copyright by all rights reserved Quintessenz

Structure and Composition of Candidate Phyla Radiation in Supragingival Plaque of Caries Patients

Song JIANG[1], Jie NIE[1], Yu Xing CHEN[1], Xiao Yan WANG[1], Feng CHEN[2]

Objective: *To investigate the composition and abundance of candidate phyla radiation (CPR) in the oral cavity in caries patients and a healthy population.*

Methods: *The raw macrogenomic sequencing data for a total of 88 subjects were downloaded from the National Centre for Biotechnology Sequence Read Archive (NCBI SRA) public database according to the public data usage specifications. Trimmomatic (Department for Metabolic Networks, Potsdam, Germany) and Bowtie 2 (University of Maryland, College Park, MD, USA) were used to quality control and dehost the host sequences. Species annotation was made using Kraken2 (Johns Hopkins University, Baltimore, MD, USA) and Bracken (Johns Hopkins University) based on the reference database. According to the results of the species annotation, the species-significant differences and species correlation of caries and healthy oral microbiota in species composition and microbiota diversity were analysed to study the distribution and abundance differences of CPR in the oral environment.*

Results: *Proteobacteria, Firmicutes, Bacteroidetes, Actinobacteria and Fusobacteria were the main components. The relative abundance of TM7 (Candidatus Saccharibacteria) and GN02 (Candidatus Gracilibacteria) of CPR is second only to the aforementioned five bacteria, indicating that CPR is an important part of the oral microbiota. TM7 and GN02 were common to both the caries patients and healthy patients and were detected in all samples, suggesting that CPR is the 'core microbiome'. There was a correlation between CPR and a variety of oral microbiota, among which the positive correlation with Capnocytophaga was the strongest, suggesting that Capnocytophaga might be the potential host bacteria of CPR.*

Conclusion: *CPR is an indispensable part of the oral microbiota. It is the 'core microflora' of the oral cavity and may play an important role in the stability and function of the oral microecological environment. Capnocytophaga may be the potential host bacteria of CPR.*

Key words: *candidate phyla radiation, caries, core microbiome, metagenomics, oral microbiota*

Chin J Dent Res 2022;25(2):107–118; doi: 10.3290/j.cjdr.b3086339

1 Department of Cariology and Endodontology, Peking University School and Hospital of Stomatology, National Centre of Stomatology, National Clinical Research Centre for Oral Diseases, Beijing, P.R. China.

2 Central Laboratory, Peking University School and Hospital of Stomatology, Beijing, P.R. China.

Corresponding authors: Dr Xiao Yan WANG, Department of Cariology and Endodontology, Peking University School and Hospital of Stomatology, #22 Zhongguancun South Avenue, Haidian District, Beijing 100081, P.R. China. Tel: 86-10-82195525; Fax: 86-10-82195525. Email: wangxiaoyan@pkuss.bjmu.edu.cn

Dr Feng CHEN, Central Laboratory, Peking University School and Hospital of Stomatology, #22 Zhongguancun South Avenue, Haidian District, Beijing 100081, P.R. China. Tel: 86-10 82195773; Fax: 86-10-82195773. Email: chenfeng2011@hsc.pku.edu.cn

This study was supported by grants-in-aid from the National Science Foundation China (grant no. 81991501) and the open project of State Key Laboratory of Oral Diseases (grant no. SKLOD2021OF03).

copyright by all rights reserved Quintessenz

Dental caries is a common and frequently occurring oral disease that affects the health of residents in China. According to statistics from the Global Burden of Disease Study (GBD) from 1990 to 2016, it was estimated that there were 2.4 billion untreated dental caries patients worldwide, with the highest prevalence of oral disease in the world[1]. Dental caries is an infectious oral disease, and its occurrence and development are closely related to oral microorganisms. In recent years, as research into the microbiota and metabolomics has progressed, scholars have begun to recognise that the microorganisms in the dental plaque biofilm on the human tooth surface as a whole participate in the occurrence and development of dental caries and have proposed the "ecological plaque hypothesis"[2,3]. This hypothesis suggests that the changes in dental plaque microecology are the causes of the onset and pathogenesis of dental caries. The microbial community in a healthy oral environment is maintained in a relatively balanced state, and microorganisms and the environment interact and restrict each other to avoid the occurrence of diseases; however, when the microbial community changes beyond its resistance threshold, the relative balance is broken, causing ecological imbalance of the flora and finally leading to the occurrence of dental caries, periodontal disease and other oral diseases. Dental caries is a disease caused by multiple factors, and its pathogenic bacteria are not a single bacterium, but multiple microorganisms[4]. Understanding and elucidating the pathogenic mechanism of oral microorganisms in dental caries is of great importance for the development of disease prevention and treatment techniques.

Since the oral cavity is an open environment and in continuous contact with the external environment, the composition of the oral microbiota is diverse. A large number of complex microorganisms colonize different sites in the mouth. Candidate phyla radiation (CPR) is a newly discovered large flora including more than 75 phylum bacteria, and accounts for an important part of one quarter of the whole bacterial domain[5]. An increasing number of studies have shown that CPR is also common in oral microbial communities, including TM7 (*Candidatus Saccharibacteria*), GN02 (*Candidatus Gracilibacteria*) and SR1 (*Candidatus Absconditabacteria*)[6,7]. CPR cells are small in size and lack a large number of essential genes related to bacterial survival, which makes them difficult to obtain by conventional laboratory culture. Very little is known about their biological characteristics and role in the occurrence and development of dental caries. CPR may have a potentially pathogenic association with a variety of oral diseases.

At present, research on oral CPR is focused mainly on the oral microbial community of periodontitis. The abundance of oral CPR in periodontitis patients is significantly higher than that in the healthy population and is related to the severity of periodontitis[8-15]. Previous studies related to caries and CPR are rare, and all of them used amplicon sequencing of 16S rRNA for their analysis and some suggested that CPR might be related to the occurrence of dental caries[16-18]. He et al[19] obtained the first culturable strain of TM7 by co-culture and proved the characteristics of its parasitic bacteria. Studies have shown that there is a complex interaction between TM7 and its host bacteria, which can enhance the biofilm formation ability of the host bacteria and help it to avoid the immune system of the human body. These characteristics of CPR may also be related to the occurrence and development of dental caries, but research is still insufficient. Previous studies on the aetiology and mechanism of dental caries have not focused on this newly discovered CPR, and a large number of second-generation sequenced genomic data have not been explored fully[20-22]. As interest in the exploration of unknown microorganisms has grown, ab initio analysis and reuse of metagenomic sequencing raw data have become important research topics, greatly improving understanding of the microbial world. The application of metagenomic data to explore the correlation between CPR in the oral cavity and the occurrence and development of dental caries further deserves to be studied in greater depth.

Materials and methods

Data acquisition

Medline/PubMed database (up to 2020.1) was searched with "metagenomic AND caries", "Candidate Phyla Radiation AND caries" and "Saccharibacteria AND caries" as the search terms. A total of 75 studies related to metagenomic and caries research were reviewed. Of these, 20 were excluded during preliminary screening. After browsing the literature abstracts and excluding studies using 16S rRNA sequencing, three studies that used shotgun metagenomics to study caries microorganisms were finally obtained[20-22]. The authors browsed the data availability statement of the literature, in which the original sequencing data of one literature was not public and could not be obtained[21]. The original sequencing data from the other two studies were uploaded to the NCBI database for open use. The literature that could not be obtained was excluded, and the full text of the

copyright all rights reserved Quintessence

remaining two studies was read. The results of one of the studies did not report the existence of CPR[22]. Another study reported the existence of two types of CPR, TM7 and GN02[20], but was not examined further. This study was selected as the source of the original metagenomic sequencing data[20].

The original metagenomic sequencing data of a total of 88 subjects were downloaded from the NCBI SRA public database, and the accession numbers were SRR6865436 to SRR6865523 in accordance with the public data usage specifications. Dental plaque samples were collected from participants of the University of Adelaide Craniofacial Biology Research Group (CBRG) and the Murdoch Children's Research Institute (MCRI)'s Peri/Postnatal Epigenetic Twins Study (PETS)[20]. The PETS (n = 193) and CBRG (n = 292) cohorts were composed of twins aged 5 to 11 years old. The number of pairs selected for metagenomic sequencing was constrained by budget; thus, a subset of the broader clinical cohort was subsampled[20]. In the original study, oral examination of the subjects was conducted by trained and experienced clinicians. The International Caries Detection and Assessment System (ICDAS II) was used to evaluate caries and was employed to diagnose caries lesions from the initial stage of enamel caries lesions to dentine caries lesions and the final stage of the latter. Subjects with enamel or dentine caries lesions were diagnosed as caries patients, whereas those without enamel or dentine caries were considered as healthy patients. The subjects did not brush their teeth the night before or on the day of plaque collection. A sterile swab was used to wipe the plaque samples thoroughly along the gingival edge and tooth. The sterile swab was then placed in a microcentrifuge tube containing 500 µl bacterial RNA protectant, and the samples were immediately stored in a refrigerator at –80°C for DNA extraction. The DNA in the samples was then extracted with lysozyme. Phenol/chloroform isoamyl alcohol was used for extraction and ethanol was employed for precipitation. The metagenomic sequencing library was prepared using a NEBNext DNA kit for Illumina (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's instructions, then 300 cycles of the metagenomic library were sequenced using a NextSeq 500 kit (Illumina, La Jolla, CA, USA) according to the manufacturer's specified standards to obtain the original metagenomic sequencing data.

Data quality control

The SRA Toolkit (version 2.10.8) was used to convert the downloaded raw data (*.sra files) into the data format required for subsequent metagenomic analysis routines (*.fastq files) to obtain raw reads. The quality of raw reads was assessed using FastQC (version 0.11.7, Babraham Bioinformatics, Cambridge, UK). Trimmomatic[23] (version 0.36), Department for Metabolic Networks, Potsdam, Germany) was used to cut adapters and primers and filter low-quality sequences from raw reads of each sample (parameter set to ILLUMINACLIP: adapters/TruSeq2-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36). Bowtie 2 (version 2.3.5.1) University of Maryland, College Park, MD, USA) was used to compare the sequences after quality control with the human reference genome hg38, and all sequences from human hosts were filtered to obtain clean reads containing only all microbial genome information[24]. FastQC (version 0.11.7) was used again to assess the quality of clean reads, and MultiQC[25] (version 1.9, Stockholm University, Stockholm, Sweden) was employed to summarise raw and clean reads quality, respectively.

Analysis of species composition

According to the Kraken2 reference database, clean reads from all samples were annotated using Kraken2[26] (version 2.0.8) (parameter set to --confidence 0.2). After classifying the species annotation results obtained from Kraken2, Bracken[27] (version 2.5.3) was used to estimate the abundance of each taxonomic level of metagenomic samples by Bayesian reestimation[28]. Using R language (R Core Team, R Foundation for Statistical Computing, Vienna, Austria), the bar charts of two groups of species composition at the phylum level and species level were drawn based on the Bracken corrected results. Using R, Venn diagrams were drawn based on species annotation results (reads threshold > 1000). Endemic or common species at species level were analysed for the caries group and healthy group. The top 20 species with relative abundance were selected for sample clustering analysis, and sample clustering was realised according to the species composition of the samples. R language was used to draw the clustering heat map of the samples, so as to study the similarities between them.

Analysis of bacterial diversity

According to the species abundance-sample matrix, principal component analysis (PCA) scatter plots and 3D maps were drawn using R language. Principal coordinates analysis (PCOA) was carried out on species abundance based on Bray-Curtis. R was used to draw scatter plots and 3D maps, and analysis of similarities

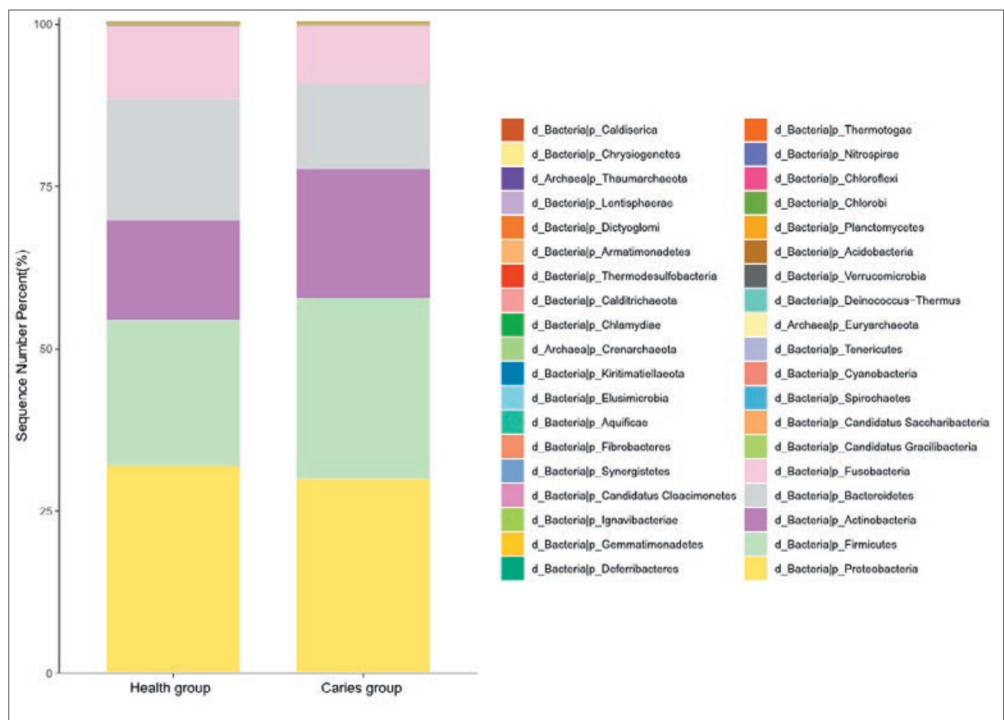


Fig 1 Relative distribution of the caries group and healthy group at one level. *Proteobacteria*, *Firmicutes*, *Bacteroidetes*, *Actinobacteria* and *Fusobacteria* accounted for more than 99% of the total sequence.

(ANOSIM) was employed to calculate β diversity between groups to compare the differences in microbial community composition.

Significant analysis of species differences

Linear discriminant analysis (LDA) was used to calculate the differences in the abundance of species at different taxonomic levels between the two groups. LDA effect size (LEfSE) was used to determine the labelled species of the two groups. The LDA score threshold for difference identification was set as 2.0, and the evolutionary branch diagram and LDA bar chart were drawn using R[29].

Analysis of correlation network

The species abundance for the caries group and healthy group was counted, and the top 40 species with the highest abundance in the two groups were screened for a Spearman correlation analysis. When correlation coefficient $|R| > 0.4$ and $P < 0.01$, it is believed that there is correlation between the two bacteria, and the correlation network diagram is drawn using R.

Results

Data overview

Espinoza et al[20] performed metagenomic shotgun sequencing on a total of 88 dental plaque samples from 44 pairs of twins and uploaded the raw data to the NCBI SRA database for open use. Of the 88 subjects, 58 with caries were classified into the caries group and 30 without caries were classified into the healthy group.

The original sequencing data were downloaded from the NCBI database according to the accession numbers SRR6865436 to SRR6865523 (in line with the standards for the use of data in public databases). After format conversion, a total of 96 Gb of paired double-ended sequence data were obtained, and each sample included an average of 5.52 million sequence fragments (1.1 Gb). After removal of low-quality sequences, primers and linkers and screening of human gene sequences, 47.8% of sequence data remained.

Species composition

Bacteria, archaea and viruses were detected in 88 samples, of which bacteria accounted for 99.90%, including 38 phyla. The main bacteria in the caries group were the same as those in the healthy group, with *Proteo-*

copyright all rights reserved Quintessenz

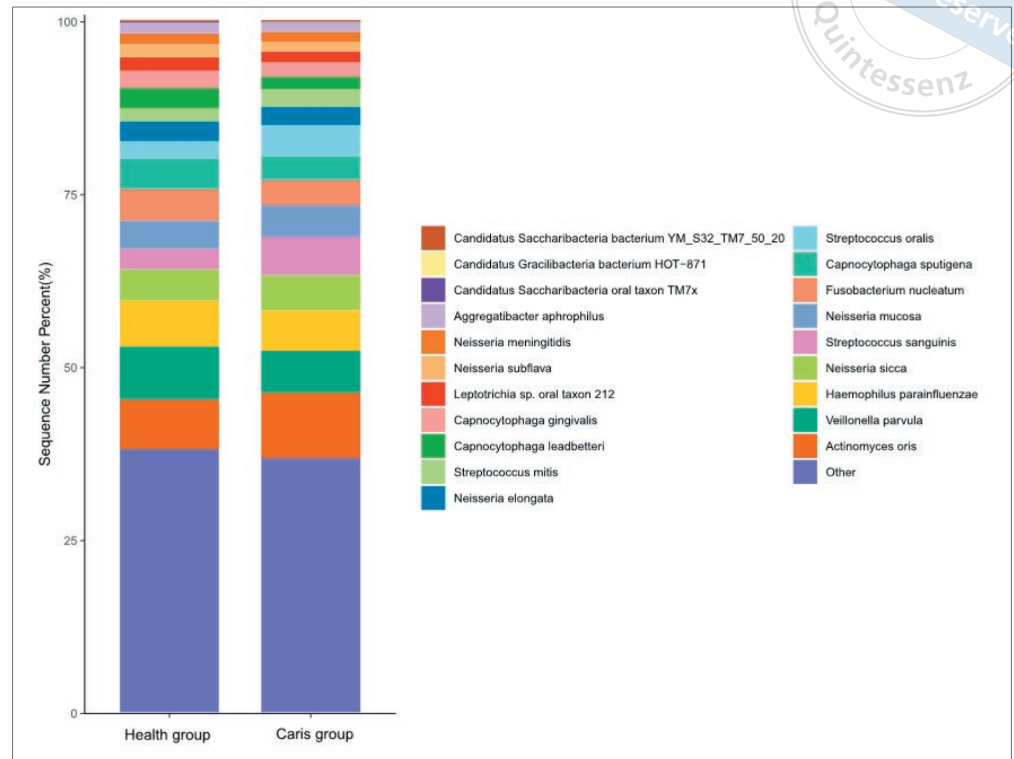


Fig 2 Relative distribution of the caries group and healthy group at the species level. *Actinomyces oris*, *Veillonella parvula*, *Haemophilus parainfluenzae* and *Neisseria sicca* are the species with relatively high abundance in both groups.

bacteria, *Firmicutes*, *Bacteroidetes*, *Actinobacteria* and *Fusobacteria* as the main bacteria, accounting for more than 99% of the total sequences. The percentages of TM7 and GN02 belonging to CPR were second to those of the aforementioned five phyla; however, we did not find another bacterium SR1 belonging to CPR.

In the caries group, the dominant phylum was *Proteobacteria* (relative abundance 29.67%), followed by *Firmicutes* (27.96%), *Actinobacteria* (19.78%), *Bacteroidetes* (13.14%) and *Fusobacteria* (8.93%), GN02 (0.15%) and TM7 (0.12%). In the healthy group, the dominant phylum was Proteobacteria (relative abundance 31.70%), followed by *Firmicutes* (22.56%), *Bacteroidetes* (18.72%), *Actinobacteria* (15.37%), *Fusobacteria* (11.10%), TM7 (0.18%) and GN02 (0.09%). Figure 1 shows the horizontal relative distribution for the two groups.

At species level, the five most dominant bacteria relative to each other in the caries group were *Actinomyces oris* (relative abundance 9.42%), *Veillonella parvula* (6.11%), *Haemophilus parainfluenzae* (5.82%), *Streptococcus sanguinis* (5.52%) and *Neisseria sicca* (5.12%). In addition, we found three species of CPR bacteria belonging to the TM7 and GN02 phylum, namely TM7x, *Candidatus Saccharibacteria* bacterium YM_S32_TM7_50_20 and *Candidatus Gracilibacteria* bacterium HOT-871. The relative abundance of *TM7x* belonging to the TM7 phylum was 0.12%, and that of *Candidatus Gracilibacteria* bacterium HOT-871 belonging to the GN02 phylum was 0.15%. The five most dominant bacteria in the healthy group were *Actinomyces oris* (relative abundance 7.14%), *Veillonella parvula* (7.67%), *Haemophilus parainfluenzae* (6.64%), *Neisseria sicca* (4.57%) and *Fusobacterium nucleatum* (4.49%). The relative abundance of *TM7x*

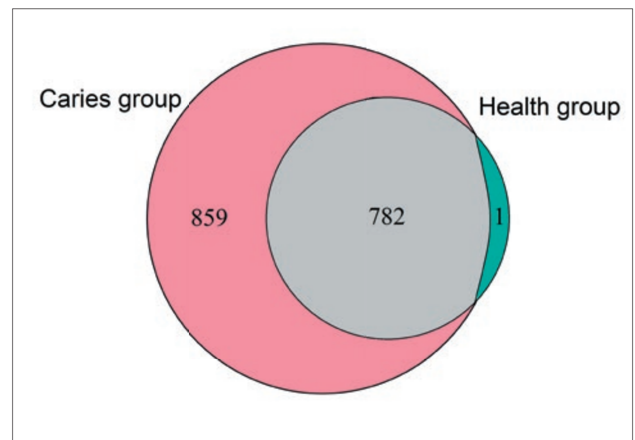


Fig 3 Common species at the species level in the caries group and healthy group. The overlapping part indicates the species composition shared by the groups, and the non-overlapping part indicates the species composition unique to each group.

copyright by
all rights reserved
Quintessenz

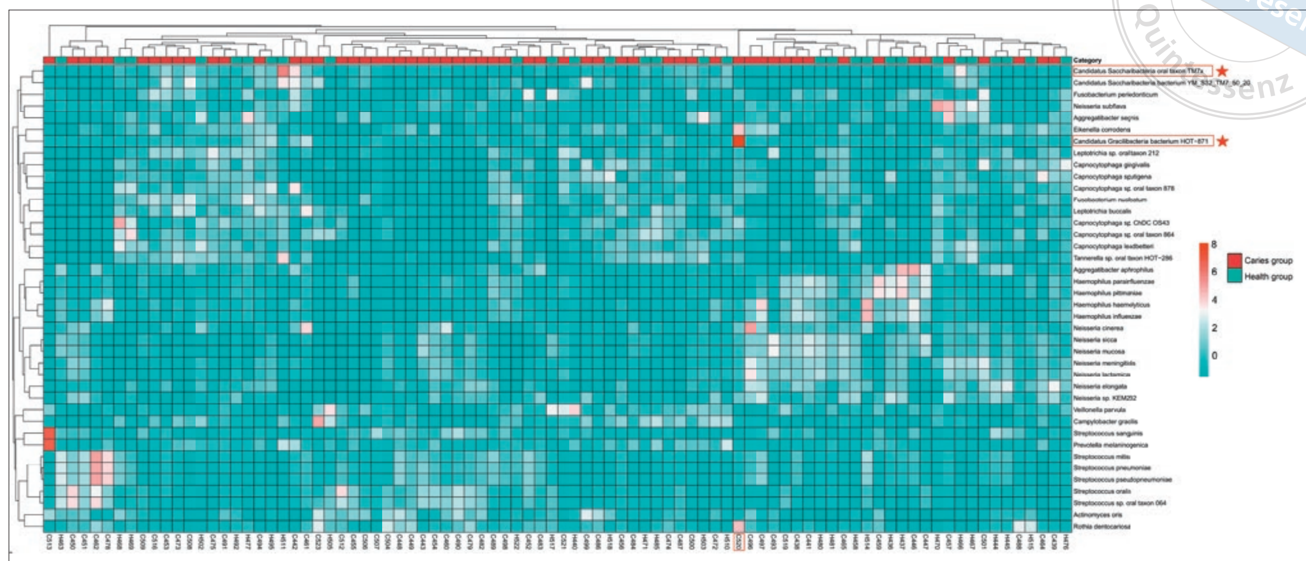


Fig 4 Results of heat map clustering at the species classification level. The horizontal axis is sample name information and also includes grouping information, and the vertical axis is species annotation name (species level in this figure). The cluster tree on the left is the similarity cluster of species distribution in each sample, the sample cluster tree on the top, and the heat map in the middle is the heat map of the relative abundance of species.

was 0.17% and that of *Gracilibacteria bacterium HOT-871* was 0.09%. Figure 2 shows the horizontal relative distribution of two groups of species.

The species in the caries group and the healthy group were analysed and a Venn diagram (Fig 3) was drawn to find the endemic and common species between the two groups. The results showed that there was a total of 782 species in the caries group and the healthy group, of which *TM7x* and *Graciliberia bacterium HOT-871* were species in common in the two groups. There were 859 endemic species in the caries group, among which *Streptococcus influenzalis*, *Arthrobacter sp. ATCC 21022* and *Streptococcus phage Dp-1* were more abundant, whereas only one endemic species, *Enterobacteriaceae bacterium ENNIH2*, was found in the healthy group.

Clustering analysis was performed on the first 20 bacteria with higher abundance in the samples (Fig 4). The results showed that there was a good clustering relationship between the samples in the caries group and the healthy group. *Neisseria*, *Streptococcus* and *Capnocytophaga* showed a good clustering relationship, respectively. Some of the caries samples showed a high abundance of CPR. It is noteworthy that the abundance of *Graciliberia bacterium HOT-871* in the C520 sample of the caries group was extremely high. The patients in this sample developed dentine caries and were not treated, suggesting that *Graciliberia bacterium HOT-871* might be related to caries activity.

Structural similarity of the microbiome

PCA was used for ordination analysis based on the original species composition matrix (Fig 5) and PCoA was used for analysis based on the distance matrix calculated from the species composition (Fig 6) to assess the similarity of the bacterial community structures between the two groups. The point distributions between the caries group and the healthy group showed a trend of cross-coverage, and there was no significant difference between the groups. The species composition of individual samples in the caries group was significantly different to that of other samples, but this did not affect the similarity of the overall flora structure of the two groups.

Significant differences in bacteria

LEfSe helped to identify the bacteria that were most likely to show significant differences at each classification level between the caries group and the healthy group (Figs 7 and 8). *Firmicutes* were the marker bacteria at the gate level in the caries group, whereas in the healthy group it was *Bacteroidetes*. The taxonomic abundance at the level of one class, two orders, two families and two genera in the caries group was relatively high. At the species level, *Streptococcus sanguinis*, *Streptococcus oralis* and other *Streptococcus* bacteria were enriched significantly in the caries group. In the healthy group, the abundance was higher at the level of

copyright
all rights reserved
Quintessenz

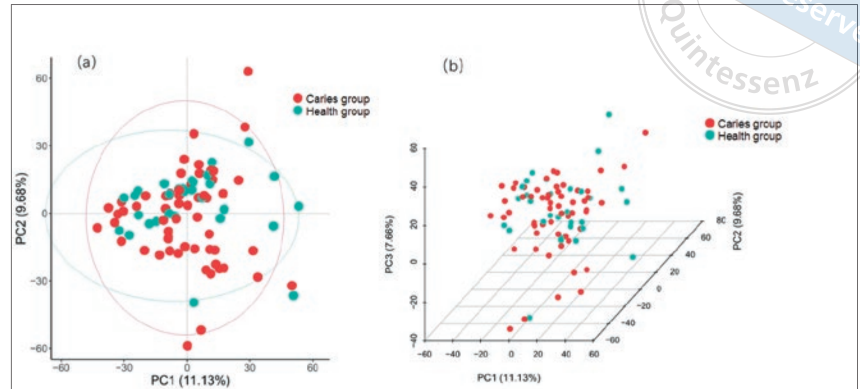


Fig 5 PCA scatter plot for the caries group and healthy group. **(a)** PCA two-dimensional scatter plot. Each point represents a sample; the red points represent caries group samples and the green points represent healthy group samples. PC1 is the first principal component, indicating that it contributes 11.13% to the sample difference, and PC2 is the second principal component, indicating that it contributes 9.66% to the sample difference. **(b)** PCA 3D scatter plot with PC3 as the third principal component, indicating that the contribution to sample difference is 7.66%.

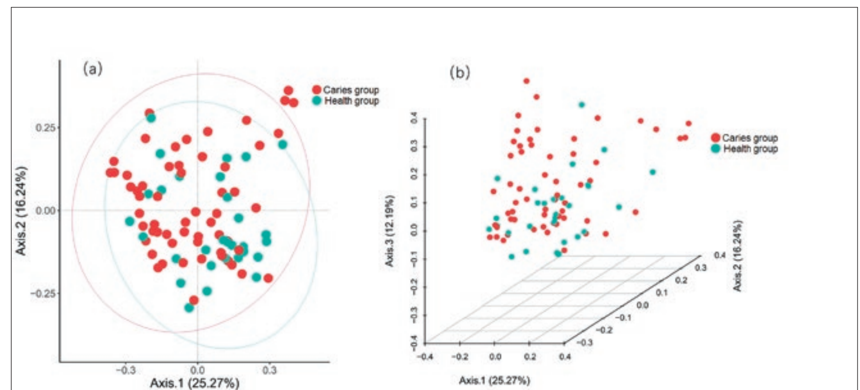


Fig 6 PCoA scatter plot for the caries group and healthy group. **(a)** PCoA two-dimensional scatter plot shows that each point represented a sample, with the red points representing the samples of the caries group and the green points representing the samples of the healthy group. Axis.1 was the first principal component, representing 25.27% and Axis.2 was the second principal component, representing 16.24% of the sample difference. **(b)** PCoA 3D scatter plot with Axis.3 as the third principal component, representing a contribution to the sample difference of 12.19%.

one class, two orders, five families and nine genera in the horizontal classification. In addition, the abundance of *Capnocytophaga* such as *Capnocytophaga leadbetteri* and *Capnocytophaga sputigena* at the species level was high.

Correlation network of bacteria

The results of the correlation analysis of the interaction between different microorganisms in the two groups are shown in Fig 9. Bacteria in the caries group and the healthy group showed different species correlations. The correlation between bacteria in the healthy group was more complex and aggregated. The larger the node, the

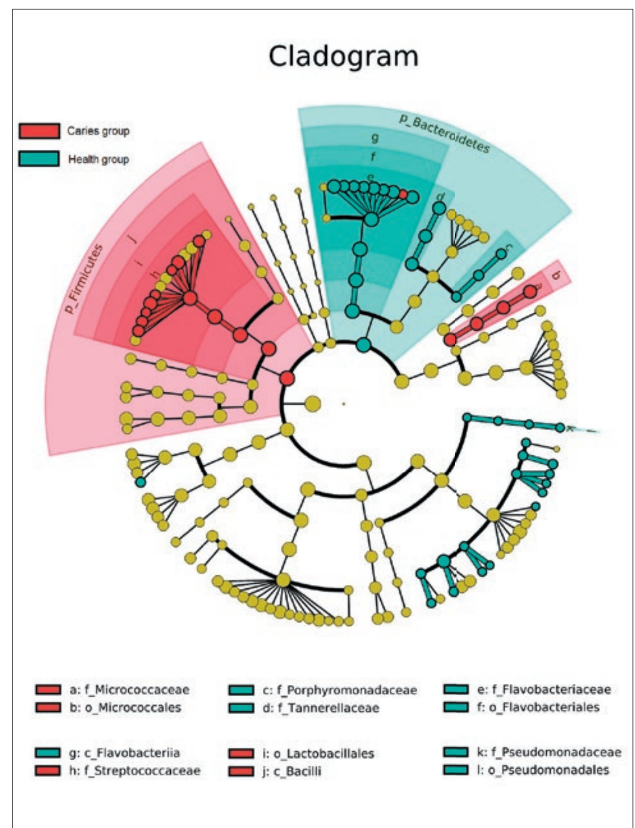


Fig 7 Cladogram for the caries group and healthy group. The coloured nodes from the inner ring to the outer ring represent the level of classification from gate to species. Different colours indicate that there were significantly different taxa between the two groups. Red indicates significantly increased bacteria in the caries group, and green indicates significantly increased bacteria in the healthy group.

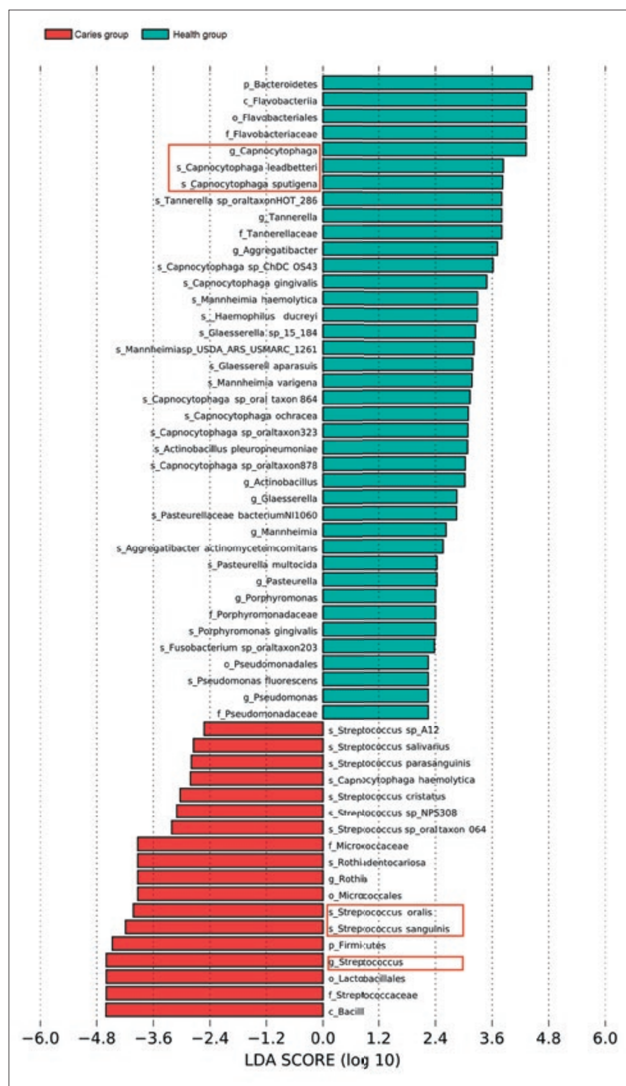


Fig 8 The column diagram of LDA was analysed by LEfSE in the caries group and healthy group. The bacteria in both groups were ranked according to the LDA score of different species at each classification level.

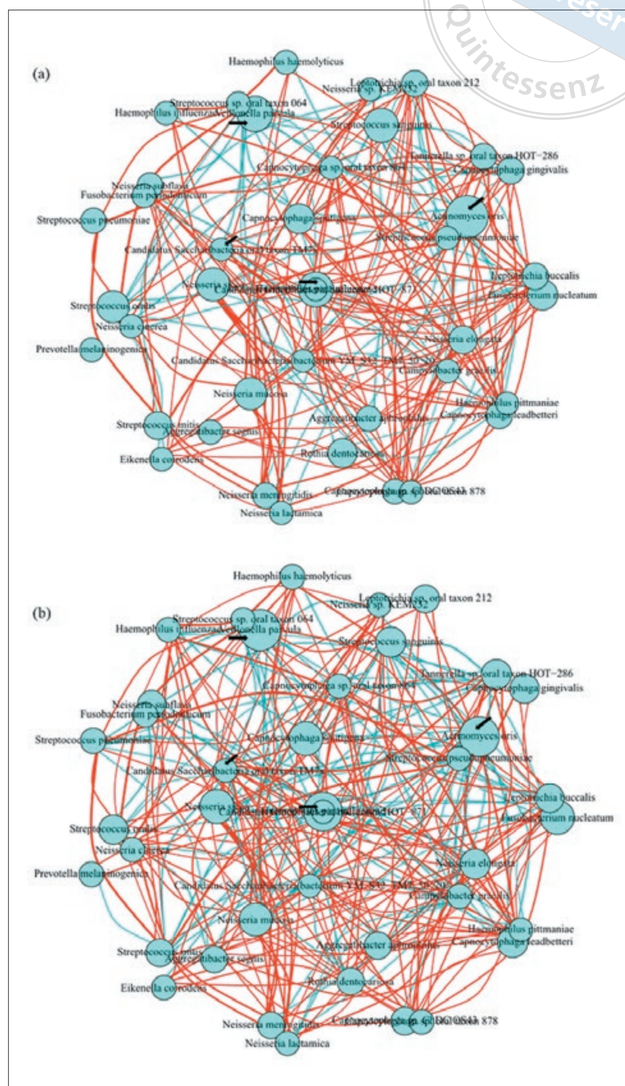


Fig 9 Flora correlation network in the **(a)** caries group and **(b)** healthy group. Nodes represent a type of bacteria. The larger the node, the more bacteria it is related to. Each link indicates a significant correlation between the two bacteria (correlation coefficient $|R| > 0.4$, $P < 0.01$, Spearman correlation analysis). Red indicates a positive correlation and blue indicates a negative correlation.

more connections it has with other nodes, indicating that the corresponding bacteria have more correlation with other bacteria. The results of this study showed that two bacteria, *Actinomyces oris* and *Veillonella parvula*, showed the most complex association with other bacteria in the caries group and the healthy group, and these were also the two bacteria with the highest relative abundance. There was a strong positive correlation between the bacteria of TM7 and GN02, and a complex correlation with a variety of other bacteria of the *Streptococcus*, *Capnocytophaga*, *Fusobacterium* and *Leptotrichia*.

Discussion

CPR is an important component of the oral flora

In this study, we used metagenomics to analyse the microbial composition and structure of caries patients and healthy patients, which was more accurate than previous studies using amplicon sequencing. All the samples detected the presence of three types of microorganisms, namely bacteria, archaea and viruses, including 38 phyla of bacteria, mainly *Proteobacteria*, *Firmicutes*,

copyright all rights reserved Quintessence

Bacteroidetes, *Actinobacteria* and *Fusobacteria*, which were similar to previous studies[16,30,31] and showed that the oral microflora of the human body was relatively stable.

The results of this study showed that the abundance of TM7 and GN02 was second only to that of the major microflora, indicating that CPR were an important component of the oral microflora. The results also indicated that the CPR at species level were mainly *TM7x* of TM7 and *Candidatus Gracilibacteria bacterium HOT-871* of GN02, both of which had high taxonomic abundance at the species level of the oral microflora, further proving the importance of CPR in the oral microbial community. *TM7x* and *Candidatus Gracilibacteria bacterium HOT-871* were detected in all samples, with a 100% detection rate indicating that CPR is common in the oral cavity and constitutes an important component of the oral microbiota. Studies have shown that CPR is a highly diverse bacterium that exists in all walks of life in nature and can exist in a variety of natural environments around the world, including fresh water, seawater, hot springs, swamps[32] and soil[33]. CPR is also detected in many parts of the human body, including the skin[34], the distal oesophagus[35] and the intestine[36], and is particularly common in the oral cavity[37]. CPR may have the ability to maintain growth in states of both health and critical illness[9].

The clustering analysis of samples in this study showed that the samples from the caries group and the healthy group showed a good clustering relationship, respectively. A similar bacterial cluster was observed between some samples from the caries group and the healthy group, indicating that the disease state might not significantly affect the bacterial species composition. *Neisseria*, *Streptococcus* and *Capnocytophaga* showed a good clustering relationship, respectively, indicating the similarity of survival conditions between the bacteria of the same genus. There was a difference in the relative abundance of species between the samples from the caries group and the healthy group, and the abundance of CPR in some samples from the caries group was significantly higher than that in other samples from the healthy group, indicating that CPR might have a potential pathogenic relationship with caries. Affected by the disease state, living habits, diet and drinking water conditions, the relative abundance of oral CPR in individuals will differ. This result of species classification with similar composition and different abundance supports the 'ecological plaque hypothesis', and caries is caused by the destruction of microbial homeostasis in the body, rather than the activities of specific microorganisms[38].

CPR is the 'core microbiome' of the oral cavity

The present results showed that around 50% bacteria in the caries group were common to the healthy group at the species level. Compared with the healthy group, there were more specific bacteria in the caries group, while only one specific bacterium was found in the healthy group. In general, a 'core microbiome' is defined as a group of members shared by a microbiota present in all or most of the different tissue sites in a human[39]. Analysis of species common to all samples using a Venn diagram is a basic method for finding the 'core microbiome'[40,41]. The oral 'core microbiome' was initially discovered by analysing multiple oral sites in three adults using 454 pyrosequencing technology[42]. Through further studies, scholars have confirmed the existence of the core microbiome and studied and expanded their overall composition[43,44]. *Firmicutes* is considered the 'core microbiome' in the oral cavity. The difference analysis of LEfSe in this study showed that bacteria of *Firmicutes* and *Streptococcus* in the caries group were significantly enriched, which was in line with the results of a previous study[45]. *Streptococcus* uses its sugar metabolism acid-producing and acid-resistant ability to change its microbial ecological environment and affect the colony structure, which may further lead to the occurrence and development of dental caries[45]. In addition, the results suggest that bacteria of the *Rothia* and *Micrococcaceae* may also be associated with dental caries.

The results of this study showed that TM7 and GN02 were common species in the caries group and the healthy group, and were detected in all samples, indicating that TM7 and GN02 belong to the 'core microbiome' of the oral cavity and might play an important role in the stability and function of the oral microecological environment, which was consistent with the results of previous studies[20,46,47]. Whether they are the core microbiota of caries remains to be further studied. Based on the level of the existing database annotations, the LEfSE difference analysis shows that there is no significant difference in the bacteria of TM7 and GN02 in the caries group and the healthy group at the overall level; however, it was found that the abundance of CPR was significantly increased in some samples. In this study, after in-depth analysis of the CPR obtained by genomic assembly and box division, it was revealed that there was a difference in the level of unknown strains of CPR between the caries group and the healthy group.

copyright by all rights reserved Quintessence

There is a complex correlation between CPR and a variety of oral microorganisms

Correlation network analysis reveals the potential correlation between the oral microflora and indicates the possible synergistic and antagonistic interactions between different microorganisms. Unique bacterial interactions are exhibited in oral colonies from different disease states. Analysis of the correlation between the microbiota in the caries group and the healthy group clearly revealed the complex relationship between the oral microbiota. The bacteria in the healthy group showed a richer and more complex correlation, suggesting that some microbial relationships in the caries group might be destroyed and lead to an ecological imbalance in the process of dental caries. This is consistent with the view expressed in a previous study that acidic conditions are produced due to an increase in fermentable carbohydrates, resulting in a change in the nutritional status of the microbial colony and disruption of microbial interactions that maintain the balance of microbial communities in a healthy state[48].

The results of this study showed a strong positive correlation between TM7 and GN02 in the caries group and the healthy group, indicating that there was a close survival relationship between the CPR. TM7 and GN02 in the two groups had a positive correlation with a variety of *Capnocytophaga*. *Capnocytophaga* is the 'core microbiome' of the oral cavity and exists widely in oral colonies. Previous studies have shown that *Capnocytophaga* are involved in the formation of dental plaque biofilms[49], which are significantly enriched in patients with gingivitis[50], dental caries[51] and oral cancer[52]. The strong positive correlation between CPR and *Capnocytophaga* suggested that *Capnocytophaga* might be a potential host bacterium of CPR, and the ability of CPR to assist host bacteria in avoiding the human immune system and promoting the formation of host biofilm might increase the role of their host bacteria in the occurrence and development of diseases[53]; however, whether the CPR truly have a parasitic relationship with the *Capnocytophaga* needs to be verified by isolation and culture in the laboratory. In addition, there was a complex correlation between *Actinomyces spp.*, *Veillonella parvula* and other bacteria. Meanwhile, *Actinomyces spp.* and *Veillonella parvula* were the most abundant bacteria in the two groups, suggesting that they might be important for oral bacterial interaction and homeostasis.

In this study, the diversity of colony structure showed that the abundance and diversity of bacterial communities in the caries group were similar to those in the healthy group, which was consistent with the results of some previous studies[31,54]. Other studies have shown that there are significant differences in the diversity of oral microorganisms between healthy controls and caries patients, indicating that there may be differences between oral flora ecosystems in different disease states[16,55,56]. The results of the colony structure may be influenced by sample size, sequencing method, individual differences and a number of other factors. Metagenomics methods can achieve greater sequencing depth than 16S amplicon sequencing methods, while increasing the number of species annotated and relatively reducing the difference in colony diversity between subgroups. In view of the lack of research on the relationship between metagenomics and caries microorganisms, more studies are needed to further clarify the structural characteristics of the colonies.

Conclusion

In this study, we analysed the difference in the composition, distribution and abundance of CPR in the oral environment of caries patients and healthy people by metagenomic bioinformatics, based on the problem of the correlation between CPR and the potential pathogenesis of dental caries. CPR is a common species in caries and healthy oral cavities with a high detection rate and relative abundance, indicating that it is an important component of the oral microflora and the 'core microflora' of the oral cavity, which may play an important role in the stability and function of the oral microecological environment. There was a correlation between CPR and a variety of oral microorganisms, and the positive correlation with *Capnocytophaga* was strongest, suggesting that *Capnocytophaga* might be the potential host bacteria of CPR. TM7 and GN02 were significantly increased in some caries samples, but the correlation between CPR bacteria and caries and its possible pathogenic mechanism still need to be studied in greater depth. Further development of culture mediums suitable for CPR and acquisition of more clinical isolates to explore the mechanism of CPR participating in the pathogenesis of dental caries will be one of the research directions of oral CPR in the future.

Conflicts of interest

The authors declare no conflicts of interest related to this study.

copyright all rights reserved Quintessenz

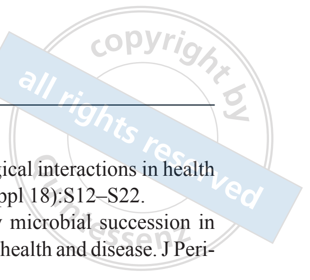
Author contribution

Dr Song JIANG developed the experimental strategy, analysed the data and drafted the manuscript; Dr Yu Xing CHEN assisted in analysing the data and writing the manuscript; Drs Jie NIE, Xiao Yan WANG and Feng CHEN refined the strategy and mentored the work. All authors approved the final manuscript.

(Received May 05, 2021; accepted Jun 15, 2021)

References

1. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: A systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388:1545–1602.
2. Marsh PD. In sickness and in health-What does the oral microbiome mean to us? An ecological perspective. Adv Dent Res 2018;29:60–65.
3. Marsh PD. Are dental diseases examples of ecological catastrophes? Microbiology (Reading) 2003;149(Pt 2):279–294.
4. Fejerskov O. Changing paradigms in concepts on dental caries: consequences for oral health care. Caries Res 2004;38:182–191.
5. Castelle CJ, Banfield JF. Major new microbial groups expand diversity and alter our understanding of the tree of life. Cell 2018;172:1181–1197.
6. Segata N, Haake SK, Mannon P, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol 2012;13:R42.
7. Zhou Y, Gao H, Mihindukulasuriya KA, et al. Biogeography of the ecosystems of the healthy human body. Genome Biol 2013;14:R1.
8. Paster BJ, Russell MK, Alpagot T, et al. Bacterial diversity in necrotizing ulcerative periodontitis in HIV-positive subjects. Ann Periodontol 2002;7:8–16.
9. Brinig MM, Lepp PW, Ouverney CC, Armitage GC, Relman DA. Prevalence of bacteria of division TM7 in human subgingival plaque and their association with disease. Appl Environ Microbiol 2003;69:1687–1694.
10. Rylev M, Bek-Thomsen M, Reinholdt J, Ennibi OK, Kilian M. Microbiological and immunological characteristics of young Moroccan patients with aggressive periodontitis with and without detectable Aggregatibacter actinomycetemcomitans JP2 infection. Mol Oral Microbiol 2011;26:35–51.
11. Liu B, Faller LL, Klitgord N, et al. Deep sequencing of the oral microbiome reveals signatures of periodontal disease. PLoS One 2012;7:e37919.
12. Kistler JO, Booth V, Bradshaw DJ, Wade WG. Bacterial community development in experimental gingivitis. PLoS One 2013;8:e71227.
13. Camelo-Castillo AJ, Mira A, Pico A, et al. Subgingival microbiota in health compared to periodontitis and the influence of smoking. Front Microbiol 2015;6:119.
14. Sousa V, Nibali L, Spratt D, et al. Peri-implant and periodontal microbiome diversity in aggressive periodontitis patients: A pilot study. Clin Oral Implants Res 2017;28:558–570.
15. Nowicki EM, Shroff R, Singleton JA, et al. Microbiota and metatranscriptome changes accompanying the onset of gingivitis. mBio 2018;9:e00575-18.
16. Xiao C, Ran S, Huang Z, Liang J. Bacterial diversity and community structure of supragingival plaques in adults with dental health or caries revealed by 16S pyrosequencing. Front Microbiol 2016;7:1145.
17. Eriksson L, Lif Holgerson P, Esberg A, Johansson I. Microbial complexes and caries in 17-year-olds with and without Streptococcus mutans. J Dent Res 2018;97:275–282.
18. Rupf S, Laczny CC, Galata V, et al. Comparison of initial oral microbiomes of young adults with and without cavitated dentin caries lesions using an in situ biofilm model. Sci Rep 2018;8:14010.
19. He X, McLean JS, Edlund A, et al. Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. Proc Natl Acad Sci U S A 2015;112:244–249.
20. Espinoza JL, Harkins DM, Torralba M, et al. Supragingival plaque microbiome ecology and functional potential in the context of health and disease. mBio 2018;9:e01631-18.
21. Al-Hebshi NN, Baraniya D, Chen T, et al. Metagenome sequencing-based strain-level and functional characterization of supragingival microbiome associated with dental caries in children. J Oral Microbiol 2018;11:1557986.
22. Wang Y, Wang S, Wu C, et al. Oral microbiome alterations associated with early childhood caries highlight the importance of carbohydrate metabolic activities. mSystems 2019;4:e00450-19.
23. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–2120.
24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.
25. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;32:3047–3048.
26. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol 2019;20:257.
27. Lu J, Salzberg SL. Ultrafast and accurate 16S rRNA microbial community analysis using Kraken 2. Microbiome 2020;8:124.
28. Buttigieg PL, Ramette A. A guide to statistical analysis in microbial ecology: A community-focused, living review of multivariate data analyses. FEMS Microbiol Ecol 2014;90:543–550.
29. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011;12:R60.
30. Jiang W, Zhang J, Chen H. Pyrosequencing analysis of oral microbiota in children with severe early childhood dental caries. Curr Microbiol 2013;67:537–542.
31. Chen L, Qin B, Du M, et al. Extensive description and comparison of human supra-gingival microbiome in root caries and health. PLoS One 2015;10:e0117064.
32. Rheims H, Spröer C, Rainey FA, Stackebrandt E. Molecular biological evidence for the occurrence of uncultured members of the actinomycete line of descent in different environments and geographical locations. Microbiology (Reading) 1996;142(Pt 10): 2863–2870.
33. Hugenholtz P, Tyson GW, Webb RI, Wagner AM, Blackall LL. Investigation of candidate division TM7, a recently recognized major lineage of the domain Bacteria with no known pure-culture representatives. Appl Environ Microbiol 2001;67:411–419.
34. Dinis JM, Barton DE, Ghadiri J, et al. In search of an uncultured human-associated TM7 bacterium in the environment. PLoS One 2011;6:e21280.
35. Pei Z, Bini EJ, Yang L, Zhou M, Francois F, Blaser MJ. Bacterial biota in the human distal esophagus. Proc Natl Acad Sci U S A 2004;101:4250–4255.
36. Kuehbacher T, Rehman A, Lepage P, et al. Intestinal TM7 bacterial phylogenies in active inflammatory bowel disease. J Med Microbiol 2008;57(Pt 12):1569–1576.
37. Weyrich LS, Duchene S, Soubrier J, et al. Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. Nature 2017;544:357–361.
38. Kidd E, Fejerskov O. Changing concepts in cariology: Forty years on. Dent Update 2013;40:277–278, 280–282, 285–286.



39. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007;449: 804–810.
40. Lazarevic V, Whiteson K, Hernandez D, François P, Schrenzel J. Study of inter- and intra-individual variations in the salivary microbiota. BMC Genomics 2010;11:523.
41. Shade A, Handelsman J. Beyond the Venn diagram: The hunt for a core microbiome. Environ Microbiol 2012;14:4–12.
42. Zaura E, Keijser BJ, Huse SM, Crielaard W. Defining the healthy "core microbiome" of oral microbial communities. BMC Microbiol 2009;9:259.
43. Ling Z, Liu X, Luo Y, et al. Pyrosequencing analysis of the human microbiota of healthy Chinese undergraduates. BMC Genomics 2013;14:390.
44. Hu YJ, Shao ZY, Wang Q, et al. Exploring the dynamic core microbiome of plaque microbiota during head-and-neck radiotherapy using pyrosequencing. PLoS One 2013;8:e56343.
45. Costa Oliveira BE, Ricomini Filho AP, Burne RA, Zeng L. The route of sucrose utilization by Streptococcus mutans affects intracellular polysaccharide metabolism. Front Microbiol 2021;12:636684.
46. Bor B, Bedree JK, Shi W, McLean JS, He X. Saccharibacteria (TM7) in the human oral microbiome. J Dent Res 2019;98:500–509.
47. Baker JL, Morton JT, Dinis M, et al. Deep metagenomics examines the oral microbiome during dental caries, revealing novel taxa and co-occurrences with host molecules. Genome Res 2021;31:64–74.
48. Marsh PD, Zaura E. Dental biofilm: Ecological interactions in health and disease. J Clin Periodontol 2017;44(suppl 18):S12–S22.
49. Teles FR, Teles RP, Uzel NG, et al. Early microbial succession in redeveloping dental biofilms in periodontal health and disease. J Periodontal Res 2012;47:95–104.
50. Acharya A, Chen T, Chan Y, Watt RM, Jin L, Mattheos N. Species-level salivary microbial indicators of well-resolved periodontitis: A preliminary investigation. Front Cell Infect Microbiol 2019;9:347.
51. Johansson I, Witkowska E, Kaveh B, Lif Holgerson P, Tanner ACR. The microbiome in populations with a low and high prevalence of caries. J Dent Res 2016;95:80–86.
52. Perera M, Al-Hebshi NN, Perera I, et al. Inflammatory bacteriome and oral squamous cell carcinoma. J Dent Res 2018;97:725–732.
53. Bedree JK, Bor B, Cen L, et al. Quorum sensing modulates the epibiotic-parasitic relationship between Actinomyces odontolyticus and its Saccharibacteria epibiont, a Nanosynbacter lyticus strain, TM7x. Front Microbiol 2018;9:2049.
54. Jiang S, Gao X, Jin L, Lo ECM. Salivary microbiome diversity in caries-free and caries-affected children. Int J Mol Sci 2016;17:1978.
55. Li Y, Ku CYS, Xu J, Saxena D, Caufield PW. Survey of oral microbial diversity using PCR-based denaturing gradient gel electrophoresis. J Dent Res 2005;84:559–564.
56. Preza D, Olsen I, Aas JA, Willumsen T, Grinde B, Paster BJ. Bacterial profiles of root caries in elderly patients. J Clin Microbiol 2008;46:2015–2021.