UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ODONTOLOGIA DE PIRACICABA

THAYNAN ESCARIÃO DA NOBREGA

IN SILICO ANALYSES OF GENETIC VARIANTS ASSOCIATED WITH PERIODONTAL DISEASES LOCATED IN CANDIDATE *CIS*-REGULATORY ELEMENTS

ANÁLISES *IN SILICO* DE VARIANTES GENÉTICAS ASSOCIADAS ÀS DOENÇAS PERIODONTAIS LOCALIZADAS EM ELEMENTOS *CIS*-REGULATÓRIOS CANDIDATOS

Piracicaba – Sao Paulo State – Brazil

2023

THAYNAN ESCARIÃO DA NOBREGA

IN SILICO ANALYSES OF GENETIC VARIANTS ASSOCIATED WITH PERIODONTAL DISEASES LOCATED IN CANDIDATE *CIS*-REGULATORY ELEMENTS

ANÁLISES *IN SILICO* DE VARIANTES GENÉTICAS ASSOCIADAS ÀS DOENÇAS PERIODONTAIS LOCALIZADAS EM ELEMENTOS *CIS*-REGULATÓRIOS CANDIDATOS

Dissertation presented to the Piracicaba Dental School of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Oral Biology, in Histology and Embryology area.

Dissertação apresentada à Faculdade de Odontologia de Piracicaba da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Biologia Buco-Dental, na Área de Histologia e Embriologia.

Orientador: Prof. Dr. Marcelo Rocha Marques.

Este exemplar corresponde à versão final da Dissertação defendida pelo aluno Thaynan Escarião da Nobrega, e orientado pelo Prof. Dr. Marcelo Rocha Marques.

Piracicaba – Sao Paulo State – Brazil

2023

CATALOG CARD

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Odontologia de Piracicaba
Marilene Girello - CRB 8/6159

N669i

Nobrega, Thaynan Escarião da, 1993-
In silico analyses of genetic variants associated with periodontal diseases located in candidate *cis*-regulatory elements / Thaynan Escarião da Nobrega. – Piracicaba, SP : [s.n.], 2023.

Orientador: Marcelo Rocha Marques.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Odontologia de Piracicaba.

1. Doenças periodontais. 2. Estudo de associação genômica ampla. 3. Variação genética. 4. Polimorfismo de nucleotídeo único. 5. Regulação da expressão gênica. I. Marques, Marcelo Rocha, 1976-. II. Universidade Estadual de Campinas. Faculdade de Odontologia de Piracicaba. III. Título.

Informações Complementares

Título em outro idioma: Análises *in silico* de variantes genéticas associadas às doenças periodontais localizadas em elementos *cis*-regulatórios candidatos
Palavras-chave em inglês:
Periodontal diseases
Genome-wide association study
Genetic variation
Polymorphism, single nucleotide
Gene expression regulation
Área de concentração: Histologia e Embriologia
Titulação: Mestre em Biologia Buco-Dental
Banca examinadora:
Marcelo Rocha Marques [Orientador]
Marco Antonio Dias da Silva
Sergio Roberto Peres Line
Data de defesa: 19-07-2023
Programa de Pós-Graduação: Biologia Buco-Dental

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-3362-4827
- Currículo Lattes do autor: http://lattes.cnpq.br/6107891454462285

APPROVEMENT SHEET

UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Odontologia de Piracicaba

A Comissão Julgadora dos trabalhos de Defesa de Dissertação de Mestrado, em sessão pública realizada em 19 de julho de 2023, considerou o candidato THAYNAN ESCARIÃO DA NOBREGA aprovado.

PROF. DR. MARCELO ROCHA MARQUES

PROF. DR. MARCO ANTONIO DIAS DA SILVA

PROF. DR. SERGIO ROBERTO PERES LINE

A Ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

DEDICATORY

I dedicate this work for those who believe that science must be accessible for everyone, as well as education, which is our strongest weapon against human ignorance.

ACKNOWLEDGMENT

My Family, as I always say, my biggest laboratory on Earth. That is the place where I have been learning in how to be a better person and it have been shaping me throughout my life. Thanks to my mother, grandmother, uncles, aunts, cousins, and for those that feels family in person or spiritually.

Through this journey, there are several friends that supported me, and I met in different scenarios in my life. I would like to thank them all for connecting with me in that specific moment, and for stay still.

The new friends that I made, or already have, both keep helping me to create roots in Piracicaba during the past two years. Home is where you live!

To Prof. *Marcelo R. Marques* for strongly guiding me through our projects during my master's degree. You have advised me in how to be more patient and you have strongly contributed to my intellectual growth in molecular biology with your exceptional ideas.

To all the Professors that I had classes at *FOP-UNICAMP* Prof. *Ana P. Souza*, Prof. *Sérgio R.P. Line*, Prof. *Ricardo D. Coletta*, Prof. *Jacks J. Junior*. Also, to Prof. *Renato C.V. Casarin* and Prof. *Juliana S. Neves* for their contributions on my qualification exam. Finally, Prof. *Andresa C. Pereira* and Prof. *Marco A.D. Silva* for fomenting my base.

I would like to thank you *Arthur A.C. Bezerra* for rendering the graphical abstract.

To the *Serviço de Apoio ao Estudante (SAE)* for assisting me at the end of my master's degree. *Vanilda S. Santos* thank you for helping me through the process.

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES)* – Finance Code 001.

ABSTRACT

Introduction: Periodontal diseases (PDs) can be defined as multifactorial inflammation of the periodontal tissues that can result in tooth loss due to supporting periodontium destruction. Genetic factors have been associated with PDs pathogenesis. However, little is known about how genetic variants (GVs) regulate certain genes that participate in PDs evolution process. GVs can alter regions of cis-regulatory elements (CREs) and these directly interfere in gene expression. **Objective:** To quantify GVs in non-coding region associated with PDs in sequences of candidate CREs. **Material and Methods:** Six selected traits namely periodontitis, aggressive periodontitis, periodontal measurement, periodontal pocket, gingival bleeding, and gingival disease were used to search GVs associated with PDs on NHGRI-EBI GWAS catalog. Also, an integrative review was performed on PubMed applying the keywords "(periodontitis or periodontal or periodontal disease or gingival or gingival disease or gingival bleeding) and (GWAS)". Subsequently, a qualitative analysis was carried out on Genome Browser (GB) to categorize GVs found in the searches according to their location having as a reference both protein-coding (PC) and non-coding RNA (NCR) genes. GVs located in non-coding region were analyzed for the presence of regulatory elements on GB using ENCODE cCRES and GeneHancer databases. Furthermore, an interaction evaluation between regulatory elements and probable target genes was made, as well as tissue expression of these genes. Finally, JASPAR TF binding site database was used to predicted binding sites for TF and Tomtom motif comparison tool to quantify motif alteration of these binding sites. **Results:** 148 articles were read entirety to verify the eligibility according to pre-established criteria. 82 articles were included for further analysis and 66 articles were excluded. 788 GVs were obtained from eligible articles and *in silico* analyses were performed on GB. More than 10% of the GVs were in exonic region. For PC genes it was found that 30 GVs were in regulatory regions by both regulatory databases. Three GVs in intronic region of NCR gene were in a candidate regulatory element. Genes that were previously described to be involved with PDs pathogenesis, such as *IL37* (PC) and *CDKN2B-AS1* (NCR), were predicted in interaction analysis conducted between CREs and genes. Motif analysis of rs904310 (G>A) which is associated with NFYB TF binding site revealed better affinity for alternate allele A ($p = 6.45x10^{-4}$) than G ($p = 3.72x10^{-3}$). This intronic variation is related to both *SNRPN* (PC) and *ENSG00000286110* (NCR) genes. It is expected that functional validation assays can confirm whether the presence of these GVs

regulate expression of predicted genes. **Conclusion:** It is concluded there are 30 GVs in non-coding region associated with PDs in sequences of candidate CREs.

Keywords: Periodontal Diseases; Genome-Wide Association Study; Genetic Variation; Polymorphism, Single Nucleotide; Gene Expression Regulation.

RESUMO

Introdução: As doenças periodontais (DPs) podem ser definidas como inflamações de origem multifatorial dos tecidos periodontais que pode resultar na perda dentária devido a destruição do periodonto de suporte. Fatores genéticos têm sido associados à patogênese das DPs. No entanto, pouco é sabido sobre como variantes genéticas (VGs) regulam determinados genes que participam do processo de evolução das DPs. VGs podem alterar regiões de elementos *cis*-regulatórios (ECRs) e estes interferem diretamente na expressão gênica**. Objetivo:** Quantificar VGs em regiões não codificantes associadas às DPs em sequências de ECRs candidatos. **Material e Métodos:** Seis *traits* foram selecionados *periodontitis, aggressive periodontitis, periodontal measurement, periodontal pocket, gingival bleeding,* e *gingival disease* para averiguar VGs associadas às DPs no catálogo *NHGRI-EBI GWAS*. Além disto, uma revisão integrativa foi realizada no *PubMed* aplicando as palavras-chave "(*periodontitis or periodontal or periodontal disease or gingival or gingival disease or gingival bleeding) and (GWAS)"*. Posteriormente, uma análise qualitativa foi realizada no *Genome Browser* (GB) para categorizar as VGs das buscas de acordo com sua localização tendo como referência genes codificantes de proteína (CP) e de RNA não codificante (RNC). As VGs localizadas em região não codificante foram analisadas quanto à presença de elementos regulatórios no GB usando as bases de dados *ENCODE cCRES* e *GeneHancer*. Foi efetuada uma avaliação da interação entre elementos regulatórios e prováveis genes alvo, bem como a expressão tecidual destes genes. Por último, a base de dados para sítios de ligação de FTs JASPAR foi utilizada para predizer sítios de FTs e a ferramenta de comparação de motif Tomtom para quantificar alteração de motif destes sítios. **Resultados:** 148 artigos foram lidos na íntegra para verificar a elegibilidade de acordo com critérios pré-estabelecidos. 82 artigos foram incluídos e 66 artigos foram excluídos. 788 VGs foram obtidas de artigos elegíveis e análises *in silico* foram realizadas no GB. Mais de 10% das VGs estavam em região exônicas. Para genes CP observou-se que 30 VGs estavam em região regulatória em ambas as bases de dados. Três VGs em região intrônica de gene RNC estavam localizadas em elemento regulatório candidato. Genes que foram previamente descritos associados com patogênese da DPs, como *IL37* (CP) e *CDKN2B-AS1* (RNC) foram preditos na análise de interação conduzida entre ECRs e genes. Análise de motif do rs904310 (G>A), que está associado com o sítio de ligação do FT NFYB, revelou melhor afinidade pelo alelo alternativo A ($p = 6.45x10^{-4}$) que o G ($p = 3.72x10^{-3}$). Esta variante intrônica está associada com os genes *SNRPN* (CP) e *ENSG00000286110* (RNC). Espera-se que ensaios de validação

funcional possam confirmar se a presença destas VGs regula a expressão dos genes preditos. **Conclusão:** Conclui-se que existem 30 VGs em regiões não codificantes associadas às DPs em sequências de ECRs candidatos.

Palavras-chave: Doenças periodontais; Estudo de Associação Genômica Ampla; Variação Genética; Polimorfismo de Nucleotídeo Único; Regulação da Expressão Gênica.

ILLUSTRATION LIST

TABLE LIST

ABBREVIATIONS AND ACRONYMS LIST

AgP – Aggressive periodontitis

bp – Base Pair

CAD – Coronary Artery Disease

cCRES – Registry of Candidate *cis*-Regulatory Elements

CDKN2B-AS1 – Cyclin Dependent Kinase Inhibitor 2B antisense RNA 1

CP – Chronic periodontitis

CpG – Cytosine-phosphate-Guanine

CRE – C*is*-regulatory element

CTCF – CCCTC binding factor

dbSNP – Database of Single Nucleotide Polymorphisms

dbVar – Database of Structural Variation

DHSs – DNase I Hypersensitive Sites

DNA – Deoxyribonucleic Acid

EBI – European Bioinformatics Institute

ENCODE – Encyclopedia of DNA Elements

eRNA – Enhancer Ribonucleic Acid

H3K4me1 – Histone H3 Lysine 4 Mono-methylation

H3K4me3 – Histone H3 Lysine 4 Tri-methylation

H3K27ac – Histone H3 Lysine 27 Acetylation

H3K27me3 – Histone H3 Lysine 27 Tri-methylation

GV – Genetic Variant

GWAS – Genome Wide Association Studies

Indel – Insertion/Deletion

LD – Linkage Disequilibrium

lncRNA – long non-coding RNA

MAF – Minor Allele Frequency

mRNA – Messenger Ribonucleic Acid

miRNA – Micro Ribonucleic Acid

NCBI – National Center for Biotechnology Information

NCR – Non-coding Ribonucleic Acid

NHGRI – National Human Genome Research Institute

PC – Protein-coding

PDs – Periodontal Diseases

P. gingivalis – *Porphyromonas gingivalis*

PWS – Prader-Willi Syndrome

RNA – Ribonucleic Acid

RNAPII – Ribonucleic Acid Polymerase II

SNRPN – *Small Nuclear Ribonucleoprotein Polypeptide N*

SPD – Severe Periodontal Disease

SNP – Single Nucleotide Polymorphism

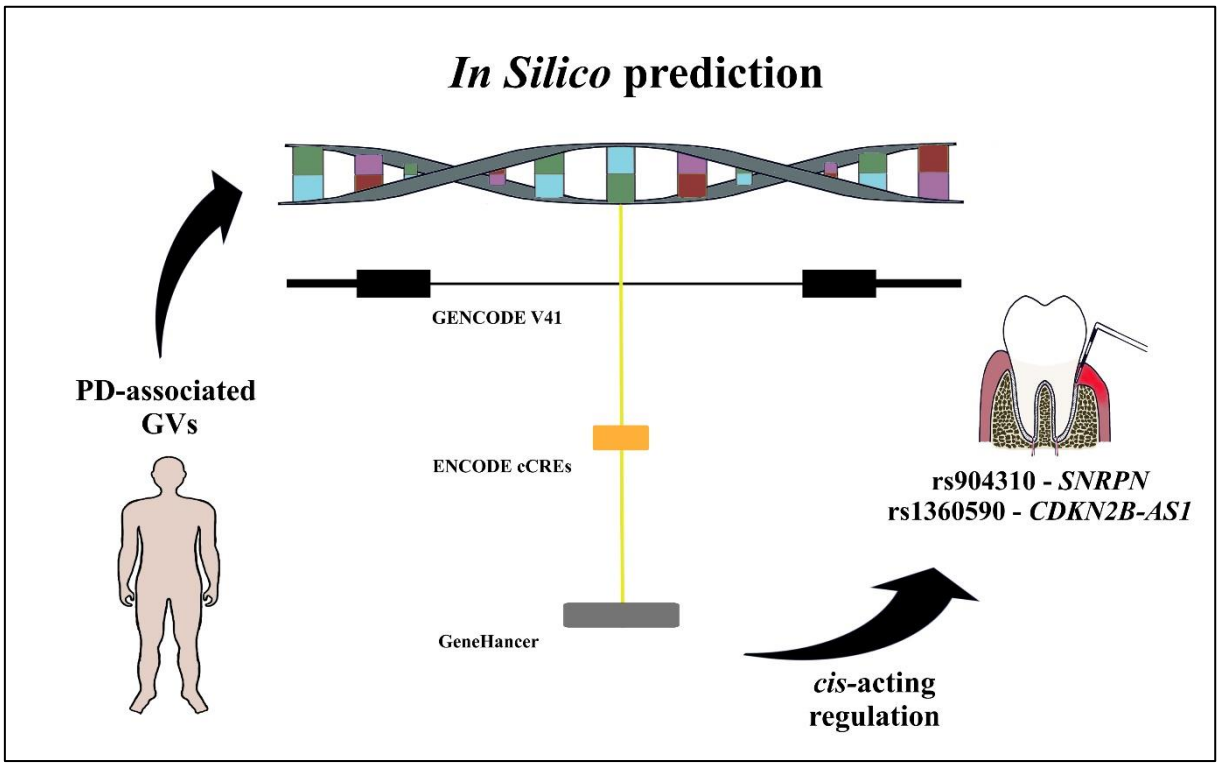
SNV – Single Nucleotide Variant

TF – Transcription Factor

TSS – Transcription Start Site

Summary

Graphical Abstract



Highlights

- An integrative review on PubMed and on NHGRI-EBI GWAS catalog detected 788 Genetic Variants associated with periodontal diseases
- *In silico* analyses found 30 GVs located in both non-coding region of genes and regulatory region
- Interaction analysis predicted regulatory regions modulating *SNRPN* and *CDKN2B-AS1* expression

1 Introduction

It is well known that disturbances caused in homeostasis between periodontal tissues, oral microbiota and hosts immune system lead to inflammation of these tissues (Marchesan, 2020; Silva et al., 2021). Periodontitis, one of the Periodontal Diseases (PDs), is a non-resolving chronic inflammation of periodontal tissues that in late stages can results in tooth loss due to destruction of supporting periodontium (Loos and Dyke, 2020).

Currently, periodontitis classification is based on severity and complexity of management (stages I-IV) and according to disease progression (grades A-C) (Caton et al., 2018). The worst scenario described is related to stage III/IV with rapid progression (grade C) (Coo et al., 2021).

Periodontitis manifestations have different causes being the sum of individuals genotype and environmental factors. Severe stages with rapidly progression have greater heritability, which is higher chances of the genetic component affecting the disease in a population (Nibali et al., 2019). PDs do not have a causal factor because they are multifactorial diseases. However, a better understanding of the genetic factors can improve individual treatment based on a precise medicine (Nashef et al., 2020; Schaefer, 2022).

There are ways to infer how individuals genotype may be related to a moderate or severe disease phenotype. Candidate gene studies, interactions between gene and environmental risk factors, animal models and Genome Wide Association Studies (GWAS) are some examples (Schaefer, 2022).

Among these models used to identify the weight of genetic in PDs the GWAS is highlighted. This type of study inquires how a Genetic Variant (GV) in a population can influence phenotypic variation and disease risk (Sollis et al., 2022). After study design and quality control, statistical analyses are performed to investigate the association between GV and phenotype with a significance level of $<5 \times 10^{-8}$ (Wang et al., 2019).

The data set generated by the GWAS is organized in a catalog and the NHGRI-EBI GWAS catalog is the most important platform in the area. Also, it is the repository recommended by the UK Biobank and the journals of Nature (Sollis et al., 2022). On this platform, it is possible to identify causal variants, understand the disease mechanisms and establish targets to the development of new therapies (Buniello et al., 2019).

The genotyping of the population analyzed in GWAS reveals the main GVs associated with the phenotype in question. GVs are alterations that occur in the genome sequence modifying base pairs type and quantity. These can be classified into structural variants or short genetic variants (1000 Genomes Project Consortium et al., 2015).

Depending on GV location, it can significantly modulate gene expression and consequently the phenotype when a causal relationship exists (Zhu et al., 2016). For this reason, GVs located in exonic regions are widely studied because these regions are responsible for coding protein.

Although intronic and intergenic regions do not code protein, GVs located in these areas can modify a regulatory region which plays an essential role in gene expression. These GVs may influence disease risk through regulatory mechanisms that affect gene expression or messenger Ribonucleic Acid (mRNA) splicing (Mountjoy et al., 2021).

In the human genome, regulatory regions control if transcripts will be synthesized by the Ribonucleic Acid Polymerase II (RNAPII) which associated with general Transcription Factors (TFs) form the transcription pre-initiation complex (Andersson and Sandelin, 2020). These regions are composed of regulatory elements which can be exemplified as promoters, insulators, and enhancers (1000 Genomes Project Consortium et al., 2015).

The Encyclopedia of DNA Elements (ENCODE) consortium aims to delineate specific segments of human genome that have a biological function. Specifically, the registry of candidate *cis*-regulatory elements (ENCODE cCRES) seeks to identify candidate CREs emphasizing promoters and enhancers (ENCODE Project Consortium et al., 2020a). Phase III project has nearly one million cCRES annotations constituting an important resource for scientific community (ENCODE Project Consortium et al., 2020b).

CREs are base pair sequences in the Deoxyribonucleic Acid (DNA) that regulate genes on the same chromosome they are located. These elements may have a promoter activity when they recruit RNAPII and transcribe an mRNA molecule; also, they have enhancer activity as they positively influence transcription initiation of another regulatory element through Enhancer Ribonucleic Acid (eRNA) molecule (Andersson and Sandelin, 2020). Another example is insulator that can be repressed by CCCTC binding factor (CTCF), which together with cohesin is involved with the three-dimensional architecture of DNA creating loops that mediates chromatin insulation and enhancer-promoter interaction (Rowley and Corces, 2018; Song et al., 2022).

These elements act in a coordinated manner, on accessible chromatin, regulating gene expression. Therefore, no matter how small changes in the sequences of these elements are, they have a direct impact on transcription process of the most varied RNAs (Klemm et al., 2019). In addition, the nuances of gene regulation by CRE are responsible for phenotypic variation and disease risk (Zaugg et al., 2022).

A recent study demonstrated that Single Nucleotide Polymorphism (SNP) rs55705857 is the causal variant of glioma. Interestingly, this polymorphism resides in an enhancer and the

change in the motif of the TF OCT2/4 was responsible for increasing the enhancer interaction with *MYC* promoter. *MYC* is associated with carcinogenesis. Thus, it becomes relevant to study how GVs can alter enhancer function and regulate the expression of disease-causing genes (Yanchus et al., 2022).

There are SNPs in enhancers associated with disease genes (Nasser et al., 2021). Moreover, enhancer-promoter affinity could be controlled by specific TFs which can have their binding site motifs modify by GV such as SNPs (Bergman et al., 2022). In this way, it is a plausible biological hypothesis that a single base variation at the human genome sequence might propel changes on motifs of TF biding sites, enhancer-promoter compatibility, transcription, and disease risk (Arensbergen et al., 2019; Carrasco et al., 2020).

2 Literature Review

2.1 Periodontal Diseases

The periodontium comprises the tissues that surround the teeth, being the gingival component formed by oral epithelium with the underlying lamina propria and the supporting periodontium by the triad cementum, periodontal ligament and alveolar bone (Nanci, 2019).

Etiopathogenesis of periodontal diseases has oral microbiome as an initial trigger, followed by host immune system response at periodontal tissues (Divaris et al., 2013; Marchesan, 2020). The diseases complexity is also due to polygenic interference and covariants such as smoker status (Freitag-Wolf et al., 2019; Nibali et al., 2019). PD can be an inflammation in the gingival component (gingivitis) or a supporting periodontium destructive form (periodontitis).

The Global Burden of Disease 2015 study estimated the Severe Periodontal Disease (SPD) affects more than 538 million people across 195 countries. Data comparison from 1990 to 2015 attest that Tropical Latin America was the only region with high prevalence and incidence of SPD as compared with global average in 2015 (Kassebaum et al., 2017).

In the past, periodontitis was classified in Chronic Periodontitis (CP) and Aggressive Periodontitis (AgP). While the former has slow progression of the disease and is associated with adults, the latter has rapid progression of periodontal tissue destruction in young individuals (Freitag-Wolf et al., 2014).

Additionally, periodontitis could be localized or generalized. Although both classifications must have at least 4mm of probing pocket depth according to the American Academy of Periodontology, generalized periodontitis has more than 30% of the sites with the disease (Armitage et al., 2004).

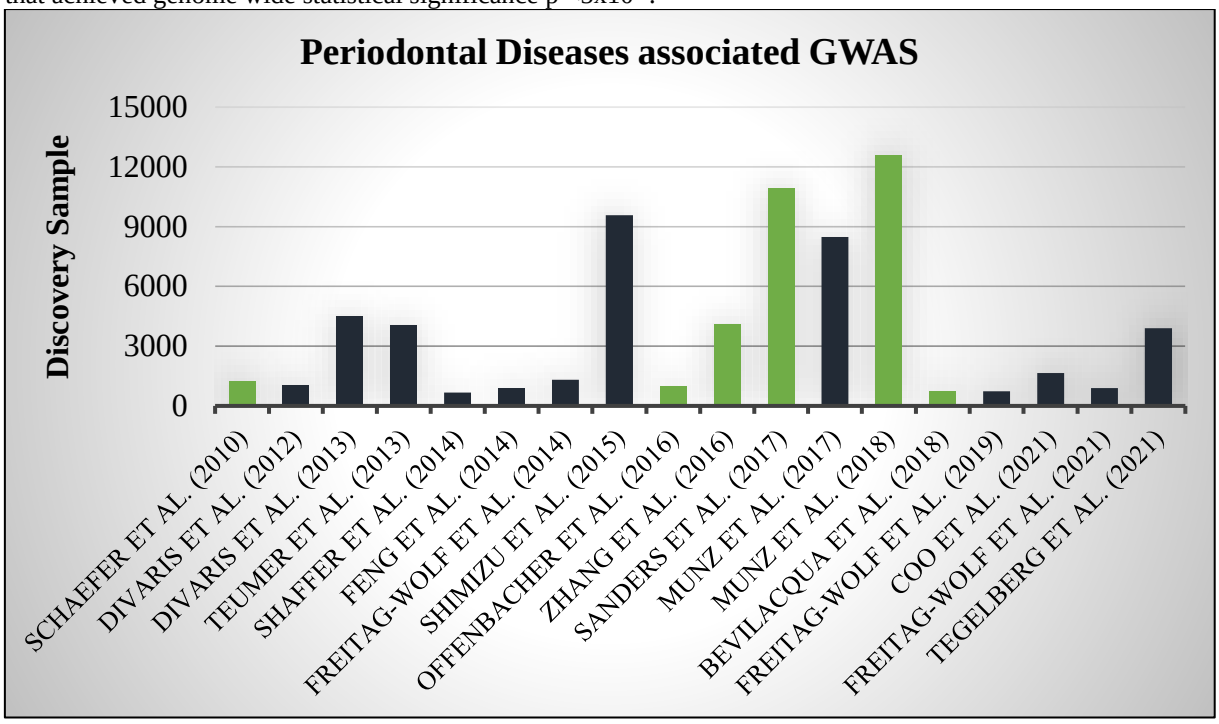
Currently, periodontitis is categorized in stages I-IV according to severity and complexity of management. Also, a second classification consider the disease progression ranging from grades A to C (Caton et al., 2018). Notably, young individuals with severe form (stage III/IV) and rapid progression (grade C) might have a similar phenotype to AgP in the old classification (Coo et al., 2021).

PDs are multifactorial complex systemic diseases affected by genetic variants. Löe et al (1986) narrates a natural history of periodontal diseases in a 15-years longitudinal study (1970-1985). The authors observed three distinct phenotypes based on individuals periodontitis progression in rapid (8.0%), moderate (81.0%) and no progression (11.0%). This study emphasizes the variety of pathways that individuals genotype can influence such disease risk.

Due to PD etiopathogenesis, candidate genes associated with immune response, microbial colonization and the epithelial barrier function are known as the probable biological mechanisms associating candidate genes to PDs risk (Marchesan, 2020). Nevertheless, in the last 14 years, GWAS related to PDs have been founding GV nearby genes there are not related to any of these routes.

There are 18 periodontal diseases associated GWAS from 2010-2021 and only six studies reached genome wide statistical significance p $<5x10^{-8}$ (Figure 1). These six studies found three associations related to periodontitis trait (Schaeffer et al., 2010; Bevilacqua et al., 2018; Munz et al., 2018), twelve to periodontal measurement (Offenbacher et al., 2016; Sanders et al., 2017) and one to gingival bleeding (Zhang et al., 2016).

Figure 1 – The 18 periodontal diseases associated GWAS that had at least one association. In green, the studies that achieved genome wide statistical significance p $<5x10^{-8}$.



Source: Own authorship.

2.2 Genetic Variant

Individuals of the same species have differences in their DNA sequence and that occurrence is the genetic diversity (Ellegren and Galtier, 2016). For instance, human genome differs roughly 0.1% (1000 Genomes Project Consortium et al., 2015). DNA replication errors and damage caused by UV radiation propel new alleles appearance through mutation (Ellegren and Galtier, 2016).

Significantly, GVs are mutations that might affect the direction of adaptive evolution both among species and within genomes. Although natural selection is the only of the four population genetic processes that directly interfere in organismal adaptation as a standard evolutionary theory, mutations can interact with selection under specific conditions (Svensson and Berger, 2019).

GVs are alterations that occur in the genome sequence modifying base pairs type and quantity. These can be classified into structural variants or short genetic variants (1000 Genomes Project Consortium et al., 2015).

Structural variants are variations that comprises more than 50 base pairs of the genome and can be classified, for example, as large insertions/deletions (indels), duplications, inversions, among other variations. The Database of Structural Variation (dbVar), part of the National Center for Biotechnology Information (NCBI), has more than 7 million records (Sayers et al., 2022).

On the other hand, short genetic variants have less than 50 base pairs in length. This group includes single nucleotide variant (SNV), short indels and microsatellites (Sayers et al., 2019). The 1000 Genomes Project Consortium (2015) states that SNVs and short indels accounted for more than 99.9% of the GVs in this consortium.

The Database of Single Nucleotide Polymorphisms (dbSNP) is the main repository for information on short human genetic variants, which is part of the NCBI, and currently has more than 1 billion records (Sayers et al., 2022). dbSNP is also linked to the University of California Santa Cruz Genome Browser (GB), a website that contains the complete sequencing of human genome available online (Lee et al., 2022).

When SNVs represent a Minor Allele Frequency (MAF) of at least 1% in the studied population, they are called SNPs. There are many consortia (1000Genomes and TOPMED) that after DNA population sequencing gives the percentage of reference allele (most common) and alternate allele (variant). Thus, the variant with a relative common frequency in population, polymorphism, can be associated with disease risk and phenotypic alteration through GWAS.

2.3 *Cis*-regulatory elements

CREs can be defined as DNA sequences at genome that guide the biding of TFs and associated coregulators at their sites, acting on the same chromosome they are located, as a DNA-protein complex that regulates gene transcription. They are classically divided into promoters, enhancers, silencers, and insulators based on how they interfere on transcription (Shukla et al., 2022).

The first element is promoter which was generally described as a part of a gene, however, currently it should be classified as an activating regulatory element. Promoters have the capacity to initiate gene transcription when they recruit RNAPII and cofactors to transcribe an RNA molecule that codes proteins (mRNA) (Andersson and Sandelin, 2020; ENCODE Project Consortium et al., 2020a).

Promoter activity is distinctly characterized by its proximity with transcription start site (TSS) ~200 base pair (bp) and accessible chromatin determined by DNase I hypersensitive sites (DHSs). Moreover, histone H3 lysine 4 tri-methylation (H3K4me3) is an epigenetic mark that defines promoter activity (ENCODE Project Consortium et al., 2020a). Also, promoters are well known for their CpG (cytosine-phosphate-guanine) island content that roughly overlaps 50% of this element region (Andersson and Sandelin, 2020).

Another activating regulatory element is enhancer that can also positively influence transcription initiation by acting on gene promoters after recruiting TFs, cofactors and further RNAPII (Shukla et al., 2022). Other interesting aspects is their distant capacity of regulating genes (Schoenfelder and Fraser, 2020), especially the super-enhancers (or stretch-enhancer) which are large genomic regions with enhancer activity >10 kb in length (Chen and Liang, 2020).

This type of CREs transcribe eRNAs which are mostly short, bidirectional, unspliced, non-polyadenylated and unstable molecules that regulates gene transcription in a *cis*-action. However, there are eRNAs with features similar to mRNA being long, unidirectional, spliced, polyadenylated and normally acting in *trans* (Sartorelli and Lauberth, 2020).

Enhancer activity is characterized by accessible chromatin determined by DHSs and active histone modifications such as histone H3 lysine 27 acetylation (H3K27ac) and histone H3 lysine 4 mono-methylation (H3K4me1) (Andersson and Sandelin, 2020; ENCODE Project Consortium et al., 2020a).

On the other hand, silencers are CREs that act repressing promoter or enhancer activity and are associated with epigenetic peaks of histone H3 lysine 27 tri-methylation (H3K27me3) (Pang

and Snyder, 2020; Segert et al., 2021). Although this mark is not precise because silencers are not well characterized as enhancers (Gisselbrecht et al., 2020; Huang and Ovcharenko, 2022; Pang et al., 2022), it is still the best mechanism to identify gene repression by silencer activity via chromatin interactions (Cai et al., 2021).

The least CREs are insulators, another type of repressive element that acts as a physical barrier weaken enhancer-promoter communication by chromatin alteration. It is also called architectural proteins due to loop formation especially guided by CTCF/cohesion proteins (Chen and Lei, 2019). Thus, CTCF insulates the promoter-enhancer interaction exclusively to the loop domain (Islam et al., 2023). Although insulators are majority directly related to CTCF protein, not all its binding sites act as insulators (Huang et al., 2021).

It is worthy emphasize that some CREs has a transient ability having a dual role as transcriptional enhancer-silencers depending on tissue type or cellular context in *Drosophila melanogaster* (Gisselbrecht et al., 2020) as well as in human cells (Huang and Ovcharenko, 2022). This bifunctionality was also described between enhancer-promoter (Andersson and Sandelin, 2020) and insulator-enhancer/promoter interaction (Song et al., 2022).

Despite of all conventional classification, our current definition of CREs may need an updated focused on its ability (i.e., to enhance, to silence) at a cellular biological stage (developmental, differentiation) and cell/tissue-specific manner (Huang and Ovcharenko, 2022; Shukla et al., 2022).

3 Justification

PDs are complex diseases that affects populations in the whole globe. It is related to multiple components such as oral microorganisms, host immune system, and genetic factors. GV is well studied due to its association with disease risk and phenotypic variation. Even though there are substantial GVs associated with PDs, the biological mechanisms still are understudied.

Seeking for GV in non-coding regions on the human genome is one of the innovative parts of this work because most research is focused on promoter and exonic regions. Since the latter region is directly involved in coding protein, chances are that phenotypic alteration occurs. Looking for non-probable regions, in other words intergenic and intronic regions, make this study remarkable on PDs.

These *in silico* analyses come as a predicted method for future functional assays that might validate *in vitro* the results achieved on a computer. It is expected that some of the GVs found in this study can increase or reduce expression levels of PDs related genes. In this way, it is worthy to presume GVs located in candidate CREs that interact in proximity or distance with genes and which cell/tissue these genes are most expressed since this information will also support the cell lineage that will be applied in culture cell assay.

Moreover, another aspect that will further contribute to the literature is GV pleiotropism which is when several diseases are associated with the same variant or targeted gene. Thus, this research can lead to new hypothesis of how complex diseases can interact in a molecular and cellular level in forthcoming *in vitro* and *in vivo* analyses.

4 Proposition

The propose of this study was to quantify genetic variants in non-coding region associated with periodontal diseases in sequences of candidate *cis*-regulatory element.

.

5 Material and Methods

The first methodological procedure performed was searching for GVs associated with PDs in the Genomic Wide Association Studies catalog, the NHGRI-EBI GWAS catalog (https://www.ebi.ac.uk/gwas/). Six traits were selected because they include terms associated with periodontal tissue, namely periodontitis, aggressive periodontitis, periodontal measurement, periodontal pocket, gingival bleeding, and gingival disease. All GVs associated with the selected traits were tabulated, as well as reference articles PubMed ID.

The second search for GVs was through an integrative review conducted on PubMed (https://pubmed.ncbi.nlm.nih.gov/). In this platform the following terms "(periodontitis or periodontal or periodontal disease or gingival or gingival disease or gingival bleeding) and (GWAS)" were applied and the PubMed ID of all articles published until 07/01/2022 were annotated. Duplicate articles were excluded.

For qualitative evaluation, articles from other areas and articles that despite of being related to subject did not have any description of GV were excluded. Eligible articles were qualitatively assessed and GVs associated with PDs and their pathogens were tabulated.

After removing GVs duplicates, a search was executed on Genome Browser (GB) (https://genome.ucsc.edu/) (human genome version GRCh38/h38 American server) to categorize them according to their genomic location related to both protein-coding (PC) and non-coding RNA (NCR) genes – GENCODE V41 – (Frankish et al., 2021). For PC genes, GVs were classified into exonic, intronic, intergenic, 5' UTR and 3' UTR regions – dbSNP 153 – (Sayers et al., 2019). Subsequently, the same search was carried out considering NCR genes and GVs were classified into exonic, intronic and intergenic regions.

On GWAS catalog, a third classification was made based on the most severe consequence of each GV. In this case, the platform considers the most severe interference that a given GV can cause in a gene. In addition, all traits related to the selected GVs were also annotated for pleiotropic association between PDs and other diseases.

GVs located in intronic and intergenic regions of PC gene, and GVs in intronic region of NCR gene were both analyzed for the presence of regulatory elements on GB. Two regulatory databases were used: ENCODE cCRES – Encyclopedia of DNA Elements registry of candidate *cis*-regulatory elements – (ENCODE Project Consortium et al., 2020a) and GeneHancer – GeneHancer Regulatory Elements and Gene Interactions – (Fishilevich et al., 2017).

It was only assessed GVs that had an enhancer mark on both ENCODE cCRES and GeneHancer. Furthermore, an interaction evaluation between enhancer and probable target

genes was made (GeneHancer), as well as tissue expression of these genes – GTEx Gene V8 – (GTEx Consortium, 2020).

The final analysis investigates if the selected GVs could change motifs of TF binding sites. JASPAR transcription factor binding site database was used to access predicted binding sites for TF (Castro-Mondragon et al., 2022). Lastly, Tomtom motif comparison tool (https://meme-suite.org/meme/tools/tomtom) quantified similarities between motifs and compared them against databases of known motifs (Gupta et al., 2007).

6 Results

A total of 166 articles were selected which 18 came from the search conducted on GWAS catalog and 148 on PubMed. During duplicate removal, it was observed that the articles found on GWAS catalog were included in the search performed on PubMed. Therefore, 148 articles were read entirety to verify the eligibility according to pre-established criteria (Figure 2).

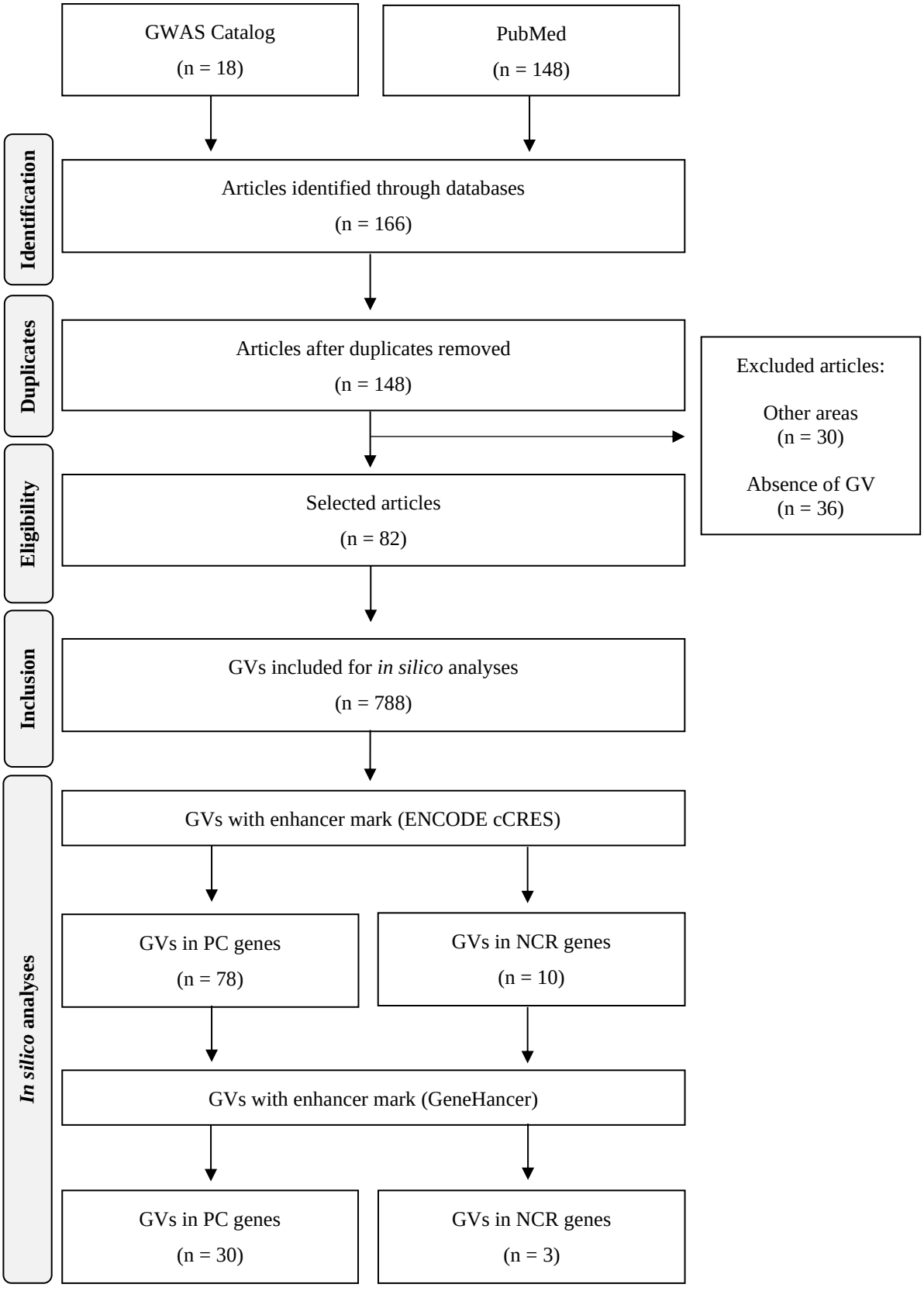
During qualitative evaluation 66 articles were excluded, which 30 were related to other areas and 36 did not have GV related to PDs and their pathogens. Thus, 82 articles were included for further analysis. 788 GVs were obtained from eligible articles and *in silico* analyses were performed on GB (Figure 2).

In both analyses carried out on GB to categorize the GVs according to their genomic location, it was observed the majority were in non-coding region both in PC genes (89.2%) and NCR genes (98.6%). Regarding exonic regions, more than 10% of the GVs were in this region when a PC gene was the reference, whereas for NCR genes this value represented 1.4% (Graphic 1 and Graphic 2).

The analysis performed on GWAS catalog concerning the most severe consequence revealed that out of the 788 GVs evaluated, 442 were not found on this platform (56.1%), 181 were classified as an intronic variant (23.0%) and 82 as an intergenic variant (10.4%). The missense, regulatory region, non-coding transcript exon, 5'UTR, 3'UTR, TF binding site, synonymous and stop gained variants together totalized 10.5% of the GVs evaluated (Graphic 3).

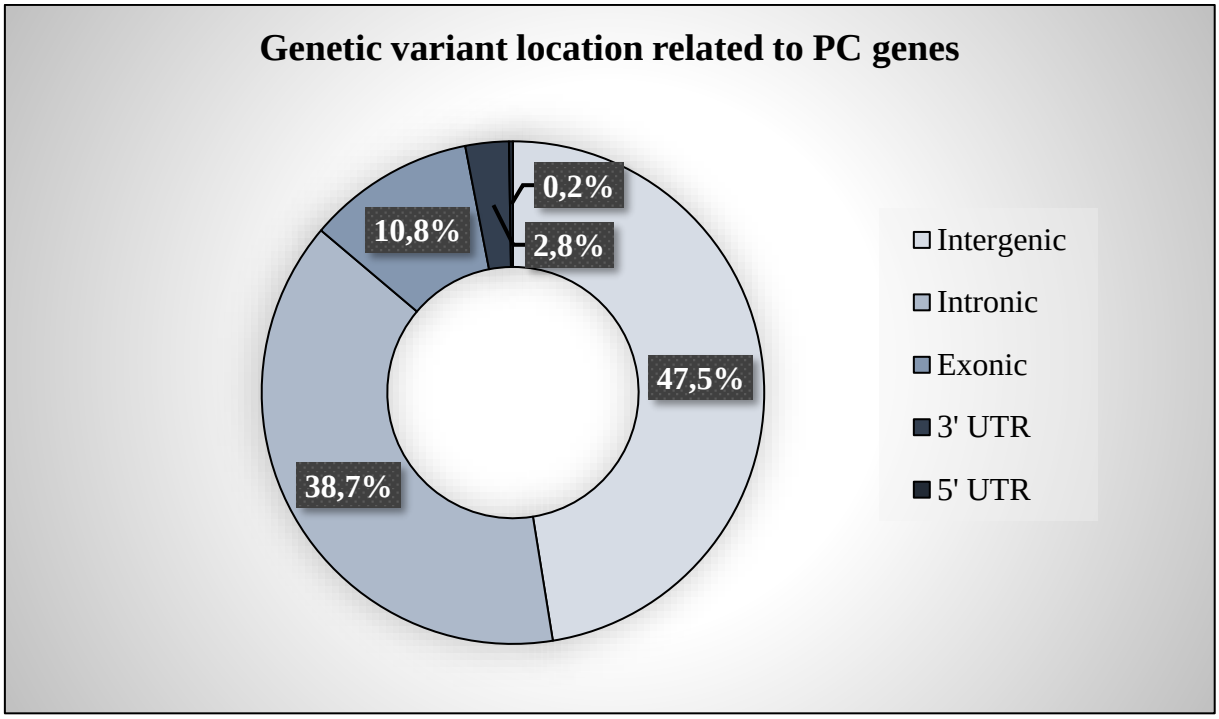
It was also observed that of the GVs that were available on GWAS catalog, the vast majority were related to at least one of the six pre-established traits in the initial search (56.7%), followed by traits related to dental caries and dentures (13, 0%), Alzheimer's disease (4.0%), interleukin-1 beta measurement (4.0%) and type 2 diabetes mellitus (3.8%). Several other traits were identified, but separately they presented values lower than 2% and for this reason they were grouped in the classification “others” (18.5%) (Graphic 4).

Figure 2 – Study design flowchart applied to select genetic variants associated with periodontal diseases and to analyze them *in silico*.



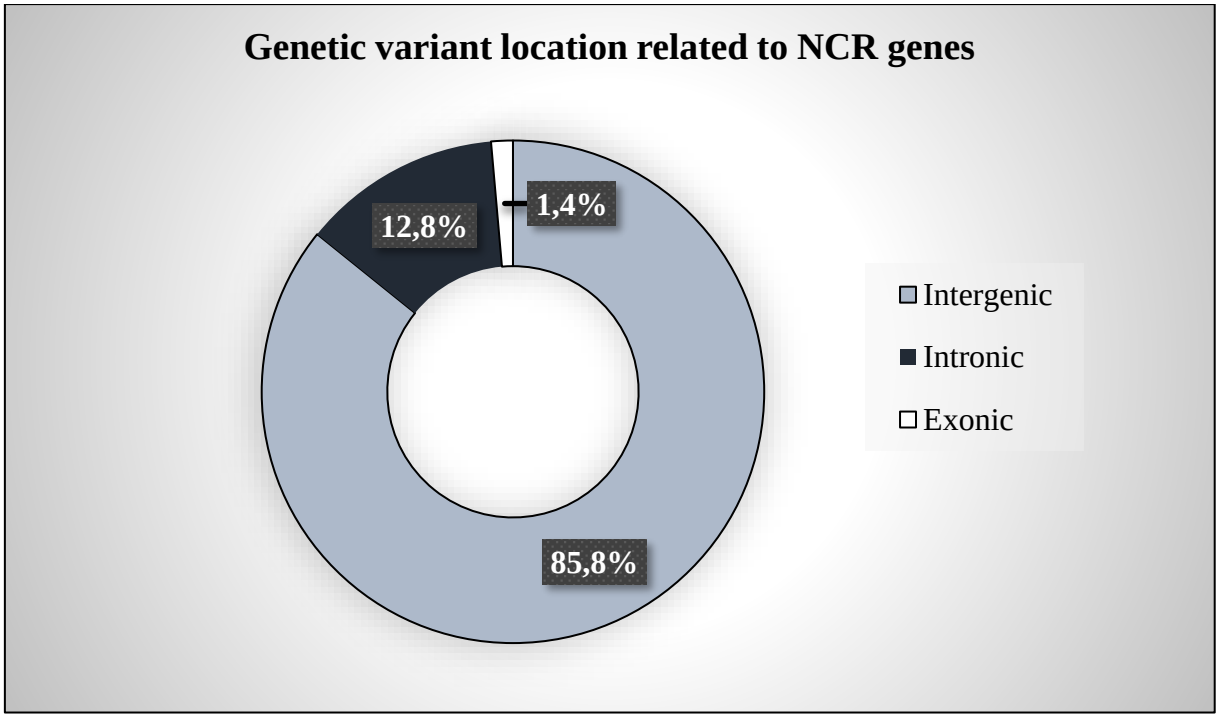
Source: Own authorship.

Graphic 1 – Genetic variants categorized according to their genomic location related to protein-coding genes on Genome Browser.



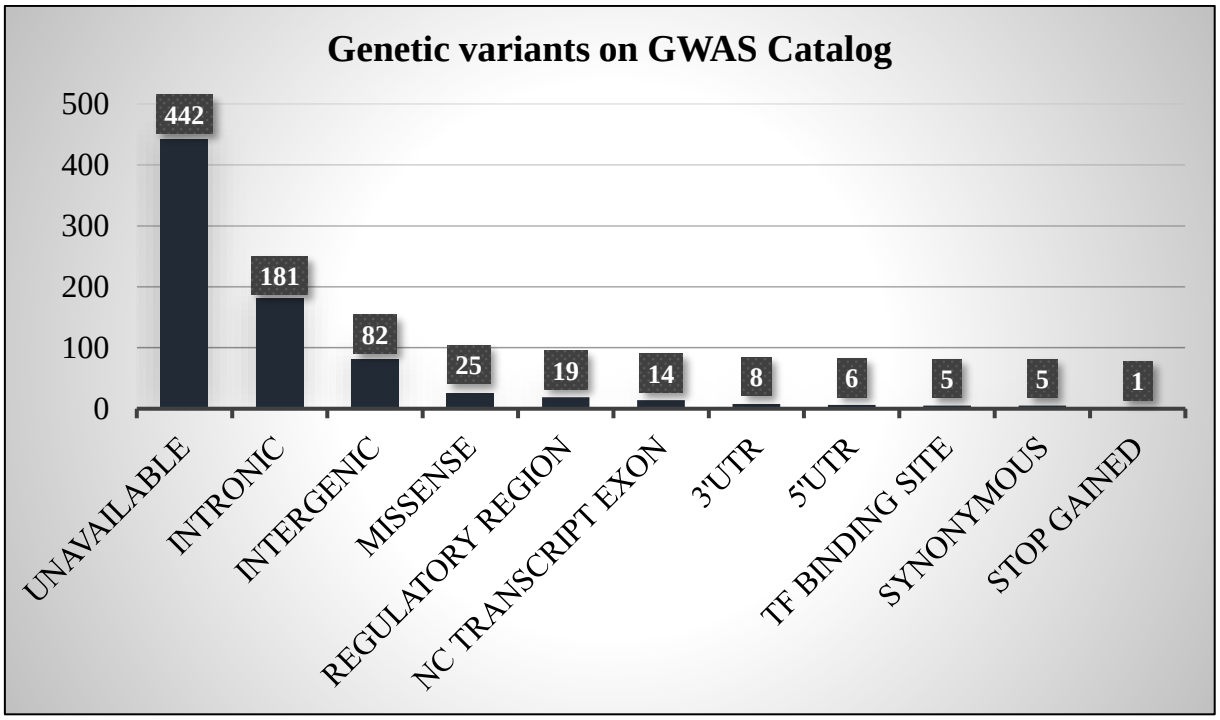
Source: Own authorship.

Graphic 2 – Genetic variants categorized according to their genomic location related to non-coding RNA genes on Genome Browser.



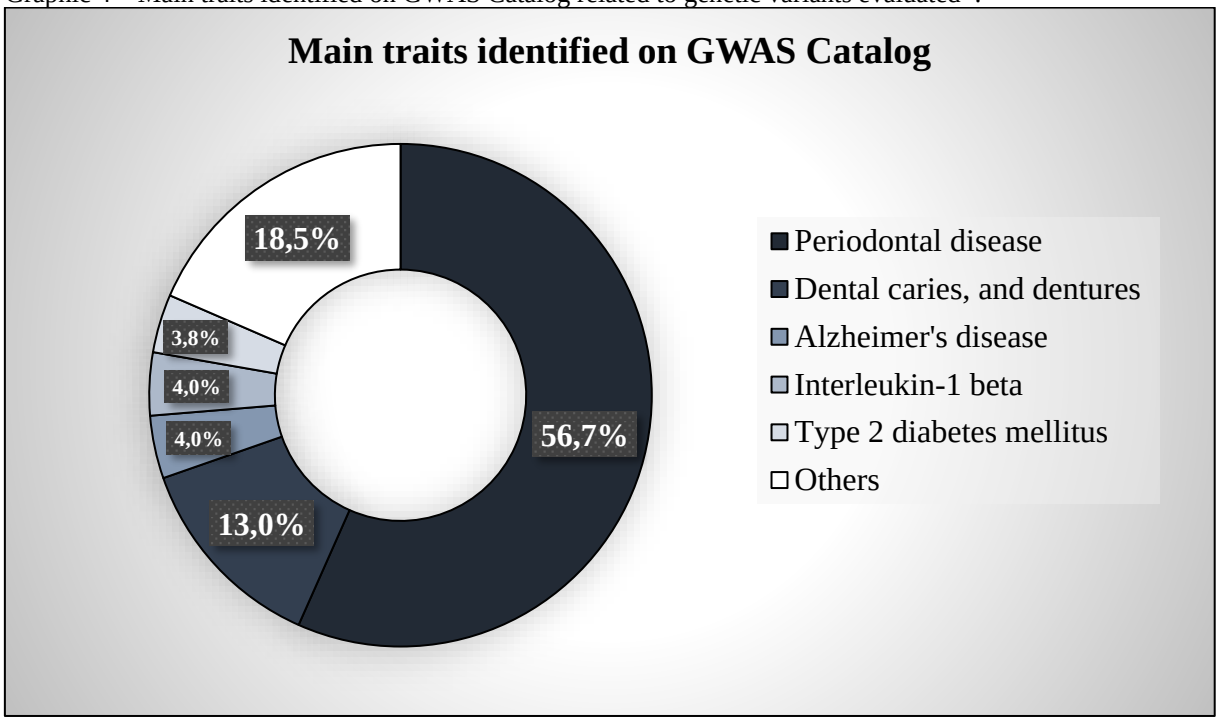
Source: Own authorship.

Graphic 3 – Categorization of genetic variants according to the most severe consequence on GWAS catalog [1].



Source: Own authorship.

Graphic 4 – Main traits identified on GWAS Catalog related to genetic variants evaluated [2].



Source: Own authorship.

[1] TF = transcription factor; NC = non-coding.

[2] Periodontal Disease refers to periodontitis, aggressive periodontitis, periodontal measurement, periodontal pocket, gingival bleeding and gingival disease traits.

Among 788 GV, 679 were in intronic and intergenic regions of PC genes. When evaluated by ENCODE cCRES database, it was observed that 78 GVs in the human genome were specifically located in enhancer candidate on GB. These 78 GVs were also analyzed by GeneHancer to verify regulatory element presence by a second regulation database. After the second analysis, 30 GVs were selected because they were in regulatory regions in both regulatory databases. The 30 GVs, their genomic location, GeneHancer ID which is the interaction between gene and regulatory element, as well as targeted genes and their respective tissue expressions (GTEx Gene V8) were tabulated (Table 1).

Regarding the GVs that were in intronic region of a NCR gene, it was verified that out of 101 GVs in this region, three were located in a regulatory element in both ENCODE cCRES and GeneHancer. The same information for PC genes was noted for these three variants (Table 2).

Table 1 – 30 genetic variants located in intronic or intergenic regions of protein-coding genes and in regulatory sequence detected by the ENCODE cCRES and GeneHancer databases with their respective ID, genomic location, GeneHancer ID, target gene and tissue expression.

(to be continued)

Genetic variant	Genomic location	GeneHancer ID	Gene	Cell/Tissue expression
rs3795391	Intronic	GH01J153384	*S100A8* *S100A9* *S100A12*	Esophagus - Mucosa Esophagus - Mucosa Whole blood
rs61815643	Intronic	GH01J206782	*IL19* *IL10*	Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes
rs3820640	Intronic	GH01J226671	*ITPKB*	Brain
rs1894399	Intronic	GH02J112781	*IL1A* *IL37*	Testis Skin
rs1878318	Intergenic	GH02J112786	*IL1B*	Whole blood
rs4848303	Intergenic	GH02J112792	*IL1B* *IL1A*	Whole blood Testis
rs4848304	Intergenic	GH02J112792	*IL1B* *IL1A*	Whole blood Testis
rs1143637	Intronic	GH02J112826	*IL1B* *IL1A*	Whole blood Testis
rs1143629	Intronic	GH02J112826	*IL1B* *IL1A*	Whole blood Testis

Table 1 – 30 genetic variants located in intronic or intergenic regions of protein-coding genes and in regulatory sequence detected by the ENCODE cCRES and GeneHancer databases with their respective ID, genomic location, GeneHancer ID, target gene and tissue expression.

(continuation)

rs2227306	Intronic	GH04J073739	*PF4V1*	Cells – Cultured fibroblasts
rs9366651	Intergenic	GH06J026335	*BTN3A2* *H4C8*	Spleen Prostate
rs2071591	Intronic	GH06J031545	*NFKBIL1* *MICA* *DDX39B*	Ovary Ovary Uterus
rs9271058	Intergenic	GH06J032602	*HLA-DQA1* *HLA-DQA2* *HLA-DQB1* *HLA-DQB2* *HLA-DRB1* *HLA-DRB5* *HLA-DPB1*	Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes
rs9271850	Intergenic	GH06J032621	*HLA-DQA1* *HLA-DQA2* *HLA-DQB1* *HLA-DQB2* *HLA-DRB1* *HLA-DRB5* *BRD2*	Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes Tireoide
rs7740539	Intronic	GH06J106379	*CRYBG1*	Esophagus - Mucosa

Table 1 – 30 genetic variants located in intronic or intergenic regions of protein-coding genes and in regulatory sequence detected by the ENCODE cCRES and GeneHancer databases with their respective ID, genomic location, GeneHancer ID, target gene and tissue expression.

(conclusion)

rs17875294	Intronic	GH06J170582	*PDCD2* *PSMB1*	Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes
rs9331896	Intronic	GH08J027610	*CLU*	Brain
rs1360590	Intergenic	GH09J022039	*CDKN2B* *CDKN2B-AS1* *UBA52P6*	Esophagus - Mucosa Colon – Transvers Testis
rs2282010	Intronic	GH09J131529	*UCK1*	Brain
rs73443078	Intronic	GH10J119416	*GRK5*	Artery – Aorta
rs73389468	Intergenic	GH10J132481	*INPP5A*	Artery – Tibial
rs7933202	Intergenic	GH11J060168	*MS4A6A*	Spleen
rs2771342	Intergenic	GH14J020988	*METTL17*	Uterus
rs8008037	Intronic	GH14J091360	*CCDC88C* *PPP4R3A*	Cells – EBV – transformed lymphocytes Cells – EBV – transformed lymphocytes
rs904310	Intronic	GH15J024856	*SNRPN*	Brain
rs9905793	Intronic	GH17J048538	*HOXB2* *HOXB3* *HOXB4* *HOXB5* *HOXB6* *HOXB7* *HOXB9* *HOXB-AS1* *HOXB-AS2* *HOXB-AS3* *SKAP1* *CDK5RAP3*	Colon – Sigmoid Colon – Sigmoid Fallopian tube Kidney – Medulla Kidney – Medulla Kidney – Medulla Colon – Transvers Testis Colon – Sigmoid Colon – Sigmoid Cells – EBV – transformed lymphocytes Pituitary
rs2902925	Intergenic	GH19J055544	*SBK3*	Heart
rs149546760	Intronic	GH19J055579	*ZNF579* *ZNF580*	Brain Thyroid
rs3803890	Intronic	GH19J055598	*FIZ1* *ZNF524* *ZNF580*	Brain Brain Thyroid
rs79821641	Intergenic	GH20J025190	*ENTPD6*	Brain

Table 2 – Three genetic variants located in intronic region of non-coding RNA genes and in regulatory sequence detected by the ENCODE cCRES and GeneHancer databases with their respective ID, genomic location, GeneHancer ID, target gene and tissue expression.

Genetic variant	Genomic location	GeneHancer ID	Gene	Cell/Tissue expression
rs904310	Intronic	GH15J024856	*SNRPN*	Brain
rs1360590	Intronic	GH09J022039	*CDKN2B*	Esophagus - Mucosa
			CDKN2B-AS1	Colon – Transvers
			UBA52P6	Testis
rs9905793	Intronic	GH17J048538	*HOXB2*	Colon – Sigmoid
			HOXB3	Colon – Sigmoid
			HOXB4	Fallopian tube
			HOXB5	Kidney – Medulla
			HOXB6	Kidney – Medulla
			HOXB7	Kidney – Medulla
			HOXB9	Colon – Transvers
			HOXB-AS1	Testis
			HOXB-AS2	Colon – Sigmoid
			HOXB-AS3	Colon – Sigmoid
			SKAP1	Cells – EBV – transformed lymphocytes
			CDK5RAP3	Pituitary

25 GVs were within TF motif according to JASPAR database. However, only eight GVs had information about the risk allele in the reference studies. Motif comparison analysis revealed that rs61815643 had higher affinity for TF GLI2 than KLF2, as well as rs2771342 that changed affinity from EN1 to PAX4. Additionally, polymorphism rs904310 (G>A), which is associated with NFYB TF binding site, and had a better affinity for alternate allele A ($p = 6.45x10^{-4}$) than G ($p = 3.72x10^{-3}$) (Figure 3). This intronic variation is related to both *SNRPN* (PC) and *ENSG00000286110* (NCR) genes (Figure 4).

Figure 3 – Motif analysis of rs904310 (G>A) which had a better affinity for alternate allele A ($p = 6.45x10^{-4}$) than G ($p = 3.72x10^{-3}$). Figure obtained on Tomtom motif comparison tool.

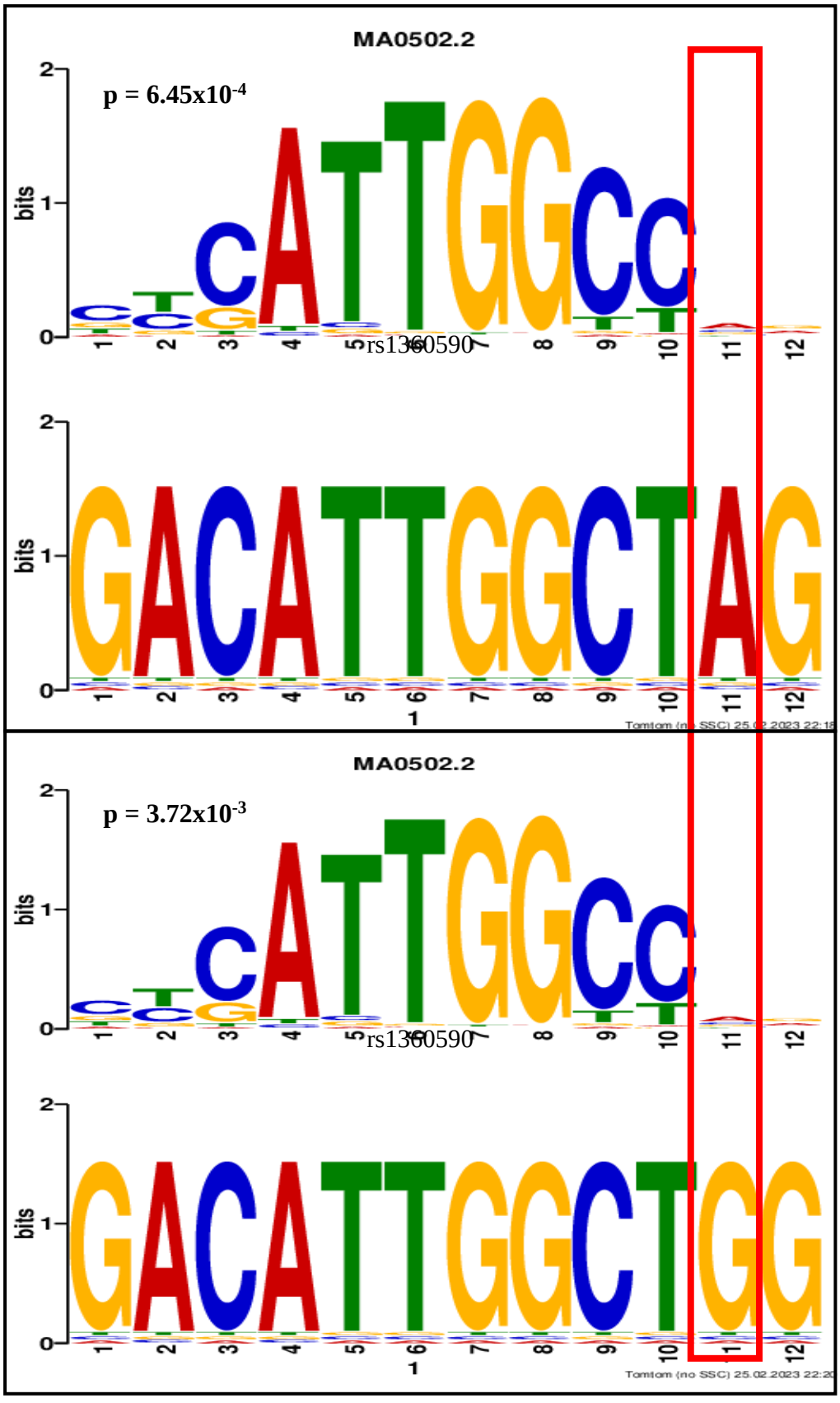
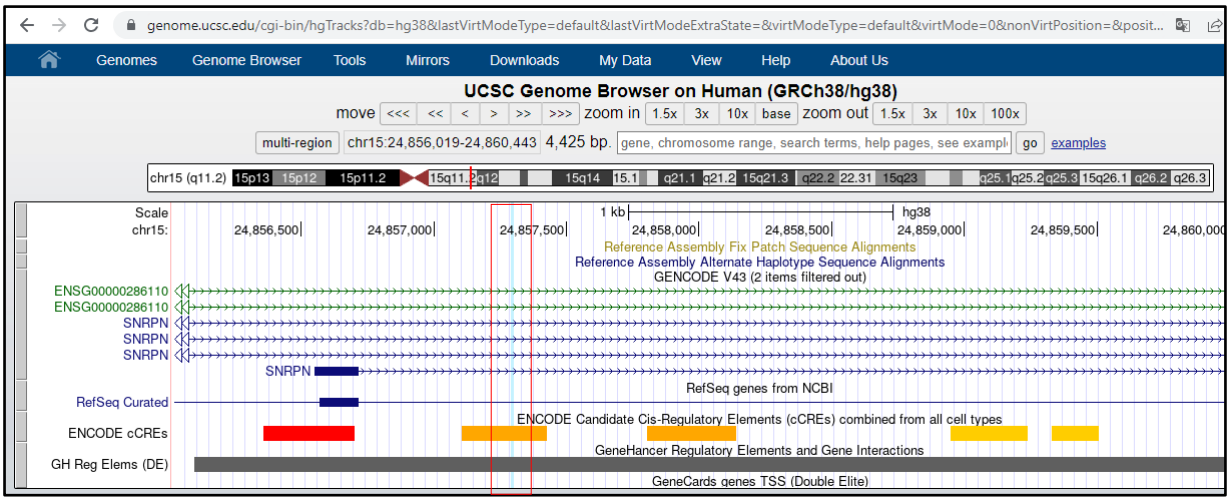


Figure 4 – The intronic variation rs904310 (light blue) related to both *SNRPN* (dark blue) and *ENSG00000286110* (green) genes. Also, the variation is in a candidate enhancer predicted by ENCODE cCREs (orange box) and GeneHancer (grey box). Figure obtained on Genome Browser website.



7 Discussion

This study gives insights in how to predict *in silico* PD-associated GVs located in CREs. These GVs were obtained through an integrative review on PubMed and on NHGRI-EBI GWAS catalog which covered the main reliable resources of information in this field. After *in silico* analyses, 30 GVs fulfilled the proposition of this study being PD-associated GVs located in CREs.

All GVs came from GWAS because it is usually seen as an unbiased approach since it analyzes each SNP individually across the entire genome (Shaffer et al., 2014). Nevertheless, several authors criticize the significant p-value threshold on GWAS ($p < 5x10^{-8}$). Wang et al. (2019) affirms that is an over conservative method because it tests each SNP individually, but SNPs function in Linkage Disequilibrium (LD) rather than as casual variant. Divaris et al. (2013) and Shaffer et al. (2014) emphasize that the threshold is extremely strict and reduces the power of detecting associations.

Currently, there are 18 PDs associated GWAS from 2010-2021 and only six studies reached genome wide statistical significance. These six studies found three associations related to periodontitis trait (Schaeffer et al., 2010; Bevilacqua et al., 2018; Munz et al., 2018), twelve to periodontal measurement (Offenbacher et al., 2016; Sanders et al., 2017) and one to gingival bleeding (Zhang et al., 2016).

It is emphasized that GWAS identify many variants associated with complex phenotypes, however, the biggest challenge is to discover causal variants (Mountjoy et al., 2021). Another problem encountered is specifically distinguish causal variant among many variants in proximity in LD (GTEx Consortium, 2020; Vaithilingam et al., 2014). Finally, understand how GVs influences complex traits such as PDs, considering there are differences on gene expression levels of individuals affected by a GV and its expression varies between cell/tissue (Zhu et al., 2016).

Our study demonstrates that rs904310 is an intronic variant of both *Small Nuclear Ribonucleoprotein Polypeptide N* (*SNRPN)* and *ENSG00000286110* genes located in candidate CRE. A GWAS confirmed this variant is statistically significant associated with periodontal disease related phenotype. The authors supplemented the classical GWAS with information of periodontal pathogens and local inflammatory response. *SNRPN* was in the group that had the highest level of *Porphyromonas gingivalis* (*P. gingivalis*) along with genes the code structural proteins of desmosome, thus, directly effecting epithelial barrier function (Offenbacher et al., 2016).

Interestingly, our interaction analysis demonstrates association between the candidate CRE affected by this GV and *SNRPN*. This gene is part of a subclass of mammalian genes called imprinted gene which are expressed for only one parental allele (Juan et al., 2022).

A recent study described that imprinted genes *Grb10* and *Ddc* were associated with an insulator guided by CTCF protein. The allele specific assemble was responsible for enhancer-promoter interaction that results in changes in the developing heart (Juan et al., 2022). Hence, imprinted genes might have different profiles of expression due to CREs interactions in two possible outcomes depending on parental expression.

Alteration in *SNRPN* is associated with parental imprint that cause Prader-Willi Syndrome (PWS) from imprinted and expressed paternal allele. Two case reports had shown association between PWS and PDs. Greenwood and Small (1990) and Yanagita et al. (2011) reported young individuals presenting PWS with extensive periodontitis. Sixteen out of thirty patients with PWS evaluated by Munné-Miralvés et al. (2020) had gingivitis.

Motif analysis of rs904310 (G>A), which is associated with NFYB TF binding site, reveals a better affinity for alternate allele A ($p = 6.45x10^{-4}$) than G ($p = 3.72x10^{-3}$). However, overexpression of *NFYB* is commonly associated with cancer and there is no evidence with PDs (Fang et al., 2018; Bezzecchi et al., 2021; Feng et al., 2021; Lu et al., 2022).

Overall, rs904310 variant is an interesting SNP for future functional validation in culture cell, due to statistic significant association with PDs, peculiar aspect of interference in a genomic imprinted region and an NCR gene, in addition to its location in a candidate CRE.

Polymorphism rs1360590 is an intronic variant of *Cyclin Dependent Kinase Inhibitor 2B antisense RNA 1* (*CDKN2B-AS1),* also designated *ANRIL/CDKN2BAS,* located in a candidate CRE. This gene is in a cluster of significant genetic susceptible locus for cardiovascular diseases, in proximity with *CDKN2A* and *CDKN2B,* genes associated with tumor suppressing and cell growth regulation respectively. *CDKN2B-AS1* transcribes a long non-coding RNA (lncRNA) that acts silencing *CDKN2A* (Schaefer et al., 2009; Aarabi et al., 2018).

Schaefer et al. (2009) confirms a shared association of Coronary Artery Disease (CAD) and AgP at Chr9p21.3 locus. *CDKN2A/CDKN2B* are the nearest coding genes of this locus. Two years later, Schaefer et al. (2011) described an association between rs1360590 and *CDKN2B-AS1*.

Aarabi et al. (2017) critically reviewed the literature regarding genetic association between PDs and CAD. The authors emphasized 9p21.3 (*CDKN2A/B/AS1* cluster) as one of the frequently loci associating the two diseases. Munz et al. (2018) in a GWAS meta-analysis highlighted an intronic variant at *CDKN2B-AS1* as a risk factor for CAD and AgP.

In vitro analyses confirms that *CDKN2B-AS1* transcripts levels were higher in healthy gingival tissues than inflamed, with higher expression in connective tissues. Moreover, *P. gingivalis* stimulation in human gingival fibroblasts had high mRNA levels of *CDKN2B-AS1* after 24h (Schaefer et al., 2011). One of our hypotheses is that a candidate CRE modulates *CDKN2B-AS1* regulation and cause disturbance on gingival connective tissues.

CDKN2B-AS1 transcribe a lncRNA that have been associated with immunosuppression. A study demonstrates that *CDKN2B-AS1* suppress atherosclerosis inflammatory response inhibiting *ADAM10* expression via DNA methylation by DNMT1 (Li et al., 2019). Lnc-RNA *CDKN2B-AS1* also relieved inflammation of ulcerative colitis by sponging micro-RNAs (miRNA) that would increase inflammatory cytokine production, barrier function and apoptosis (Tian et al., 2020).

Li et al. (2021) observed a similar pattern whereas an inflammatory disease (coronary atherosclerosis) was modulated by a competing endogenous RNA (lnc-RNA *CDKN2B-AS1)* that absorbs a proinflammatory miRNA (miR-126-5p), thus alleviating inflammation.

Moreover, *in vitro* studies have confirmed that transfection of pcDNA-*CDKN2B-AS1* in cultured cells depressed expression of proinflammatory cytokines such as TNF-α, IL-1β, IL-8, and INF-γ (Tian et al., 2020; Li et al., 2021). TNF-α triggers NF-κB activation and upregulation of *CDKN2B-AS1*. The linear transcript interacts with YY1 TF forming a protein complex that regulates *IL-6* and *IL-8* expression (Aarabi et al., 2018).

These proinflammatory cytokines are produced due to bacterial dysbiosis stimulation leading to innate and adaptive immune response. Several cells are involved such as macrophage, NK cell, neutrophil, and dendritic cell (innate) additionally to T and B cells after antigen presentation (adaptive) (Kinane et al., 2017).

Furthermore, motif analysis on GB using JASPAR revealed that rs1360590 is located at the binding site of STAT5b TF which is a master regulator of key biological pathways involved in hemopoietic (Smith et al., 2023). Moreover, a deep learning-based autoencoder predicted the main immunosuppressing genes associated with periodontitis pathogenesis. Afterwards, STAT5b was one of the six leading TFs involved with those genes that might mediate immunosuppressing (Ning et al., 2021).

Collectively, rs1360590 is another interesting variant to be validate because it is located in a candidate CRE and its target gene transcribe a lncRNA *CDKN2B-AS1* with pleiotropic action in inflammatory diseases upon downregulation of *IL-6* and *IL-8* via TNF-α stimulus. Also, this variant is not cataloged on GWAS catalog and there are hundreds GVs in *CDKN2B-AS1* statistically associated with CAD.

The limitation of this study is the lack of functional validation *in vitro* to confirm some of the hypotheses *in silico*. Nonetheless, we want to test if the CRE located in rs904310 modulate *SNRPN* expression, as well as rs1360590 might modulate *CDKN2B-AS1*. Finally, understand how these regulation expression changes are related to biological mechanisms related to PDs pathogenesis.

It still a challenge solves the puzzles that leads to the role of regulatory elements in multifactorial diseases such as PDs. However, few studies have been demonstrating the power of enhancer disturbance in diseases etiopathogenesis (enhanceropathies) in monogenic diseases, cancer, and common systemic diseases (Zaugg et al., 2022).

8 Conclusion

It is concluded there are 30 genetic variants in non-coding region associated with periodontal diseases in sequences of candidate *cis*-regulatory element.

References[3]

1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. Nature. 2015 Oct 1;526(7571):68-74. doi: 10.1038/nature15393.

Aarabi G, Zeller T, Heydecke G, Munz M, Schäfer A, Seedorf U. Roles of the Chr.9p21.3 *ANRIL* Locus in Regulating Inflammation and Implications for Anti-Inflammatory Drug Target Identification. Front Cardiovasc Med. 2018 May 18;5:47. doi: 10.3389/fcvm.2018.00047.

Aarabi G, Zeller T, Seedorf H, Reissmann DR, Heydecke G, Schaefer AS, et al. Genetic Susceptibility Contributing to Periodontal and Cardiovascular Disease. J Dent Res. 2017 Jun;96(6):610-617. doi: 10.1177/0022034517699786.

Andersson R, Sandelin A. Determinants of enhancer and promoter activities of regulatory elements. Nat Rev Genet. 2020 Feb;21(2):71-87. doi: 10.1038/s41576-019-0173-8.

Arensbergen JV, Pagie L, FitzPatrick VD, Haas MD, Baltissen MP, Comoglio F, et al. High-throughput identification of human SNPs affecting regulatory element activity. Nat Genet. 2019 Jul;51(7):1160-1169. doi: 10.1038/s41588-019-0455-2.

Armitage GC. Periodontal diagnoses and classification of periodontal diseases. Periodontol 2000. 2004;34:9-21. doi: 10.1046/j.0906-6713.2002.003421.x.

Bergman DT, Jones TR, Liu V, Ray J, Jagoda E, Siraj L, et al. Compatibility rules of human enhancer and promoter sequences. Nature. 2022 Jul;607(7917):176-184. doi: 10.1038/s41586-022-04877-w.

Bevilacqua L, Navarra CO, Pirastu N, Lenarda RD, Gasparini P, Robino A. A genome-wide association study identifies an association between variants in EFCAB4B gene and

[3] According to FOP/UNICAMP norms, based on standardization of the International Committee of Medical Journal Editors - Vancouver Group. Abbreviation of journals according to PubMed.

periodontal disease in an Italian isolated population. J Periodontal Res. 2018 Dec;53(6):992-998. doi: 10.1111/jre.12598.

Bezzecchi E, Bernardini A, Ronzio M, Miccolo C, Chiocca S, Dolfini D, Mantovani R. NF-Y Subunits Overexpression in HNSCC. Cancers (Basel). 2021 Jun 16;13(12):3019. doi: 10.3390/cancers13123019.

Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019 Jan 8;47(D1):D1005-D1012. doi: 10.1093/nar/gky1120.

Cai Y, Zhang Y, Loh YP, Tng JQ, Lim MC, Cao Z, et al. H3K27me3-rich genomic regions can function as silencers to repress gene expression via chromatin interactions. Nat Commun. 2021 Jan 29;12(1):719. doi: 10.1038/s41467-021-20940-y.

Carrasco Pro S, Bulekova K, Gregor B, Labadorf A, Fuxman Bass JI. Prediction of genome-wide effects of single nucleotide variants on transcription factor binding. Sci Rep. 2020 Oct 19;10(1):17632. doi: 10.1038/s41598-020-74793-4.

Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Lemma RB, Turchi L, Blanc-Mathieu R, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2022 Jan 7;50(D1):D165-D173. doi: 10.1093/nar/gkab1113.

Caton JG, Armitage G, Berglundh T, Chapple ILC, Jepsen S, Kornman KS, et al. A new classification scheme for periodontal and peri-implant diseases and conditions - Introduction and key changes from the 1999 classification. J Clin Periodontol. 2018 Jun;45 Suppl 20:S1-S8. doi: 10.1111/jcpe.12935.

Chen D, Lei EP. Function and regulation of chromatin insulators in dynamic genome organization. Curr Opin Cell Biol. 2019 Jun;58:61-68. doi: 10.1016/j.ceb.2019.02.001.

Chen H, Liang H. A High-Resolution Map of Human Enhancer RNA Loci Characterizes Super-enhancer Activities in Cancer. Cancer Cell. 2020 Nov 9;38(5):701-715.e5. doi: 10.1016/j.ccell.2020.08.020.

Coo A, Cruz R, Quintela I, Herrera D, Sanz M, Diz P, et al. Genome-wide association study of stage III/IV grade C periodontitis (former aggressive periodontitis) in a Spanish population. J Clin Periodontol. 2021 Jul;48(7):896-906. doi: 10.1111/jcpe.13460.

Divaris K, Monda KL, North KE, Olshan AF, Lange EM, Moss K, et al. Genome-wide association study of periodontal pathogen colonization. J Dent Res. 2012 Jul;91(7 Suppl):21S-28S. doi: 10.1177/0022034512447951.

Divaris K, Monda KL, North KE, Olshan AF, Reynolds LM, Hsueh WC, et al. Exploring the genetic basis of chronic periodontitis: a genome-wide association study. Hum Mol Genet. 2013 Jun 1;22(11):2312-24. doi: 10.1093/hmg/ddt065.

Ellegren H, Galtier N. Determinants of genetic diversity. Nat Rev Genet. 2016 Jul;17(7):422-33. doi: 10.1038/nrg.2016.58.

ENCODE Project Consortium, Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature. 2020a Jul;583(7818):699-710. doi: 10.1038/s41586-020-2493-4.

ENCODE Project Consortium, Snyder MP, Gingeras TR, Moore JE, Weng Z, Gerstein MB, Ren B, et al. Perspectives on ENCODE. Nature. 2020b Jul;583(7818):693-698. doi: 10.1038/s41586-020-2449-8.

Fang Z, Gong C, Yu S, Zhou W, Hassan W, Li H, et al. NFYB-induced high expression of E2F1 contributes to oxaliplatin resistance in colorectal cancer via the enhancement of CHK1 signaling. Cancer Lett. 2018 Feb 28;415:58-72. doi: 10.1016/j.canlet.2017.11.040.

Feng L, Xu X, Zhao K. NFYB potentiates STK33 activation to promote cisplatin resistance in diffuse large B-cell lymphoma. Leuk Res. 2021 Dec;111:106708. doi: 10.1016/j.leukres.2021.106708.

Feng P, Wang X, Casado PL, Küchler EC, Deeley K, Noel J, et al. Genome wide association scan for chronic periodontitis implicates novel locus. BMC Oral Health. 2014 Jul 9;14:84. doi: 10.1186/1472-6831-14-84.

Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, Rosen N, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford). 2017 Jan 1;2017:bax028. doi: 10.1093/database/bax028.

Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland JE, Mudge JM, et al. GENCODE 2021. Nucleic Acids Res. 2021 Jan 8;49(D1):D916-D923. doi: 10.1093/nar/gkaa1087.

Freitag-Wolf S, Dommisch H, Graetz C, Jockel-Schneider Y, Harks I, Staufenbiel I, et al. Genome-wide exploration identifies sex-specific genetic effects of alleles upstream NPY to increase the risk of severe periodontitis in men. J Clin Periodontol. 2014 Dec;41(12):1115-21. doi: 10.1111/jcpe.12317.

Freitag-Wolf S, Munz M, Junge O, Graetz C, Jockel-Schneider Y, Staufenbiel I, et al. Sex-specific genetic factors affect the risk of early-onset periodontitis in Europeans. J Clin Periodontol. 2021 Nov;48(11):1404-1413. doi: 10.1111/jcpe.13538.

Freitag-Wolf S, Munz M, Wiehe R, Junge O, Graetz C, Jockel-Schneider Y, et al. Smoking Modifies the Genetic Risk for Early-Onset Periodontitis. J Dent Res. 2019 Nov;98(12):1332-1339. doi: 10.1177/0022034519875443.

Gisselbrecht SS, Palagi A, Kurland JV, Rogers JM, Ozadam H, Zhan Y, et al. Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts. Mol Cell. 2020 Jan 16;77(2):324-337.e8. doi: 10.1016/j.molcel.2019.10.004.

Greenwood RE, Small IC. Case report of the Prader-Willi syndrome. J Clin Periodontol. 1990 Jan;17(1):61-3. doi: 10.1111/j.1600-051x.1990.tb01049.x.

GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020 Sep 11;369(6509):1318-1330. doi: 10.1126/science.aaz1776.

Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. Genome Biol. 2007;8(2):R24. doi: 10.1186/gb-2007-8-2-r24.

Huang D, Ovcharenko I. Enhancer-silencer transitions in the human genome. Genome Res. 2022 Mar;32(3):437-448. doi: 10.1101/gr.275992.121.

Huang H, Zhu Q, Jussila A, Han Y, Bintu B, Kern C, et al. CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains. Nat Genet. 2021 Jul;53(7):1064-1074. doi: 10.1038/s41588-021-00863-6.

Islam Z, Saravanan B, Walavalkar K, Farooq U, Singh AK, Radhakrishnan S, et al. Active enhancers strengthen insulation by RNA-mediated CTCF binding at chromatin domain boundaries. Genome Res. 2023 Jan;33(1):1-17. doi: 10.1101/gr.276643.122.

Juan AM, Foong YH, Thorvaldsen JL, Lan Y, Leu NA, Rurik JG, et al. Tissue-specific Grb10/Ddc insulator drives allelic architecture for cardiac development. Mol Cell. 2022 Oct 6;82(19):3613-3631.e7. doi: 10.1016/j.molcel.2022.08.021.

Kassebaum NJ, Smith AGC, Bernabé E, Fleming TD, Reynolds AE, Vos T, et al. Global, Regional, and National Prevalence, Incidence, and Disability-Adjusted Life Years for Oral Conditions for 195 Countries, 1990-2015: A Systematic Analysis for the Global Burden of Diseases, Injuries, and Risk Factors. J Dent Res. 2017 Apr;96(4):380-387. doi: 10.1177/0022034517693566.

Kinane DF, Stathopoulou PG, Papapanou PN. Periodontal diseases. Nat Rev Dis Primers. 2017 Jun 22;3:17038. doi: 10.1038/nrdp.2017.38.

Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. Nat Rev Genet. 2019 Apr;20(4):207-220. doi: 10.1038/s41576-018-0089-8.

Lee BT, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2022 update. Nucleic Acids Res. 2022 Jan 7;50(D1):D1115-D1122. doi: 10.1093/nar/gkab959.

Li H, Han S, Sun Q, Yao Y, Li S, Yuan C, et al. Long non-coding RNA CDKN2B-AS1 reduces inflammatory response and promotes cholesterol efflux in atherosclerosis by inhibiting ADAM10 expression. Aging (Albany NY). 2019 Mar 29;11(6):1695-1715. doi: 10.18632/aging.101863.

Li J, Chen J, Zhang F, Li J, An S, Cheng M, et al. LncRNA CDKN2B-AS1 hinders the proliferation and facilitates apoptosis of ox-LDL-induced vascular smooth muscle cells via the ceRNA network of CDKN2B-AS1/miR-126-5p/PTPN7. Int J Cardiol. 2021 Oct 1;340:79-87. doi: 10.1016/j.ijcard.2021.08.009.

Löe H, Anerud A, Boysen H, Morrison E. Natural history of periodontal disease in man. Rapid, moderate and no loss of attachment in Sri Lankan laborers 14 to 46 years of age. J Clin Periodontol. 1986 May;13(5):431-45. doi: 10.1111/j.1600-051x.1986.tb01487.x.

Loos BG, Van Dyke TE. The role of inflammation and genetics in periodontal disease. Periodontol 2000. 2020 Jun;83(1):26-39. doi: 10.1111/prd.12297.

Lu T, Li C, Xiang C, Gong Y, Peng W, Hou F, et al. Over-expression of NFYB affects stromal cells reprogramming and predicts worse survival in gastric cancer patients. Aging (Albany NY). 2022 Sep 22;14(19):7851-7865. doi: 10.18632/aging.204294.

Marchesan JT. Inflammasomes as contributors to periodontal disease. J Periodontol. 2020 Oct;91 Suppl 1(Suppl 1):S6-S11. doi: 10.1002/JPER.20-0157.

Mountjoy E, Schmidt EM, Carmona M, Schwartzentruber J, Peat G, Miranda A, et al. An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. Nat Genet. 2021 Nov;53(11):1527-1533. doi: 10.1038/s41588-021-00945-5.

Munné-Miralvés C, Brunet-Llobet L, Cahuana-Cárdenas A, Torné-Durán S, Miranda-Rius J, Rivera-Baró A. Oral disorders in children with Prader-Willi syndrome: a case control study. Orphanet J Rare Dis. 2020 Feb 10;15(1):43. doi: 10.1186/s13023-020-1326-8.

Munz M, Richter GM, Loos BG, Jepsen S, Divaris K, Offenbacher S, et al. Genome-wide association meta-analysis of coronary artery disease and periodontitis reveals a novel shared risk locus. Sci Rep. 2018 Sep 12;8(1):13678. doi: 10.1038/s41598-018-31980-8.

Munz M, Willenborg C, Richter GM, Jockel-Schneider Y, Graetz C, Staufenbiel I, et al. A genome-wide association study identifies nucleotide variants at SIGLEC5 and DEFA1A3 as risk loci for periodontitis. Hum Mol Genet. 2017 Jul 1;26(13):2577-2588. doi: 10.1093/hmg/ddx151.

Nanci A. Ten cate - Histologia Oral. 9a edição. Editora GEN Guanabara Koogan. 2019.

Nashef A, Matthias M, Weiss E, Loos BG, Jepsen S, van der Velde N, et al. Translation of mouse model to human gives insights into periodontitis etiology. Sci Rep. 2020 Mar 17;10(1):4892. doi: 10.1038/s41598-020-61819-0.

Nasser J, Bergman DT, Fulco CP, Guckelberger P, Doughty BR, Patwardhan TA, et al. Genome-wide enhancer maps link risk variants to disease genes. Nature. 2021 May;593(7858):238-243. doi: 10.1038/s41586-021-03446-x.

Nibali L, Bayliss-Chapman J, Almofareh SA, Zhou Y, Divaris K, Vieira AR. What Is the Heritability of Periodontitis? A Systematic Review. J Dent Res. 2019 Jun;98(6):632-641. doi: 10.1177/0022034519842510.

Ning W, Acharya A, Sun Z, Ogbuehi AC, Li C, Hua S. Deep Learning Reveals Key Immunosuppression Genes and Distinct Immunotypes in Periodontitis. Front Genet. 2021 Mar 12;12:648329. doi: 10.3389/fgene.2021.648329.

Offenbacher S, Divaris K, Barros SP, Moss KL, Marchesan JT, Morelli T, et al. Genome-wide association study of biologically informed periodontal complex traits offers novel insights into the genetic basis of periodontal disease. Hum Mol Genet. 2016 May 15;25(10):2113-2129. doi: 10.1093/hmg/ddw069.

Offenbacher S, Jiao Y, Kim SJ, Marchesan J, Moss KL, Jing L, et al. GWAS for Interleukin-1β levels in gingival crevicular fluid identifies IL37 variants in periodontal inflammation. Nat Commun. 2018 Sep 11;9(1):3686. doi: 10.1038/s41467-018-05940-9.

Pang B, Snyder MP. Systematic identification of silencers in human cells. Nat Genet. 2020 Mar;52(3):254-263. doi: 10.1038/s41588-020-0578-5.

Pang B, van Weerd JH, Hamoen FL, Snyder MP. Identification of non-coding silencer elements and their regulation of gene expression. Nat Rev Mol Cell Biol. 2022 Nov 7. doi: 10.1038/s41580-022-00549-9.

Rowley MJ, Corces VG. Organizational principles of 3D genome architecture. Nat Rev Genet. 2018 Dec;19(12):789-800. doi: 10.1038/s41576-018-0060-8.

Sanders AE, Sofer T, Wong Q, Kerr KF, Agler C, Shaffer JR, et al. Chronic Periodontitis Genome-wide Association Study in the Hispanic Community Health Study / Study of Latinos. J Dent Res. 2017 Jan;96(1):64-72. doi: 10.1177/0022034516664509.

Sartorelli V, Lauberth SM. Enhancer RNAs are an important regulatory layer of the epigenome. Nat Struct Mol Biol. 2020 Jun;27(6):521-528. doi: 10.1038/s41594-020-0446-0.

Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2019 Jan 8;47(D1):D23-D28. doi: 10.1093/nar/gky1069.

Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20-D26. doi: 10.1093/nar/gkab1112.

Schaefer AS. Complementary Experimental Methods in Genetics Open Up New Avenues of Research to Elucidate the Pathogenesis of Periodontitis. Adv Exp Med Biol. 2022;1373:209-227. doi: 10.1007/978-3-030-96881-6_11.

Schaefer AS, Richter GM, Dommisch H, Reinartz M, Nothnagel M, Noack B, et al. CDKN2BAS is associated with periodontitis in different European populations and is activated by bacterial infection. J Med Genet. 2011 Jan;48(1):38-47. doi: 10.1136/jmg.2010.078998.

Schaefer AS, Richter GM, Groessner-Schreiber B, Noack B, Nothnagel M, El Mokhtari NE, et al. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet. 2009 Feb;5(2):e1000378. doi: 10.1371/journal.pgen.1000378.

Schaefer AS, Richter GM, Nothnagel M, Manke T, Dommisch H, Jacobs G, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet. 2010 Feb 1;19(3):553-62. doi: 10.1093/hmg/ddp508.

Schoenfelder S, Fraser P. Long-range enhancer-promoter contacts in gene expression control. Nat Rev Genet. 2019 Aug;20(8):437-455. doi: 10.1038/s41576-019-0128-0.

Segert JA, Gisselbrecht SS, Bulyk ML. Transcriptional Silencers: Driving Gene Expression with the Brakes On. Trends Genet. 2021 Jun;37(6):514-527. doi: 10.1016/j.tig.2021.02.002.

Shaffer JR, Polk DE, Wang X, Feingold E, Weeks DE, Lee MK, et al. Genome-wide association study of periodontal health measured by probing depth in adults ages 18-49 years. G3 (Bethesda). 2014 Feb 19;4(2):307-14. doi: 10.1534/g3.113.008755.

Shimizu S, Momozawa Y, Takahashi A, Nagasawa T, Ashikawa K, Terada Y, et al. A genome-wide association study of periodontitis in a Japanese population. J Dent Res. 2015 Apr;94(4):555-61. doi: 10.1177/0022034515570315.

Shukla V, Cetnarowska A, Hyldahl M, Mandrup S. Interplay between regulatory elements and chromatin topology in cellular lineage determination. Trends Genet. 2022 Oct;38(10):1048-1061. doi: 10.1016/j.tig.2022.05.011.

Silva LM, Doyle AD, Greenwell-Wild T, Dutzan N, Tran CL, Abusleme L, et al. Fibrin is a critical regulator of neutrophil effector function at the oral mucosal barrier. Science. 2021 Dec 24;374(6575):eabl5450. doi: 10.1126/science.abl5450.

Smith MR, Satter LRF, Vargas-Hernández A. STAT5b: A master regulator of key biological pathways. Front Immunol. 2023 Jan 23;13:1025373. doi: 10.3389/fimmu.2022.1025373.

Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. 2022 Nov 9:gkac1010. doi: 10.1093/nar/gkac1010.

Song Y, Liang Z, Zhang J, Hu G, Wang J, Li Y, et al. CTCF functions as an insulator for somatic genes and a chromatin remodeler for pluripotency genes during reprogramming. Cell Rep. 2022 Apr 5;39(1):110626. doi: 10.1016/j.celrep.2022.110626.

Svensson EI, Berger D. The Role of Mutation Bias in Adaptive Evolution. Trends Ecol Evol. 2019 May;34(5):422-434. doi: 10.1016/j.tree.2019.01.015.

Tegelberg P, Leppilahti JM, Ylöstalo A, Tervonen T, Kettunen J, Suominen AL, et al. Genome-wide association study of periodontal pocketing in Finnish adults. BMC Oral Health. 2021 Nov 30;21(1):611. doi: 10.1186/s12903-021-01964-8.

Teumer A, Holtfreter B, Völker U, Petersmann A, Nauck M, Biffar R, et al. Genome-wide association study of chronic periodontitis in a general German population. J Clin Periodontol. 2013 Nov;40(11):977-85. doi: 10.1111/jcpe.12154.

Tian Y, Cui L, Lin C, Wang Y, Liu Z, Miao X. LncRNA CDKN2B-AS1 relieved inflammation of ulcerative colitis via sponging miR-16 and miR-195. Int Immunopharmacol. 2020 Nov;88:106970. doi: 10.1016/j.intimp.2020.106970

Vaithilingam RD, Safii SH, Baharuddin NA, Ng CC, Cheong SC, Bartold PM, et al. Moving into a new era of periodontal genetic studies: relevance of large case-control samples using severe phenotypes for genome-wide association studies. J Periodontal Res. 2014 Dec;49(6):683-95. doi: 10.1111/jre.12167.

Wang MH, Cordell HJ, Van Steen K. Statistical methods for genome-wide association studies. Semin Cancer Biol. 2019 Apr;55:53-60. doi: 10.1016/j.semcancer.2018.04.008.

Yanagita M, Hirano H, Kobashi M, Nozaki T, Yamada S, Kitamura M, et al. Periodontal disease in a patient with Prader-Willi syndrome: a case report. J Med Case Rep. 2011 Jul 28;5:329. doi: 10.1186/1752-1947-5-329.

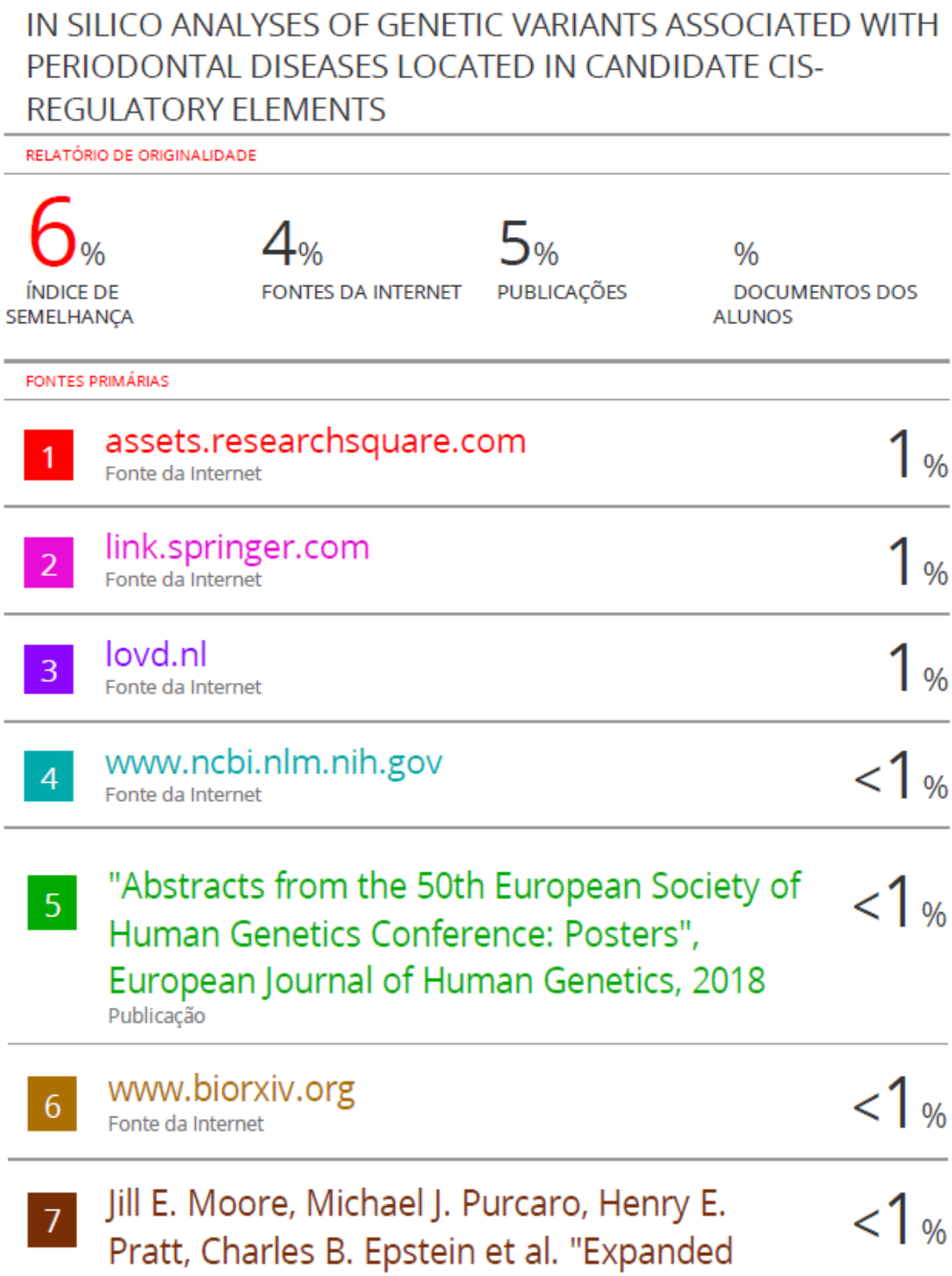
Yanchus C, Drucker KL, Kollmeyer TM, Tsai R, Winick-Ng W, Liang M, et al. A noncoding single-nucleotide polymorphism at 8q24 drives IDH1-mutant glioma formation. Science. 2022 Oct 7;378(6615):68-78. doi: 10.1126/science.abj2890.

Zaugg JB, Sahlén P, Andersson R, Alberich-Jorda M, de Laat W, Deplancke B, et al. Current challenges in understanding the role of enhancers in disease. Nat Struct Mol Biol. 2022 Dec;29(12):1148-1158. doi: 10.1038/s41594-022-00896-3.

Zhang S, Divaris K, Moss K, Yu N, Barros S, Marchesan J, et al. The Novel *ASIC2* Locus is Associated with Severe Gingival Inflammation. JDR Clin Trans Res. 2016 Jul;1(2):163-170. doi: 10.1177/2380084416645290.

Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016 May;48(5):481-7. doi: 10.1038/ng.3538.

Attachment 1 – Plagiarism prevention and originality verification



IN SILICO ANALYSES OF GENETIC VARIANTS ASSOCIATED WITH PERIODONTAL DISEASES LOCATED IN CANDIDATE CIS-REGULATORY ELEMENTS

RELATÓRIO DE ORIGINALIDADE

6%	4%	5%	%
ÍNDICE DE SEMELHANÇA	FONTES DA INTERNET	PUBLICAÇÕES	DOCUMENTOS DOS ALUNOS

FONTES PRIMÁRIAS

1	assets.researchsquare.com Fonte da Internet	1 %
2	link.springer.com Fonte da Internet	1 %
3	lovd.nl Fonte da Internet	1 %
4	www.ncbi.nlm.nih.gov Fonte da Internet	<1 %
5	"Abstracts from the 50th European Society of Human Genetics Conference: Posters", European Journal of Human Genetics, 2018 Publicação	<1 %
6	www.biorxiv.org Fonte da Internet	<1 %
7	Jill E. Moore, Michael J. Purcaro, Henry E. Pratt, Charles B. Epstein et al. "Expanded	<1 %