



OPEN ACCESS

EDITED BY

Digvijay Verma,
Babasaheb Bhimrao Ambedkar University, India

REVIEWED BY

J. Christopher Fenno,
University of Michigan, United States
Miroslaw Ksiazek,
Jagiellonian University, Poland

*CORRESPONDENCE

Cristian Cortez
✉ cristian.cortez@pucv.cl
Juan P. Cárdenas
✉ juan.cardenas@umayor.cl

[†]These authors have contributed equally to this work

RECEIVED 20 May 2023

ACCEPTED 04 July 2023

PUBLISHED 19 July 2023

CITATION

Morales-Olavarria M, Nuñez-Belmar J, González D, Vicencio E, Rivas-Pardo JA, Cortez C and Cárdenas JP (2023) Phylogenomic analysis of the *Porphyromonas gingivalis* - *Porphyromonas gulae* duo: approaches to the origin of periodontitis. *Front. Microbiol.* 14:1226166. doi: 10.3389/fmicb.2023.1226166

COPYRIGHT

© 2023 Morales-Olavarria, Nuñez-Belmar, González, Vicencio, Rivas-Pardo, Cortez and Cárdenas. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Phylogenomic analysis of the *Porphyromonas gingivalis* - *Porphyromonas gulae* duo: approaches to the origin of periodontitis

Mauricio Morales-Olavarria^{1†}, Josefa Nuñez-Belmar^{1†}, Dámariz González¹, Emiliano Vicencio², Jaime Andres Rivas-Pardo^{1,3}, Cristian Cortez^{2*} and Juan P. Cárdenas^{1,3*}

¹Centro de Genómica y Bioinformática, Facultad de Ciencias, Ingeniería y Tecnología, Universidad Mayor, Santiago, Chile, ²Escuela de Tecnología Médica, Facultad de Ciencias, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, ³Escuela de Biotecnología, Facultad de Ciencias, Ingeniería y Tecnología, Universidad Mayor, Santiago, Chile

Porphyromonas gingivalis is an oral human pathogen associated with the onset and progression of periodontitis, a chronic immune-inflammatory disease characterized by the destruction of the teeth-supporting tissue. *P. gingivalis* belongs to the genus *Porphyromonas*, which is characterized by being composed of Gram-negative, asaccharolytic, non-spore-forming, non-motile, obligatory anaerobic species, inhabiting niches such as the oral cavity, urogenital tract, gastrointestinal tract and infected wound from different mammals including humans. Among the *Porphyromonas* genus, *P. gingivalis* stands out for its specificity in colonizing the human oral cavity and its keystone pathogen role in periodontitis pathogenesis. To understand the evolutionary process behind *P. gingivalis* in the context of the *Porphyromonas* genus, in this study, we performed a comparative genomics study with publicly available *Porphyromonas* genomes, focused on four main objectives: (A) to confirm the phylogenetic position of *P. gingivalis* in the *Porphyromonas* genus by phylogenomic analysis; (B) the definition and comparison of the pangenomes of *P. gingivalis* and its relative *P. gulae*; and (C) the evaluation of the gene family gain/loss events during the divergence of *P. gingivalis* and *P. gulae*; (D) the evaluation of the evolutionary pressure (represented by the calculation of Tajima-D values and dN/dS ratios) comparing gene families of *P. gingivalis* and *P. gulae*. Our analysis found 84 high-quality assemblies representing *P. gingivalis* and 14 *P. gulae* strains (from a total of 233 *Porphyromonas* genomes). Phylogenomic analysis confirmed that *P. gingivalis* and *P. gulae* are highly related lineages, close to *P. loveana*. Both organisms harbored open pangenomes, with a strong core-to-accessory ratio for housekeeping genes and a negative ratio for unknown function genes. Our analyses also characterized the gene set differentiating *P. gulae* from *P. gingivalis*, mainly associated with unknown functions. Relevant virulence factors, such as the FimA, Mfa1, and the hemagglutinins, are conserved in *P. gulae*, *P. gingivalis*, and *P. loveana*, suggesting that the origin of those factors occurred previous to the *P. gulae* - *P. gingivalis* divergence. These results suggest an unexpected evolutionary relationship between the *P. gulae* - *P. gingivalis* duo and *P. loveana*, showing more clues about the origin of the role of those organisms in periodontitis.

KEYWORDS

Porphyromonas gingivalis, virulence factors, phylogenomics, orthogroups, dN/dS, Tajima D value, gene gain/loss model

1. Introduction

The *Porphyromonas* genus comprises Gram-negative, asaccharolytic, non-spore-forming, non-motile, and obligate anaerobic species (Paster et al., 1994; Summanen and Finegold, 2015). Its ecological niche is mainly the oral cavity of humans and other mammals, therefore being described by some authors as indigenous to the oral microbiome. It has also been isolated from the urogenital and gastrointestinal tracts and infected wound sites in humans and other mammals (Summanen et al., 2009; Sakamoto and Ohkuma, 2013; Rajilić-Stojanović and de Vos, 2014; Sakamoto et al., 2015; Sato et al., 2015; Zamora-Cintas et al., 2018; Guilloux et al., 2021). There are several officially described *Porphyromonas* species (Gibson and Genco, 2006). Of these species, *Porphyromonas gingivalis* is one of the most studied because of its key role in the pathogenesis of human periodontitis. It is a chronic, non-communicable, immunoinflammatory disease characterized by the irreversible destruction of dental supporting tissues, collectively known as the periodontium. Periodontitis is caused by the constant challenge of the dysbiotic subgingival biofilm, where *P. gingivalis* acts as a keystone pathogen. Therefore, the uncontrolled immune response against these pathogens is responsible for the destructive nature of this disease (Hajishengallis and Lamont, 2014).

Due to its prevalent preference for the oral cavity, *Porphyromonas* evolution could be influenced by niche construction, a process in which the organisms periodically change the local resource distribution in ways that enhance their own fitness. As a keystone pathogen, the large impact on the structure and function of microbial communities can be influenced by perturbations in the microenvironment caused by niche construction, with a concomitant effect on the evolution of itself and the community (Laland et al., 1999). For example, *P. gingivalis* is closely related to *P. gulae*, a member of the oral microbiome of various mammalian hosts (Fournier et al., 2001); additionally, some virulence factors found in *P. gingivalis*, such as the fimbrial proteins, are also present in *P. gulae*, showing high sequence conservation (Hamada et al., 2008; Fujiwara-Takahashi et al., 2020).

P. gingivalis, as a key member of the dysbiotic subgingival microbiota, is capable to colonize, invade, and destroy tooth-supporting tissues (periodontium) via the expression of several virulence factors (Hajishengallis and Lamont, 2014; Zenobia and Hajishengallis, 2015). Despite having a relatively low abundance in the subgingival biofilm, this pathogen has keystone abilities that cause microbial dysbiosis and alter the host's immune response, promoting the pathologic destructive inflammatory environment and potentially influencing the course of other inflammatory disorders, even causing systemic complications (Olsen et al., 2017).

This pathogen synthesizes, expresses, and secretes several virulence factors contributing to its pathogenesis, including lipopolysaccharide (LPS), gingipains, hemagglutinins, and fimbriae, among others (Nakayama and Ohara, 2017). All those factors

contribute to the damage of periodontal tissue and exacerbate inflammatory responses. For instance, LPS released after bacterial lysis and secreted through outer membrane vesicles (OMVs) activates Toll-like receptor 4 (TLR4) (Waller et al., 2016), promoting destructive periodontal inflammation (Zhang D. et al., 2015). The cysteine protease family members, the gingipains, play crucial roles in complex host-pathogen interactions. In multispecies heterotypic communities, its protease activity degrades ligands linked to interspecific aggregation and coaggregation, influencing the structure and composition of biofilms (Jia et al., 2019). Moreover, the degradation of host proteins as a pathogenicity strategy enables the bacterium to evade the immune response (Chen et al., 2023), modulate signaling pathways that increase its virulence (Lourbakos et al., 2001), and acquire essential nutrients during infection (Guo et al., 2010). Additionally, fimbriae play a key role in host tissue adhesion (Chen et al., 2023) as well as bacterial aggregation and coaggregation, two processes necessary for the formation of oral biofilms (Hasegawa and Nagano, 2021). The two main fimbriae systems in *P. gingivalis*, FimA, and Mfa1, are involved in the interaction with the host-epithelial surface, the immune response, autoaggregation, and the aggregation of other oral bacteria, such as *Treponema* spp. and *Streptococcus gordonii* (Chung et al., 2000; Yamada et al., 2005; Hajishengallis et al., 2006; Harokopakis et al., 2006; Eskin et al., 2007; Kuboniwa et al., 2009). The *fim* and *mfa* gene clusters contain a diversity of polymorphisms that could be associated with different genotypes and are often utilized to classify *P. gingivalis* strains (Hasegawa and Nagano, 2021). In the case of periodontitis-like diseases in other mammals, *P. gulae* adhesion systems have been reported as key players in the development of periodontitis; moreover, *P. gulae* can invade human gingival epithelial cells *in vitro*, and its efficiency could be associated with the *fimA* genotype (Inaba et al., 2019). In cats, types B and C fimbriae were frequently detected in subjects with periodontitis, which increased the use of *fimA* as a marker related to virulences in periodontitis (Iwashita et al., 2019).

The current knowledge available indicates that there are several *Porphyromonas* species and that only a few species are involved in periodontal disease, being *P. gingivalis* and *P. gulae*. The availability of genomic information for different *Porphyromonas* strains and metagenome-assembled genomes (MAGs) raises the opportunity for the genomic analysis of several aspects of evolution and development of the pathogenic properties of *P. gingivalis* in the context of its genus and its close relatedness with *P. gulae*. Despite there are some genomic studies comparing *P. gingivalis* strains in order to make associations between genetic content and its pathogenic properties (Chen et al., 2004; Lenzo et al., 2016; Chen et al., 2017; Romero-Lastra et al., 2017; Mendez et al., 2019), currently, there are only a few studies covering possible relationships between evolution and gene content for *P. gingivalis* in comparison with other microorganisms. One remarkable study performed with 32 isolates from 18 species (O'Flynn et al., 2015) showed that *P. gingivalis* genomes contained some distinctive features, such as genes involved in iron transport and heme utilization, whereas lacked other genes (involved in processes such as

protoporphyrin biosynthesis), when they were compared with sequences from other *Porphyromonas* species. Another study (Endo et al., 2015) compared *P. gingivalis* genomes with the sequences from other members of the so-called “red complex” (*Treponema denticola* and *Tannerella forsythia*), a set of organisms involved in the development of periodontal disease (Mohanty et al., 2019); this study predicted metabolic complementations in the fatty acid biosynthesis pathways between those organisms, suggesting a potential effect of pathway complementation in the development of their roles in disease.

Despite the availability of previous studies and genomic information, the relationship between genetic diversity, pathogenesis, and evolution of the adaptation in *P. gingivalis* to its role in human periodontitis remains to be clarified. Considering the potential evolutionary history of *P. gingivalis* in the context of its genus, and the existence of other species such as *P. gulae* exhibiting similar characteristics in terms of niche and mechanism of infection in other mammals, a comparative genomics study with an evolutionary perspective can help to understand how *P. gingivalis* acquired its features to be a keystone pathogen. Since virulence is the product of complex pathogen-host interactions (Diard and Hardt, 2017); the evolutionary analysis of virulence factors could help to explain the acquirement of the lifestyle of a keystone pathogen, is an important milestone to understand the evolution of some members of *Porphyromonas*. In addition to virulence factors, other accessory proteins in the machinery of *Porphyromonas* could show evidence of adaptation to the oral microbiome. Understanding the evolutionary process behind the evolution of *P. gingivalis* as a keystone pathogen could give signs of how other *Porphyromonas* had evolved in a different way to potentially become a keystone pathogen, like potentially could be *P. gulae*. In this study, we describe *Porphyromonas* pangenome and explore the phylogenetic relation of *P. gingivalis* virulence with other members of the genus, especially *P. gulae*, the closest species to *P. gingivalis*.

2. Methodology

2.1. Genome dataset definition

Porphyromonas genomes were selected from NCBI Genbank FTP site (August 2022). The dataset was filtered using two criteria: their taxonomic genomic identity and degree of completeness and contamination. The taxonomic identity was established by using the program ‘classify_wf’ of the GTDB-TK program, version 2.1.0 (Chaumeil et al., 2019), using the database release 207 as the reference, selecting all genomes classified into the *Porphyromonas* genus (“g_Porphyromonas” or “g_Porphyromonas_A”). Genome completeness and contamination were calculated using the program ‘lineage_wf’ from CheckM version 1.1.3 (Parks et al., 2015); only those genomes with completeness equal to or higher than 90%, and contamination below 5%, were selected, as previously suggested for “high-quality drafts” (Bowers et al., 2017).

2.2. Definition of species groups

In order to detect the genomic species represented among the selected genomes, we combined the prediction from GTDB-TK (see above) with the prediction of clusters defined by average nucleotide identity (ANI) values. All genomes were compared in an all-vs-all

manner using FastANI version 1.32 (Jain et al., 2018) with default parameters. The raw pairwise comparison data was filtered, discarding all ANI values below 95%, the classical intra-species boundary for microbial genomes (Richter and Rosselló-Móra, 2009). Filtered pairwise comparisons were analyzed by the MCL program, creating putative genomic species clusters as observed in network clustering (van Dongen and Abreu-Goodger, 2012).

2.3. Annotation, identification of orthogroups, and phylogenomic tree

All members of the final dataset were annotated *de novo* using Prokka, version 1.11 (Seemann, 2014) (relevant parameters: *-metagenome -kingdom Bacteria -addgenes*). Orthogroups from the set of the final dataset were calculated by Orthofinder version 2.5.5 (Emms and Kelly, 2019) with the ‘-og’ parameter. For phylogenomic tree elaboration, a concatenated multiple sequence alignment was constructed from a set of 38 single-copy conserved orthogroups by using MAFFT v. 7.490 (parameters: *-maxiterate 1,000 -localpair*) (Katoh et al., 2019). The alignment was used by *iqtree* v. 2.1.4 (Nguyen et al., 2015) (parameters: *-m TEST -alrt 1,000*) to generate a maximum likelihood-based tree with an aLRT with 1,000 replicates as the branch support test. The phylogenomic tree was visualized using the *Toytree* python package (Eaton, 2020), or with the tool FigTree v.1.4.4¹ when it corresponds. The genomic information from *Tannerella forsythia* (assembly GCA_000238215.1) was utilized as the outgroup for the *Porphyromonas* genus tree due to its relatedness with the genus (Summanen and Finegold, 2015).

2.4. Pangenome analyses and definition

The pangenome represents the collection of all groups of orthologous genes (orthogroups) from a set of genomes (Moldovan and Gelfand, 2018). To analyze the pangenome of *P. gingivalis* and *P. gulae*, separated Orthofinder executions were performed with the proteomes from each cluster without any outgroup (parameters: *-M msa -y*). The orthogroup matrix (including unassigned orthogroups) was obtained for each run and utilized for different pangenome metrics. Pangenome curves were created using the *panplots* function in R (created by SioStef)² using 1,000 permutations. The *alpha* values for Heap’s law equation were calculated using the *curve_fit* function of the *scipy* package in Python, using as objective the equation “ $a * (x ** b)$.” The determination of shell, cloud, “soft-core” and core components of the pangenome was deduced from the complete orthogroup matrix by using Python scripts with the *pandas* package, considering *core* gene families as the orthogroups present in 100% of the strains, *soft-core* groups as present in between 90 and 99.9% of the strains, *shell* as groups present between 89 and 15% of strains, and *cloud* as the gene families present in between 14% and the equivalent to two strains. Unique groups can be deduced from the set of “species-specific orthogroups,” and the “unassigned genes,” both reported by

¹ <https://github.com/rambaut/figtree>

² <https://github.com/SioStef/panplots/>

Orthofinder. Figures were created with *ggplot2* and the *ggarrangement* R packages.

2.5. Predictive proteomic analysis

For predicted proteomes was made by combining the prediction of EggNOG mapper version 2.1.6 (Cantalapiedra et al., 2021) (parameters: “--tax_scope_mode narrowest --tax_scope prokaryota_broad --go_evidence experimental”).

2.6. Gene gain/loss model for *Porphyromonas gingivalis* - *Porphyromonas gulae* divergence tree

In order to create a general gene gain/loss model, all proteomes from genomes from *P. gingivalis* and *P. gulae*, in addition to one genome of *P. loveana* as the outgroup were compared using Orthofinder as previously mentioned. The generated orthogroup matrix was utilized for the generation of a phylogenomic tree. This tree and the binary version of the complete orthogroup matrix were used by the software Count (Csurös, 2010) for the calculation of gene gain/loss rates following the Wagner parsimony model, using convergence criteria were set to a likelihood delta of 0.05 with a maximum of 1,000 rounds, and the same penalty score (equal to 1) for gains and losses. The final tree was represented using the ETE3 Python package (Huerta-Cepas et al., 2016).

2.7. Evolutionary metrics: Tajima D pairwise dN/dS ratios and selection among sites

Tajima's D value is a statistical test that infers rates of rare alleles and assigns scores to the orthogroups, therefore, detecting variation from a neutral model of molecular evolution (Zhang Q. et al., 2015). In order to calculate Tajima's D value, the nucleotide sequences from the coding sequences of the single-copy conserved orthogroups of *P. gingivalis*-only or *P. gulae*-only, when corresponding, were utilized. Sequences were aligned with MAFFT, as mentioned above. Nucleotide multiple alignments were utilized to calculate Tajima's D values using the *tajima.test* function from *pegas* R package.³ Graphs were created with *ggplot2*.

Since that dN/dS is more suitable for comparisons between sequences with considerable distances (Rocha et al., 2006), dN/dS values were computed from the nucleotide sequence of the single-copy orthogroups conserved in both *P. gingivalis* and *P. gulae* genomes, using *P. loveana* as outgroup for both comparisons. Each orthogroup set was aligned using MAFFT (G-INS-i mode); calculation of dN, dS, and ω was made by using CODEML program from the PAML package (Yang, 2007), using the following parameters: “*runmode* = -2, *seqtype* = 1, *CodonFreq* = 3, *model* = 1, *NSsites* = 0, *icode* = 0, *fix_kappa* = 0, *kappa* = 1, *fix_omega* = 0, *omega* = 0.5.” Pairwise comparisons with distances equal to zero, dN/dS > 5, and dS > 10 were

discarded from the analysis. For all pairwise comparisons, the statistical significance among groups was evaluated using the Kolmogorov-Smirnov (*ks.test* in R) test (value of *p* < 0.01).

2.8. Phylogenetic analyses of virulence factors of *Porphyromonas*

Hidden Markov Model (HMM) profile search (using HMMER 3.1b) were performed for the following Pfam entries to search for the following virulence factors: PF06321 (“*P. gingi_FimA*,” detectable in both FimA and FimC proteins), PF15495 (“*Major fimbrial subunit protein type IV, Fimbrillin, C-terminal*,” found in Mfa1 proteins), PF07675 (“*Cleaved Adhesin Domain*”; present in both gingipains and hemagglutinins), PF10365 (*DUF2436*, as a confirmatory hit for gingipains and hemagglutinins), PF01364 (“*Peptidase family C25*,” found in gingipains but not in hemagglutinins), and PF04371 (“*Porphyromonas-type peptidyl-arginine deiminase*”), found in the peptidyl-arginine deiminase (PAD) enzyme associated with protein citrullination in *P. gingivalis* (Gabarrini et al., 2015). Proteins from the *Porphyromonas* dataset were retrieved and aligned with MAFFT (G-INS-i mode). Phylogenetic trees of each PFAM were inferred using *iqtree* v. 2.1.4 (Nguyen et al., 2015) with the maximum likelihood (ML) method with an aLRT with 1,000 replicates as the branch support test (main parameters: *-m TEST --alrt 1,000*). For tree visualization, the FigTree v.1.4.4 software (see text footnote 1) was utilized.

3. Results

3.1. An up-to-date phylogeny of the *Porphyromonas* genus and representation of species groups

The *Porphyromonas* genus comprises a set of several organisms from both human and animal sources (Summanen and Finegold, 2015). Using the content available in the NCBI Genbank Genome repository, with the additional filter using taxonomic and completeness/contamination standards, the *Porphyromonas* dataset utilized in this study comprises 233 genomes, representing 36 genomic species clusters (Figure 1; Supplementary Figure S1; Supplementary Tables S1, S2). The five most represented genomic species in the dataset were *P. gingivalis*, *P. pasteri*, *P. levii*, *P. gulae*, *P. macacae* and *P. cangingivalis*, with 84, 34, 20, 14, 7, and 7 genomes, respectively. Of the set of genomic species, 19 clusters do not have an official species name; in most of those cases, only the GTDB-TK assignment was shown. Moreover, some strains, such as *Porphyromonas* sp. SUG530 (GCA_022772085.1), uncultured *Porphyromonas* sp. SRR2034640_bin.3_metaWRAP_v1.1_MAG (GCA_915070105.1) or uncultured *Porphyromonas* sp. RxaearOaYr_bin.43.MAG (GCA_943912835.1), do not contain any species-level classification in GTDB-TK (Supplementary Table S1). The *P. uenonis* 60-3 strain (isolated from a vaginal sample) is part of a different species cluster than the canonical *P. uenonis* (DSM 23387, isolated from a sacral decubitus ulcer (Finegold et al., 2004)). This latter finding suggests that strain 60-3 could correspond to an undescribed species. The undescribed species (taxa with magenta names in Figure 1) cover an important fraction of the *Porphyromonas* tree,

³ <https://cran.r-project.org/web/packages/pegas/index.html>

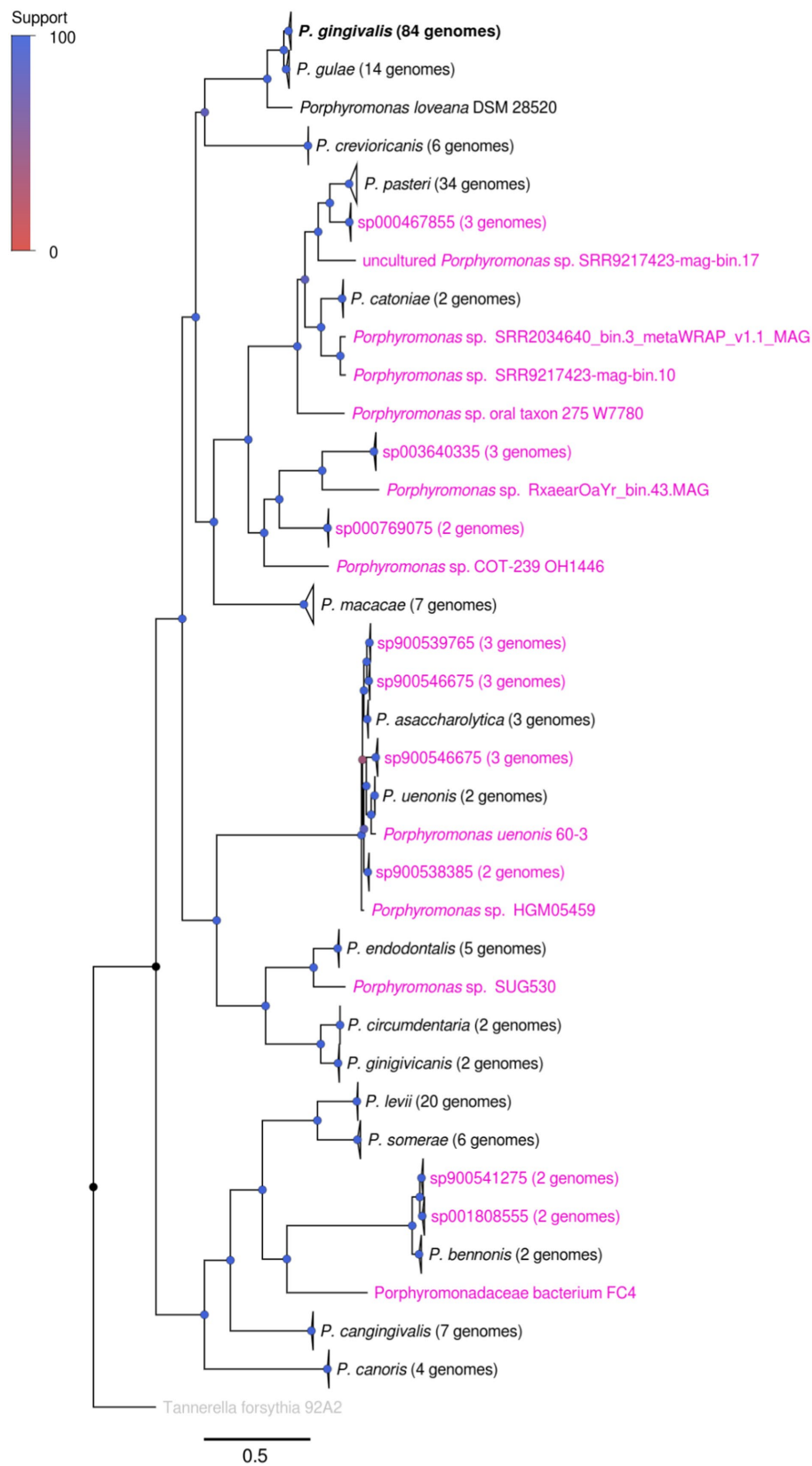


FIGURE 1

Phylogenomic tree showing the relationship between different *Porphyromonas* genomic species. The tree was created from the alignment of 38 single-copy conserved protein families, using the maximum likelihood method in IQTREE, with the use of the approximate likelihood ratio test (aLRT) as the branch support test. Taxa names colored in magenta represented genomes or species groups without an official name classification, according

(Continued)

FIGURE 1 (Continued)

to GTDB-TK and manual review (see [Supplementary Table S2](#)); taxa names in black represented species groups with an official name. Branch support is represented by colors in each node, according to the *Support* color scale.

suggesting an extensive uncharacterized genetic content among the genus.

This phylogenomic tree, made from 38 conserved orthogroups ([Figure 1](#)), confirmed several previously reported phylogenetic associations between different members of the *Porphyromonas* genus. For example, the closed relationship between *P. gingivalis* and *P. gulae* has been previously shown in different studies ([Fournier et al., 2001](#)). Interestingly, this study also found a strong relatedness between the *P. gulae* - *P. gingivalis* duo with *P. loveana*, an organism isolated from the oral cavity of a musky rat kangaroo, a marsupial ([Bird et al., 2016](#)). In addition to this finding, the close relationship between *P. uenonis* and *P. asaccharolytica* was also shown previously ([Finogold et al., 2004](#)), as well as the close relationship between *P. levii* and *P. somerae* ([Summanen et al., 2009](#)). However, from now on, this study will be focused on the genome set of 84 *P. gingivalis* and 14 *P. gulae* strains.

3.2. Comparing the pangenomes from *Porphyromonas gingivalis* and *Porphyromonas gulae*

Both the mutually highly related *P. gingivalis* and *P. gulae* are involved in periodontal diseases in their respective hosts ([Lenzo et al., 2016](#)). Therefore, studying evolutionary relationships between those organisms could be important to analyze to explain the origin of the *P. gingivalis* phenotype. Given that we found 84 high-quality genomes for *P. gingivalis* and 14 for *P. gulae*, we first compared them by analyzing each pangenome separately. For this purpose, the Orthofinder output was used to generate a gene matrix to calculate core, “persistent,” shell, cloud, and unique gene families, as well as to calculate the pangenome accumulation curves ([Figures 2A,B; Table 1](#)). Additionally, the percentages of different COG categories (assigned to each gene by using EggNOG mapper) were plotted among the different fractions ([Figures 2C,D](#)), and the Log₂ of the ratios between the Core and the “Accessory” (the combination of “persistent,” shell, cloud, and unique gene fractions) among the different genomes were represented into boxplots ([Figures 2E,F](#)). The pangenome accumulation curves, fitted to the Power Law equation ([Figures 2C,D](#)), suggest that both pangenomes are open ($\gamma > 0$, see [Table 1](#)), a feature that may reflect the contact of *P. gingivalis* and *P. gulae* with several other microorganisms in a complex environment ([Rouli et al., 2015](#)), as the mammalian oral cavity. Comparing the different functional COG categories suggests that categories associated with biosynthesis and metabolism (represented by letters C to H) are more represented in the Core and persistent fractions in both *P. gingivalis* and *P. gulae* pangenomes. Moreover, the calculation of the Log₂ ratios between the Core and “Accessory” fractions for both pangenomes ([Figures 2E,F](#)) showed that categories C, G, H, I, J, and M (representing the COG categories “Energy production and conversion,” “Carbohydrate transport and metabolism,” “Coenzyme transport and metabolism,” “Lipid transport and metabolism,”

“Translation, ribosomal structure and biogenesis,” and “Cell wall/membrane/envelope biogenesis,” respectively) have a median Log₂ (Core/Accessory) higher than 1. In the case of the *P. gingivalis* pangenome, categories P (“Inorganic ion transport and metabolism”), O (“Posttranslational modification, protein turnover, chaperones”), and E (“Amino acid transport and metabolism”) have also higher Log₂(Core/Accessory) ratios. COG categories with higher Log₂ (Core/Accessory) ratios in *P. gulae* only were D (“Cell cycle control, cell division, chromosome partitioning”), F (“Nucleotide transport and metabolism”), and Q (“Secondary metabolites biosynthesis, transport and catabolism”). In all those cases, the comparison between the Core and Accessory groups was significant ($p < 0.01$, two-sided Wilcoxon Rank Sum test in R). These results suggest that, in general, housekeeping and central metabolism/transport functions were more represented in the Core than in the Accessory gene set. This feature has been observed in other studies with other bacterial models ([Gaba et al., 2020; Hyun et al., 2022](#)). In contrast, genes without any COG hit (marked as “@” category in [Figure 2](#)) were most prevalent in the accessory fraction of the pangenome in both *P. gingivalis* and *P. gulae*. In the case of *P. gulae*, category X (“Mobilome: prophages, transposons”) was also more represented in the accessory than in the core gene fraction of the pangenome. In *P. gingivalis*, genes from category V (“Defense mechanisms”) had more negative Log₂(Core/Accessory) ratios, whereas category X was not suitable for Log₂ calculation, since the Core fraction had zero genes in that category ([Figure 2C](#)). It is worth noticing that differences from data from categories X, V, and without COG between Core and Accessory groups were significant as well ($p < 0.01$, two-sided Wilcoxon Rank Sum test).

The core gene fractions from *P. gingivalis* and *P. gulae* contained common and differentiative features. In order to compare them to define which functions were exclusive to the core fractions in each organism, a bidirectional comparison was performed with representatives from each gene from the core between those two species ([Figure 3](#)). This comparison showed a set of 1,199 bidirectional hits, and a total of 136 and 168 genes without a clear bidirectional hit in the core pangenome components of *P. gingivalis* and *P. gulae*. Respectively. Among the functional categories observed in the common and differential core contents, two cases are worth noticing between *P. gingivalis* and *P. gulae*. First, in the first one, there are more genes associated with category V (“Defense mechanisms”) in *P. gulae* than in *P. gingivalis*; some of those functions are associated with a Cas-Cmr CRISPR system, several genes for multidrug efflux pump AcrA/B proteins and ABC-type antimicrobial peptide transport system, permease component (SalY)- encoding genes. In counterpart, categories H and O (“Coenzyme transport and metabolism” and “Posttranslational modification, protein turnover, chaperones,” respectively) are more present in the core of *P. gingivalis* than in *P. gulae*; some of those genes encode for ABC systems for siderophores, a set of serine proteases and a cation-transporting P-type ATPase (data not shown).

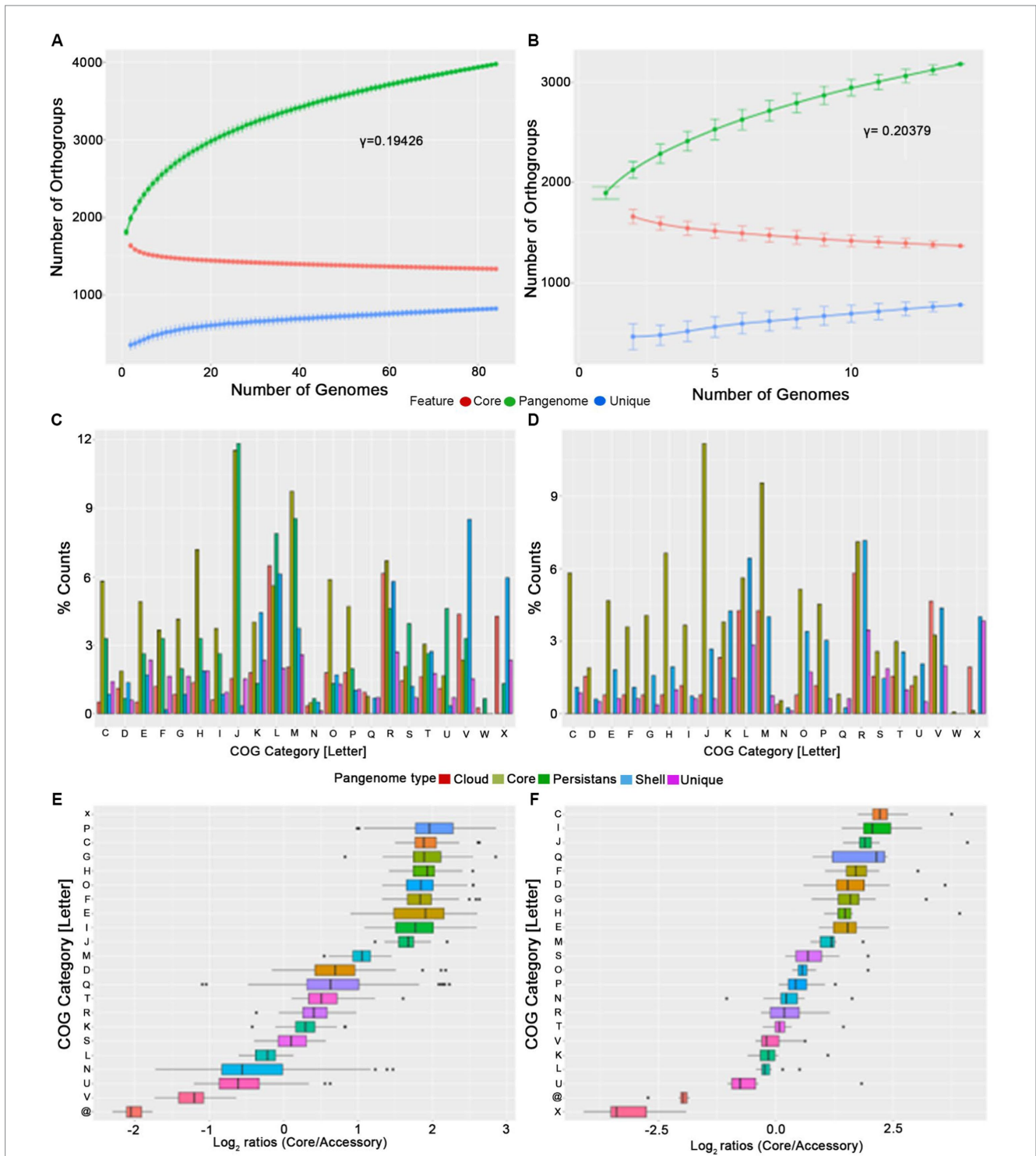


FIGURE 2
 Pangenome analysis of *P. gingivalis* (A,C,E) and *P. gulae* (B,D,F). The analysis included pangenome accumulation curves (A,B), a barplot of the relative percentage of counts for each COG category (C,D), and the Log₂ ratios of the percentage of core vs. accessory genes for each category (E,F).

3.3. Gene gain/loss model in the *Porphyromonas gingivalis*/*Porphyromonas gulae* ancestry

Our phylogenomic analysis of the *Porphyromonas* genus confirmed that *P. gingivalis* and *P. gulae* are strongly related, as well as

that *P. loveana* is also directly related to them (Figure 1). In order to check about the genetic gain and loss events during divergence between *P. gingivalis* and *P. gulae*, the orthogroup and phylogenetic information from 84 *P. gingivalis* and 14 *P. gulae* genomes (including one *P. loveana* genome as an outgroup) were used to create a gene gain-loss model to represent the number of genes present in the

predicted common ancestors across the phylogeny, by using the tool Count with Wagner parsimony analysis. The gain-loss model tree for this set of genomes is shown in Figure 4A (detailed version in Supplementary Figure S2), where each node has the number of genes analyzed, as well as the number of genes gained and lost. A set of 1,662 gene families is predicted in the ancestor of the divergence between *P. gulae* and *P. gingivalis* lineages; in this context, the last common ancestor (LCA) of all 84 *P. gingivalis* genomes contained 1,661 genes, involving the loss of 95 genes and the acquisition of 94 genes in comparison with the previous ancestor. In the LCA of *P. gulae*, 1,633 genes were found in the predicted ancestor, with a gain of 35 genes and a loss of 65 genes. Across the intraspecific branches in the *P. gulae/P. gingivalis* tree, relative gene gain/loss events were relatively

low, occurring with a higher frequency of gain and loss events when nodes are closer to the terminal leaves (Supplementary Figure S2).

In order to understand which classes of functions were acquired in the differentiation process between *P. gingivalis* and *P. gulae*, the functions of the genes found in the LCAs of all *P. gingivalis* and/or *P. gulae* were annotated with EggNOG mapper using COGs (Clusters of Orthologous Genes) categories at the root level. Considering the predicted contents of the LCAs for *P. gingivalis* and *P. gulae*, the next step is to define the content acquired and missed during the formation of those predicted ancestors. According to COGs assignments, the gained genes in the LCA for *P. gulae* were mainly associated with “K” and “L” categories (Figures 4B,C), related to “Transcription” and “Replication, recombination, and repair,” with 0.18 and 0.18%, respectively. However, this LCA also suffered the loss of several genes involved in the “L” category (0.61%), and in other more generic categories such as “S” (“Function unknown”), as well as genes without any category. Some categories, such as “E” (Amino acid transport and metabolism), were more acquired in this predicted ancestor. Some of the categories with only loss were “O” and “P” (Posttranslational modification, protein turnover, chaperones, and Inorganic ion transport and metabolism), with one and four genes.

All the 94 acquired genes in the LCA of *P. gingivalis* belonged to the Unknown function category (Figures 4D,E); on the other hand, the missed genes with no COG category in this node corresponded to 20 OGs. As well as *P. gulae* “K” and “L” were the won categories with high values (0.42 and 0.60%). Interestingly, category “E” (related to amino acids) has more OGs loss compared to those won it and with *P. gulae*. Some categories with only won are the “D” (Cell cycle control, cell division, chromosome partitioning) and “J” (Translation, ribosomal structure, and biogenesis). As the *P. gulae* LCA, categories “O” and “P” have been lost (0.12 and 0.36%).

A BlastP analysis against SwissProt database for the genes acquired by *P. gulae* and *P. gingivalis* LCAs was performed to obtain more clues about the function of those genes. Results showed that 34 out of 36 proteins for *P. gulae* and 85 out of 94 for *P. gingivalis*, were predicted

TABLE 1 Main statistics from pangenome analysis for *P. gingivalis* and *P. gulae*.

	<i>P. gingivalis</i>	<i>P. gulae</i>
Total genomes	84	14
Total Pangenome Genes	3,976	3,179
Core genes (in 100% of genomes)	1,335	1,367
Soft-core genes (<100%, more than 95%)	145	0
Shell genes (<95%, more than 15%)	550	786
Cloud genes (<15%, at least in two genomes)	1,124	247
Unique genes (in just one genome)	822	779
gamma value (power law equation)	0.19426	0.20379

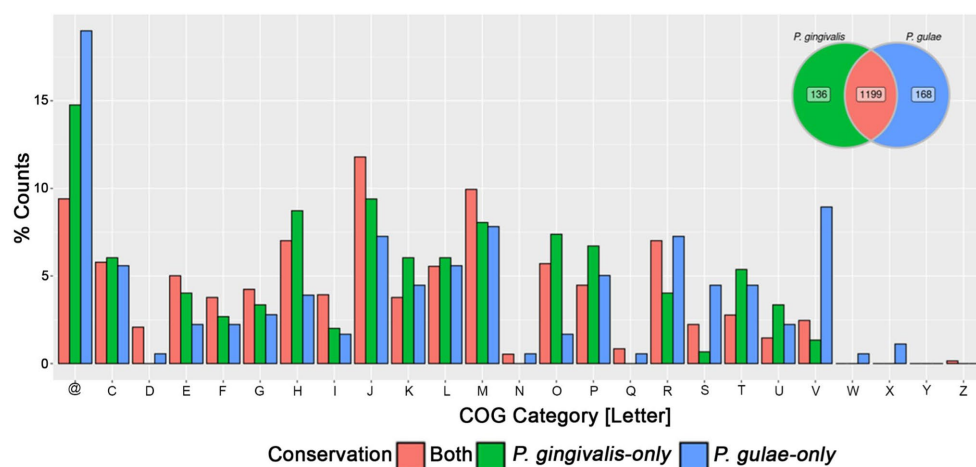
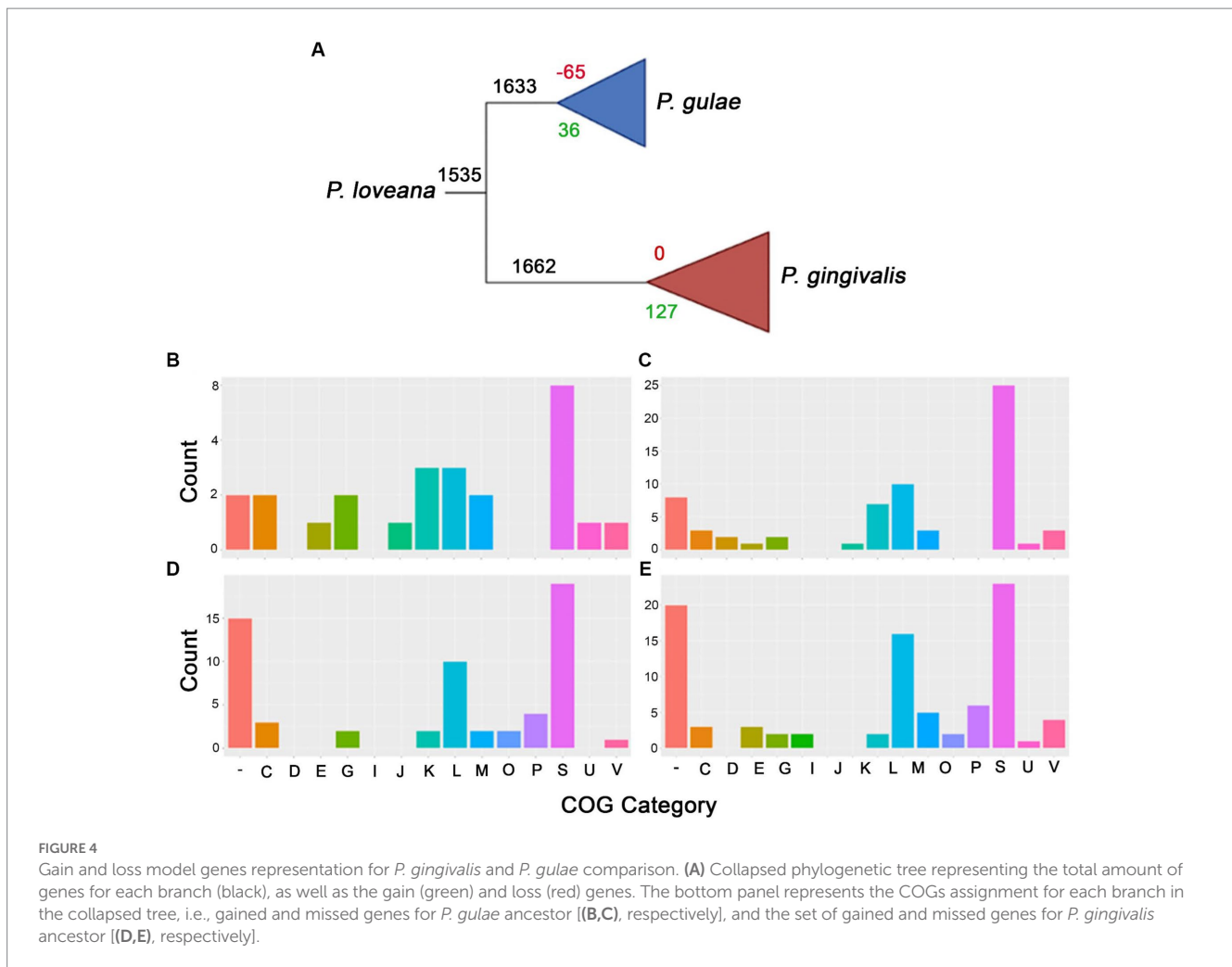


FIGURE 3

Comparison of the pangenome core of *P. gingivalis* and *P. gulae*. Bidirectional hits between the pangenome cores of each organism were represented in the Venn diagram (upper right), as well as the gene representatives without bidirectional hits in either *P. gingivalis* or *P. gulae*. The barplot of the relative percentage of counts for each COG category represented the genes found in the core from both organisms or found only in each pangenome core. @: Genes without any COG.



as hypothetical proteins. Global inspections of this protein and their functions for *P. gulae* (data not shown) show that 19 of those 34 sequences were hypothetical proteins; the rest corresponded to genes associated with secretion systems, transcriptional regulators, post-translational modifications, enzymes involved in metabolism, unknown protein domains, mobilome proteins, and DNA interaction proteins. For *P. gingivalis*, 46 of those 85 genes were hypothetical proteins; the rest encoded genes associated with secretion systems, transcriptional regulation, post-translational modifications, stress response and pathogenicity, DNA plasmid partition, and phage interaction.

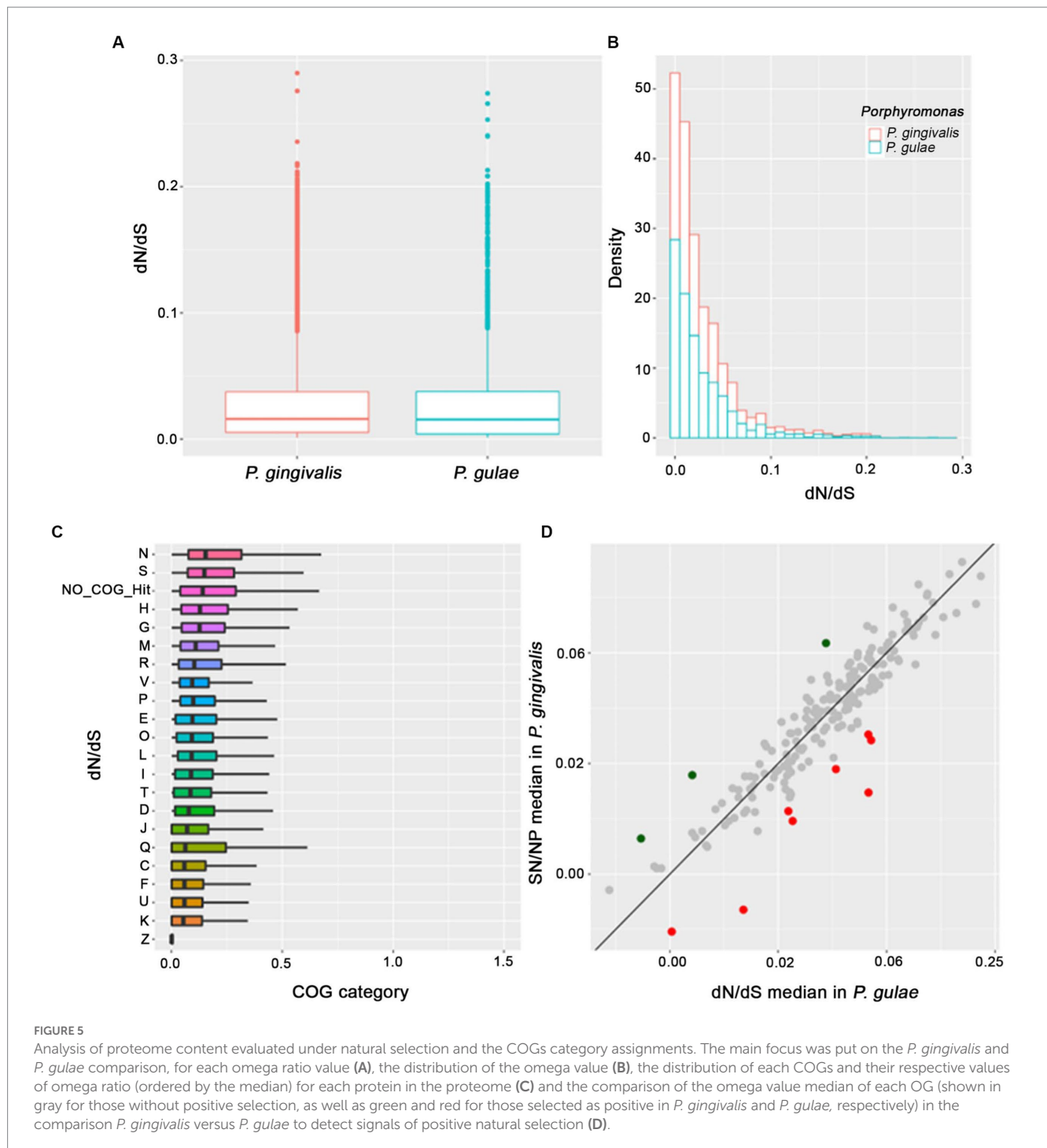
3.4. Evolutive pressure in the shared core orthogroups from *Porphyromonas gingivalis* and *Porphyromonas gulae*

The presence of shared genes in the *P. gulae* and *P. gingivalis* lineages, raises the question about the potential effect of natural selection in the different species. In order to evaluate this possibility, the shared orthogroups of *P. gingivalis* and *P. gulae* were calculated (via Orthofinder) using the 99 genomes from these organisms (and the *P. loveana* genomes as the outgroup) and the common single-copy

core orthogroups were utilized for the calculation of the evolutionary selection parameters (dN, dS, t, and dN/dS ratio).

The distribution of the dN/dS ratio between *P. gingivalis* and *P. gulae* showed an expected behavior of a long proportion of genes with dN/dS < 1 (Figures 5A,B), which is a signal of purifying selection in most of the genes. The mean for the dN/dS ratio was 0.028, compared to the mean of 15.185 of dS, showing a higher rate of synonyms mutations in the OGs set (data not shown for dS alone version and Figure 5). The means of the ratios between *P. gingivalis* and *P. gulae* are similar (0.0282 and 0.0273, respectively, with no significant difference according to the Kolmogorov–Smirnov test), however, in the same way, the mean of the dS between the two species are similar (15 and 16.1, respectively), this result shows that most of the orthogroups between both species do not have strong evidence of positive selection. If the comparisons are divided by the functional category of their genes, COGs categories with the highest dN/dS ratio medians can be found (Figure 5C); the analysis of the pairwise distributions showed that the N and S COG categories (“Cell motility” and “Unknown function,” respectively) harbored the highest median for comparisons. By contrast, K and Z categories (“Transcription” and “Cytoskeleton,” respectively) possessed the lower medians for dN/dS values.

Considering this feature, we perform an additional analysis where each orthogroup is compared between their values for the



P. gingivalis - *P. loveana* and *P. gulae* - *P. loveana* pairwise calculations. In Figure 5C, the scatterplot represents the median of each orthogroup considering their values for the *P. gingivalis* - *P. loveana* and *P. gulae* - *P. loveana* calculations. This comparison found 11 orthogroups with differentially higher ratios (3 for *P. gingivalis* and 9 for *P. gulae*). Using a Blastp analysis versus the SwissProt database, a more detailed description of each orthogroup was obtained (Table 2). Those genes harbored different functions, including a variety of housekeeping functions. For example, in *P. gingivalis* those genes were related to anaerobic respiration, glycogen storage, and DNA damage recognition, as 4-hydroxybutyryl-CoA dehydratase/vinyl acetyl-CoA-Delta-isomerase, glycogen synthase and Uvr ABC system protein B,

respectively. On the other hand, *P. gulae* has protein related to cell wall maintenance, ion membrane translocation, anaerobic respiration, signal recognition receptor, as well as amino acid, and fatty acid metabolism.

3.5. Tajima values for *Porphyromonas gingivalis* and *Porphyromonas gulae* core orthogroups

Tajima D value is a statistical test that uses DNA data to detect natural selection (Korneliusson et al., 2013). In order to detect general

TABLE 2 List of orthogroups found under positive selection with their respective protein function association in the comparison between *P. gingivalis* and *P. gulae*.

Species feature	OG	Function
Higher in <i>P. gingivalis</i>	OG0000556	4-hydroxybutyryl-CoA dehydratase/vinylacetyl-CoA-Delta-isomerase
Higher in <i>P. gulae</i>	OG0000576	D-alanine--D-alanine ligase
Higher in <i>P. gulae</i>	OG0000350	Ion-translocating oxidoreductase complex subunit G
Higher in <i>P. gulae</i>	OG0000351	Na(+)-translocating ferredoxin:NAD(+) oxidoreductase complex subunit E
Higher in <i>P. gingivalis</i>	OG0000702	Glycogen synthase
Higher in <i>P. gulae</i>	OG0000267	Signal recognition particle receptor
Higher in <i>P. gulae</i>	OG0000713	Acetoacetyl-CoA:acetate/butyrate CoA transferase alpha subunit
Higher in <i>P. gingivalis</i>	OG0000394	UvrABC system protein B
Higher in <i>P. gulae</i>	OG0000720	Acyl-CoA dehydrogenase, short-chain specific

trends in the conserved CDS in both *P. gingivalis* or *P. gulae*, the set of single-copy core genes obtained from the previously analyzed *P. gingivalis* and *P. gulae* pangenomes separately, we performed the test with 1,335 core families of *P. gingivalis* and 1,367 core families from *P. gulae*. Figure 6A shows the histogram curves densities of D value for two core orthogroups. *P. gingivalis* has a peak between 0 and -1.5 . The negative values are associated with an excess of rare variation, which is interpreted as population growth or positive selection (Carlson et al., 2005). *P. gulae* Tajima's value distribution is more shifted to the negative but closer to 0 than *P. gingivalis*. Overall, *P. gingivalis* D-value distribution has a higher negative shift than genes from *P. gulae* and it was significant (p -value $< 2 \times 10^{-16}$, Welch Two Sample t -test; Figure 6B). *P. gingivalis* may exhibit lower values either because of positive selection (more rare variants) or because a population is recovering or expanding after a bottleneck effect. The bottleneck effect refers to a reduction in the population size, resulting in the loss of genetic diversity of the species (Peery et al., 2012). In contrast, the more positive distribution of *P. gulae* values may reflect a possible population contraction or allele frequency stabilization associated with the emergence of this organism.

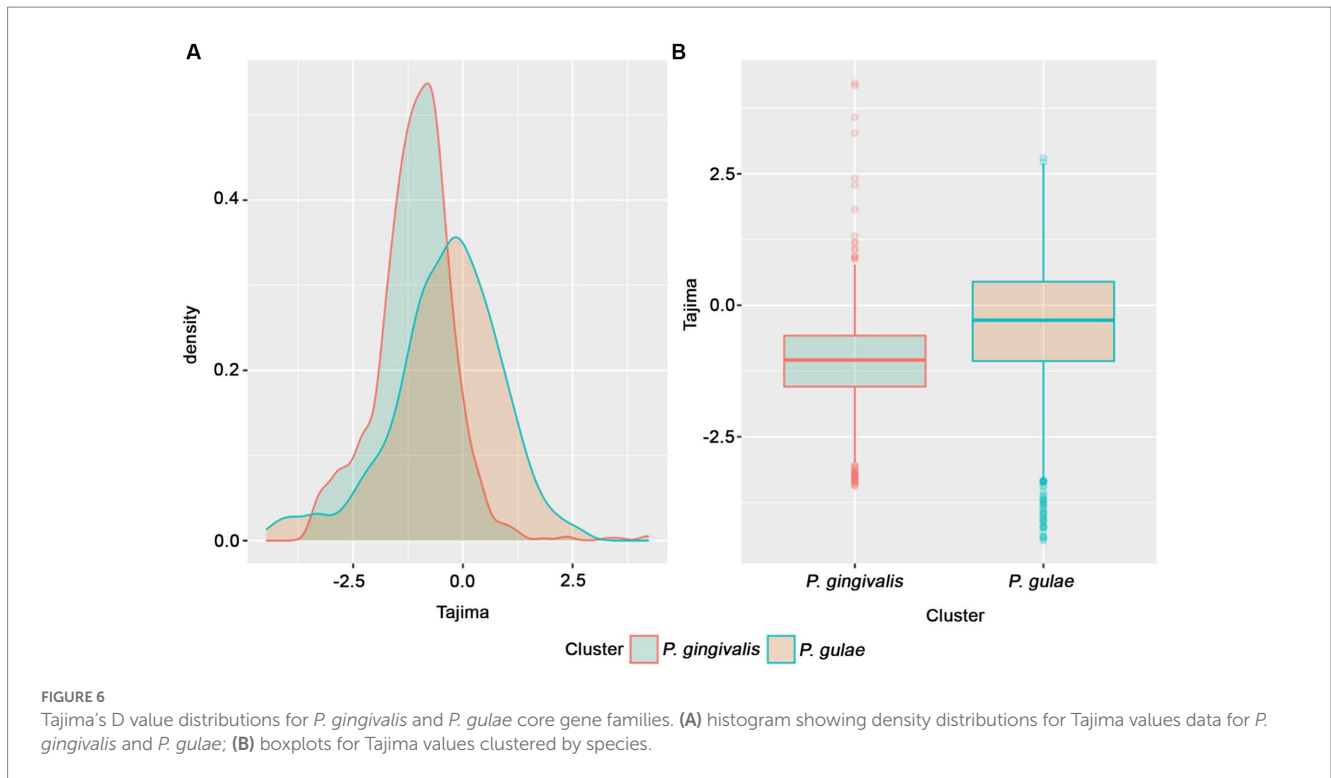
To observe these results in more detail, the top ten OGs with the highest and lowest Tajima values of *P. gingivalis* were retrieved (Table 3). The most positive Tajima value was 4,222; which corresponds to a gene that encodes a putative RagB, which is part of an outer membrane transporter protein that functions as a peptide importer (Madej et al., 2020). Other genes found encoded proteins with functions such as TonB-dependent receptor P3, putative tyrosine-protein kinase in *cps* locus, dipeptide and tripeptide permease B, UDP-glucose 6-dehydrogenase TuaD, Peptide chain release factor 2, Major fibrium tip subunit FimD, and polysaccharide biosynthesis/

export protein, in addition with some hypothetical proteins. These sets of genes have the most neutral/balanced selection in *P. gingivalis*. They are involved in multiple cellular processes, like heme metabolism homeostasis, heme, iron, and amino acid uptake from the host, regulation, and transport of capsular polysaccharides, protein synthesis, and function of bacterial fimbriae.

On the other hand, the top ten orthogroups with the lowest Tajima's D value (Table 4), are genes with an exacerbated proportion of rare alleles (values between -3.22 and -3.095). The most negative one encodes a hypothetical protein, that matches (by blast analysis in the NCBI platform) with another TonB-dependent receptor, which is involved in iron and heme uptake from the host. The other proteins are: calcium-transporting ATPase 1, DUF4248 domain-containing protein, cAMP-activated global transcriptional regulator CRP, type VI-B CRISPR-associated RNA-guided ribonuclease Cas13b, 10 kDa chaperonin, transcriptional regulatory protein QseB, Acyl-CoA dehydrogenase C-terminal domain-containing protein, peptidoglycan-N-acetylglucosamine deacetylase, and GNAT family N-acetyltransferase. Their functions are related overall to iron and heme uptake from the host, calcium, and cAMP intracellular levels homeostasis, a prokaryotic defense mechanism against foreign genetic elements, regulation of biofilm formation, fatty acid metabolism, and cell wall modifications.

3.6. Virulence factor conservation and evolution among *Porphyromonas* genomes

P. gingivalis is known to produce a set of virulence factors involved in the formation of the periodontal disease (Aleksijević et al., 2022; Nuñez-Belmar et al., 2022). Among them, FimA and Mfa1 are filamentous structures essential for the interaction of *P. gingivalis* with other bacteria of the biofilm and with host cells (Nagano et al., 2018). *P. gulae* also was shown to exhibit these adhesive molecular systems (O'Flynn et al., 2015). The major and minor fimbriae exhibit adhesion properties that are suggested to play a central role in periodontitis development (Mendez et al., 2019). Other virulence factors are the gingipains and the Hemagglutinins (Hag). Hag proteins are bacterial surface glycoproteins associated with bacterial adhesion to erythrocytes and host cells (Mendez et al., 2019). For *P. gingivalis* metabolism the heme acquisition is essential, thus, hemagglutinins adhesive domain function is related to this process, facilitating the acquisition of heme. Hag-encoding genes (*hagA*, *hagB*, and *hagC*) are described as relevant virulence factors of *P. gingivalis* (Nakayama and Ohara, 2017). Gingipains are lysine- and arginine-specific proteinases secreted by *P. gingivalis*, and they are considered as one of the most important virulence factors in this organism; their principal function is the degradation of host proteins and the processing of the fimbriae subunits (Lunar Silva and Cascales, 2021). Finally, peptidyl-arginine deiminase (PDA) are enzymes previously reported in *P. gingivalis* capable of producing protein citrullination associated with their capabilities in microbe-host interaction and a role in some diseases such as arthritis (Maresz et al., 2013). The differential degree of presence/absence of orthogroups among *Porphyromonas* may also occur in the known virulence factors associated with *P. gingivalis* in periodontitis and other diseases. To check for other hits for a variety of the main virulence factors, a *hmmmer* search for different Pfam domains (Supplementary Table S3) was performed, searching for



families associated with previously reported virulence factors, with a posterior phylogenetic analysis of those hits. From the PFAM database, this analysis utilized the families identified as PF06321, containing sequences for FimA (genotypes I, II, and IV) and FimC; PF15495 containing Mfa1 and other related proteins; and PF010365, containing hemagglutinin and other proteins such as Lys-gingipains and Arg-gingipains.

These analyses suggested that the number of genes containing several of those domains exhibited a differentiative pattern in the group of *P. gingivalis* - *P. gulae* - *P. loveana*. For example, there are more genes containing the PF06321 domain (“*P_gingi_FimA*”) in those organisms than in the rest of the Porphyromonads. Hits for PF06321 and PF15495 (“*Major fimbrial subunit protein type IV, Fimbrillin, C-terminal*”) domains were widely distributed among *Porphyromonas* genomes, being detected in multiple copies in *P. gingivalis*, *P. gulae*, *P. loveana*, and *P. macacae*. A phylogenetic tree based on amino acid sequences of FimA (Figure 7) and Mfa1 (Figure 8) from the selected *Porphyromonas* genomes. The original PF06321 FimA phylogenetic tree (Supplementary Figure S3) showed two main branches: one of them represented the FimC (proteins from *P. gingivalis*, *P. gulae*, *P. loveana*, *P. levii*, and *P. crevioricaris*), and the second clade represented FimA from several *Porphyromonas*. The FimA clade is colored by each genotype (type I, type II, and type IV in Figure 7), and represented proteins from *P. cangingivalis*, *P. canoris*, *P. levii*, *P. uenonis*, *P. endodontalis*, *P. macacae*, *P. levii*, *P. gingivalis*, *P. gulae*, and *P. loveana*. Genotype II FimA proteins were detected in the trifecta *P. gingivalis*, *P. gulae*, and *P. loveana*, whereas genotypes I and IV were found in *P. gingivalis* and *P. gulae*. Interestingly, FimA homologs were also detected in other members of the genus, such as *P. endodontalis*, *P. levii*, *P. cangingivalis*, *P. uenonis* and *P. canoris*; moreover, the most related to the ancestral branches were found in *P. cangingivalis* and

P. canoris, suggesting that this branch could predate the origin of the *P. gulae* - *P. gingivalis* duo.

The minor fimbriae Mfa1 was also present in a variety of species of the *Porphyromonas* genus (Figure 8), being detected in *P. gingivalis*, *P. gulae*, *P. loveana*, *P. macacae*, *P. levi*, *P. uenonis*, *P. crevioricaris*, *P. gingivicanis*, *P. circumdentaria*, *P. endodontalis*, *P. somerae*, and *P. asaccharolytica*. The original tree PF15495 phylogenetic tree (Supplementary Figure S4) showed two main branches, one of them clearly composed by Mfa1 homologs. In this subtree, two lineages of *P. gingivalis* - *P. gulae* - *P. loveana* were detected: One is associated with *P. macacae* and *P. levi*, which each are isolated from dog oral cavity and cow skin lesions. The other lineage is associated with *P. endodontalis*, which is isolated from human subgingival biofilm and saliva, besides has been closely associated with periapical lesions of patients with periodontitis (Lombardo Bedran et al., 2012). This phylogenetic configuration suggests that, in *P. gulae* and *P. gingivalis*, the *mfa1* genes could have two origins, dictated by their transfer from other members of the *Porphyromonas* genus.

The phylogenetic tree with protein sequences filtered by PFAM PF010365 from the *Porphyromonas* dataset (Figure 9) included proteins such as the hemagglutinins and the gingipains [divided into lysine-gingipains, arginine-gingipains, proteases with an adhesive hemagglutinin-like domain (Mysak et al., 2014)], as well as a set of hypothetical proteins. Lysine-gingipains are exclusively associated with *P. gingivalis*, whereas, the Arginine-gingipains were associated with *P. loveana* and *P. gingivalis*. The hemagglutinin A lineage was represented by two branches: one with *P. gulae* and *P. loveana* proteins, and the other with *P. gingivalis* and *P. gulae* proteins. The presence of hemagglutinins and gingipains in *P. loveana* may suggest that this species could have a similar pathogenic phenotype as observed in *P. gingivalis*.

TABLE 3 List of orthogroups covering the 10 highest Tajima values of in the *P. gingivalis* core, including their protein names and functions ($p < 0.05$).

Top position	OG	Tajima value	Protein	Function
1	OG0000294	4.2226932	Hypothetical protein (RagB)	RagB is an outer membrane transporter protein that functions as a peptide importer in <i>P. gingivalis</i> .
2	OG0000293	4.17820696	TonB-dependent receptor P3	Protein involved in the acquisition of essential nutrients, such as iron and heme, from the host
3	OG0000559	3.574256432	Hypothetical protein	Unknown
4	OG0001241	3.348000822	Hypothetical protein	Unknown
5	OG0000454	3.275760793	Putative tyrosine-protein kinase in cps region	The cps (capsular polysaccharide synthesis) region is responsible for the biosynthesis of the capsular polysaccharides
6	OG0001351	2.715270824	Dipeptide and tripeptide permease B	Protein involved in the uptake of peptides that are used by the bacterium for the synthesis of various virulence factors, such as gingipains
7	OG0000872	2.418843779	UDP-glucose 6-dehydrogenase TuaD	Biosynthesis of the bacterial cell surface polysaccharide
8	OG0000873	2.287357949	Peptide chain release factor 2	Termination of protein synthesis during bacterial translation
9	OG0001381	1.90044368	Major fimbrium tip subunit FimD	Assembly and function of the bacterial fimbriae
10	OG0000455	1.826818797	Hypothetical protein (polysaccharide biosynthesis/export protein)	Assembly of the polysaccharide chain into a repeating unit, as well as its transport across the cell membrane

For the case of the PAD proteins, a phylogenetic tree using those sequences showed that *Porphyromonas* PF04371 hits were distributed in two lineages (Supplementary Figure S5). One comprises the PAD proteins from *P. gingivalis*, *P. gulae*, and *P. loveana*. The other lineage contains the putative agmatine deiminases from nearly all members of the genus. This two-branch divergence suggests that the previous protein citrullination activity involved in pathogenic features in *P. gingivalis* could also be present in *P. gulae* and *P. loveana*.

4. Discussion

4.1. The uncovered diversity of the *Porphyromonas* genus and its relationship with the *Porphyromonas gingivalis* - *Porphyromonas gulae* duo

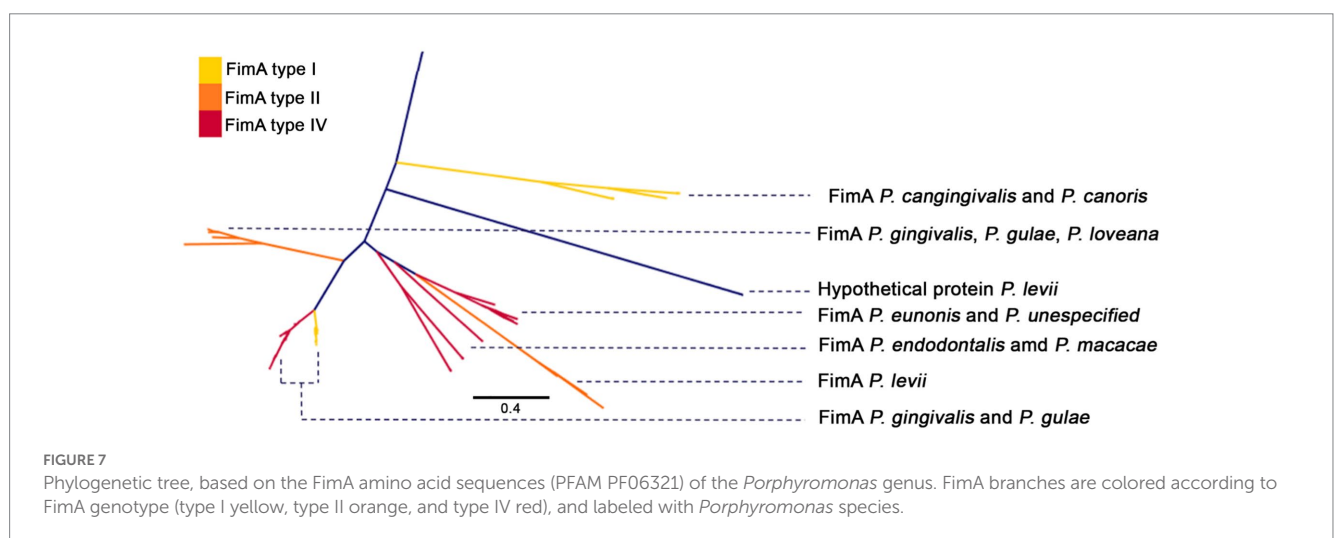
In order to investigate how *P. gingivalis* could become a pathogen in the context of the *Porphyromonas* genus, the first aspect we needed to analyze is to validate its position in relation with its relatives. We performed with a set of 233 high-quality *Porphyromonas* genomes a combination of a phylogenomic analysis with the use of genomic metrics for the establishment of genomic species clusters (Maturana and Cárdenas, 2021). The phylogenetic structure of the genus (Figure 1) showed that there are some clades mostly composed of undescribed species, most of them already classified by the GTDB (Supplementary Table S1). This feature has become common in

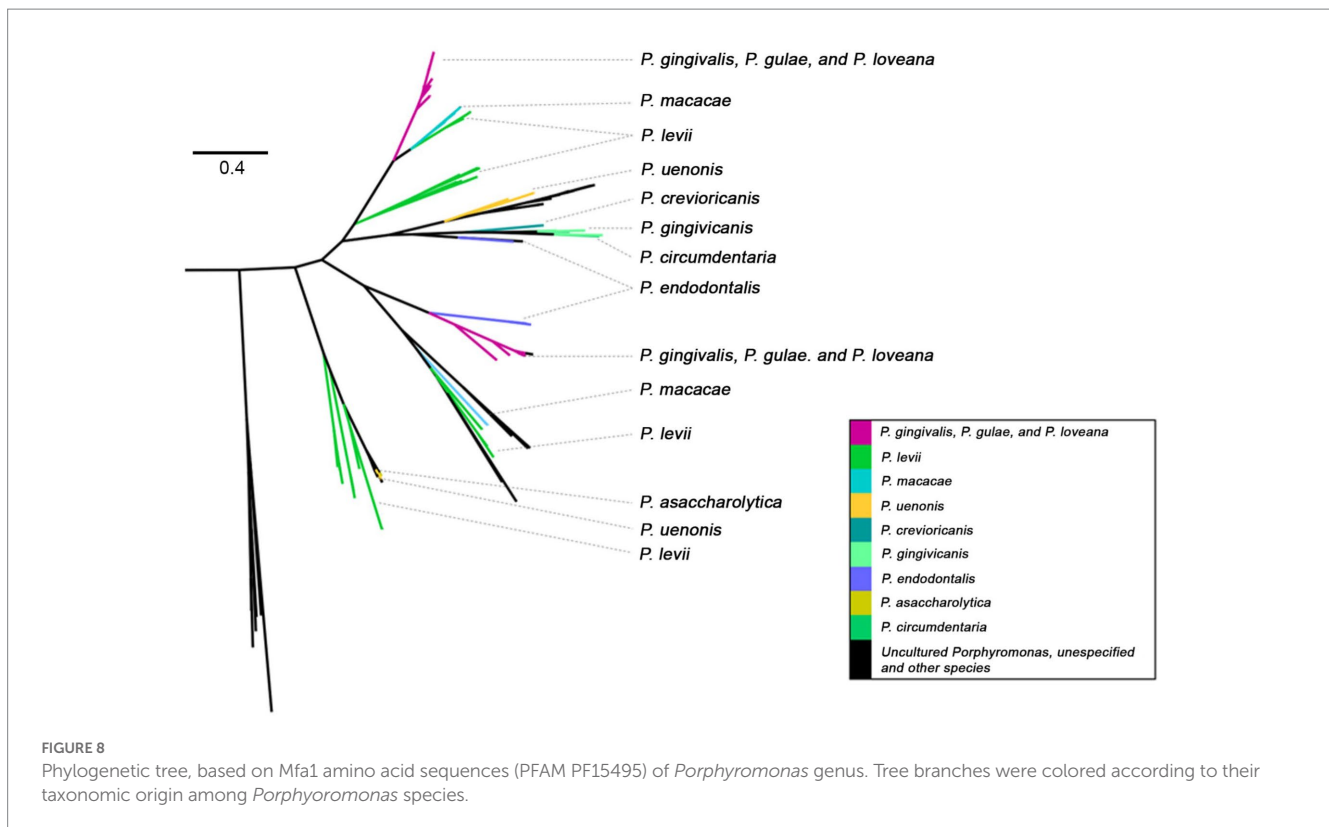
several genera among bacterial and archaeal species as a reflection of their novelty in specialized databases (Chaumeil et al., 2022). Additionally, the phylogenomic analysis may propose new species or genera from previously described and undescribed strains (Sawana et al., 2014; Gupta et al., 2020). In this case, several undescribed *Porphyromonas* species could be representatives of different oral cavity inhabitants, as well as colonizers from other body sites, with their roles in host health and disease yet to be understood. For example, two undescribed species (mentioned as “oral taxon 278” and “oral taxon 275”) detected in plaque samples (Mei et al., 2020), and the misnamed strain 60–3 (named as a *P. uenonis* strain, despite it is not part of the species) may represent some relevant new species remaining to be officially named and described.

The phylogenomic analysis of the *Porphyromonas* genus confirmed the strong relatedness with *P. gulae* [formerly a *P. gingivalis* subspecies associated with non-human hosts, see (Fournier et al., 2001)], but also showed a strong relatedness with *P. loveana*, a strain isolated from the oral cavity of a musky rat kangaroo (Bird et al., 2016). The exact role of this latter species in oral health and disease in its host remains unclear. However, as seen in later results (see Section 3.6), several virulence factors were also conserved in this species, as well as the *P. gingivalis* - *P. gulae* duo. This may suggest a potential role in the periodontal disease of *P. loveana* in their marsupial host. This may also suggest two potential scenarios: (A) the divergence between *P. gingivalis*/*P. gulae* duo and *P. loveana* are the potential origins of the periodontal disease and it is as old as the divergence of early mammals and marsupials [estimated to occur between 130 and 170 million years

TABLE 4 List of orthogroups covering the 10 lowest Tajima values of in the *P. gingivalis* core, including their protein names and functions ($p < 0.05$).

Top position	OG	Tajima value	Protein	Function
1	OG0001008	-3.223375698	Hypothetical protein (TonB-dependent receptor)	Protein involved in the acquisition of essential nutrients, such as iron and heme, from the host
2	OG0000740	-3.211357054	Calcium-transporting ATPase 1	Ion pump involved in calcium homeostasis and tolerance to high concentrations of calcium
3	OG0000475	-3.206282293	Hypothetical protein (DUF4248 domain-containing protein)	Unknown, it is suggest that is involved in protein-protein interactions, DNA binding, or enzymatic activities
4	OG0001040	-3.204872637	cAMP-activated global transcriptional regulator CRP	Regulation of respond to chages in intracellular levels of the signaling molecule cyclic AMP (cAMP)
5	OG0000392	-3.186069821	Hypothetical protein (type VI-B CRISPR-associated RNA-guided ribonuclease Cas13b)	Prokaryotic defense mechanism against foreign genetic elements such as viruses and plasmids
6	OG0000511	-3.158209672	10 kDa chaperonin	Folding of newly synthesized proteins
7	OG0000636	-3.155269224	Transcriptional regulatory protein QseB	Response regulator that is activated by the QseC sensor kinase, involved in the regulation of biofilm formation and antibiotic resistance
8	OG0000667	-3.137577035	Hypothetical protein (Acyl-CoA dehydrogenase C-terminal domain-containing protein)	Catalyze the first step in the breakdown of fatty acids
9	OG0001389	-3.121169508	Peptidoglycan-N-acetylglucosamine deacetylase	Modification of peptidoglycan, an structural molecule of bacterial cell wall
10	OG0001157	-3.095009015	Hypothetical protein (GNAT family N-acetyltransferase)	Modification of a variety of cell wall components, including lipopolysaccharides and capsular polysaccharides





ago (Cifelli and Davis, 2003)]; or (B) periodontal disease is potentially rooted in the mammals, regardless of the date of origin of the *Porphyromonas* species associated to it. To address these affirmations, it will be necessary to sequence more periodontitis-associated Porphyromonads in several other mammals.

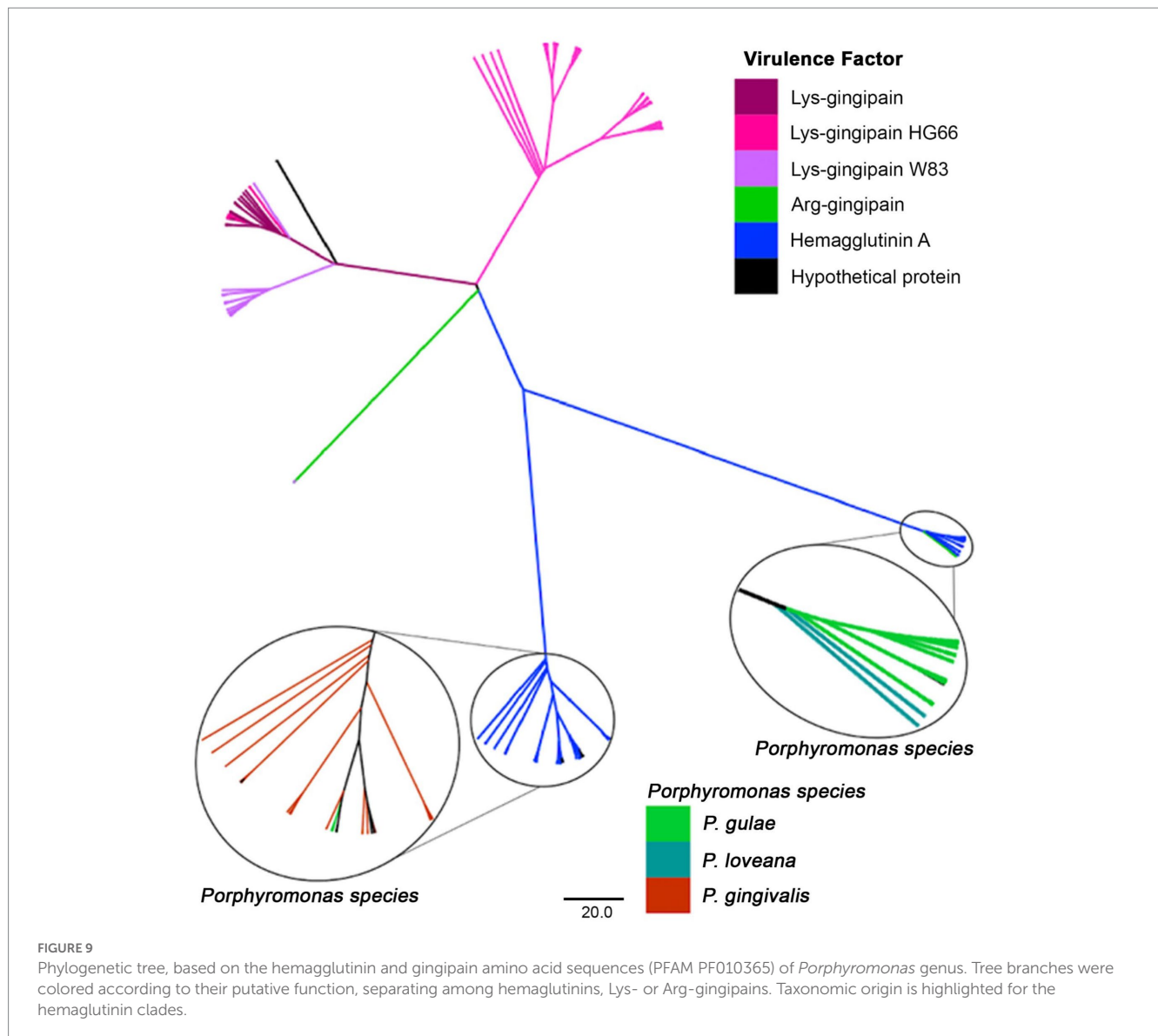
4.2. Pangenome openness of *Porphyromonas gingivalis* and *Porphyromonas gulae* and the potential role of gained genes among LCAs during species divergence

Pangenome approaches allow us to know which genetic repertoire of a bacterial group, especially discover essential genes for the survival or even all the virulence factors in a specific bacterial species (Bosi et al., 2015; Rouli et al., 2015). Our analysis shows that the pangenome of *P. gingivalis* and *P. gulae* are in an openness stage (Figure 2; Table 1), suggesting that the gene content from both species will be accepting more new gene families while new genomes are added to the analysis. The genomic core for both organisms was enriched housekeeping functional categories whereas the accessory fraction of the pangenomes was enriched by genes without COG and associated with mobile elements and defense mechanisms. This adaptation of the accessory core in both species is a previously observed feature, reflecting a recent history of mobile element transfer and the utilization of fitness factors (Brockhurst et al., 2019). The presence of unknown genes Interestingly, some proteins present in the *P. gulae* were related to phage defense like CRISPR-Cas system associate proteins. Recent studies have suggested the presence of different phages in *P. gingivalis*, raising the potential for cross-phage infection

between relative species like *P. gulae* and *P. gingivalis* (Matrishin et al., 2022).

The gene gain/loss analysis showed a comparable number of gene gain and gene loss events during the divergences between *P. gingivalis*, *P. gulae*, and *P. loveana*, with the number of gene gain/loss events after the speciation process. Since that previous evidence showed that selection could be a significant driver of gene loss and reductive genome evolution in bacterial species (Koskiniemi et al., 2012) and that gene content reconstruction can reflect the evolution of microbial pangenomes in a specific taxon (Iranzo et al., 2019; Maturana and Cárdenas, 2021), events observed during *P. gingivalis* - *P. gulae* bifurcation suggested a discrete role of gene gain/loss in this species delimitation, as a signal that those two species are very similar in the evolutionary pressure produced by their respective environments.

To our knowledge, there is no previous report about gain and/or loss events describing *Porphyromonas* evolution, even if it is focused only on *P. gingivalis*. The fact that several genes gained by the LCA of *P. gingivalis* were related to metabolic enzymes and transcriptional regulators, genes with these specific characteristics were categorized as essential for fitness in *P. gingivalis* (Miller et al., 2017). Moreover, one of the genes was involved in the nitro-groups (nitroreductase family protein). In the oral microenvironment, nitrosative stress could be challenging due to the high intake of dietary nitrate, some genes are necessary for the growth of this oral pathogen (Lewis et al., 2012). Some of the gained genes by both LCAs were related directly to the mobilome of this pathogen, which involved genes responsible for relaxase/mobilization nuclease, transposase, plasmid mobilization relaxosome, integrase, and transposons. All those elements play an important role in genetic rearrangement and are considered driving forces of bacterial diversification (Olsen et al., 2018). Two proteins that were identified as T9SS type A were gained in this node, T9SS



translocates proteins to the outer membrane, which is important for virulence factors. These proteins play an important role in cell survival and fitness in response to the microenvironment (Lasica et al., 2017). One important virulence factor was reported, glycosyltransferase, which is important for capsular polysaccharide biosynthesis (Naito et al., 2008).

In the case of *P. gulae*, we found at least two type IX secretion systems which are interesting in the way that these kinds of systems are present in some bacteria species and could play two main roles, which are gliding motility or a weapon for some pathogens, as *P. gingivalis* described above. To our knowledge, this kind of system was not described previously in *P. gulae*, delving into the sequences inside the genome of some members of *Porphyromonas* could reveal some particular features. Proteins related to transcription factors process or interact with DNA as well as an energy process, where were pNresent between the genes gained by the *P. gulae* LCA. Some elements like relaxosome protein, which is a protein that allows the conjugation process, a type of horizontal transfer gene, this kind of system could play an important role in the genetic transfer of

information between members of *Porphyromonas* (Watanabe et al., 2017). Finally, the LCA of *P. gulae*, like *P. gingivalis*, also gained a nitroreductase protein. These two species are members of the oral microbiome in mammals, as we described above the oral microenvironment is rich in nitro-groups.

4.3. Differential gene selection across *Porphyromonas gingivalis* and *Porphyromonas gulae* may be part of the specialization process

The analysis of dN/dS ratios is a method to study evolutionary pressure among macroevolution in gene families (Vitti et al., 2013). The measure of those values was performed for the shared core of *P. gingivalis* and *P. gulae*, considering *P. loveana* as the outgroup. The observation that the “N” and “S” categories contained the highest median for dN/dS data, could reflect a slightly higher diversification effect for genes associated with those categories; genes associated with

category N were involved in functions such as cell cycle control, cell division, and chromosome partitioning, including proteins like ParA (Bignell and Thomas, 2001), a member of a large group of P-loop ATPases with a deviant Walker A motif, involved in DNA partitioning (Mishra and Srinivasan, 2022). Proteins like this could be under the effect of positive selection due to they are essential for proper DNA partitioning and cell division, in the same way, the effect of natural selection could give advantages to ensuring accurate DNA and improve fitness (Liu et al., 2008). Another protein found in category N was part of a complex involved in the translocation of lipopolysaccharide (LPS), the permease protein Lpt (Lpt). The translocation of LPS to the outer membrane could be related to cell motility and membrane dynamics (Veronika et al., 2011). LPS is an important virulence factor for *P. gingivalis* and/or *P. gulae*, the effect of positive selection could improve the impact of this virulence factor in the oral microbiome (Dashper et al., 2017). The S category is the second with the highest median; genes from this category were generally annotated as membrane proteins and lipoproteins. Further investigation is necessary to characterize those genes.

The *P. gulae* and *P. gingivalis* lineages could have different evolutionary pressures. The comparison of their respective values for the set of each shared orthogroup showed that only a few gene families contained a considerable signal making them candidates to have a more positive selection in one lineage than in another: three belonged to *P. gingivalis* and six to *P. gulae*. The function of some of these gene families could give some clues about the specialization of each lineage. For example, genes under positive selection in *P. gingivalis* were 4-hydroxy butyryl-CoA dehydratase/vinyl acetyl-CoA-Delta-isomerase, Glycogen synthase, and UvrABC system protein B. The first protein is involved in the oxidation of fatty acids mostly in anaerobic microorganisms, this protein has a relationship with *P. gingivalis*, *Clostridium aminobutyricum*, and *Archaeoglobus fulgidus* (Gerhardt et al., 2000) and its positive selection may respond to the need to improve its fitness to synthesize lipids in their environment, in comparison with *P. gulae*. Glycogen synthesis could influence the survival mechanism of microorganisms (Park, 2015); previous studies in other pathogens such as *Vibrio cholerae* showed that glycogen biosynthesis is a fitness factor under nutrient-limiting conditions, such as the aquatic environment (Kamp et al., 2013). The UvrABC system protein B is part of a complex involved in DNA repair and with a relevant role in damage resistance in *P. gingivalis* during oxidative stress (Henry et al., 2012). The differentially higher positive selection may respond to the need to improve fitness by diversifying selection in those genes when the *P. gingivalis* lineage is compared with *P. gulae*.

4.4. Tajima's D values suggest evidence of a bottleneck effect in the human host

As mentioned earlier, *P. gingivalis* and *P. gulae* are closely related phylogenetically (O'Flynn et al., 2015; Fujiwara-Takahashi et al., 2020). Additionally, the closest relative to them is *P. loveana*, a species isolated from the oral cavity of Australian marsupials (Bird et al., 2016). *P. gingivalis* and *P. gulae* emerged from a common ancestor, and although both are associated with periodontitis, they differ in host specificity. *P. gingivalis* is isolated from human samples and *P. gulae*

from other mammal hosts (Fujiwara-Takahashi et al., 2020). Such different host specialties could be associated with changes in neutrality among core genes.

We performed the Tajima D statistical test with core gene families of *P. gingivalis* and *P. gulae* and generated curves and boxplots with all the Tajima's scores for gene families of those organisms (Figure 6B). The differences among the Tajima values between species were significant ($p < 2 \times 10^{-16}$). *P. gingivalis* distribution presented a distribution towards lower Tajima's D values than *P. gulae*, exhibiting less neutrality in their core orthologs (higher frequencies of rare alleles). This observation may be explained due to an expansion after a bottleneck effect.

The bottleneck effect refers to a past event that causes the reduction of the population since the outcome is the loss of genetic diversity, with impacts on the appearance of rare alleles (Peery et al., 2012). The foundation effect is caused when a small group of individuals establish a new population. This causes the emergence of a new population with less genetic diversity and an increment in the frequency of rare alleles (Nielsen, 2001). In the case of *P. gingivalis*, this may be explained due to its specialization as the only Porphyromonad causing periodontitis in the human oral cavity, exhibiting an important site, host, and activity specificity (Bostanci and Belibasakis, 2012). In contrast, its relative *P. gulae*, shows similar characteristics and a role in periodontal disease but is associated with a wider host range, as it is isolated from different mammals. Considering the pathogenicity and specificity of the hosts of both species, as well as the findings of our study, we dare to hypothesize that *P. gingivalis* experienced the bottleneck effect, followed by an amplification of rare alleles, during its colonization and establishment in the human oral cavity. We speculate that both *P. gingivalis* and *P. gulae*, from their same common ancestor, contained the virulence factors necessary for periodontal infection and after a speciation event, they evolved to adapt to their hosts, being *P. gingivalis* the one whose evolved to a different species in the human body. However, since the numbers of *P. gingivalis* and *P. gulae* genomes were unequal, sequencing of more *P. gulae* genomes and their posterior analysis can help to confirm these allele frequency differences.

4.5. Virulence factor conservation among *Porphyromonas* species shows a common pattern in lineage emergence for periodontal disease

Phylogenetic analyses for some of the most relevant virulence factors involved in periodontitis showed patterns that could be explained by the evolution of *Porphyromonas* species, as well as by their association with their hosts. For example, phylogenies for the major fimbriae (FimA) and the minor fimbriae (Mfa1) showed the co-clustering of *P. gulae*, *P. gingivalis* and *P. loveana* sequences, in agreement with the common evolutionary origin for those organisms. In the case of the different genotypes (I, II, and IV), the potentially closest clade to the root of the group is the type II genotype, which also includes the ortholog from *P. loveana*. The furthest branch contained FimA genotypes I and IV, encoded by *P. gingivalis*, and *P. gulae* strains. Current evidence has suggested that type I genotype is associated with samples from patients with periodontal health,

whereas genotypes II and IV are associated with samples from patients diagnosed with periodontitis (Wang et al., 2020). This may suggest that the rise of the attenuated type I could be posterior to the emergence of the pathogenic type II, aspect that could ensure the early capability of *P. gingivalis* and *P. gulae* to succeed in their role in periodontal disease in their respective niches. In contrast, the late adaptation into genotypes I and IV may respond to dynamics of host specificity adaptations as long as *P. gulae* and *P. gingivalis* are adjusting to their new niches. Previous studies suggested that genotype IV fimbriae are associated to increased adhesion efficiency in comparison with genotype I (Mendez et al., 2019); this behavior may reflect a late adaptation to increase adhesion abilities after the formation of the “soft” genotype I phenotype.

The presence of FimA orthologs in other Porphyromonads (*P. cangingivalis*, *P. canoris*, *P. levii*, *P. uenonis*, *P. endodontalis*, *P. macacae*, *P. levi*) expand the previous findings from (Fujiwara-Takahashi et al., 2020), which only considered *P. gulae* and *P. gingivalis*. In fact, one early clade in the FimA phylogeny contained orthologs from *P. cangingivalis* and *P. canoris*, species isolated from canine oral cavities. *P. cangingivalis* is isolated from health and early periodontitis-associated subgingival plaque, and *P. canoris* from subgingival plaque associated to periodontitis (Collins et al., 1994; Love et al., 1994). This finding suggests that the origin of FimA may predate the adaptation of the role of Porphyromonads in periodontal disease.

A similar “trifecta-enriched” phylogenetic pattern was also observed with Mfa1, a protein expressed by *P. gingivalis* on its surface. The primary function of the fimbrial systems is to adhere to host epithelial cells and colonize subgingival tissue. However, Mfa1 is also involved in microorganism auto-aggregation, microbe-microbe interactions, and suppression of the host immune response (Arjunan et al., 2016), via the evasion of autophagy-lysosome function of myeloid dendritic cell (DC) when it internalizes *P. gingivalis* (El-Awady et al., 2015). Mfa1 is also able to induce osteoclastogenesis by interacting with the Toll-like receptor of murine macrophage (RAW264 cells) *in vitro* (Suzuki et al., 2022). Osteoclastogenesis and consequent bone resorption is the hallmark of periodontitis (Hajishengallis and Lamont, 2016). Additionally, the *in vitro* interaction between *P. gingivalis* Mfa1 and the streptococcal antigen I/II receptor (SspA/B) of the early colonizer *Streptococcus gordonii* was demonstrated (Roky et al., 2020).

A phylogenetic tree of Mfa1 proteins was constructed from the PF15495 hits found in *Porphyromonas* genomes. Congruent Mfa1 homologs were detected in *P. gingivalis*, *P. gulae*, and *P. loveana*, in addition to *P. macacae*, *P. levi*, *P. uenonis*, *P. crevioricanis*, *P. gingivicanis*, *P. circumdentaria*, *P. endodontalis*, and *P. asaccharolytica*. Whereas *P. macacae* and *P. cangingivalis* were identified in the oral cavity of a variety of mammals (dogs, cats, monkeys, and bovines) (O’Flynn et al., 2015; Gabarrini et al., 2018). Additionally, *P. uenonis*, and *P. asaccharolytica* have been detected in the cervicovaginal tract and appendicitis- or peritonitis-associated samples (Finegold et al., 2004; Lithgow et al., 2022) or in diabetic patients with foot infections (Imirzalioglu et al., 2014). Another interesting aspect of the Mfa1 phylogeny is that homologs from the trio *P. gingivalis* - *P. gulae* - *P. loveana* are distributed in two separated clades suggesting that in those organisms, the origin of this protein was carried out twice. The separation of those clades is a resemblance of the existence of two proposed genotypes in the Mfa1 genes observed in *P. gingivalis* and *P. gulae*, the so-called 70-kDa and 53-kDa types (Fujiwara-Takahashi

et al., 2020). Structural and more detailed evolutionary studies of those separated clades are worth a more profound next study.

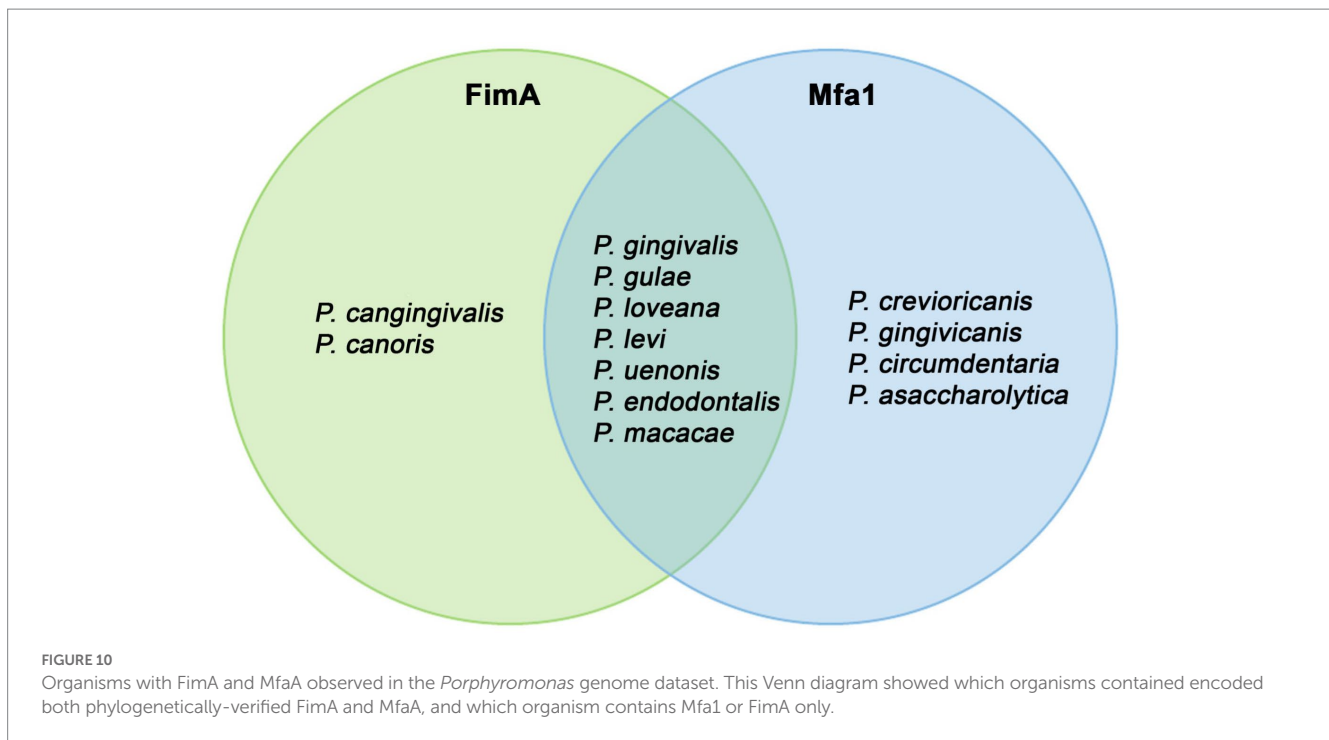
Organisms with detected FimA and Mfa1 homologs can be grouped (Figure 10). More species harbored Mfa1 compared to FimA. Two species, *P. cangingivalis* and *P. canoris*, contained FimA but not Mfa1, and four organisms (*P. crevioricanis*, *P. gingivicanis*, *P. circumdentaria*, and *P. asaccharolytica*) contained Mfa1 and not FimA. *P. crevioricanis* and *P. gingivicanis* were isolated from the gingival crevicular fluids of dogs (Hirasawa and Takada, 1994), whereas *P. circumdentaria* and *P. asaccharolytica* were isolated from the gingival margins of cats (Love et al., 1992), and from a variety of nonoral clinical samples (Shah and Collins, 1988). These differences in gene content may imply that those organisms evolved to contain the minor fimbria to function in other roles outside of the periodontal disease. A more profound study can help to elucidate the properties of those Mfa1 homologs.

The evolutionary patterns of the hemagglutinin-containing proteins are more straightforward. The phylogenetic compartmentalization between arginine-gingipains (RgpA), lysine-gingipains (Kgp), and hemagglutinins are clear. Interestingly, this tree only included proteins from the trifecta *P. gingivalis* - *P. gulae* - *P. loveana*. Gingipains were found only in *P. gingivalis* and in very few strains of *P. gulae*, whereas hemagglutinins were found in the three related organisms. The tree divided the Lys-gingipains into two lineages, associated with HG66 and W83 strains, respectively. The HG66 is an avirulent *P. gingivalis* strain exhibiting the type I genotype FimA, whereas the strain W83 is a more virulent variant, exhibiting the type IV genotype FimA (Mendez et al., 2019). The divergence between both lineages may be associated with differences in the virulence of the *P. gingivalis* strains, as a signal of a more specific adaptation to human host as a disrupting virulence factor. On the other hand, hemagglutinins (HagA or HagB) are well conserved in the trifecta, as a virulence factor important for heme acquisition, periodontal inflammation, and tissue breakdown. The evolutionary pattern of those hemagglutinins is almost completely species-specific, suggesting a strong vertical transference pattern, and potentially, another host-specific adaptation.

5. Conclusion

The following study, aimed to understand the evolution of *P. gingivalis* in the context of the *Porphyromonas* genus, could confirm the high relatedness with *P. gulae* but also found an unexpected relatedness with the marsupial-associated *P. loveana*; in coincidence with this relationship, several marker genes are more conserved in this group in comparison with the rest of Porphyromonads, suggesting that this group of three organisms was involved in the origin of the phenotype causing periodontal disease. Since those organisms have a characteristic presence in mammals, periodontal disease could be as old as the origin of the mammals.

The phylogenomic approach also found an interesting set of described and undescribed species, offering new opportunities for posterior studies. The phylogenetic analysis of the markers also opens the opportunity to study the properties of those virulence factors in the non-periodontal *Porphyromonas* species.



Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository or repositories and accession number(s) can be found in the article / [Supplementary material](#).

Author contributions

JC and CC conceived the study. MM-O, JN-B, DG, and JC analyzed the data. MM-O, JN-B, DG, EV, JR-P, CC, and JC collaboratively elaborated, edited, and corrected the text and figures in the manuscript. All authors contributed to the article and approved the submitted version.

Funding

JC is supported by ANID Fondecyt project 11200209, CC is supported by ANID Fondecyt project 11190073, and JR-P is supported by ANID Fondecyt project 1221064. JN-B is supported by the Integrative Genomics Ph.D. program (Universidad Mayor, Chile).

References

- Aleksijević, L. H., Aleksijević, M., Škrlec, I., Šram, M., Šram, M., and Talapko, J. (2022). *Porphyromonas gingivalis* virulence factors and clinical significance in periodontal disease and coronary artery diseases. *Pathogens* 11:1173. doi: 10.3390/pathogens11101173
- Arjunan, P., El-Awady, A., Dannebaum, R. O., Kunde-Ramamoorthy, G., and Cutler, C. W. (2016). High-throughput sequencing reveals key genes and immune homeostatic pathways activated in myeloid dendritic cells by *Porphyromonas gingivalis* 381 and its fimbrial mutants. *Mol Oral Microbiol* 31, 78–93. doi: 10.1111/omi.12131
- Bignell, C., and Thomas, C. M. (2001). The bacterial ParA-ParB partitioning proteins. *J. Biotechnol.* 91, 1–34. doi: 10.1016/s0168-1656(01)00293-0

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1226166/full#supplementary-material>

- Bird, P. S., Trott, D. J., Mikkelsen, D., Milinovich, G. J., Hillman, K. M., Burrell, P. C., et al. (2016). *Porphyromonas loveana* sp. nov., isolated from the oral cavity of Australian marsupials. *Int. J. Syst. Evol. Microbiol.* 66, 3771–3778. doi: 10.1099/ijsem.0.000898
- Bosi, E., Fani, R., and Fondi, M. (2015). Defining orthologs and pangenome size metrics. *Methods Mol. Biol.* 1231, 191–202. doi: 10.1007/978-1-4939-1720-4_13
- Bostanci, N., and Belibasakis, G. N. (2012). *Porphyromonas gingivalis*: an invasive and evasive opportunistic oral pathogen. *FEMS Microbiol. Lett.* 333, 1–9. doi: 10.1111/j.1574-6968.2012.02579.x
- Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., et al. (2017). Minimum information about a single amplified genome

- (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35, 725–731. doi: 10.1038/nbt.3893
- Brockhurst, M. A., Harrison, E., Hall, J. P. J., Richards, T., McNally, A., and MacLean, C. (2019). The ecology and evolution of pangenomes. *Curr. Biol.* 29, R1094–R1103. doi: 10.1016/j.cub.2019.08.012
- Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* 38, 5825–5829. doi: 10.1093/molbev/msab293
- Carlson, C. S., Thomas, D. J., Eberle, M. A., Swanson, J. E., Livingston, R. J., Rieder, M. J., et al. (2005). Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Res.* 15, 1553–1565. doi: 10.1101/gr.4326505
- Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848
- Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2022). GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38, 5315–5316. doi: 10.1093/bioinformatics/btac672
- Chen, W. A., Dou, Y., Fletcher, H. M., and Boskovic, D. S. (2023). Local and systemic effects of *Porphyromonas gingivalis* infection. *Microorganisms* 11, 470. doi: 10.3390/microorganisms11020470
- Chen, T., Hosogi, Y., Nishikawa, K., Abbey, K., Fleischmann, R. D., Walling, J., et al. (2004). Comparative whole-genome analysis of virulent and avirulent strains of *Porphyromonas gingivalis*. *J. Bacteriol.* 186, 5473–5479. doi: 10.1128/JB.186.16.5473-5479.2004
- Chen, T., Siddiqui, H., and Olsen, I. (2017). In silico comparison of 19 *Porphyromonas gingivalis* strains in genomics, phylogenetics, Phylogenomics and functional genomics. *Front. Cell. Infect. Microbiol.* 7:28. doi: 10.3389/fcimb.2017.00028
- Chung, W. O., Demuth, D. R., and Lamont, R. J. (2000). Identification of a *Porphyromonas gingivalis* receptor for the *Streptococcus gordonii* SspB protein. *Infect. Immun.* 68, 6758–6762. doi: 10.1128/IAI.68.12.6758-6762.2000
- Cifelli, R. L., and Davis, B. M. (2003). Paleontology: Marsupial origins. *Science* 302, 1899–1900. doi: 10.1126/science.1092272
- Collins, M. D., Love, D. N., Karjalainen, J., Kanervo, A., Forsblom, B., Willems, A., et al. (1994). Phylogenetic analysis of members of the genus *Porphyromonas* and description of *Porphyromonas cangingivalis* sp. nov. and *Porphyromonas cansulci* sp. nov. *Int. J. Syst. Bacteriol.* 44, 674–679. doi: 10.1099/00207713-44-4-674
- Csurös, M. (2010). Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* 26, 1910–1912. doi: 10.1093/bioinformatics/btq315
- Dashper, S. G., Mitchell, H. L., Seers, C. A., Gladman, S. L., Seemann, T., Bulach, D. M., et al. (2017). *Porphyromonas gingivalis* uses specific domain rearrangements and allelic exchange to generate diversity in surface virulence factors. *Front. Microbiol.* 8:48. doi: 10.3389/fmicb.2017.00048
- Diard, M., and Hardt, W.-D. (2017). Evolution of bacterial virulence. *FEMS Microbiol. Rev.* 41, 679–697. doi: 10.1093/femsre/fux023
- Eaton, D. A. R. (2020). Toytree: a minimalist tree visualization and manipulation library for python. *Methods Ecol. Evol.* 11, 187–191. doi: 10.1111/2041-210X.13313
- El-Awady, A. R., Miles, B., Scisci, E., Kurago, Z. B., Palani, C. D., Arce, R. M., et al. (2015). *Porphyromonas gingivalis* evasion of autophagy and intracellular killing by human myeloid dendritic cells involves DC-SIGN-TLR2 crosstalk. *PLoS Pathog.* 10:e1004647. doi: 10.1371/journal.ppat.1004647
- Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238. doi: 10.1186/s13059-019-1832-y
- Endo, A., Watanabe, T., Ogata, N., Nozawa, T., Aikawa, C., Arakawa, S., et al. (2015). Comparative genome analysis and identification of competitive and cooperative interactions in a polymicrobial disease. *ISME J.* 9, 629–642. doi: 10.1038/ismej.2014.155
- Eskan, M. A., Hajishengallis, G., and Kinane, D. F. (2007). Differential activation of human gingival epithelial cells and monocytes by *Porphyromonas gingivalis* fimbriae. *Infect. Immun.* 75, 892–898. doi: 10.1128/IAI.01604-06
- Finegold, S. M., Vaisanen, M.-L., Rautio, M., Eerola, E., Summanen, P., Molitoris, D., et al. (2004). *Porphyromonas uenonis* sp. nov., a pathogen for humans distinct from *P. asaccharolytica* and *P. endodontalis*. *J. Clin. Microbiol.* 42, 5298–5301. doi: 10.1128/JCM.42.11.5298-5301.2004
- Fournier, D., Mouton, C., Lapiere, P., Kato, T., Okuda, K., and Ménard, C. (2001). *Porphyromonas gulae* sp. nov., an anaerobic, gram-negative coccobacillus from the gingival sulcus of various animal hosts. *Int. J. Syst. Evol. Microbiol.* 51, 1179–1189. doi: 10.1099/00207713-51-3-1179
- Fujiwara-Takahashi, K., Watanabe, T., Shimogishi, M., Shibasaki, M., Umeda, M., Izumi, Y., et al. (2020). Phylogenetic diversity in FIM and MFA gene clusters between *Porphyromonas gingivalis* and *Porphyromonas gulae*, as a potential cause of host specificity. *J. Oral Microbiol.* 12:1775333. doi: 10.1080/20002297.2020.1775333
- Gaba, S., Kumari, A., Medema, M., and Kaushik, R. (2020). Pan-genome analysis and ancestral state reconstruction of class halobacteria: probability of a new super-order. *Sci. Rep.* 10:21205. doi: 10.1038/s41598-020-77723-6
- Gabarrini, G., Chlebawicz, M. A., Vega Quiroz, M. E., Veloo, A. C. M., Rossen, J. W. A., Harmsen, H. J. M., et al. (2018). Conserved citrullinating exoenzymes in *Porphyromonas* species. *J. Dent. Res.* 97, 556–562. doi: 10.1177/0022034517747575
- Gabarrini, G., de Smit, M., Westra, J., Brouwer, E., Vissink, A., Zhou, K., et al. (2015). The peptidylarginine deiminase gene is a conserved feature of *Porphyromonas gingivalis*. *Sci. Rep.* 5:13936. doi: 10.1038/srep13936
- Gerhardt, A., Cinkaya, I., Linder, D., Huisman, G., and Buckel, W. (2000). Fermentation of 4-aminobutyrate by clostridium aminobutyricum: cloning of two genes involved in the formation and dehydration of 4-hydroxybutyryl-CoA. *Arch. Microbiol.* 174, 189–199. doi: 10.1007/s002030000195
- Gibson, F. C., and Genco, C. A. (2006). “The Genus *Porphyromonas*,” In *The Prokaryotes*, (Eds.) M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer and E. Stackebrandt (New York, NY: Springer New York), 428–454.
- Guilloux, C.-A., Lamoureux, C., Beuruelle, C., and Héry-Arnaud, G. (2021). *Porphyromonas*: a neglected potential key genus in human microbiomes. *Anaerobe* 68:102230. doi: 10.1016/j.anaerobe.2020.102230
- Guo, Y., Nguyen, K.-A., and Potempa, J. (2010). Dichotomy of gingipains action as virulence factors: from cleaving substrates with the precision of a surgeon's knife to a meat chopper-like brutal degradation of proteins. *Periodontol.* 54, 15–44. doi: 10.1111/j.1600-0757.2010.00377.x
- Gupta, R. S., Patel, S., Saini, N., and Chen, S. (2020). Robust demarcation of 17 distinct bacillus species clades, proposed as novel *Bacillaceae* genera, by phylogenomics and comparative genomic analyses: description of *Robertmurraya kyonggiensis* sp. nov. and proposal for an emended genus bacillus limiting it only to the members of the subtilis and cereus clades of species. *Int. J. Syst. Evol. Microbiol.* 70, 5753–5798. doi: 10.1099/ijsem.0.004475
- Hajishengallis, G., and Lamont, R. J. (2014). Breaking bad: manipulation of the host response by *Porphyromonas gingivalis*. *Eur. J. Immunol.* 44, 328–338. doi: 10.1002/eji.201344202
- Hajishengallis, G., and Lamont, R. J. (2016). Dancing with the stars: how choreographed bacterial interactions dictate nososymbiosis and give rise to keystone pathogens, accessory pathogens, and pathobionts. *Trends Microbiol.* 24, 477–489. doi: 10.1016/j.tim.2016.02.010
- Hajishengallis, G., Tapping, R. I., Harokopakis, E., Nishiyama, S., Ratti, P., Schifferle, R. E., et al. (2006). Differential interactions of fimbriae and lipopolysaccharide from *Porphyromonas gingivalis* with the toll-like receptor 2-centred pattern recognition apparatus. *Cell. Microbiol.* 8, 1557–1570. doi: 10.1111/j.1462-5822.2006.00730.x
- Hamada, N., Takahashi, Y., Watanabe, K., Kumada, H., Oishi, Y., and Umamoto, T. (2008). Molecular and antigenic similarities of the fimbrial major components between *Porphyromonas gulae* and *P. gingivalis*. *Vet. Microbiol.* 128, 108–117. doi: 10.1016/j.vetmic.2007.09.014
- Harokopakis, E., Albzreh, M. H., Martin, M. H., and Hajishengallis, G. (2006). TLR2 transmodulates monocyte adhesion and transmigration via Rac1- and PI3K-mediated inside-out signaling in response to *Porphyromonas gingivalis* fimbriae. *J. Immunol.* 176, 7645–7656. doi: 10.4049/jimmunol.176.12.7645
- Hasegawa, Y., and Nagano, K. (2021). *Porphyromonas gingivalis* FimA and Mfa1 fimbriae: current insights on localization, function, biogenesis, and genotype. *Jpn. Dent Sci Rev* 57, 190–200. doi: 10.1016/j.jdsr.2021.09.003
- Henry, L. G., McKenzie, R. M. E., Robles, A., and Fletcher, H. M. (2012). Oxidative stress resistance in *Porphyromonas gingivalis*. *Future Microbiol.* 7, 497–512. doi: 10.2217/fmb.12.17
- Hirasawa, M., and Takada, K. (1994). *Porphyromonas gingivicanis* sp. nov. and *Porphyromonas crevioricanis* sp. nov., isolated from beagles. *Int. J. Syst. Bacteriol.* 44, 637–640. doi: 10.1099/00207713-44-4-637
- Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* 33, 1635–1638. doi: 10.1093/molbev/msw046
- Hyun, J. C., Monk, J. M., and Palsson, B. O. (2022). Comparative pangenomics: analysis of 12 microbial pathogen pangenomes reveals conserved global structures of genetic and functional diversity. *BMC Genomics* 23:7. doi: 10.1186/s12864-021-08223-8
- Imirzalioglu, C., Sethi, S., Schneider, C., Hain, T., Chakraborty, T., Mayser, P., et al. (2014). Distinct polymicrobial populations in a chronic foot ulcer with implications for diagnostics and anti-infective therapy. *BMC. Res. Notes* 7:196. doi: 10.1186/1756-0500-7-196
- Inaba, H., Nomura, R., Kato, Y., Takeuchi, H., Amano, A., Asai, F., et al. (2019). Adhesion and invasion of gingival epithelial cells by *Porphyromonas gulae*. *PLoS One* 14:e0213309. doi: 10.1371/journal.pone.0213309
- Iranzo, J., Wolf, Y. I., Koonin, E. V., and Sela, I. (2019). Gene gain and loss push prokaryotes beyond the homologous recombination barrier and accelerate genome sequence divergence. *Nat. Commun.* 10:5376. doi: 10.1038/s41467-019-13429-2
- Iwashita, N., Nomura, R., Shirai, M., Kato, Y., Murakami, M., Matayoshi, S., et al. (2019). Identification and molecular characterization of *Porphyromonas gulae* fimA types among cat isolates. *Vet. Microbiol.* 229, 100–109. doi: 10.1016/j.vetmic.2018.12.018
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9

- Jia, L., Han, N., Du, J., Guo, L., Luo, Z., and Liu, Y. (2019). Pathogenesis of important virulence factors of *Porphyromonas gingivalis* via toll-like receptors. *Front. Cell. Infect. Microbiol.* 9:262. doi: 10.3389/fcimb.2019.00262
- Kamp, H. D., Patimalla-Dipali, B., Lazinski, D. W., Wallace-Gadsden, F., and Camilli, A. (2013). Gene fitness landscapes of *Vibrio cholerae* at important stages of its life cycle. *PLoS Pathog.* 9:e1003800. doi: 10.1371/journal.ppat.1003800
- Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinformatics* 20, 1160–1166. doi: 10.1093/bib/bbx108
- Korneliusson, T. S., Moltke, I., Albrechtsen, A., and Nielsen, R. (2013). Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinformatics* 14:289. doi: 10.1186/1471-2105-14-289
- Koskineniemi, S., Sun, S., Berg, O. G., and Andersson, D. I. (2012). Selection-driven gene loss in bacteria. *PLoS Genet.* 8:e1002787. doi: 10.1371/journal.pgen.1002787
- Kuboniwa, M., Amano, A., Hashino, E., Yamamoto, Y., Inaba, H., Hamada, N., et al. (2009). Distinct roles of long/short fimbriae and gingipains in homotypic biofilm development by *Porphyromonas gingivalis*. *BMC Microbiol.* 9:105. doi: 10.1186/1471-2180-9-105
- Laland, K. N., Odling-Smee, F. J., and Feldman, M. W. (1999). Evolutionary consequences of niche construction and their implications for ecology. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10242–10247. doi: 10.1073/pnas.96.18.10242
- Lasica, A. M., Ksiazek, M., Madej, M., and Potempa, J. (2017). The type IX secretion system (T9SS): highlights and recent insights into its structure and function. *Front. Cell. Infect. Microbiol.* 7:215. doi: 10.3389/fcimb.2017.00215
- Lenzo, J. C., O'Brien-Simpson, N. M., Orth, R. K., Mitchell, H. L., Dashper, S. G., and Reynolds, E. C. (2016). *Porphyromonas gulae* has virulence and immunological characteristics similar to those of the human periodontal pathogen *Porphyromonas gingivalis*. *Infect. Immun.* 84, 2575–2585. doi: 10.1128/IAI.01500-15
- Lewis, J. P., Yanamandra, S. S., and Anaya-Bergman, C. (2012). HcpR of *Porphyromonas gingivalis* is required for growth under nitrosative stress and survival within host cells. *Infect. Immun.* 80, 3319–3331. doi: 10.1128/IAI.00561-12
- Lithgow, K. V., Buchholz, V. C. H., Ku, E., Korschuh, S., D'Aubeterre, A., and Sycuro, L. K. (2022). Protease activities of vaginal *Porphyromonas* species disrupt coagulation and extracellular matrix in the cervicovaginal niche. *NPJ Biofilms Microbiomes* 8:8. doi: 10.1038/s41522-022-00270-7
- Liu, J., Zhang, Y., Lei, X., and Zhang, Z. (2008). Natural selection of protein structural and functional properties: a single nucleotide polymorphism perspective. *Genome Biol.* 9:R69. doi: 10.1186/gb-2008-9-4-r69
- Lombardo Bedran, T. B., Marcantonio, R. A. C., Spin Neto, R., Alves Mayer, M. P., Grenier, D., Spolidorio, L. C., et al. (2012). *Porphyromonas endodontalis* in chronic periodontitis: a clinical and microbiological cross-sectional study. *J. Oral Microbiol.* 4:10123. doi: 10.3402/jom.v4i0.10123
- Lourbakos, A., Yuan, Y. P., Jenkins, A. L., Travis, J., Andrade-Gordon, P., Santulli, R., et al. (2001). Activation of protease-activated receptors by gingipains from *Porphyromonas gingivalis* leads to platelet aggregation: a new trait in microbial pathogenicity. *Blood* 97, 3790–3797. doi: 10.1182/blood.v97.12.3790
- Love, D. N., Bailey, G. D., Collings, S., and Briscoe, D. A. (1992). Description of *Porphyromonas circumdentaria* sp. nov. and reassignment of *Bacteroides salivovus* (Love, Johnson, Jones, and Calverley 1987) as *Porphyromonas* (Shah and Collins 1988) *salivosa* comb. nov. *Int. J. Syst. Bacteriol.* 42, 434–438. doi: 10.1099/00207713-42-3-434
- Love, D. N., Karjalainen, J., Kanervo, A., Forsblom, B., Sarkiala, E., Bailey, G. D., et al. (1994). *Porphyromonas canoris* sp. nov., an asaccharolytic, black-pigmented species from the gingival sulcus of dogs. *Int. J. Syst. Bacteriol.* 44, 204–208. doi: 10.1099/00207713-44-2-204
- Lunar Silva, I., and Cascales, E. (2021). Molecular strategies underlying *Porphyromonas gingivalis* virulence. *J. Mol. Biol.* 433:166836. doi: 10.1016/j.jmb.2021.166836
- Madej, M., White, J. B. R., Nowakowska, Z., Rawson, S., Scavenius, C., Enghild, J. J., et al. (2020). Structural and functional insights into oligopeptide acquisition by the RagAB transporter from *Porphyromonas gingivalis*. *Nat. Microbiol.* 5, 1016–1025. doi: 10.1038/s41564-020-0716-y
- Maresz, K. J., Hellvard, A., Sroka, A., Adamowicz, K., Bielecka, E., Koziel, J., et al. (2013). *Porphyromonas gingivalis* facilitates the development and progression of destructive arthritis through its unique bacterial peptidylarginine deiminase (PAD). *PLoS Pathog.* 9:e1003627. doi: 10.1371/journal.ppat.1003627
- Matrishin, C. B., Haase, E. M., Dewhirst, F. E., Mark Welch, J. L., Miranda-Sanchez, F., MacFarland, D. C., et al. (2022). Phages are important unrecognized players in the ecology of the oral pathogen *Porphyromonas gingivalis*. *BioRxiv*, 12:30.519816. doi: 10.1101/2022.12.30.519816
- Maturana, J. L., and Cárdenas, J. P. (2021). Insights on the evolutionary genomics of the Blautia genus: potential new species and genetic content among lineages. *Front. Microbiol.* 12:660920. doi: 10.3389/fmicb.2021.660920
- Mei, M. L., Yan, Z., Duangthip, D., Niu, J. Y., Yu, O. Y., You, M., et al. (2020). Effect of silver diamine fluoride on plaque microbiome in children. *J. Dent.* 102:103479. doi: 10.1016/j.jdent.2020.103479
- Mendez, K. N., Hoare, A., Soto, C., Bugueño, I., Olivera, M., Meneses, C., et al. (2019). Variability in genomic and virulent properties of *Porphyromonas gingivalis* strains isolated from healthy and severe chronic periodontitis individuals. *Front. Cell. Infect. Microbiol.* 9:246. doi: 10.3389/fcimb.2019.00246
- Miller, D. P., Hutcherson, J. A., Wang, Y., Nowakowska, Z. M., Potempa, J., Yoder-Himes, D. R., et al. (2017). Genes contributing to *Porphyromonas gingivalis* fitness in abscess and epithelial cell colonization environments. *Front. Cell. Infect. Microbiol.* 7:378. doi: 10.3389/fcimb.2017.00378
- Mishra, D., and Srinivasan, R. (2022). Catching a walker in the act—DNA partitioning by ParA family of proteins. *Front. Microbiol.* 13:856547. doi: 10.3389/fmicb.2022.856547
- Mohanty, R., Asopa, S. J., Joseph, M. D., Singh, B., Rajguru, J. P., Saidath, K., et al. (2019). Red complex: polymicrobial conglomerate in oral flora: a review. *J. Family Med. Prim. Care* 8, 3480–3486. doi: 10.4103/jfmpc.jfmpc_759_19
- Moldovan, M. A., and Gelfand, M. S. (2018). Pangenomic definition of prokaryotic species and the phylogenetic structure of *Prochlorococcus* spp. *Front. Microbiol.* 9:428. doi: 10.3389/fmicb.2018.00428
- Mysak, J., Podzimek, S., Sommerova, P., Lyuya-Mi, Y., Bartova, J., Janatova, T., et al. (2014). *Porphyromonas gingivalis*: major periodontopathogenic pathogen overview. *J. Immunol. Res.* 2014:476068. doi: 10.1155/2014/476068
- Nagano, K., Hasegawa, Y., Iijima, Y., Kikuchi, T., and Mitani, A. (2018). Distribution of *Porphyromonas gingivalis* fimA and mfa1 fimbrial genotypes in subgingival plaques. *PeerJ* 6:e5581. doi: 10.7717/peerj.5581
- Naito, M., Hirakawa, H., Yamashita, A., Ohara, N., Shoji, M., Yukitake, H., et al. (2008). Determination of the genome sequence of *Porphyromonas gingivalis* strain ATCC 33277 and genomic comparison with strain W83 revealed extensive genome rearrangements in *P. gingivalis*. *DNA Res.* 15, 215–225. doi: 10.1093/dnares/dsn013
- Nakayama, M., and Ohara, N. (2017). Molecular mechanisms of *Porphyromonas gingivalis*-host cell interaction on periodontal diseases. *Jpn. Dent. Sci. Rev.* 53, 134–140. doi: 10.1016/j.jdsr.2017.06.001
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Nielsen, R. (2001). Statistical tests of selective neutrality in the age of genomics. *Heredity* 86, 641–647. doi: 10.1046/j.1365-2540.2001.00895.x
- Núñez-Belmar, J., Morales-Olavarría, M., Vicencio, E., Vernal, R., Cárdenas, J. P., and Cortez, C. (2022). Contribution of -omics technologies in the study of *Porphyromonas gingivalis* during periodontitis pathogenesis: a minireview. *Int. J. Mol. Sci.* 24:620. doi: 10.3390/ijms24010620
- O'Flynn, C., Deusch, O., Darling, A. E., Eisen, J. A., Wallis, C., Davis, I. J., et al. (2015). Comparative genomics of the genus *Porphyromonas* identifies adaptations for Heme synthesis within the prevalent canine Oral species *Porphyromonas gingivalis*. *Genome Biol. Evol.* 7, 3397–3413. doi: 10.1093/gbe/evv220
- Olsen, I., Chen, T., and Tribble, G. D. (2018). Genetic exchange and reassignment in *Porphyromonas gingivalis*. *J. Oral Microbiol.* 10:1457373. doi: 10.1080/20002297.2018.1457373
- Olsen, I., Lambris, J. D., and Hajishengallis, G. (2017). *Porphyromonas gingivalis* disturbs host-commensal homeostasis by changing complement function. *J. Oral Microbiol.* 9:1340085. doi: 10.1080/20002297.2017.1340085
- Park, K.-H. (2015). Roles of enzymes in glycogen metabolism and degradation in *Escherichia coli*. *J. Appl. Glycosci.* 62, 37–45. doi: 10.5458/jag.jag-2015_005
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenoltz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Paster, B. J., Dewhirst, F. E., Olsen, I., and Fraser, G. J. (1994). Phylogeny of *Bacteroides*, *Prevotella*, and *Porphyromonas* spp. and related bacteria. *J. Bacteriol.* 176, 725–732. doi: 10.1128/jb.176.3.725-732.1994
- Peery, M. Z., Kirby, R., Reid, B. N., Stoelting, R., Doucet-Béer, E., Robinson, S., et al. (2012). Reliability of genetic bottleneck tests for detecting recent population declines. *Mol. Ecol.* 21, 3403–3418. doi: 10.1111/j.1365-294X.2012.05635.x
- Rajilić-Stojanović, M., and de Vos, W. M. (2014). The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol. Rev.* 38, 996–1047. doi: 10.1111/1574-6976.12075
- Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106
- Rocha, E. P. C., Smith, J. M., Hurst, L. D., Holden, M. T. G., Cooper, J. E., Smith, N. H., et al. (2006). Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J. Theor. Biol.* 239, 226–235. doi: 10.1016/j.jtbi.2005.08.037
- Roky, M., Trent, J. O., and Demuth, D. R. (2020). Identification of functional domains of the minor fimbrial antigen involved in the interaction of *Porphyromonas gingivalis* with oral streptococci. *Mol. Oral Microbiol.* 35, 66–77. doi: 10.1111/omi.12280
- Romero-Lastra, P., Sánchez, M. C., Ribeiro-Vidal, H., Llama-Palacios, A., Figuero, E., Herrera, D., et al. (2017). Comparative gene expression analysis of *Porphyromonas gingivalis* ATCC 33277 in planktonic and biofilms states. *PLoS One* 12:e0174669. doi: 10.1371/journal.pone.0174669
- Rouli, L., Merhej, V., Fournier, P. E., and Raoult, D. (2015). The bacterial pangenome as a new tool for analysing pathogenic bacteria. *New Microbes New Infect.* 7, 72–85. doi: 10.1016/j.nmni.2015.06.005

- Sakamoto, M., Li, D., Shibata, Y., Takeshita, T., Yamashita, Y., and Ohkuma, M. (2015). *Porphyromonas pasteri* sp. nov., isolated from human saliva. *Int. J. Syst. Evol. Microbiol.* 65, 2511–2515. doi: 10.1099/ijms.0.000294
- Sakamoto, M., and Ohkuma, M. (2013). *Porphyromonas crevioricanis* is an earlier heterotypic synonym of *Porphyromonas cansulci* and has priority. *Int. J. Syst. Evol. Microbiol.* 63, 454–457. doi: 10.1099/ijms.0.042531-0
- Sato, T., Tomida, J., Naka, T., Fujiwara, N., Hasegawa, A., Hoshikawa, Y., et al. (2015). *Porphyromonas bronchialis* sp. nov. isolated from intraoperative bronchial fluids of a patient with non-small cell lung cancer. *Tohoku J. Exp. Med.* 237, 31–37. doi: 10.1620/tjem.237.31
- Sawana, A., Adeolu, M., and Gupta, R. S. (2014). Molecular signatures and phylogenomic analysis of the genus Burkholderia: proposal for division of this genus into the emended genus Burkholderia containing pathogenic organisms and a new genus *Paraburkholderia* gen. Nov. harboring environmental species. *Front. Genet.* 5:429. doi: 10.3389/fgene.2014.00429
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shah, H. N., and Collins, M. D. (1988). Proposal for reclassification of *Bacteroides asaccharolyticus*, *Bacteroides gingivalis*, and *Bacteroides endodontalis* in a new genus. *Porphyromonas*. *Int. J. Syst. Bacteriol.* 38, 128–131. doi: 10.1099/00207713-38-1-128
- Summanen, P., and Finegold, S. M. (2015). “*Porphyromonas*” in *Bergey’s Manual of Systematics of Archaea and Bacteria*. eds. W. B. Whitman, F. Rainey, P. Kämpfer, M. Trujillo, J. Chun and P. DeVos et al. (Wiley Online Library), 1–14.
- Summanen, P. H., Lawson, P. A., and Finegold, S. M. (2009). *Porphyromonas bennonis* sp. nov., isolated from human clinical specimens. *Int. J. Syst. Evol. Microbiol.* 59, 1727–1732. doi: 10.1099/ijms.0.001909-0
- Suzuki, Y., Kikuchi, T., Goto, H., Takayanagi, Y., Kawamura, S., Sawada, N., et al. (2022). *Porphyromonas gingivalis* fimbriae induce Osteoclastogenesis via toll-like receptors in RAW264 cells. *Int. J. Mol. Sci.* 23:15293. doi: 10.3390/ijms232315293
- van Dongen, S., and Abreu-Goodger, C. (2012). Using MCL to extract clusters from networks. *Methods Mol. Biol.* 804, 281–295. doi: 10.1007/978-1-61779-361-5_15
- Veronika, M., Welsch, R., Ng, A., Matsudaira, P., and Rajapakse, J. C. (2011). Correlation of cell membrane dynamics and cell motility. *BMC Bioinformatics* 12:S19. doi: 10.1186/1471-2105-12-S13-S19
- Vitti, J. J., Grossman, S. R., and Sabeti, P. C. (2013). Detecting natural selection in genomic data. *Annu. Rev. Genet.* 47, 97–120. doi: 10.1146/annurev-genet-111212-133526
- Waller, T., Kesper, L., Hirschfeld, J., Dommisch, H., Kölpin, J., Oldenburg, J., et al. (2016). *Porphyromonas gingivalis* outer membrane vesicles induce selective tumor necrosis factor tolerance in a toll-like receptor 4- and mTOR-dependent manner. *Infect. Immun.* 84, 1194–1204. doi: 10.1128/IAI.01390-15
- Wang, H., Zhang, W., Wang, W., and Zhang, L. (2020). The prevalence of fimA genotypes of *Porphyromonas gingivalis* in patients with chronic periodontitis: a meta-analysis. *PLoS One* 15:e0240251. doi: 10.1371/journal.pone.0240251
- Watanabe, T., Shibasaki, M., Maruyama, F., Sekizaki, T., and Nakagawa, I. (2017). Investigation of potential targets of *Porphyromonas* CRISPRs among the genomes of *Porphyromonas* species. *PLoS One* 12:e0183752. doi: 10.1371/journal.pone.0183752
- Yamada, M., Ikegami, A., and Kuramitsu, H. K. (2005). Synergistic biofilm formation by *Treponema denticola* and *Porphyromonas gingivalis*. *FEMS Microbiol. Lett.* 250, 271–277. doi: 10.1016/j.femsle.2005.07.019
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Zamora-Cintas, M., Marín, M., Quiroga, L., Martínez, A., Fernández-Chico, M. A., Bouza, E., et al. (2018). Identification of *Porphyromonas* isolates from clinical origin using MALDI-TOF mass spectrometry. *Anaerobe* 54, 197–200. doi: 10.1016/j.anaerobe.2018.06.017
- Zenobia, C., and Hajishengallis, G. (2015). *Porphyromonas gingivalis* virulence factors involved in subversion of leukocytes and microbial dysbiosis. *Virulence* 6, 236–243. doi: 10.1080/21505594.2014.999567
- Zhang, D., Li, S., Hu, L., Sheng, L., and Chen, L. (2015). Modulation of protease-activated receptor expression by *Porphyromonas gingivalis* in human gingival epithelial cells. *BMC Oral Health* 15:128. doi: 10.1186/s12903-015-0105-8
- Zhang, Q., Tyler-Smith, C., and Long, Q. (2015). An extended Tajima’s D neutrality test incorporating SNP calling and imputation uncertainties. *Stat. Interface* 8, 447–456. doi: 10.4310/SII.2015.v8.n4.a4