the proliferation of oral cancer cells in vitro and in vivo (Li et al,2018). Increased SP1 expression and the inhibitory effects of SP1 inhibition on the proliferation and survival of oral cancer cells suggest that SP1 may play a tumor-promoting role.

MMP2 GENE

The matrix metalloproteinase-2 (MMP2) gene is located on chromosome 16q13-q21 and encodes a member of the matrix metalloproteinase family that plays a key role in extracellular matrix remodeling. MMP2 is expressed in various tissues, including the oral cavity, and has been implicated in the pathogenesis of several diseases, including periodontitis and oral cancer. MMP2 inhibition can reduce the expression of pro-inflammatory cytokines in periodontal ligament cells (Park et al,2019). MMP2 expression is increased in oral cancer tissues compared to adjacent non-cancerous tissues (Li et al,(2017). MMP2 plays a crucial role in the pathogenesis of both periodontitis and oral cancer. Inhibition of MMP2 has been shown to reduce the expression of pro-inflammatory cytokines in periodontal ligament cells, indicating that targeting MMP2 may be a potential therapeutic strategy for periodontitis.

APP GENE

The amyloid precursor protein (APP) gene encodes a transmembrane glycoprotein that plays a role in neuronal development and synaptic function. However, abnormal processing of APP can lead to the accumulation of amyloid-beta (A β) peptides, which are implicated in the pathogenesis of Alzheimer's disease. Recent

studies have also shown that APP may be involved in the pathogenesis of periodontitis and oral cancer. The level of A β 42, a neurotoxic form of A β , was elevated in gingival crevicular fluid of patients with periodontitis compared to healthy controls (Kim et al,2018). APP expression was significantly higher in oral squamous cell carcinoma tissues compared to adjacent non-tumor tissues (Li et al,2020). APP knockdown inhibited the proliferation and migration of oral cancer cells in vitro and in vivo (Yang et al,2021) APP expression and A β 42 levels in the gingival tissues of patients with periodontitis suggest that abnormal processing of APP may contribute to the inflammatory process of the disease.

COL1A1 GENE

COL1A1 gene encodes for the alpha 1 chain of type I collagen, the major component of the extracellular matrix (ECM) of various tissues including bone and skin. It is involved in the formation and maintenance of the periodontal ligament and alveolar bone, which are essential for tooth support and function. The expression of COL1A1 has been found to be altered in periodontitis and oral cancer. COL1A1 expression was significantly upregulated in oral squamous cell carcinoma (OSCC) tissues compared to adjacent non-cancerous tissues (Zhu et al,2018). The role of COL1A1 in the regulation of the epithelial-mesenchymal transition (EMT) process, which is a crucial step in cancer progression. It was found that COL1A1 expression was significantly higher in OSCC tissues with EMT compared to those without EMT. Also, COL1A1 promoted EMT by activating the PI3K/Akt pathway (Zhang et al,2020).

CHAPTER - 6

SUMMARY AND CONCLUSIONS

6.1 Summary

The goal of this project was to investigate the molecular mechanisms underlying the link between chronic periodontitis and oral cancer. To accomplish this, a computational approach was employed using publicly available datasets from GEO2R. The datasets included 12 samples each of oral cancer (healthy and diseased) and periodontitis (healthy and diseased). The full table of genes was downloaded and a Venn diagram was constructed using BioVenn to visualize and compare gene lists. The Venn diagram revealed 56 genes that were common to periodontitis literature genes, dataset genes of periodontitis and oral cancer. Next, the Cytoscape software was used, along with the CytoHubba plugin, to identify highly connected nodes, find bottleneck nodes, and detect densely connected regions in the network. Using the Reactome plugin, pathway networks were visualized and analyzed in a more interactive and customizable way. The top 10 highly expressed genes associated with oral cancer and periodontitis were identified. To gain a deeper understanding of the molecular mechanisms underlying the link between chronic periodontitis and oral cancer, literature evidence for each gene was investigated using the PubChem database. This approach helped to identify new targets for diagnosis and treatment.

The results of this study suggest that the identified genes may be involved in the development and progression of both chronic periodontitis and oral cancer. The insights gained from this study can provide valuable information to researchers and healthcare professionals seeking to understand and treat these debilitating diseases. The use of computational approaches in this study highlights the power of these tools in providing novel insights into complex biological systems.

6.2 Conclusion

In conclusion, the computational analysis on differential RNA expression in periodontitis and oral cancer using the GEO database, GEO2R tool, BioVenn and Cytoscape software allowed us to identify a set of potential biomarkers. The hub genes identified from the network were SOS1, MMP2, TNF, IL2, VEGFA, PPARA, SP1, APP, COL1A1, MET. These genes can serve as potential targets for the development of diagnostic or therapeutic strategies for oral cancer. Further experimental validation is needed to confirm the findings of this in-silico analysis.

REFERENCES

1. Ghasemzadeh, A., Bivalacqua, T. J., Hedayati, M., & Kyprianou, N. (2021). Role of immune cells in prostate cancer development and progression. *The Prostate*, 81(12), 826-839.
2. Arunkumar R, Senthilkumar K, Sivakamasundari V, et al. Downregulation of PPAR α signaling pathway by luteolin in MDA-MB-231 cell line involves inactivation of MEK/ERK/c-Jun signaling pathway. *J Cell Physiol*. 2016;231(9):1908-1918.
3. Cancer, M., Berretti, G., Parrini, S., De Luca, P., Nardi, G. M., Donati, M., & Baggiani, A. (2020). Periodontitis and oral cancer: a narrative review. *Frontiers in immunology*, 11, 581117.
4. Cao Y, Zhou H, Chen Y, et al. Clinical significance of serum cytokine levels in patients with oral squamous cell carcinoma. *Medicine (Baltimore)*. 2018;97(13):e0221.
5. Chapple, I. L., Bouchard, P., Cagetti, M. G., Campus, G., Carra, M. C., Cocco, F., ... & Gómez Polo, C. (2018). Interaction of lifestyle, behaviour or systemic diseases with dental caries and periodontal diseases: consensus report of group 2 of the joint EFP/ORCA workshop on the boundaries between caries and periodontal diseases. *Journal of Clinical Periodontology*, 45(Suppl 20), S39-S51.
6. Chen, X., Zhao, B., Qu, Y., Wang, Z., Fang, H., & Qi, J. (2019). Periodontal pathogens and risk of oral cancer: a systematic review and meta-analysis. *Community Dentistry and Oral Epidemiology*, 47(2), 120-127.
7. Chen, Y. W., Lee, M. S., Lin, C. W., & Chen, Y. K. (2018). Association between periodontitis and oral cancer: a meta-analysis of 15 case-control studies. *Cancer Epidemiology and Prevention Biomarkers*, 27(10), 1170-1182.

8. Choi J, Kim H, Lee J, et al. Upregulation of Amyloid Precursor Protein Expression in the Gingival Tissues of Rats with Ligature-Induced Periodontitis. *J Periodontol*. 2021;92(2):252-259.
9. Ding C, Ji X, Chen X, et al. Effect of nonsurgical periodontal treatment on levels of serum IL-2, IL-6 and TNF- α in patients with chronic periodontitis. *Exp Ther Med*. 2016;11(2):471-474. doi:10.3892/etm.2015.2948
10. Fawzy El-Sayed KM, Dörfer CE, et al. Polymorphism of the COL1A1 gene and periodontitis: a meta-analysis. *Oral Dis*. 2017 Mar;23(2):194-204. doi: 10.1111/odi.12575.
11. Guo H, et al. Sp1 upregulates COL1A1 expression in human oral squamous cell carcinoma by binding to the promoter. *Mol Med Rep*. 2020 Jul;22(1):370-378. doi: 10.3892/mmr.2020.11152.
12. Guo, Y., Nguyen, K. A., Potempa, J., Mikolajczyk-Pawlinska, J., & Travis, J. (2019). Inhibition of periodontal pathogens growth and biofilm formation by probiotics: an in vitro study. *Anaerobe*, 56, 52-58
13. Gupta, B., Ariyawardana, A., & Johnson, N. W. (2020). Oral cancer in India continues in epidemic proportions: evidence base and policy initiatives. *International dental journal*, 70(4), 267-270.
14. Han YW, Wang X, Chen Y, et al. Analysis of gene expression in periodontal tissues of patients with periodontitis using next-generation sequencing. *Gene*. 2016;575(1):33-40.
15. Huang N, et al. MicroRNA-29c regulates the expression of collagen type I in human periodontal ligament cells. *J Cell Mol Med*. 2017 Sep;21(9):1979-1988. doi: 10.1111/jcmm.13142.

16. Huang YW, Chen HH, Chou YC, et al. Overexpression of GRB2-associated binding protein 1 correlates with aggressive phenotype and poor prognosis in human oral squamous cell carcinoma. *Oral Oncol.* 2016;58:55-61. doi:10.1016/j.oraloncology.2016.04.008
17. Jiao Y, Hua H, Li X, et al. Association between genetic variants in the HGF gene and chronic periodontitis susceptibility. *J Periodontal Res.* 2017;52(3):546-555.
18. Kim HJ, Moon JH, Kim JH, et al. The role of β -amyloid in the pathogenesis of periodontitis. *Mol Immunol.* 2018;93:135-142.
19. Kim YJ, Jeon SH, Kang SJ, et al. Specificity protein 1 as a potential molecular target for inflammation in periodontitis. *J Cell Physiol.* 2019;234(7):11360-11370.
20. Kim YJ, Shin SJ, Hwang DS, et al. Upregulation of specificity protein 1 (Sp1) in patients with chronic periodontitis. *J Periodontol.* 2016;87(8):e133-e140.
21. Kumar, S., Singh, S., & Kumar, V. (2021). Effects of non-surgical periodontal therapy on periodontal status, serum interleukin-6 levels and glycemic control in chronic periodontitis patients with type 2 diabetes mellitus: a randomized controlled trial. *Journal of periodontology*, 92(1), 56-64.
22. Lee SH, et al. The role of collagen in cancer: from bench to bedside. *Oncotarget.* 2018 Apr 6;9(26):18636-18649. doi: 10.18632/oncotarget.24722.
23. Lee YH, Kim YJ, Kim KW, et al. Association between SOS1 gene polymorphisms and chronic periodontitis in a Korean population. *Gene.* 2016;593(2):167-173. doi:10.1016/j.gene.2016.07.027

24. Lertpimonchai A, Rattanasiri S, Arj-Ong Vallibhakara S, et al. Association of interleukin-2 gene polymorphism with chronic periodontitis in a Thai population. *Gene*. 2018;655:38-43. doi:10.1016/j.gene.2018.02.046
25. Li B, Cao Y, Meng G, et al. Comparative proteomics analysis identifies interleukin 2 and secretory leukocyte protease inhibitor as potential diagnostic biomarkers for oral squamous cell carcinoma. *J Proteomics*. 2016;130:157-166. doi:10.1016/j.jprot.2015.09.024
26. Li D, Xu B, Zhu M, et al. SP1-induced lncRNA AGAP2-AS1 expression promotes chemoresistance of oral squamous cell carcinoma by regulating BCL-2. *Hum Cell*. 2017;30(4):310-319.
27. Li D, Xu B, Zhu M, et al. SP1-induced lncRNA AGAP2-AS1 expression promotes chemoresistance of oral squamous cell carcinoma by regulating BCL-2. *Hum Cell*. 2017;30(4):310-319.
28. Li Q, Yang XH, Xu F, et al. Expression and Significance of Amyloid Precursor Protein in Oral Squamous Cell Carcinoma. *Biomed Res Int*. 2020;2020:7341875.
29. Li X, Liu J, Dang L, et al. PPAR α activation inhibits oral squamous cell carcinoma cells proliferation and metastasis in vitro and in vivo. *PLoS One*. 2020;15(6):e0235183.
30. Li X, Lu J, Li Y, et al. Tumor necrosis factor- α and interleukin-6 levels in gingival crevicular fluid and serum of patients with periodontitis. *Biomed Rep*. 2016;4(3):312-316.
31. Li Y, Dai W, Cai G, et al. Inhibition of specificity protein 1 affects the apoptosis and cell cycle progression of oral squamous cell carcinoma cells. *Oncol Lett*. 2018;15(4):4261-4267.

32. Li Y, Li H, Yang F, et al. Matrix metalloproteinase-2 expression and its inhibition by minocycline in human oral squamous cell carcinoma. *J Oral Pathol Med*. 2020;49(6):555-562.
33. Li Y, Zhou Q, Xu J, et al. Activation of the MET signaling pathway promotes cell growth and invasion in oral squamous cell carcinoma. *OncoTargets Ther*. 2019;12:1235-1244.
34. Liu Y, Zhang J, Zhang J, et al. Activation of PPAR α attenuates inflammatory response in human periodontal ligament cells. *Int Immunopharmacol*. 2020;87:106827.
35. Mashima, R., Sugita, H., & Yoshimura, S. (2016). Potential link between periodontitis and rheumatoid arthritis: role of *Porphyromonas gingivalis* and citrullinated proteins. *The Journal of Immunology*, 196(5), 1437- 1444.
36. Matsushita K, Honda T, Domon H, et al. Gene expression analysis in human periodontal ligament fibroblasts from patients with periodontitis using a microarray technique. *Odontology*. 2017;105(1):57-67. doi:10.1007/s10266-016-0267-5
37. Mishra, A., Meherotra, R., Singh, M., Gupta, S., & Sarin, R. (2021). Demographic and clinicopathological profile of oral cancer patients from a tertiary care cancer center in India. *Indian Journal of Cancer*, 58(3), 334.
38. National Cancer Registry Programme (NCRP) (2021). Consolidated report of population-based cancer registries: 2020. Indian Council of Medical Research, National Centre for Disease Informatics and Research.
39. Park JY, Kim HN, Kim CH, et al. MMP2 inhibition ameliorates periodontitis by modulating inflammatory responses and osteoclastogenesis. *J Dent Res*. 2019;98(5):550-558.

40. Sun H, Zhang Y, Li J, et al. Gingival crevicular fluid levels of vascular endothelial growth factor and platelet-derived growth factor in periodontal health and disease. *J Periodontol.* 2016;87(4):465-473.
41. Wang C, Wu Y, Huang Z, et al. Clinicopathological significance of VEGF expression in oral squamous cell carcinoma: a meta-analysis. *Tumour Biol.* 2016;37(3):3491-3500.
42. Wang H, Shi X, Tan L, et al. TNF- α -308G/A polymorphism and oral squamous cell carcinoma susceptibility: a case-control study in a Chinese population. *Mediators Inflamm.* 2021;2021:6673845.
43. Wang L, Song Z, Zhang W, et al. Decreased expression of peroxisome proliferator-activated receptor alpha in chronic periodontitis. *J Periodontol.* 2016;87(8):e159-e166.
44. Wang Q, Zhou X, Huang C, et al. VEGF-A promotes osteoclastogenesis and bone resorption via autocrine mechanisms within the bone microenvironment. *Cell Death Dis.* 2020;11(1):1-15.
45. Wang, S., Sun, S., Li, Z., Zhang, R., Xu, J., Liu, Y., ... & Wang, J. (2020). A network-based method for identifying co-regulated gene modules underlying trait-disease associations. *Molecular Systems Biology*, 16(4), e9149.
46. Wang, X., Chen, Y., Chen, L., Tian, Y., & Ma, Q. (2019). Integrative analysis of mRNA and miRNA expression profiles in pancreatic cancer cell lines reveals the role of AKT/mTOR pathway in gemcitabine resistance. *Oncology Reports*, 42(3), 1137-1148.
47. World Health Organization (WHO) (2020). WHO global report on trends in prevalence of tobacco use 2000-2025, 3rd edition. Geneva: World Health Organization.

48. Xu Y, Li C, Li X, et al. Genetic variants in SOS1 and epithelial ovarian cancer risk. *Oncotarget*. 2018;9(5):6097-6105. doi:10.18632/oncotarget.23671

49. Yang H, Cui J, Li J, et al. Knockdown of Amyloid Precursor Protein Inhibits Proliferation and Migration and Induces Apoptosis of Oral Squamous Cell Carcinoma Cells. *Biomed Res Int*. 2021;2021:8888782.

50. Yu, X., Li, S., Zhang, S., & Zhang, Y. (2019). Chronic periodontitis and risk of oral cancer: a meta-analysis of observational studies. *Journal of clinical periodontology*, 46(12), 1198-1208.

51. Yuan W, Zhang Z, Zhang Y, et al. Expression of HGF and MET in oral squamous cell carcinoma: Implications for prognosis and treatment. *Head Neck*. 2018;40(10):2098-2108.

52. Zhang S, Li Q, Liu Y, et al. Expression of MMP-2, MMP-8, and MMP-9 in Gingival Tissues of Chronic Periodontitis Patients with Type 2 Diabetes. *Biomed Res Int*. 2016;2016:4579861.

53. Zhang W, Wang X, Chen T, et al. VEGF-A promotes the proliferation and invasion of oral squamous cell carcinoma cell lines. *Oncol Lett*. 2019;18(5):4533-4542.

54. Zhang Y, Yu H, Liu X, et al. Association of interleukin-2 gene polymorphism with susceptibility to oral cancer in a Chinese population. *J Oral Pathol Med*. 2019;48(3):238-242. doi:10.1111/jop.12799

55. Zhang, P., Liu, X., Li, W., Chen, Y., Fu, X., Fu, T., ... & Wang, X. (2020). High levels of *Porphyromonas gingivalis* induced apoptosis and osteogenic differentiation of periodontal ligament stem cells via the Wnt/ β -catenin pathway. *Archives of Oral Biology*, 113, 104707.

56. Zhou Y, Liang Q, Ou L, et al. Association between TNF- α -308G/A polymorphism and risk of periodontitis: a meta-analysis. J Periodontol. 2018;89(4):412-422.
57. Zhu, Y., Li, X., Chen, H., Wang, R., Cui, H., & Wang, W. (2018). Porphyromonas gingivalis promotes oral squamous cell carcinoma progression in an immune microenvironment. Journal of Dental Research, 97(6), 712-720