



Population-Genomic Insights into Variation in *Prevotella intermedia* and *Prevotella nigrescens* Isolates and Its Association with Periodontal Disease

Yifei Zhang^{1*†}, Min Zhen^{2†}, Yalin Zhan², Yeqing Song¹, Qian Zhang¹ and Jinfeng Wang^{3*}

¹ Central Laboratory, Peking University School and Hospital of Stomatology, Beijing, China, ² Department of Periodontology, Peking University School and Hospital of Stomatology, Beijing, China, ³ Computational Genomics Lab, Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing, China

OPEN ACCESS

Edited by:

Ingar Olsen,
University of Oslo, Norway

Reviewed by:

Tsute Chen,
The Forsyth Institute, United States
Renate Lux,
School of Dentistry, University of
California, Los Angeles, United States

*Correspondence:

Yifei Zhang
wingsflying2000@163.com
Jinfeng Wang
wangjf@biols.ac.cn

[†]These authors have contributed
equally to this work.

Received: 01 June 2017

Accepted: 04 September 2017

Published: 21 September 2017

Citation:

Zhang Y, Zhen M, Zhan Y, Song Y,
Zhang Q and Wang J (2017)
Population-Genomic Insights into
Variation in *Prevotella intermedia* and
Prevotella nigrescens Isolates and Its
Association with Periodontal Disease.
Front. Cell. Infect. Microbiol. 7:409.
doi: 10.3389/fcimb.2017.00409

High-throughput sequencing has helped to reveal the close relationship between *Prevotella* and periodontal disease, but the roles of subspecies diversity and genomic variation within this genus in periodontal diseases still need to be investigated. We performed a comparative genome analysis of 48 *Prevotella intermedia* and *Prevotella nigrescens* isolates that from the same cohort of subjects to identify the main drivers of their pathogenicity and adaptation to different environments. The comparisons were done between two species and between disease and health based on pooled sequences. The results showed that both *P. intermedia* and *P. nigrescens* have highly dynamic genomes and can take up various exogenous factors through horizontal gene transfer. The major differences between disease-derived and health-derived samples of *P. intermedia* and *P. nigrescens* were factors related to genome modification and recombination, indicating that the *Prevotella* isolates from disease sites may be more capable of genomic reconstruction. We also identified genetic elements specific to each sample, and found that disease groups had more unique virulence factors related to capsule and lipopolysaccharide synthesis, secretion systems, proteinases, and toxins, suggesting that strains from disease sites may have more specific virulence, particularly for *P. intermedia*. The differentially represented pathways between samples from disease and health were related to energy metabolism, carbohydrate and lipid metabolism, and amino acid metabolism, consistent with data from the whole subgingival microbiome in periodontal disease and health. Disease-derived samples had gained or lost several metabolic genes compared to healthy-derived samples, which could be linked with the difference in virulence performance between diseased and healthy sample groups. Our findings suggest that *P. intermedia* and *P. nigrescens* may serve as “crucial substances” in subgingival plaque, which may reflect changes in microbial and environmental dynamics in subgingival microbial ecosystems. This provides insight into the potential of *P. intermedia* and *P. nigrescens* as new predictive biomarkers and targets for effective interventions in periodontal disease.

Keywords: *Prevotella intermedia*, *Prevotella nigrescens*, periodontitis, comparative genomics, pathogenesis, metabolic function

INTRODUCTION

Periodontitis is a term describing specific inflammatory and destructive changes in the tooth-supporting tissues, such as the gingiva, periodontal ligaments, alveolar bone, and root cementum. If not treated appropriately, it can lead to tissue destruction and ultimately tooth loss. Moreover, there is increasing evidence of associations between periodontitis and cardiovascular disease (Stewart and West, 2016), diabetes (Preshaw et al., 2012), and obesity (Moura-Grec et al., 2014). The formation of dental plaque biofilms at the gingival margin is an initiating factor of periodontitis (Socransky and Haffajee, 1994).

The investigation of the microbiology of periodontal diseases spans several decades. Socransky et al. proposed a role for six major microbial complexes in the subgingival biofilm (Socransky et al., 1998) and described the relationship between the addition of species during “microbial succession,” leading to the development of gingival inflammation (Socransky and Haffajee, 2005). Yellow (*Streptococcus* species), green (*Campylobacter concisus*, *Eikenella corrodens*), purple (*Veillonella parvula*), and blue (*Actinomyces* species) complexes are associated with periodontal health, whereas red (*Porphyromonas gingivalis*, *Treponema denticola*, *Tannerella forsythia*) and orange complexes (*Fusobacterium*, *Prevotella*, and *Campylobacter* species) are closely associated with disease. Although the core members of red complexes (particularly *P. gingivalis*) are considered “keystone pathogens” (Hajishengallis et al., 2012) that can orchestrate periodontitis even at low abundances (Hajishengallis et al., 2011), they need additional changes in the local environment to reach a level to cause disease (Bogren et al., 2007). In a subgingival biofilm redevelopment study, Uzel et al. (2011) found that members of the green and orange complexes, such as *Fusobacterium* subspecies and *Prevotella intermedia*, increased much faster in periodontitis subjects than in periodontally healthy subjects. From days 7 to 21, *F. nucleatum* ssp. polymorphum, *P. intermedia*, and *Prevotella nigrescens* continued to increase, while the levels of members of the red complex did not change significantly even after 21 days. In another study that analyzed 13,261 plaque samples, it was posited that colonization of the subgingival environment by orange complex species was required for the establishment of the red complex species (Socransky et al., 1998). This reminds us that orange complex species may act as instigators to facilitate more pathogenic species of the red complex, and could be targeted as a prevention strategy, resulting in the inhibition of pathogenesis caused by dysbacteriosis.

P. intermedia and *P. nigrescens*, member of the orange complex, are among the most frequently encountered species in subgingival plaque (Zambon et al., 1981; Teanpaisan et al., 1995; Kamma et al., 2004). *P. intermedia* is a periodontitis-associated member of the subgingival microbiome, whereas *P. nigrescens* has been detected in most subjects and does not change in proportion from health to disease (Griffen et al., 2012; Abusleme, et al., 2013). Thus, it is a core species of the subgingival microbiome (Hong et al., 2015). One study showed that in natural gingivitis, at the start (gingival health) and after 14 days of experimental gingivitis, *P. nigrescens* is the

predominant microorganism (Lie et al., 2001). Gharbia et al. found that *P. intermedia* and *P. nigrescens* may have different site specificities and surface properties and thus their roles in the pathogenesis of periodontitis would be expected to differ (Gharbia et al., 1994).

The pathogenic potential of *P. intermedia* and *P. nigrescens* may vary among strains (Dorn et al., 1998). A *P. nigrescens* strain isolated from a chronic periodontitis lesion produced unique polysaccharides that may play an important role in the development of chronic inflammatory lesions (Yamane et al., 2005). Furthermore, the profile of degradative enzymes produced by the organism may vary depending on its site. Maeda et al. (1998) suggested that *P. intermedia* may increase the activity of degradative enzymes under certain conditions and support the progression of periodontitis.

Recently, much data on the relationship between the oral microbiome and periodontal disease have been generated due to advances in high-throughput sequencing technologies (Duran-Pinedo et al., 2014; Jorth et al., 2014; Li et al., 2014; Kirst et al., 2015), but there are few reports on strain-level detection. Because each bacterial strain in a microbial ecosystem can be considered a genetically defined, basic functional unit (Zhang and Zhao, 2016), focusing on new microbiome findings alone would limit our exploration of their clinical significance. It is necessary to further investigate how each strain in oral ecology, individually or in combination, contributes to the onset and progression of periodontal diseases.

In this study, we hypothesized that *P. intermedia* and *P. nigrescens* may serve as “crucial substances” in subgingival plaque, playing key roles in the local microenvironment. When the matrix is disturbed, the microbial community may become pathogenic. Specific clonal types of strains may be responsible for disease progression, and thus strains that represent relatively avirulent clones may be less effective for maintaining the biofilm. We performed a comparative genome analysis of *P. intermedia* and *P. nigrescens* strains from periodontal disease and healthy sites. We investigated genetic elements specific to each group and their associations with health and disease, and the *Prevotella* pathogenesis mechanisms that might contribute to the onset and progression of periodontal diseases.

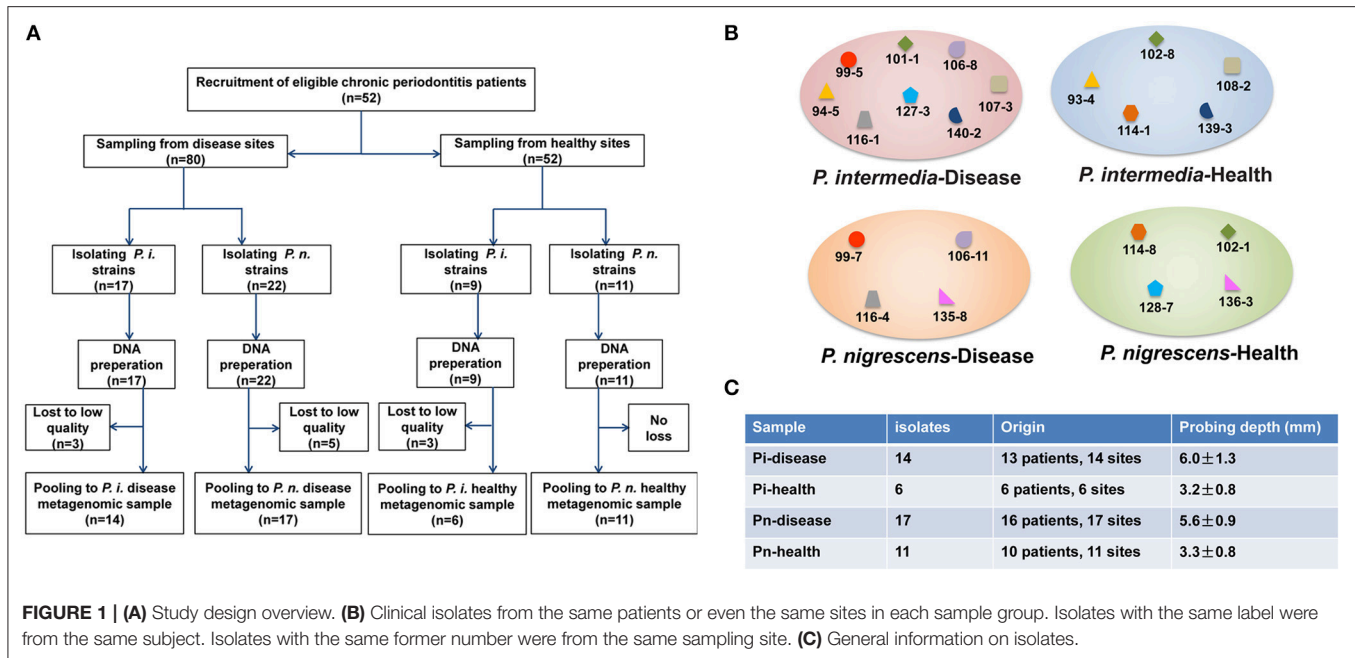
RESULTS

Clinical Strain Isolation

An overview of our study design is provided in **Figure 1**. We analyzed 48 *Prevotella* isolates (Table S1), all taken from different sites from 35 subjects: 14 *P. intermedia* isolates, derived from sites with periodontal disease (Pi-disease group); 6 *P. intermedia* isolates from healthy sites (Pi-health); 17 *P. nigrescens* isolates from diseased sites (Pn-disease); and 11 *P. nigrescens* isolates from healthy sites (Pn-health). In some cases, the same subjects contributed samples of healthy and diseased areas.

General Genome Features of Pooled Samples

Genome sequencing data were generated from four pooled samples, as described below. The general genome features of



P. intermedia and *P. nigrescens* samples from diseased and healthy sites are listed in **Table 1**. About 80% of reads of clinical samples of *P. intermedia* could be mapped to the *P. intermedia*-17 (GenBank number: NC_017860.1/NC_017861.1) reference genome, whereas the remainder had best matches in various species, such as other *Prevotella* species, *Bacteroides* species, *P. gingivalis*, *T. forsythia*, *Capnocytophaga* species, *Eubacterium* species, *Flavobacterium* species, *Alistipes* species, and *Fusobacterium* species. Major differences between Pi-disease and Pi-health sample groups were that partial reads in Pi-disease can be mapped to *Acidithiobacillus ferrivorans*, *Anaerofustis stercorihominis*, *Mycoplasma fermentans*, *Ruminococcaceae* species and *Treponema* sp., while some reads in Pi-health mapped to *Enterobacter cloacae* and *Xanthomonas fragariae*.

Regarding the *P. nigrescens* samples, ~20% of the reads mapped to species other than the reference strain *P. nigrescens* ATCC33563 (GenBank number: NZ_AFPX00000000.1), including *Prevotella* species, *Bacteroides* species, *P. gingivalis*, *T. forsythia*, *Alistipes* species, *Odoribacter splanchnicus*, *Neisseria* species, *Parabacteroides distasonis*, *Flavobacterium* species, *Tannerella* species, *Treponema* species, *Streptococcus agalactiae*, and *Faecalibacterium prausnitzii*. Major differences between Pn-disease and Pn-health samples were that some reads in the Pn-disease group can be mapped to *Bergeyella zoohelcum*, *Chryseobacterium* species, *Clostridiales* species, while some reads in healthy group mapped to *Acinetobacter baumannii* and *Saccharomyces cerevisiae*.

Open reading frames (ORFs) likely to encode proteins were predicted according to NCBI Prokaryotic Genomes Annotation Pipeline (Tatusova et al., 2016). Then all predicted proteins were analyzed for sequence similarities against the Uniprot database (The UniProt Consortium, 2015). The total predicted ORFs in each group were: Pi-disease: 8020; Pi-health: 6255; Pn-disease: 6943; and Pn-health: 7710. There were 813 orthologous genes

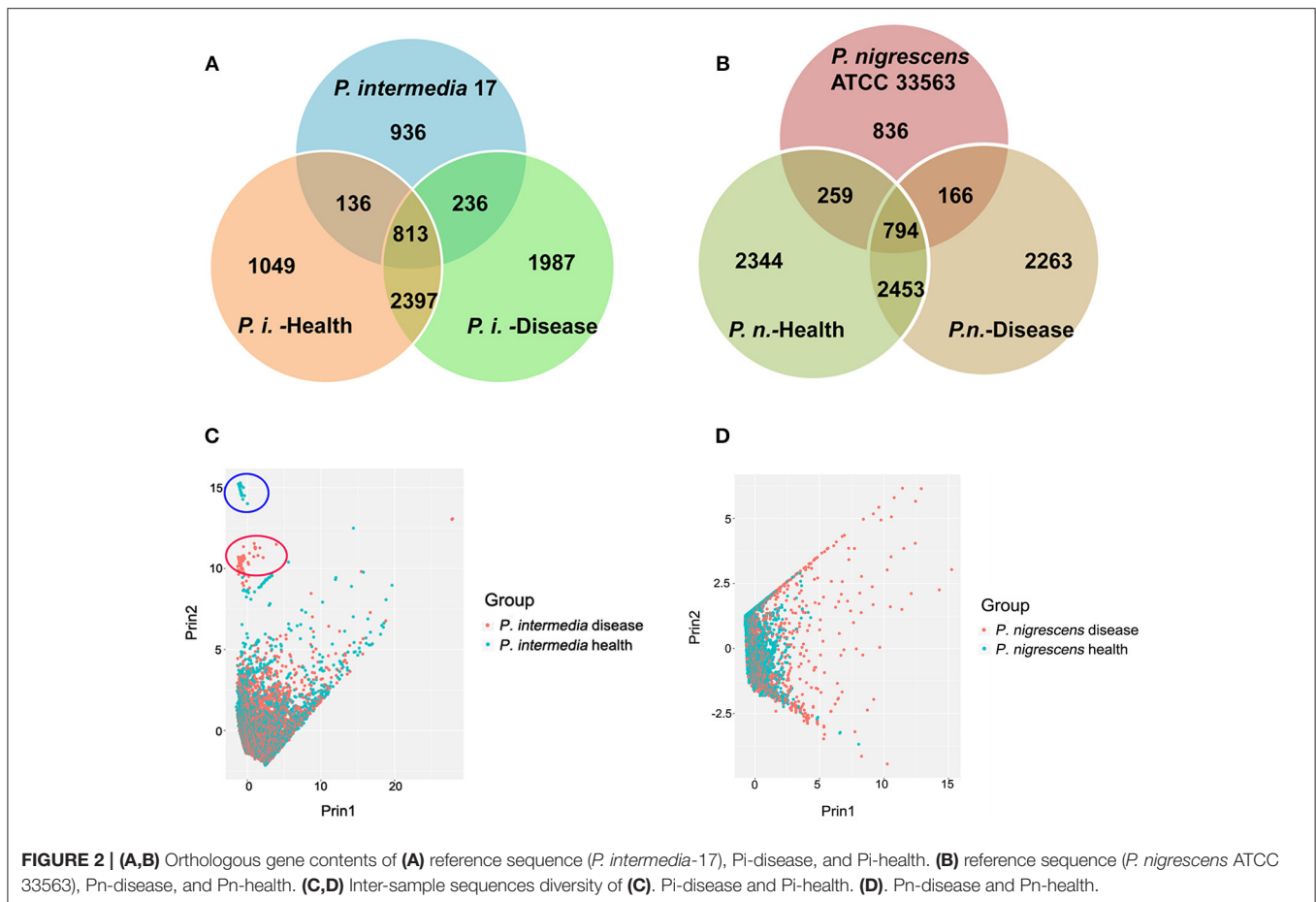
TABLE 1 | Summary of analyzed genome data.

Sample	Total reads	Mapped reads	Total bases	Mapped bases of reference	Coverage (%)
Pi-disease	40,287,600	33,417,719	3,824,344,295	2,554,975	94.65
Pi-health	23,341,144	18,767,547	2,160,552,995	2,528,402	93.66
Pn-disease	23,735,598	19,687,748	2,293,047,725	2,586,054	96.89
Pn-health	27,985,784	23,417,711	2,705,008,273	2,550,345	95.55

The reads of *P. intermedia* samples were mapped against the reference *P. intermedia* 17. The reads of *P. nigrescens* samples were mapped against reference *P. nigrescens* ATCC 33563.

shared by the Pi-disease, Pi-health, and reference sequences (**Figure 2A**). In total, 3210 genes were shared between the Pi-disease and Pi-health groups, whereas Pi-disease samples had more group-specific genes than Pi-health samples (1987 vs. 1049). Totally 794 orthologous genes were shared by the Pn-disease, Pn-health, and reference sequences. Both diseased and healthy groups of *P. nigrescens* had almost the same numbers of group-specific genes (2344 vs. 2263), and shared 3247 genes between Pn-disease and Pn-health groups (**Figure 2B**).

Based on the lengths of matched segments in sample sequences, the percent identity of matched segments to reference, the full length of the matched contig, the degree of coverage of matched segments to reference, and the degree of coverage of matched segments to contigs, we investigated inter-sample diversity by performing a principal component analysis (PCA) of Pi sample groups (**Figure 2C**) and Pn sample groups (**Figure 2D**). Between Pi-disease and Pi-health samples, most contigs could not be differentiated (**Figure 2C**), indicating that overall there were no obvious differences in genome sequences between the groups that matched the reference. However,



two additional small clusters of samples were clearly distinct (**Figure 2C**), consisting of contigs from each group. The contigs in the two distinct clusters were identified in Table S2, which were all belonged to the predicted insert sequences of Pi-disease or Pi-health samples. We also identified the annotated genes in these distinct contigs, most of which were conjugal transfer proteins (e.g., TraG, TraA, TraB, TraD etc. for Pi-disease group; TraE, TraK, TraM, and TraN for Pi-health group), phage associated proteins (both groups), recombinase (both groups), peptidase (Pi-health group), DNA-binding protein (Pi-disease group), and tetracycline resistance genes (e.g., tet Q and tet R; for Pi-disease group). Regarding the Pn group, most contigs from Pn-health samples clustered with those from Pn-disease; however, a considerable proportion of contigs in the Pn-disease group were clearly distinct from those in the Pn-health group (**Figure 2D**).

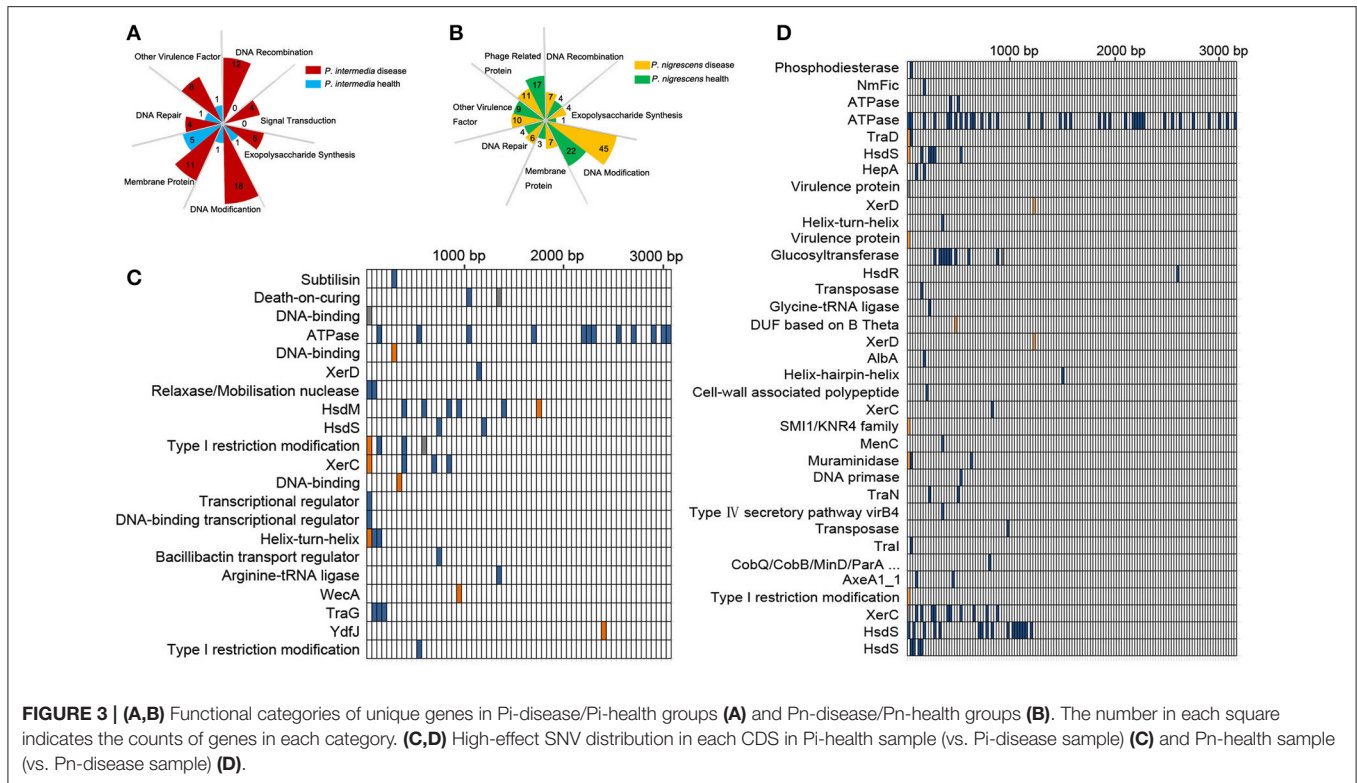
Unique Genes in Each Group

Unique Pi genes are listed in Tables S3, S4. The Pi-disease group had more unique genes involved in DNA modification, DNA recombination, DNA repair, and signal transduction, and more virulence-related factors (**Figure 3A**). Moreover, there were differences in exopolysaccharide biosynthesis and membrane proteins between the Pi-disease and Pi-health groups. Unique Pn genes are listed in Tables S5, S6. The

dominant genes differentiating the groups were related to DNA modification, phage proteins, DNA recombination, and DNA repair (**Figure 3B**). The Pn-disease group had many more specific genes that functioned in DNA modification, and more specific phage-related proteins were found in the Pn-health samples, but both groups had almost the same number of unique genes related to virulence. There were also differences in exopolysaccharide biosynthesis and membrane proteins between the groups.

Single-Nucleotide Variants in Core Regions

To investigate functional differences in orthologous genes between the Pi groups, we detected single-nucleotide variants (SNVs) in coding sequences (CDSs), and screened those whose functions were predicted to be highly affected (**Figure 3C**). The most affected regions were those encoding CDSs involved in DNA binding, DNA modification, and recombination, including HsdM, HsdS, XerC, XerD, and Type I restriction modification DNA-specific domain protein. The dominant regions that were affected by SNVs in the Pn-health (vs. Pn-disease) group were those encoding CDSs involved in ATPase, DNA modification (HsdS, HsdR), DNA recombination (XerC, XerD), conjugate transposon (TraD, TraN, TraI), and virulence protein (**Figure 3D**).



Label Sequences in Each Group

We also identified unique sequences in each sample, and tried to find any discrepant virulence gene candidates using BLASTX searches against the virulence factor database (VFDB; last accessed Nov 25, 2016; Chen et al., 2012) under a threshold *e*-value ($\leq 1e-5$). The numbers of unique sequences and matched virulence factors in each group are listed in Table S7. Most virulence factors specific to disease groups were capsule and lipopolysaccharide (LPS) synthesis-related factors. Moreover, disease groups had more unique sequences related to secretion systems, where they and the diverse proteins they secreted represent critical determinants of competitive fitness and pathogenic potential. The most abundant specific gene cluster in the Pi-disease group was the Type VI secretion system (T6SS). Pn-disease samples had specific sequences mapped to T3SS. Both the Pn-disease and Pn-healthy groups had T4SS effectors, but with different sequences. BspA and HlyB existed in all groups (health as well as disease), but with unique sequences in each sample group.

Furthermore, disease groups had more specific virulence factors annotated as proteinases and toxins. Only Pi-disease and Pn-disease samples possessed specific sequences that could match proteasome-associated ATPase (mpa gene).

Metabolic Pathway Reconstructions

We performed a comparison of metabolic networks between combined disease (Pi, Pn- disease) and combined health samples (Pi, Pn- health). We first determined the functional gene orthologs (KOs) in metabolic and regulatory pathways that

shared by Pi-disease and Pn-disease (disease-share), and by Pi-health and Pn-health (health-share). Then we identified the differential KOs involved in networks between disease-share and health-share. We also managed to define special metabolic and regulatory characteristics of each sample group. (Figures 4, 5). The metabolic pathway of Pi-disease is highly consistent with that of Pn-disease, except for a few differences that mainly involved in energy metabolism, carbohydrate metabolism, and amino acid metabolism. In regulatory pathways, the major differences were seen in signal transduction (Figures 4A, 5, Figure S1) Pi-health and Pn-health shared all the KOs so they had exactly the same metabolic and regulatory pathways (Figure 4B, Figure S1B). Disease-share had specific KOs involved in the generation of 3-hydroxy-3-methylglutaryl-CoA through valine, leucine and isoleucine degradation (e.g., E.2.3.3.10), and amino acid metabolism (e.g., murI), while health-share had specific KOs involved in pathways including Sulfur metabolism (e.g., cysI), glycolysis/gluconeogenesis (e.g., adhP), amino acid metabolism (e.g., CNDP2), amino sugar and nucleotide sugar metabolism (e.g., neuA, neuB, GME) (Figures 4, 5A). In addition, there was a distinct difference in signal transduction between disease-share and health-share. Both had unique two-component systems: e.g., algR and arcB in diseased samples and evgS, rcsC, and barA in healthy samples (Figure 5B, Figure S1).

We identified the KOs that merely exist/absent in Pi-disease or Pn-disease samples as well. Compared to other three sample groups, Pi-disease samples had gained specific genes involved in carbohydrate and lipid metabolism (e.g., lacZ, crtB, idi, and glmM), metabolism of cofactors and

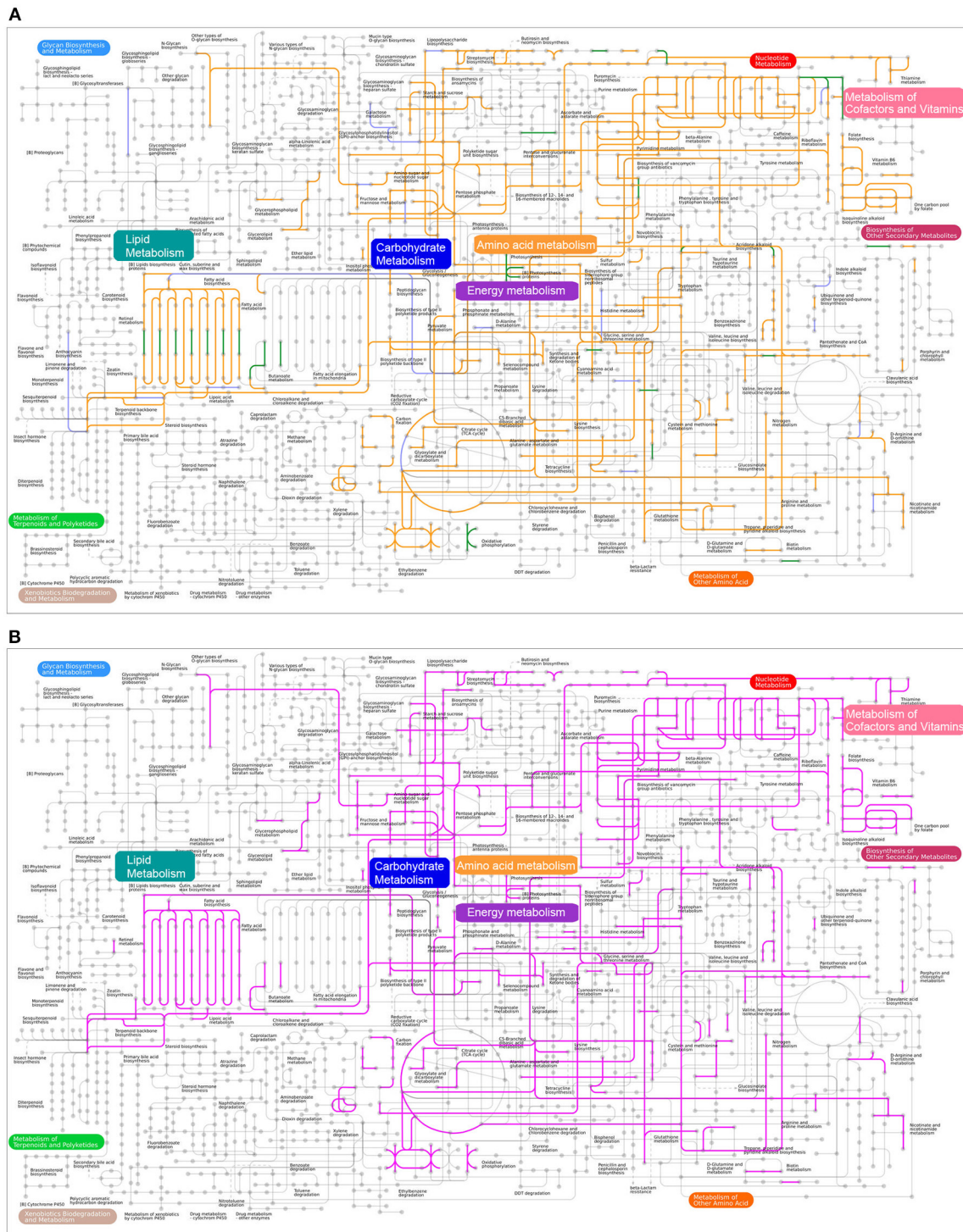
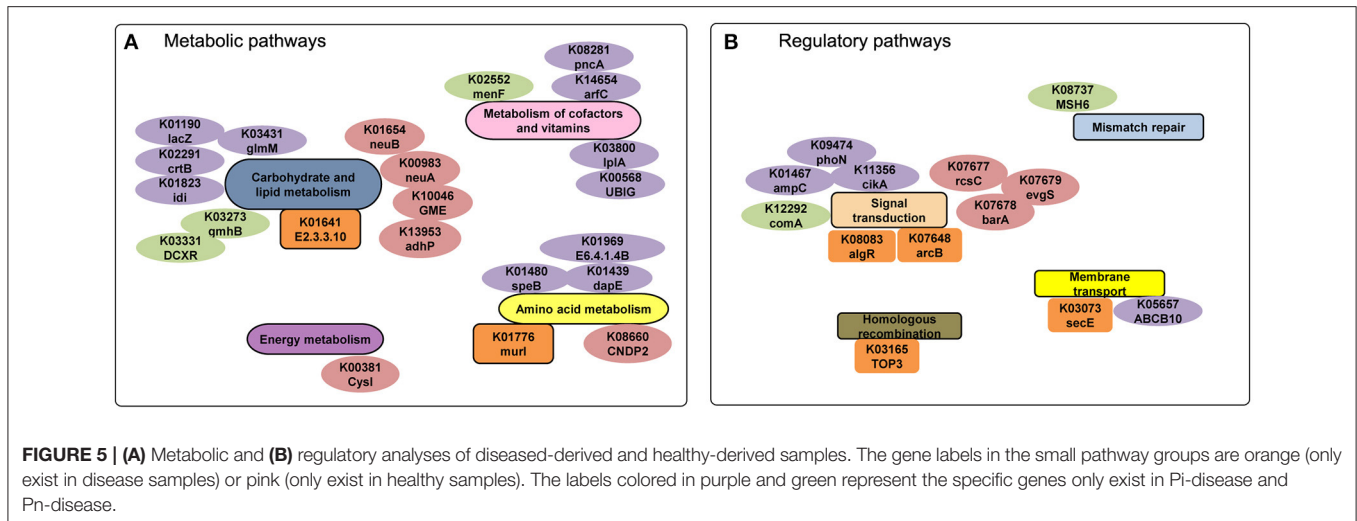


FIGURE 4 | (A) Metabolic pathway reconstructions of Pi-disease and Pn-disease. Pipelines colored in orange represent the common pathways shared by both samples. The pipelines colored in purple and green represent the specific pathway in Pi-disease and Pn-disease, respectively; **(B)** Metabolic pathway reconstructions of Pi-health and Pn-health. Pipelines colored in pink represent the common pathways shared by both samples; no specific pathway was seen in Pi-health or Pn-health.

vitamin (e.g., *arfC*, *UBIG*, *pncA*, and *lplA*) and amino acid metabolism (e.g., *speB*, *dapE*, and *E6.4.1.4B*) (Figures 4, 5), while Pn-disease samples had extra genes with specific pathways

including lipopolysaccharide biosynthesis (e.g., *gmhB*), pentose and glucuronate interconversions (e.g., *DCXR*), and ubiquinone biosynthesis (e.g., *menF*). In regulatory pathways, Pi-disease



samples had more unique genes involved in two-component system (e.g., *phoN*, *ampC*, and *cikA*) and membrane transport (e.g., *ABCB10*). Pn-disease had more unique genes involved in mismatch repair (e.g., *MSH6*) and two-component system (*comA*) (Figure 5). On the other hand, Pi-disease samples have also lost several genes, which involved in metabolism of cofactors and vitamin (e.g., K00077 *apbA*, K01719 *hemD*, K01495 *folE*, and K00652 *bioF*), fatty acid metabolism (e.g., K02372 *fabZ*, and K09478 *ACADSB*), and nucleotide metabolism (e.g., K01588 *purE* and K01493 *comEB*). There was only one gene ortholog that absent in Pn-disease samples only, which is involved in lipopolysaccharide metabolism (K02535 *LPXC*).

Function of Unique Gene Clusters

To understand the functions of different genes among the disease-derived (combined Pi,Pn- disease) and healthy-derived (combined Pi, Pn- health) isolates, protein functions were examined for clusters that were unique to isolates from diseased and healthy sites, based on the clusters of orthologous groups (COG) database (Tatusov et al., 2001) (Figure 6A). Diseased samples were enriched in 17 COGs, mostly related to gene regulation, signal transduction, biosynthesis, and function of cell wall structure, cell growth processes, and transport or metabolism functions, whereas healthy samples were enriched in only 1 COG related to energy production and conversion. The specific genes are listed in Tables S8, S9 and Figure 6. Diseased isolates had more unique factors related to cell wall and outer membrane structure and function, such as teichuronic acid biosynthesis glycosyltransferase (*TuaC*), D-alanyl-lipoteichoic acid biosynthesis protein (*DltB*), and the outer membrane cobalamin receptor protein. Others included the lipid A export ATP-binding/permease protein (*MsbA*), acetyltransferase (*EpsM*), lipid-A-disaccharide synthase, and dTDP-4-dehydrorhamnose 3,5-epimerase (*dTDP_sugar_isom*), which are involved in the pathway of LPS biosynthesis (Figure 6B).

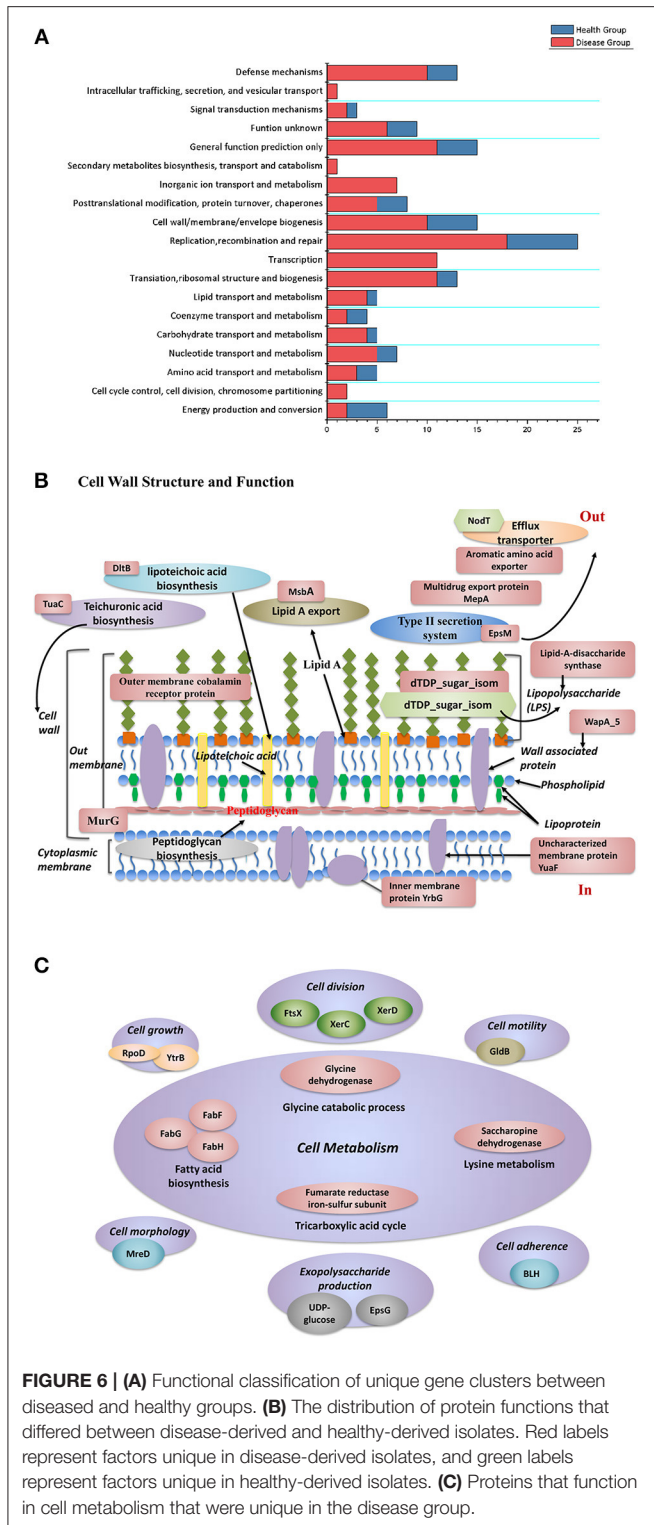
Regarding microbial growth, disease-derived samples had more unique factors that influence cell division, cell metabolism,

cell morphology, cell motility, and exopolysaccharide biosynthesis (Figure 6C). Diseased samples of both species contained an unusual *YpdA/YpdB* two-component regulatory system involved in a carbon control network whereas healthy isolates did not. In particular, isolates from diseased groups had a unique anaerobic ribonucleoside-triphosphate reductase activating protein, which is activated under anaerobic conditions, whereas healthy groups had a unique fumarate hydratase (class I, aerobic) that functions as an aerobic enzyme, involved in the citric acid cycle. Moreover, isolates from both diseased and healthy sites possessed different DNA methylases (Tables S8, S9). Diseased samples had more unique gene clusters that could aid resistance to stress factors, such as arsenate reductase and arsenic-resistance protein, which can enhance resistance to arsenate; heat-shock protein, which may help cells recover from stress; Na^+ /drug antiporter; and cold-shock DEAD box protein. Isolates from both types of isolates had different FeS cluster assembly proteins including *SufB*, which helps facilitate iron uptake from extracellular iron chelators under iron limitation.

LPSs are important factors involved in the virulence of *Prevotella* species, and we found that both diseased and healthy samples had several different factors directly or indirectly related to LPSs. Thus, we carried out a comprehensive analysis of LPS biosynthesis proteins in type culture strains from the NCBI database and clinical strains from each group of this study by reconstructing a neighbor-joining phylogenetic tree with MEGA 6 (Tamura et al., 2013) (Figure S2). LPS biosynthesis proteins from *P. intermedia* and *P. nigrescens* isolates clustered together. Notably, four disease-derived isolates clustered within a highly distinct and uniform clade.

Mobile Genetic Elements

Each group had many mobile genetic elements. We counted the number of conjugate transposons (CTNs), transposases (Tns), and integrases, and divided by the number of isolates contained in each group to calculate the mobile genetic elements (MGE)/isolates ratio (Table 2). The ratio of CTNs to isolates in



Pi-health was higher than that in Pi-disease group, while the ratio in Pn-disease and Pn-health groups were similar. The ratios of Tns and integrase to isolates were higher in healthy groups (both in Pi-health and Pn-health) than that in disease groups. We also found that each group contained several prophages or degraded

TABLE 2 | Mobile genetic elements in each group.

	Pi-disease	Pi-health	Pn-disease	Pn-health
*CTns/isolates	41/14 = 2.93	22/6 = 3.67	44/17 = 2.59	28/11 = 2.55
Tns/isolates	60/14 = 4.29	34/6 = 5.67	61/17 = 3.59	66/11 = 6
Integrase/isolates	40/14 = 2.86	24/6 = 4	56/17 = 3.29	52/11 = 4.73
# Prophages	4	3	10	2

*CTns were calculated by the number of protein that found in conjugate transposon.
Prophages included intact prophages and defective prophages.

prophages, with Pn-disease and Pi-health samples having more than Pi-disease and Pn-health samples.

CRISPR Array Sequences

The numbers of spacers in Pi-disease, Pi-healthy, Pn-disease, and Pn-healthy samples were 577, 183, 980, and 522, respectively. In Pi-disease and Pn-disease samples, 4.7 and 4.8% spacers showed significant nucleotide similarity to the phage database, slightly higher than for Pn-healthy (4.4%) and Pi-healthy (2.7%) samples. More spacers matched the plasmid database in Pi-disease (3.3%) and Pi-healthy (3.3%) samples than in Pn-disease (2.7%) and Pn-healthy (2.9%) samples (Figure S3).

DISCUSSION

The oral cavity is one of the most biologically complex sites in the body, having a natural microflora with a characteristic composition (Kurtz et al., 2004; Nasidze et al., 2009). Normally, the oral microflora exists harmoniously with the host, and will only cause disease (most commonly, caries and periodontitis) when homeostasis breaks down (Socransky and Haffajee, 2005). Clarifying the exact mechanism(s) by which the commensal microorganisms become pathogenic could open new opportunities for more effective disease prevention, treatment, and control. Most studies that have investigated the issue have focused on the diversity in the oral microbial ecology, making comparisons between patients and healthy populations (Human Microbiome Project Consortium, 2012; Yang et al., 2012; Wang et al., 2013). However, because host susceptibility is a key factor in oral disease, the impact of population variation on such results should not be ignored. Thus, in our study, we collected isolates from different sites from patients in the same population (those with periodontitis) to minimize this issue.

Full-genome sequencing of individual strain can provide more complete information indeed, but this approach is impractical for large population samples. It may be unable to understand how the genomes of these different strains are similar to or differ from each other through “pooled-sequencing,” however, our original intention was to consider the isolates that from disease or health as a whole and determined the unique functional genes (e.g., involved in metabolic pathway and virulence) that merely exist/absent in diseased or healthy sites at population level. And “pooled-sequencing” would have little impact on the presence or absence of certain genes in each group because of the high sequencing coverage. Besides, those candidate genes could be

further verified in more individual isolates. Since *P. intermedia* distributes more prevalent in diseased sites than in healthy sites, thus we obtained much less healthy-derived *P. intermedia* isolates than diseased-derived. However, our sequencing depth was higher than 800X, so all the isolates in each group would be covered. The unique genes or orthologs comparison was relied on the presence or absence, not on the quantity. Thus, we did not consider the difference in gene or orthologs abundance.

The genome of *P. intermedia* is unusually enriched in CTNs and MTNs (Belda-Ferre et al., 2012), suggesting a high level of genomic plasticity. In this study, all samples showed many mobile genetic elements, exogenous insert sequences, and proteins related to prophages, indicating a high occurrence of transfer events (e.g., conjugation, transduction, and transformation). This may be one explanation why the genomes of *P. intermedia* and *P. nigrescens* were not conserved and showed high genomic diversity. The environment under periodontal disease conditions is more diverse than that under healthy conditions (Naito et al., 2016). However, in this study, the horizontal transfer events seemed more likely to occur under healthy conditions. The matches of spacers to phage and plasmid sequences suggested that higher ratio of phage/plasmid space has more frequent encounter of exogenous extracellular DNA. Since each unique spacer in CRISPR is acquired from invasive foreign genetic elements, such as phages and plasmids, our data indicated that the host genomes of *P. intermedia* and *P. nigrescens* have acquired immunity against such phages or plasmids, and also serve as a record of their interactions with phages and plasmids. Under diseased conditions, bacterial hosts may interact more often with phages, and thus gain enhanced resistance to phage attack. However, the interaction between host and plasmid may depend on species.

We also found major differences in factors related to genome modification and recombination between disease-derived and health-derived samples in each species, indicating that the *Prevotella* isolates from disease sites may be more capable of genomic reconstruction. We also determined the unique sequences in each group. Our results indicate that *P. intermedia* and *P. nigrescens* in diseased samples had more specific virulence factors related to the cell capsule and LPS. Signal transduction was another key difference between diseased and healthy samples. The type VI secretion system (T6SS) shows intriguing versatility, targeting effector proteins to both eukaryotic cells and competitor bacterial cells (Shi et al., 2015). Many pathogens use it as a weapon to compete against rival bacterial cells by injecting them with multiple antibacterial toxins (Ho et al., 2014). In our study, only the Pi-disease group had unique sequences that matched T6SS-related proteins, indicating that Pi-disease samples may have specific pathogenic properties based on T6SS. The Type III secretion system (T3SS) has two homologous families: the flagellar T3SS drives cell motility, and the non-flagellar T3SS is a sophisticated molecular machinery of bacteria used to inject (translocate) bacterial proteins (effectors) into eukaryotic cells, a trait frequently associated with virulence (MacIntyre et al., 2010). Thus, Pn-disease samples may express unique virulence through T3SS. The effectors of secretion systems play important roles in the pathogenesis of bacteria,

and different effectors may have different functions (Enninga and Rosenshine, 2009). We found that both Pn-disease and Pn-healthy samples had T4SS effectors, but with different sequences. Whether this reflects a distinction in virulence between the two sample types may be worthy of further investigation.

Bacteroides surface protein A (BspA) is a secreted surface protein belonging to the leucine-rich repeat (LRR) family (Siemer and Dehio, 2015). It is a well-defined virulence factor of *T. forsythia* (Sharma et al., 1998) that may trigger the release of bone-resorbing pro-inflammatory cytokines from monocytes and chemokine IL-8 from gingival epithelial cells. It has also been shown to mediate bacterial adherence and invasion into epithelial cells. HlyB is thought to be closely related to multi-drug resistance. We found that the disease groups had more specific virulence factors annotated as proteinases and toxins, among which BspA and HlyB were common in all groups but with different sequences. Another virulence factor that should be highlighted is proteasome-associated ATPase (mpa gene). This protein is involved in the proteasomal Pup-dependent pathway, which is involved in protein degradation and pathogenesis (Darwin et al., 2005; Sharma, 2010). In our study, only the disease groups had specific sequences matching mpa.

Previous studies didn't show any difference in amino acid metabolism and glucose metabolism between *P. intermedia* and *P. nigrescens* (Takahashi and Yamada, 2000a,b). And our results confirmed that *P. intermedia* and *P. nigrescens* strains derived from healthy sites had common metabolic and regulatory networks from the overall aspect. However, disease-derived samples had gained or lost several metabolic genes compared to healthy-derived samples. It is reported that pathogens always require new metabolic pathways that allow them to exploit available food sources in order to thrive in new environment (Schmidt and Hensel, 2004), and the loss of non-essential metabolic functions could contribute to virulence by putting less demand on metabolic pathways (Rohmer et al., 2011). Thus, the metabolic function could be linked with the difference in virulence performance between diseased and healthy groups, and pathways could potentially be good targets for anti-microbial therapies (Rohmer et al., 2011). Besides, our results showed that Pi-disease samples gained/lost more genes than Pn-disease samples, indicating *P. intermedia* could be more adaptable to disease environment. Recent surveys on the subgingival microbiome in periodontitis suggest that the differentially represented pathways between periodontal disease and health are related to energy metabolism (e.g., aminoacyl-tRNA biosynthesis, nitrogen metabolism), carbohydrate and lipid metabolism (e.g., glycolysis/gluconeogenesis, citrate cycle, pyruvate metabolism) (Duran-Pinedo et al., 2014), and purine and pyrimidine metabolism (Kirst et al., 2015). Our data are consistent with those results regarding metabolic differences between disease and healthy samples.

The major differences in GO categories between the diseased and healthy groups were in transporters and DNA repair and recombination proteins, also consistent with a recent study (Kirst et al., 2015). Moreover, both diseased groups had several specific factors that have been reported to be more highly expressed in periodontitis subjects, such as Btu, Clp,

Fts, RpoD, and homologs to internalins (Duran-Pinedo et al., 2014). Disease-derived isolates had more unique gene clusters whose functions were related to drug and metal resistance (e.g., zinc metalloprotease, ABC transporter), also consistent with a previous report (Duran-Pinedo et al., 2014). The ability of *P. intermedia* and *P. nigrescens* to alter their metabolic characteristics depending on the nutrients available has been confirmed (Krakoff, 2006). These relationships could indicate that the functional potential of the subgingival microbiome under different health conditions might be determined according to the *P. intermedia* and *P. nigrescens* in the subgingival plaque.

The high level of plasticity and diversity of the *Prevotella* genomes has been demonstrated in other studies (Saito et al., 2001; Purushe et al., 2010). However, different strains may be affected by different exogenous factors, which could facilitate adaptation to varied environments and available nutrients, and generate virulence. This may be why *Prevotella* is associated with various infectious diseases, such as endodontic infections (Ruan et al., 2015), acute apical abscesses (Hsiao et al., 2012), apical periodontitis, dental implant failure (Siqueira and Rocas, 2013), and even cystic fibrosis (Dingsdag et al., 2016). In our study, the species-level comparative analysis of *Prevotella* species suggests that members of the “orange complex” may be able to capture important functional differences across samples and those differences could reflect changes in microbial and environmental dynamics in subgingival microbial ecosystems. We should consider *P. intermedia* and *P. nigrescens* as new biomarkers with predictive value, as well as new targets for effective interventions in periodontal disease.

METHODS

Sample Collection and Isolation of *P. intermedia* and *P. nigrescens*

Patients with chronic periodontitis were recruited from the Department of Periodontology, Peking University School and Hospital of Stomatology. Enrollment occurred between December 2013 and May 2015. The inclusion criteria were: good general health and not pregnant, aged 18 to 65 (male or female), had not taken any antibiotic or antimycotic compounds in the past 3 months or undergone any periodontal therapy within 1 year, no aggressive form of periodontal disease, and periodontal disease sites at least 5 mm probing depth and attachment loss >0 mm. Subjects were excluded if they were pregnant, had any systemic condition which could affect the progress of periodontal disease. Periodontally healthy sites require a probing depth not >4 mm and no attachment loss.

The study protocol was approved by the institutional review board of Peking University School and Hospital of Stomatology (Beijing, China) (approval number: PKUSSIRB-2012063).

For each subject, subgingival plaque samples were collected from gingival pockets with a sterile Gracey curette from the bottom of the pocket to the coronal portion of the pocket. Samples were placed in individual tubes containing 500 μ L Ringer's solution and then inoculated onto trypticase soy agar supplemented with 5% sheep blood, 1 mg/mL yeast

extract, 1 μ g/mL menadione, and 5 μ g/mL hemin. Plates were grown anaerobically (90% N₂, 5% H₂, 5% CO₂) at 37°C for 5 days, and black-pigment clones were selected and subcultured for purity. *P. intermedia* and *P. nigrescens* isolates were identified through full-length 16S rDNA sequencing (Sangon Biotech, Shanghai, China). Briefly, the full-length 16S rDNA region was amplified with a universal primer set (27forward, 5'-AGAGTTTGATCCTGGCTCAG-3'; 1492reverse, 5'-GGTTACCTTGTACGACTT-3'). The PCR condition is as follows: 5 min initial denaturation at 95°C; 35 cycles of denaturation at 95°C (30 s), annealing at 60°C (30 s), elongation at 72°C (90 s); and final extension at 72°C for 10 min. The resulting 16SrRNA sequences were searched against the National Center of Biotechnology Information (NCBI) database, and assigned to species based on the best match at greater than 97% nucleotide identity over at least 95% length of the query.

Genome Sequencing and Annotation

Genomic DNA of each isolate was extracted using the Qiagen DNA Purification Kit following the protocol. After checking quality, genomic DNA was pooled according to the study design, and then a sequencing library was built and sequenced on an Illumina HiSeq 2500 system (Biotechnology Corporation, Shanghai, China). Approximately 23–40 million paired-end reads (2 \times 125 nt multiplex) were obtained, representing >100-fold genome coverage. Genome mapping to reference genomes was performed using the Burrows-Wheeler Aligner software (ver. 0.7.12). Contigs were assembled de novo using SPAdes-3.5.0 (Bankevich et al., 2012), and filtered by deleting contigs with coverage \leq 10% and length < 500 bp. Then we performed a cluster analysis of remained contigs using CD-HIT version 4.6 (Fu et al., 2012). Contigs with similarity \geq 90% were compared, and only the longest one reserved. To understand the relationship between the contigs and the reference genome, we aligned the contigs to the reference genomes (*P. intermedia* samples to *P. intermedia* 17 and *P. nigrescens* samples to *P. nigrescens* ATCC 33563), and the lengths of matched segments in sample sequences, the percent identity of matched segments to reference, the full length of the matched contig, the degree of coverage of matched segments to reference, and the degree of coverage of matched segments to contigs were recorded for principal component analysis.

The draft genome was annotated using the NCBI Prokaryotic Genomes Annotation Pipeline (Tatusova et al., 2016). The remaining un-mapped reads of four samples were searched by BLASTN program against NCBI non-redundant nucleotide database with an E-value threshold of 1E-5.

Unique Genes Prediction

To determine the unique genes in each sample group, we first performed alignments using the pan-genomes analysis pipeline (PGAP) among those four sample groups. The chosen standard was GF, score: 40, identity: 0.6, and coverage: 0.8. We further carried out a group-to-group BLASTN search. For example, the genes obtained from Pi-disease sample group were aligned with the genome sequences in Pi-health sample by BLASTN search, and the results were further filtered by deleting genes with

identity > 30% and coverage > 30%, then we obtained a unique genes list for Pi-disease sample group.

Label Sequences Identifying in Each Group

To identify unique sequences in each sample, the sequences in each sample group were compared with all the other sample groups through MUMmer 3.23 (Kurtz et al., 2004) (for example, Pi-disease vs. Pi-health; Pi-disease vs. Pn-disease; and Pi-disease vs. Pn-health). The sequences (≥ 500 bp) that existed in Pi-disease but absent in any other groups were considered as label sequences of Pi-disease group. Label sequences in each group were further annotated by BLASTX searching against the virulence factors database (VFDB, last accessed: November 25, 2016; Chen et al., 2012) with a word size of 10 under the thresholds of the highest bit score and e-value ($\leq 1e-05$).

Clustered Regularly Interspaced Short Palindromic Repeat Analysis

Clustered regularly interspaced short palindromic repeat (CRISPR) regions were identified from contigs using the CRISPRs Finder online (Grissa et al., 2007). The sequences of CRISPR direct repeats and spacers in the four sequenced samples were extracted. To characterize a target of each spacer sequence, the spacer list was subjected to a BLASTN search with a word size of seven against the plasmid and the viral database, as described previously (Pride et al., 2012; Rho et al., 2012; Watanabe et al., 2013; Zhou et al., 2015; Wang et al., 2016). Hits were considered significant under the threshold of the e-value ($\leq 1e-02$).

Metabolic Reconstruction

The metabolic and regulatory pathways were reconstructed using 146 KEGG pathways and 22 KEGG regulatory pathways. An overview of the complete metabolism in biological systems was provided through the web-based interactive tool “Pathways Explorer” (iPath) (Yamada et al., 2011).

Mobile Genetic Elements Prediction

The assembled contigs were compared to reference genomes (*P. intermedia* 17 and *P. nigrescens* ATCC 33563) through

MUMmer 3.23 (Kurtz et al., 2004), and sequences under the thresholds of length ≥ 500 bp and sequence identity < 90% were predicted as insert sequences. The predicted insert sequences were further searched using BlastX analysis in MvirDB database (MvirDB; last accessed April 21, 2012; Zhou et al., 2007) to investigate the possible mobile genetic elements (e.g., conjugate transposons, transposases, and integrases) (E value $\leq 1e-7$). Prophages or degraded prophages were predicted using PHAge Search Tool (PHAST), which is a web server that available at (<http://phast.wishartlab.com>) (Zhou et al., 2011).

Data Access

The raw reads have been deposited in the Sequence Read Archive under the following accession numbers: Pi-disease SRR5830884; Pi-health SRR5830885; Pn-disease SRR5830886; Pn-health SRR5830889.

AUTHOR CONTRIBUTIONS

Experimental design: YFZ and JW; Clinical examination and sampling: MZ and YLZ; Experiments: YFZ, MZ, and YS; data analysis: YFZ, JW, YS, YLZ, and QZ; Manuscript revised: YFZ, JW, and MZ. All authors read and approved the final manuscript.

FUNDING

This study was supported by grants 81300880 from National Natural Science Foundation of China.

ACKNOWLEDGMENTS

We thank Aimin Qi from Shanghai Biotechnology Corporation for helpful technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fcimb.2017.00409/full#supplementary-material>

REFERENCES

- Abusleme, L., Dupuy, A. K., Dutzan, N., Silva, N., Burleson, J. A., Strausbaugh, L. D., et al. (2013). The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *ISME J.* 7, 1016–1025. doi: 10.1038/ismej.2012.174
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Belda-Ferre, P., Alcaraz, L. D., Cabrera-Rubio, R., Romero, H., Simon-Soro, A., Pignatelli, M., et al. (2012). The oral metagenome in health and disease. *ISME J.* 6, 46–56. doi: 10.1038/ismej.2011.85
- Bogren, A., Teles, R. P., Torresyap, G., Haffajee, A. D., Socransky, S. S., and Wennstrom, J. L. (2007). Clinical and microbiologic changes associated with the combined use of a powered toothbrush and a triclosan/copolymer dentifrice: a 3-year prospective study. *J. Periodontol.* 78, 1708–1717. doi: 10.1902/jop.2007.070028
- Chen, L., Xiong, Z., Sun, L., Yang, J., and Jin, Q. (2012). VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res.* 40, D641–D645. doi: 10.1093/nar/gkr989
- Darwin, K. H., Lin, G., Chen, Z., Li, H., and Nathan, C. F. (2005). Characterization of a *Mycobacterium tuberculosis* proteasomal ATPase homologue. *Mol. Microbiol.* 55, 561–571. doi: 10.1111/j.1365-2958.2004.04403.x
- Dingsdag, S., Nelson, S., and Coleman, N. V. (2016). Bacterial communities associated with apical periodontitis and dental implant failure. *Microb. Ecol. Heal. Dis.* 27:31307. doi: 10.3402/mehd.v27.31307
- Dorn, B. R., Leung, K. L., and Progulsk-Fox, A. (1998). Invasion of human oral epithelial cells by *Prevotella intermedia*. *Infect. Immun.* 66, 6054–6057.
- Duran-Pinedo, A. E., Chen, T., Teles, R., Starr, J. R., Wang, X., Krishnan, K., et al. (2014). Community-wide transcriptome of the oral microbiome in subjects with and without periodontitis. *ISME J.* 8, 1659–1672. doi: 10.1038/ismej.2014.23

- Enninga, J., and Rosenshine, I. (2009). Imaging the assembly, structure and activity of type III secretion systems. *Cell. Microbiol.* 11, 1462–1470. doi: 10.1111/j.1462-5822.2009.01360.x
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Gharbia, S. E., Haapasalo, M., Shah, H. N., Kotiranta, A., Lounatmaa, K., Pearce, M. A., et al. (1994). Characterization of *Prevotella intermedia* and *Prevotella nigrescens* isolates from periodontic and endodontic infections. *J. Periodontol.* 65, 56–61. doi: 10.1902/jop.1994.65.1.56
- Griffen, A. L., Beall, C. J., Campbell, J. H., Firestone, N. D., Kumar, P. S., Yang, Z. K., et al. (2012). Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *ISME J.* 6, 1176–1185. doi: 10.1038/ismej.2011.191
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57. doi: 10.1093/nar/gkm360
- Hajishengallis, G., Darveau, R. P., and Curtis, M. A. (2012). The keystone-pathogen hypothesis. *Nat. Rev. Microbiol.* 10, 717–725. doi: 10.1038/nrmicro2873
- Hajishengallis, G., Liang, S., Payne, M. A., Hashim, A., Jotwani, R., Eskan, M. A., et al. (2011). Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. *Cell Host. Microbe* 10, 497–506. doi: 10.1016/j.chom.2011.10.006
- Ho, B. T., Dong, T. G., and Mekalanos, J. J. (2014). A view to a kill: the bacterial type VI secretion system. *Cell Host. Microbe* 15, 9–21. doi: 10.1016/j.chom.2013.11.008
- Hong, B. Y., Furtado Araujo, M. V., Strausbaugh, L. D., Terzi, E., Ioannidou, E., and Diaz, P. I. (2015). Microbiome profiles in periodontitis in relation to host and disease characteristics. *PLoS ONE* 10:e0127077. doi: 10.1371/journal.pone.0127077
- Hsiao, W. W., Li, K. L., Liu, Z., Jones, C., Fraser-Liggett, C. M., and Fouad, A. F. (2012). Microbial transformation from normal oral microbiota to acute endodontic infections. *BMC Genomics* 13:345. doi: 10.1186/1471-2164-13-345
- Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234
- Jorth, P., Turner, K. H., Gumus, P., Nizam, N., Buduneli, N., and Whiteley, M. (2014). Metatranscriptomics of the human oral microbiome during health and disease. *MBio* 5, e01012–e01014. doi: 10.1128/mBio.01012-14
- Kamma, J. J., Nakou, M., Gmur, R., and Baehni, P. C. (2004). Microbiological profile of early onset/aggressive periodontitis patients. *Oral Microbiol. Immunol.* 19, 314–321. doi: 10.1111/j.1399-302x.2004.00161.x
- Kirst, M. E., Li, E. C., Alfant, B., Chi, Y. Y., Walker, C., Magnusson, I., et al. (2015). Dysbiosis and alterations in predicted functions of the subgingival microbiome in chronic periodontitis. *Appl. Environ. Microbiol.* 81, 783–793. doi: 10.1128/AEM.02712-14
- Krakoff, L. R. (2006). Home is where the heart (and prevention of heart disease) is. *J. Hypertens.* 24, 2345–2346. doi: 10.1097/HJH.0b013e3280113624
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12
- Li, Y., He, J., He, Z., Zhou, Y., Yuan, M., Xu, X., et al. (2014). Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. *ISME J.* 8, 1879–1891. doi: 10.1038/ismej.2014.28
- Lie, M. A., van der Weijden, G. A., Timmerman, M. F., Loos, B. G., van Steenberghe, T. J., and van der Velden, U. (2001). Occurrence of *Prevotella intermedia* and *Prevotella nigrescens* in relation to gingivitis and gingival health. *J. Clin. Periodontol.* 28, 189–193. doi: 10.1034/j.1600-051x.2001.028002189.x
- MacIntyre, D. L., Miyata, S. T., Kitaoka, M., and Pukatzki, S. (2010). The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19520–19524. doi: 10.1073/pnas.1012931107
- Maeda, N., Okamoto, M., Kondo, K., Ishikawa, H., Osada, R., Tsurumoto, A., et al. (1998). Incidence of *Prevotella intermedia* and *Prevotella nigrescens* in periodontal health and disease. *Microbiol. Immunol.* 42, 583–589. doi: 10.1111/j.1348-0421.1998.tb02328.x
- Moura-Grec, P. G., Marsicano, J. A., Carvalho, C. A., and Sales-Peres, S. H. (2014). Obesity and periodontitis: systematic review and meta-analysis. *Cien. Saude Colet.* 19, 1763–1772. doi: 10.1590/1413-81232014196.13482013
- Naito, M., Ogura, Y., Itoh, T., Shoji, M., Okamoto, M., Hayashi, T., et al. (2016). The complete genome sequencing of *Prevotella intermedia* strain OMA14 and a subsequent fine-scale, intra-species genomic comparison reveal an unusual amplification of conjugative and mobile transposons and identify a novel *Prevotella*-lineage-specific repeat. *DNA Res.* 23, 11–19. doi: 10.1093/dnares/dsv032
- Nasidze, I., Li, J., Quinque, D., Tang, K., and Stoneking, M. (2009). Global diversity in the human salivary microbiome. *Genome Res.* 19, 636–643. doi: 10.1101/gr.084616.108
- Preshaw, P. M., Alba, A. L., Herrera, D., Jepsen, S., Konstantinidis, A., Makrilakis, K., et al. (2012). Periodontitis and diabetes: a two-way relationship. *Diabetologia* 55, 21–31. doi: 10.1007/s00125-011-2342-y
- Pride, D. T., Salzman, J., Haynes, M., Rohwer, F., Davis-Long, C., White, R. A. III., et al. (2012). Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. *ISME J.* 6, 915–926. doi: 10.1038/ismej.2011.169
- Purushe, J., Fouts, D. E., Morrison, M., White, B. A., Mackie, R. I., Coutinho, P. M., et al. (2010). Comparative genome analysis of *Prevotella ruminicola* and *Prevotella bryantii*: insights into their environmental niche. *Microb. Ecol.* 60, 721–729. doi: 10.1007/s00248-010-9692-8
- Rho, M., Wu, Y. W., Tang, H., Doak, T. G., and Ye, Y. (2012). Diverse CRISPRs evolving in human microbiomes. *PLoS Genet.* 8:e1002441. doi: 10.1371/journal.pgen.1002441
- Rohmer, L., Hocquet, D., and Miller, S. I. (2011). Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. *Trends Microbiol.* 19, 341–348. doi: 10.1016/j.tim.2011.04.003
- Ruan, Y., Shen, L., Zou, Y., Qi, Z., Yin, J., Jiang, J., et al. (2015). Comparative genome analysis of *Prevotella intermedia* strain isolated from infected root canal reveals features related to pathogenicity and adaptation. *BMC Genomics* 16:122. doi: 10.1186/s12864-015-1272-3
- Saito, K., Takahashi, N., Horiuchi, H., and Yamada, T. (2001). Effects of glucose on formation of cytotoxic end-products and proteolytic activity of *Prevotella intermedia*, *Prevotella nigrescens* and *Porphyromonas gingivalis*. *J. Periodont. Res.* 36, 355–360. doi: 10.1034/j.1600-0765.2001.360602.x
- Schmidt, H., and Hensel, M. (2004). Pathogenicity islands in bacterial pathogenesis. *Clin. Microbiol. Rev.* 17, 14–56. doi: 10.1128/CMR.17.1.14-56.2004
- Sharma, A. (2010). Virulence mechanisms of *Tannerella forsythia*. *Periodontol.* 2000 54, 106–116. doi: 10.1111/j.1600-0757.2009.00332.x
- Sharma, A., Sojar, H. T., Glurich, I., Honma, K., Kuramitsu, H. K., and Genco, R. J. (1998). Cloning, expression, and sequencing of a cell surface antigen containing a leucine-rich repeat motif from *Bacteroides forsythus* ATCC 43037. *Infect. Immun.* 66, 5703–5710.
- Shi, B., Chang, M., Martin, J., Mitreva, M., Lux, R., Klokkevold, P., et al. (2015). Dynamic changes in the subgingival microbiome and their potential for diagnosis and prognosis of periodontitis. *MBio* 6, e01926–e01914. doi: 10.1128/mBio.01926-14
- Siamer, S., and Dehio, C. (2015). New insights into the role of *Bartonella* effector proteins in pathogenesis. *Curr. Opin. Microbiol.* 23, 80–85. doi: 10.1016/j.mib.2014.11.007
- Siqueira, J. F. Jr., and Rocas, I. N. (2013). Microbiology and treatment of acute apical abscesses. *Clin. Microbiol. Rev.* 26, 255–273. doi: 10.1128/CMR.00082-12
- Socransky, S. S., and Haffajee, A. D. (1994). Evidence of bacterial etiology: a historical perspective. *Periodontol.* 2000 5, 7–25. doi: 10.1111/j.1600-0757.1994.tb00016.x
- Socransky, S. S., and Haffajee, A. D. (2005). Periodontal microbial ecology. *Periodontol.* 2000 38, 135–187. doi: 10.1111/j.1600-0757.2005.00107.x
- Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C., and Kent, R. L. Jr. (1998). Microbial complexes in subgingival plaque. *J. Clin. Periodontol.* 25, 134–144. doi: 10.1111/j.1600-051X.1998.tb02419.x
- Stewart, R., and West, M. (2016). Increasing evidence for an association between periodontitis and cardiovascular disease. *Circulation* 133, 549–551. doi: 10.1161/CIRCULATIONAHA.115.020869
- Takahashi, N., and Yamada, T. (2000a). Glucose metabolism by *Prevotella intermedia* and *Prevotella nigrescens*. *Oral Microbiol. Immunol.* 15, 188–195. doi: 10.1034/j.1399-302x.2000.150307.x

- Takahashi, N., and Yamada, T. (2000b). Pathways for amino acid metabolism by *Prevotella intermedia* and *Prevotella nigrescens*. *Oral Microbiol. Immunol.* 15, 96–102. doi: 10.1034/j.1399-302x.2000.150205.x
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29, 22–28.
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., et al. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44, 6614–6624. doi: 10.1093/nar/gkw569
- Teapaisan, R., Douglas, C. W., and Walsh, T. F. (1995). Characterisation of black-pigmented anaerobes isolated from diseased and healthy periodontal sites. *J. Periodont. Res.* 30, 245–251. doi: 10.1111/j.1600-0765.1995.tb02129.x
- The UniProt Consortium (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* 43, D204–D212. doi: 10.1093/nar/gku989
- Uzel, N. G., Teles, F. R., Teles, R. P., Song, X. Q., Torresyap, G., Socransky, S. S., et al. (2011). Microbial shifts during dental biofilm re-development in the absence of oral hygiene in periodontal health and disease. *J. Clin. Periodontol.* 38, 612–620. doi: 10.1111/j.1600-051X.2011.01730.x
- Wang, J., Gao, Y., and Zhao, F. (2016). Phage-bacteria interaction network in human oral microbiome. *Environ. Microbiol.* 18, 2143–2158. doi: 10.1111/1462-2920.12923
- Wang, J., Qi, J., Zhao, H., He, S., Zhang, Y., Wei, S., et al. (2013). Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. *Sci. Rep.* 3:1843. doi: 10.1038/srep01843
- Watanabe, T., Nozawa, T., Aikawa, C., Amano, A., Maruyama, F., and Nakagawa, I. (2013). CRISPR regulation of intraspecies diversification by limiting IS transposition and intercellular recombination. *Genome Biol. Evol.* 5, 1099–1114. doi: 10.1093/gbe/evt075
- Yamada, T., Letunic, I., Okuda, S., Kanehisa, M., and Bork, P. (2011). iPath2.0: interactive pathway explorer. *Nucleic Acids Res.* 39, W412–W415. doi: 10.1093/nar/gkr313
- Yamane, K., Yamanaka, T., Yamamoto, N., Furukawa, T., Fukushima, H., Walker, C. B., et al. (2005). A novel exopolysaccharide from a clinical isolate of *Prevotella nigrescens*: purification, chemical characterization and possible role in modifying human leukocyte phagocytosis. *Oral Microbiol. Immunol.* 20, 1–9. doi: 10.1111/j.1399-302X.2004.00178.x
- Yang, F., Zeng, X., Ning, K., Liu, K. L., Lo, C. C., Wang, W., et al. (2012). Saliva microbiomes distinguish caries-active from healthy human populations. *ISME J.* 6, 1–10. doi: 10.1038/ismej.2011.71
- Zambon, J. J., Reynolds, H. S., and Slots, J. (1981). Black-pigmented *Bacteroides* spp. in the human oral cavity. *Infect. Immun.* 32, 198–203.
- Zhang, C., and Zhao, L. (2016). Strain-level dissection of the contribution of the gut microbiome to human metabolic disease. *Genome Med.* 8:41. doi: 10.1186/s13073-016-0304-1
- Zhou, C. E., Smith, J., Lam, M., Zemla, A., Dyer, M. D., and Slezak, T. (2007). MvirDB—a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. *Nucleic Acids Res.* 35, D391–D394. doi: 10.1093/nar/gkl791
- Zhou, H., Zhao, H., Zheng, J., Gao, Y., Zhang, Y., Zhao, F., et al. (2015). CRISPRs provide broad and robust protection to oral microbial flora of gingival health against bacteriophage challenge. *Protein Cell* 6, 541–545. doi: 10.1007/s13238-015-0182-0
- Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res.* 39, W347–W352. doi: 10.1093/nar/gkr485

Conflict of Interest Statement: 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.

Copyright © 2017 Zhang, Zhen, Zhan, Song, Zhang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor 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.