The Salivary Microbiome: Analysis of by Pyrosequencing and the Relationship with *Helicobacter pylori* Infection

Yingjie Ji, [a] Xiao Liang [a] and Hong Lu [a] #

[a] Division of Gastroenterology and Hepatology, Shanghai Institute of Digestive Disease; Key Laboratory of Gastroenterology and Hepatology, Ministry of Health; Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

Running Head: Relationship between salivary microbiome and *Helicobacter pylori* Infection

Address correspondence to Hong Lu, hlu@sjtu.edu.cn

Y.J.J and X.L contributed equally to this work.

Words count of abstract: 249

Words count of text: 3806

Abstracts

Backgrounds: There have been reports of *Helicobacter pylori* (*H. pylori*) in the oral cavity and it has been suggested that the oral cavity may be a reservoir for *H. pylori* reflux from the stomach.

Objectives: High-throughput pyrosequencing was used to assess the structure and composition of oral microbiota communities in individuals with or without confirmed *H. pylori* infection.

Methods: Saliva samples were obtained from 34 *H. pylori* infected and 24 *H. pylori* uninfected subjects. Bacterial genomic DNA was extracted and examined by pyrosequencing by amplification of the 16S rDNA V3-V4 hypervariable regions followed by bioinformatics analysis. Saliva sampling was repeated from 22 of the 34 *H. pylori* infected subjects 2 months after *H. pylori* eradication.

Results: High-quality sequences (2,812,659) clustered into 95,812 operational taxonomic units (OTUs; 97% identity), representing 440 independent species belonging to 138 genera, 68 families, 36 orders, 21 classes, and 11 phyla. Species richness (alpha diversity) of *H. pylori* infected subjects was similar to that of uninfected subjects. Eradication treatment decreased saliva bacterial diversity. Beta diversity analysis showed that the salivary microbial community structure differed between *H. pylori* infected and uninfected subjects both before and after *H. pylori* eradication.

Conclusions: Salivary microbiota diversity was similar in *H. pylori* infected and uninfected individuals. Antibiotic therapy was associated with a decline in salivary bacterial diversity. Both *H. pylori* infection and its eradication caused the oral microbiota alterations in community and structure. The present of *H. pylori* in oral cavity was not related with its infection status in stomach.

Trial registration: ClinicalTrials.gov, Identifier: NCT03730766

Keywords: *Helicobacter pylori*, salivary microbiota, pyrosequencing, 16S rDNA

Importance

The oral cavity plays a vital role in *Helicobacter pylori* transmission among human. High-throughput pyrosequencing of the 16S rDNA V3-V4 hypervariable regions was used to assess the structure and composition of oral microbiota communities in individuals with or without confirmed *Helicobacter pylori* infection. We show that both *Helicobacter pylori* infection and eradication cause microbiota alterations in the oral microbiota. Prior studies report detection of *Helicobacter pylori* in the oral cavity by polymerase chain reaction. We show that the presence of *Helicobacter pylori* in the oral cavity is unrelated with its infection status in the stomach.

Introduction

Helicobacter pylori (*H. pylori*) is a Gram-negative bacterium that colonizes the human gastric epithelium. It belonged to *Helicobacter* genus, *Helicobacteraceae* family, *Campylobacterales* order, *Epsilonproteobacteria* class, and *Proteobacteria* phyla. *H. pylori* infection is characterized by mucosal inflammation (gastritis) and may result in peptic ulcer disease or gastric adenocarcinoma (1). *H. pylori* is transmitted between humans by a variety of routes including gastro-oral and fecal-oral mechanisms that include contaminated water and food. It has also been postulated that oral cavity may play a role in *H. pylori* transmission and possibly act as a reservoir (2). For example, *H. pylori* has been detected in the oral cavity using the polymerase chain reaction (PCR). *H. pylori* has also been successfully cultured from saliva from individuals with positive results of both a saliva *H. pylori* antigen test and *H. pylori* flagella test (3).

The oral cavity is one of the most complex and largest microbial habitats that harbors hundreds of different bacteria which play important roles in maintaining oral homeostasis and influencing the development of both oral and systematic diseases (4). Many factors in the oral environment including intraoral pH and salivary iron concentration have been reported to have significant relationships with oral microbial communities (5). However, there was few reports about *H. pylori* and its relationship with the microbial community structure in human saliva.

Currently, most oral bacteria species cannot be cultivated *in vitro* using traditional cultivation methods requiring the use of molecular biological techniques, such as checkerboard hybridization, microarray chips, and the quantitative real-time PCR (6) to identify and classify the currently uncultivable bacteria. However, many low-abundance bacteria species still cannot be detected using

69 these approaches which impedes the comprehensive and in-depth understanding of oral bacteria
70 diversity. In this study, we used amplicon pyrosequencing of 16S rDNA V3-V4 hypervariable regions
71 to define the bacterial composition, abundance, and structure of salivary microbiome in people with
72 and without active *H. pylori* infections. In addition, we also characterized the salivary biodiversity of a
73 subgroup of subjects before and after the *H. pylori* eradication.
74

Methods

Subjects and Sample Collections

This study was performed in accordance with the recommendations of the Ethics Committee of the Renji Hospital of Shanghai Jiao Tong University. All subjects gave written informed consent in accordance with Declaration of Helsinki.

The samples were collected in Renji Hospital of Shanghai Jiao Tong University, China, from August to November in 2018. A total of 58 subjects were recruited, including 34 subjects with *H. pylori* infection and 24 uninfected subjects. We first conducted a cross-sectional study of the salivary microbiota of 34 *H. pylori* infected and 24 uninfected subjects. A prospective study was then performed in a subgroup of 22 subjects with *H. pylori* infection who underwent salivary analysis both before and after successful *H. pylori* eradication.

All subjects received both endoscopy and ^{13}C urea breath test (^{13}C-UBT) before enrollment. The *H. pylori* infection status was confirmed by positive rapid urease test (RUT), histology and ^{13}C-UBT. Absence of infection was defined as negative results for all tests (i.e., RUT, histology and ^{13}C-UBT). *H. pylori* infected subjects received eradiation therapy consisting of esomeprazole 20 mg b.i.d., bismuth potassium citrate 600 mg b.i.d, amoxicillin 1000 mg t.i.d., and metronidazole 400 mg t.i.d. for 14 days. *H. pylori* eradication was confirmed by ^{13}C-UBT at least 6 weeks after the end of treatment. Saliva samples were collected from 22 subjects both before and 2 months following successful *H. pylori* eradication. Subjects were characterized into four groups. *H. pylori* uninfected group (uninfected) (n=24), *H. pylori* infected group (infected) (n=34), and Pre-eradicated *H. pylori* infected group (pre-eradicated) (n=22) and a successful eradicated group (eradicated) (n=22). Inclusion criteria of

subjects were: age of 20-65 years old male or female, with good oral hygiene (including brushing teeth twice a day) and with no bad eating habits (7). Exclusion criteria included: 1) the presence of dental carious or any untreated cavitated carious lesions and oral abscesses, 2) periodontal disease or periodontal pockets ≥4 mm, 3) the use of antibiotics or PPI within 2 months before the study, 4) previous diagnosis of a serious systemic diseases (such as diabetes, hypertension or cardiopathy) or any diseases affecting oral health (such as Sjogren's syndrome or any disease characterized by xerostomia), 5) pregnancy of breastfeeding, and 6) smoking or alcohol drinking. The detailed clinical parameters of the 58 subjects are shown in Table S1.

Salivary sampling

Sampling was performed in the morning 2 hours after eating. Saliva samples were collected from each subject according to the Manual of Procedures for Human Microbiome Project (http://hmpdacc.org/resources/tools_protocols.php), with minor modifications. Approximately 3-4 ml of non-stimulated saliva was collected in two sterile, labeled 2 mL Eppendorf tubes, which were immediately placed on ice. Within 3 hours of collection, samples were transported on ice and stored at -80°C until use (8).

DNA Extraction and Pyrosequencing

DNA was extracted from the saliva samples using the E.Z.N.A. ® Soil DNA Kit (OMEGA, USA), following the manufacturer's instructions, and stored at −20°C prior to further analysis. PCR amplification of the bacterial 16S rDNA hypervariable V3-V4 region was performed using the forward

117 primer 338F (5'-ACTCCTACGGGAGGCAGCA-3') and the reverse primer 806R
118 (5'-GGACTACHVGGGTWTCTAAT-3'). Sample-specific 7-bp barcodes were incorporated into the
119 primers for multiplex sequencing. Details of the barcodes are shown in Table S2. PCR amplification
120 were performed on an ABI 2720 instrument (ABI, USA) with an initial denaturation at 98°C for 2
121 minutes, followed by 25 cycles of denaturation (15s at 98°C), annealing (30s at 55°C), extension (30s at
122 72◦C), and ended with a final extension (5 min at 72◦C). PCR amplicons were purified with Agencourt
123 AMPure Beads (Beckman Coulter, Indianapolis, IN) and quantified using the Quant-iT PicoGreen
124 dsDNA Assay Kit (Invitrogen, Carlsbad, CA, USA). Equimolar concentrations of purified amplicons
125 were pooled in equal amounts. Subsequently, the paired-end 2×300 bp pyrosequencing was performed
126 on the Illlumina MiSeq platform with MiSeq Reagent Kit v3 (Illlumina, USA), following the vendor's
127 standard protocols.
128
129 **Sequence Analysis**
130 The Quantitative Insights Into Microbial Ecology (QIIME) pipeline was employed to process the
131 sequencing data. Raw sequences were filtered to obtain high-quality sequences according to QIIME (9).
132 The high-quality sequences were clustered into operational taxonomic units (OTUs) at 97% sequence
133 identity by UCLUST (10). The representative sequences selected from each OTU were classified
134 taxonomically by BLAST searching against the Human Oral Microbiome Database (HOMD), which
135 provides a detailed record of the type, metabolism, and pathogenicity of oral bacteria (11). Then, an
136 OTU table was further generated to record the OTU abundance of each sample and the taxonomic
137 classification of these OTUs. Finally, to minimize the difference of sequencing depth across samples,

the OTU table was modified by removing OTUs containing less than 0.001% of total sequences across all samples for further analysis (12).

Bioinformatics and Statistical Analysis

Sequence data analyses were mainly performed using QIIME (version 1.8.0) and R packages (version 3.2.4). The alpha diversity analysis including Chao 1 richness estimator, Abundance-based Coverage Estimator (ACE) metric, Shannon diversity index, and Simpson index, were calculated at 97% identity using the QIIME (13). Ranked abundance curves were generated to compare both the richness and evenness of OTUs among samples. The beta diversity analysis including Nonmetric Multidimensional Scaling (NMDS), and unweighted UniFrac distances based principal coordinate analysis (PCoA), were performed using the R package to evaluate the similarity among various bacterial communities (14). The significance of differentiation of microbiota structure among groups was assessed by Adonis test (15). The taxonomy compositions and abundances were visualized by MEGAN (version 6.6.7) software (16). Linear discriminant analysis effect size (LEfSe) was used to compare the bacterial community structures between the samples from the patients with and without *H. pylori* infection, as well as before and after the eradication regimen, using the online Galaxy workflow framework (http://huttenhower.sph.harvard.edu/galaxy/) (17). Co-occurrence analysis among genera was performed among 50 most abundant genera. Correlations with $|RHO| > 0.6$ and $P < 0.01$ were visualized as co-occurrence network using Cytoscape (version 3.6.1) (18). Microbial functions were predicted by Phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt), based on high-quality sequences, and aligned to the Kyoto Encyclopedia of Genes and

159 Genomes (KEGG) database (19).

160

161 **Data Access**

162 All raw sequences were deposited in the NCBI Sequence Read Archive under accession number

163 SRP167714.

164

Results

Global Sequencing Data

A total of 2,812,659 high-quality sequences (representing 79% of the total sequences) were acquired from the 80 saliva samples, with an average of 35,158 sequences per sample (ranging from 19,210 to 44,310). The average sequence length was 445 bp, with the maximum length being 548 bp and the shortest length being 136 bp (Figure S1). Clustering of all high-quality sequences at 97% identity resulted in 70,489 OTUs, which were BLAST-searched against the HOMD database for taxonomic assignments. After removing the low-credibility OTUs (together contributing only 6.7% of all sequences), a modified OTU table was obtained consisting of 95,812 OTUs with an average of 1,198 OTUs per sample (ranging from 697 to 1,584).

Bacterial Abundance and Distribution

The bacterial distribution was characterized in terms of the relative taxonomic abundances. A total of 11 phyla, 21 classes, 36 orders, 68 families, 138 genera and 440 species were detected in the saliva samples. The taxonomic distributions of the predominant bacteria (relative abundance >1% of the total sequences) in subjects with and without *H. pylori* at different levels were shown in Figure 1. The 6 most abundant phyla were *Proteobacteria* (40.1% of the total sequences), *Firmicutes* (31.6%), *Bacteroidetes* (13.0%), *Actinobacteria* (7.4%), *Fusobacteria* (6.1%), and *TM7* (1.0%), together accounting for 99.2% of the total sequences. At genus level, saliva microbiota was dominated by *Neisseria*, *Streptococcus*, *Haemophilus*, *Veillonella*, and *Prevotella*, with average relative abundances of 20.2, 16.5, 10.5, 8.0, and 8.0%, respectively. The compositions in taxa of the microbial communities

186 according to the tested sample groupings are provide in Figure S2.

187

188 **Bacterial Diversity Analysis**

189 The saliva microbiota richness, measured by numbers of observed OTUs, was similar in
190 uninfected subjects and infected subjects (Figure S3A). The alpha diversity indices of Chao1, ACE,
191 Shannon, Inverse Simpson, Good's coverage, and Simpson's evenness are shown in Table 1. The
192 Shannon diversity index was higher in uninfected subjects than in infected subjects, but there was no
193 significant difference between groups by *t*-test (1417.58 vs. 1393.60, $p>0.05$). Besides, the ACE
194 richness index (1491.22 vs. 1465.97, $p>0.05$), Chao 1 richness estimator (1417.58 vs. 1393.60, $p>0.05$),
195 and the Inverse Simpson diversity index (1.02 vs. 1.02, $p>0.05$) was also higher in uninfected subjects,
196 with no significant difference, indicating the similar bacterial diversity of *H. pylori* uninfected saliva
197 compared to the infected subjects. Good's coverage estimator for each group was over 98%, indicating
198 that the current sequencing depth was sufficient to saturate the bacterial diversity of saliva. In addition,
199 Simpson's evenness index indicated that the bacterial-community distribution in two groups was
200 uneven, which was also observed in the rank-abundance curve (Figure S4).

201

202 **Bacterial Community Structures**

203 To gain insights into the similarities in the bacterial community structures among uninfected and
204 infected subjects, PCoA of beta diversity analysis was performed based on the unweighted UniFrac
205 distances, which demonstrated different community structures among two groups (PERMANOVAR,
206 pseudo-F: 1.49, $p=0.033$). As shown in Figure 2A, the overall microbial composition of infected

207 subjects deviated from that of uninfected subjects. Furthermore, the results of NMDS based on the
208 genus level classification exhibited clear segregations in community structures among groups (Figure
209 S5A).
210
211 **Differential Microbiota Compositions**
212 There were significant differences in the community compositions among two groups. As shown
213 in Figure 3, a cladogram representation of significantly different taxa among groups was performed by
214 LEfSe. The microbial composition was significantly different at the genus level, with 16 significantly
215 different genera among the two groups. *Acinetobacter*, *Ralstonia*, *Leptothrix*, *Sphingomonas*,
216 *Ochrobactrum*, *Afipia*, *Leptotrichia*, *Oribacterium*, and *Moryella* exhibited a relatively higher
217 abundance in infected subjects, and can be considered *H. pylori*-enriched genera. *Alloprevotella*,
218 *Aggregatibacter*, *Klebsiella*, *Leptotrichlaceae__G_1_*, *Fusobacterium*, *Parvimonas*, and *Peptococcus*
219 were relatively more abundant in uninfected subjects, which could be considered to be decreased in the
220 infected group. These higher or lower expressed genera in infected subjects can be considered as *H.*
221 *pylori*-associated genera (LAD >2, $p < 0.05$).
222
223 **Eradication therapy for *H. pylori* partially changed salivary microbiota**
224 To determine the effects of *H. pylori* eradication therapy on saliva microbial composition, saliva
225 samples from a subgroup of *H. pylori* infected subjects (n=22), were collected before (pre-eradicated
226 group) and 2 months after treatment; saliva collected after successful eradication were classified into
227 the eradicated group. The within-individual diversity in the samples from eradicated group was lower

than pre-eradicated group (Figure S3B). The Shannon diversity index, ACE richness index, and Chao 1 richness estimator were higher in pre-eradicated subjects than in eradicated subjects, with a significant difference between groups by *t*-test (Shannon p=0.015, ACE p=0.003, Chao 1 p=0.002), indicating significant alteration in the within-individual diversity in samples from eradicated subjects compared to their baseline samples (Table 2).

The beta diversity using unweighted UniFrac showed significant differences in the overall microbial composition between pre-eradicated and eradicated groups (PERMANOVAR, pseudo-F: 3.34, p=0.001) (Figure 2B). The NMDS also exhibited clear segregations in community structures among groups (Figure S5B).

The relative difference of *H. pylori*-associated taxa was compared before and after eradication by LEFse analysis (Figure S6). Among the *H. pylori*-enriched genera, *Ralstonia*, *Leptotrichia*, *Sphingomonas*, *Leptothrix*, *Oribacterium*, and *Acinetobacter* increased after the eradication, while *Ochrobactrum* decreased after the successful eradication ($p<0.05$, paired Wilcoxon rank-sum test). Of the genera that decreased in infected subjects, *Alloprevotella*, *Aggregatibacter*, *Leptotrichlaceae__G_1_*, *Parvimonas*, and *Fusobacterium* decreased after the eradication ($p<0.05$, paired Wilcoxon rank-sum test). Besides, we found that at phyla level, *Fusobacteria* increased after *H. pylori* eradication.

***Helicobacter species* in the oral cavity**

H. pylori was detected in 38 out of 80 saliva samples, occupying 0.0139% of all the total sequences (Figure 4). 12 of the 34 subjects in infected subjects (35.3%), 11 of 24 subjects in the

uninfected group (45.8%), and 15 of 22 subjects in the eradicated group (68.2%) were found to possess *H. pylori* in the oral cavity, respectively, (p=0.054). The *H. pylori* signature was present in the saliva of subjects with negative ^{13}C-UBT, RUT, and in *H. pylori* infected individuals after successful *H. pylori* eradication. We've also compared the prevalence of *H. pylori* both before and after *H. pylori* eradication. Pre-eradiation 7 subjects has positive saliva and 15 were negative. In 6 of the 7 with *H. pylori* pretreatment it was no longer detected by PCR after eradication. Interestingly, 9 in 15 subjects who were *H. pylori* negative before eradication had it detected post *H. pylori* eradication whereas 6 remained free of *H. pylori* after eradication.

Co-occurrence Network Analysis and Function Predictions

Co-occurrence analysis was used to discern relationships among the saliva microbiota at the genus level. As shown in the network diagram for the 50 most abundant genera (Figure S7), 12 genera displayed positive associations, and 1 genera displayed a negative association. Among them, *Atopobium* and *Solobacterium* exhibited a high degree of linkages with other genera. High rho values were found for the *Atopobium-Megasphaera* (0.83), *Filifactor-Treponema* (0.75), and *Atopobium-Prevotella* (0.74) pairs. Besides, *Oribacterium*, *Acinetobacter,* and *Ralstonia* exhibited a positive association (0.68).

To predict the functions of saliva bacterial community, PICRUSt analysis was performed based on the 16S rDNA composition data of each sample (Figure S8). A total of 41 metabolic functions were predicted in all samples with the most enrichment in membrane transport (11.8%), replication and repair (9.7%), amino acid metabolism (9.4%), carbohydrate metabolism (9.2%), translation (6.6%), and

270 energy metabolism (5.8%).

271

272

Discussion

A comprehensive and thorough investigation of the bacterial diversity of saliva microbiota is essential for understanding the how or whether *H. pylori* infection alters the salivary saliva microbiota. The technology of high-throughput pyrosequencing has provided new cognizance of the structures and compositions of microbiota communities.

By comparing the alpha diversity indexes we found that the bacterial diversity in saliva was similar among the *H. pylori* uninfected and *H. pylori* infected people. Our study is consistent with the notion that *H. pylori* in the stomach has little or no effect on the bacterial diversity of the oral cavity (20). The Shannon diversity index, ACE richness index, and Chao 1 richness estimator all declined after eradication of *H. pylori* compared to the baseline samples ($p<0.05$), which was consistent with the prior studies that use of PPIs and antibiotics may affect the oral microbiome (21, 22).

According to the beta diversity analysis based on the unweighted UniFrac distances, the community structures of saliva microbiota were different in *H. pylori* uninfected and infected individuals, which is contrary to the results of Christian's study (20). Samples from the *H. pylori* infected subjects tended to cluster together, while the microbiota in the uninfected subjects appeared to be more variable suggesting that gastric *H. pylori* infection may affect oral bacterial components. Clear segregations by the PCoA and NMDS analysis among individuals before and after *H. pylori* eradication therapy demonstrated that successful eradication or eradication therapy changed the oral bacterial components to some extent.

In addition to the presence of different bacterial members, the abundance of some bacteria also differed significantly among groups. We clearly observed that some bacteria in the saliva of *H. pylori*

infected individuals showed a significantly reduced abundance, among which *Aggregatibacter*, *Klebsiella*, *Fusobacterium*, and *Parvimonas* were pathogenic bacteria. *Aggregatibacter* is a dominant etiology of infective endocarditis (23). *Klebsiella* and *Fusobacterium* can lead to liver abscess, pneumonia, and meningitis (24). *Parvimonas* has been isolated as a causative agent in a variety of systemic infections, including meningitis, septic arthritis, chest wall abscess, spondylodiscitis, empyema, endocarditis, hepatic abscesses, and brain abscess (25). However, the abundance of other bacteria significantly increased in saliva of *H. pylori* infected individuals, most of which were oral microbiota composition, including *Sphingomonas*, *Ochrobactrum*, *Afipia*, *Leptotrichia*, *Oribacterium*, and *Moryell*, except *Acinetobacter* causing infectious diseases like pneumonia and urinary tract infections (26), and *Leptotrichia*, a potential cariogenic genera (27). While in Christian's study, no significant difference in oral communities between *H. pylori* infected and uninfected individuals were detected at genus level (20), this may be due to the different target sequencing region of 16s rDNA, sample size, or geographic location. Interestingly, most *H. pylori*-enriched genera increased after the eradication, including *Ralstonia*, *Leptotrichia*, *Sphingomonas*, *Leptothrix*, *Oribacterium*, and *Acinetobacter.* The exception was *Ochrobactrum*. However, genera low expressed in *H. pylori* infected saliva experienced a further decline after *H. pylori* eradication therapy, including *Alloprevotella*, *Aggregatibacter*, *Leptotrichlaceae__G_1_*, *Parvimonas*, and *Fusobacterium*, most of which are pathogenic bacteria. The presence of *Ralstonia* positively correlated with the presence of *Oribacterium* and *Acinetobacter*, each of which increased in patients with *H. pylori* after successful eradication. Our study suggests that *H. pylori* infection may change the saliva microbiota by reducing the number of conditional pathogenic bacteria and increasing the number of normal bacteria composition. After *H.*

315 *pylori* eradication therapy, most conditional pathogenic genera in saliva decline while most symbiotic
316 bacteria become more abundant.

317 Although the clinical significance of these alterations is not known, *H. pylori* unexpectedly and
318 clearly altered the oral microbiota composition. Previous studies have reported acid inhibition in upper
319 gastric tract may have an effect on the oral microbiome leading to alterations in the microbiota (28). In
320 addition, changes in gastric pH could also lead to an alteration in the pH of oral cavity (29). *H. pylori*
321 generates large amount of urease, an enzyme which decomposes urea into ammonia and carbon dioxide
322 and transiently reduce the acidic environment in the stomach (30). We proposed that *H. pylori* likely
323 changed the community and structure of oral microbiota through changes in the acidic environment in
324 stomach. The use of PPIs during the eradication therapy would further inhibit the pH in stomach,
325 leading to further alteration in saliva microbiota, which can partially explain why genera enriched in *H.*
326 *pylori* infected individuals would further increase and genera low expressed in *H. pylori* infected
327 individuals would decline after successful eradication. Although the precise mechanism has yet to be
328 clarified, to our knowledge this is the first study to clearly show oral microbiota alterations as a result
329 of *H. pylori* infection in a cohort of subjects. Additional studies to investigate these possible causal
330 relationships would likely provide interesting findings. Besides, by PICRUSt analysis, we predicted
331 that the saliva bacterial functions mainly enriched in membrane transport, replication and repair, amino
332 acid metabolism, carbohydrate metabolism, translation, and energy metabolism.

333 Using amplicon pyrosequencing of 16S rDNA V3-V4 hypervariable regions we detected *H. pylori*
334 in the oral cavity of almost half of the subjects regardless of whether they had gastric infection with *H.*
335 *pylori*. Subjects who did not have *H. pylori* in the oral cavity before eradication surprisingly had *H.*

pylori detected in saliva samples after *H. pylori* eradication therapy. Clearly, using these techniques the prevalence of *H. pylori* in oral cavity is not clearly associated with colonization status in the stomach which is not consistent with the notion that the oral cavity represents a secondary site for *H. pylori* colonization (31). The gastric and oral mucosa differ markedly. For example, of the two only the gastric mucosa expresses Lewis^{b} antigen, an ABO blood group antigen, which enables adherence of *H. pylori* to the epithelial surfaces. It has been proposed that *H. pylori* is a passerby in oral cavity, rather than a colonizer and it may be also be included in the material in gastroesophageal reflux. The natural history of *H. pylori* infection has been that after *H. pylori* eradication from the stomach, gastric reinfection is rare and when it occurs early it can often be shown to be recrudescence (the same genotype) whereas later reinfections are most often reinfection with a different genotype (32). The hypothesis that *H. pylori* was a common passerby rather than a colonizer would partly explain why recurrences are most common in areas with poor sanitation and a high prevalence of *H. pylori* and rare in developed countries whose frequency of *H. pylori* infection had become much lower than that of poor regions.

Strengths and Limitations

Our study showed that *H. pylori* infection and the eradication treatment resulted in alterations of oral microbiota. However, there were limitations to our study. The technique of high-throughput pyrosequencing we used in our study could detect the microbiota at genus level precisely. Metagenomics sequencing was not able to be used to detect *H. pylori*-specific virulence factors such as VacA, CagA, OipA, etc. or full sequence (33). One issue with the interpretation is that there was no control sample of *H. pylori* uninfected individuals receiving the same antimicrobial therapy which precluded determination about whether the presence of *H. pylori*, the antimicrobial therapy, or both were dominant factors in changing the within-individual diversity of the oral cavity.

Conclusions

Our study showed that bacterial diversity was similar in *H. pylori* infected and uninfected people. Eradication therapy was associated with a decline the bacterial diversity in oral cavity. Both *H. pylori* infection and eradication therapy caused alterations in community and structure of the oral microbiota. *H. pylori* is found commonly in the oral cavity with no clear relation to *H. pylori* infection of the stomach.

Acknowledgments

This study was financially supported by National Natural Science Foundation of China (81170355 and 81370592) and Clinical Research Center, Shanghai Jiao Tong University School of Medicine. We acknowledged Shanghai Personal Biotechnology Co., Ltd. For their kind help in 454 pyrosequencing and bioinformatics analysis. All authors don't have any potentially conflicting interests.

Reference

1. Malfertheiner P, Megraud F, O'Morain CA. 2017. Management of Helicobacter pylori infection-the Maastricht V/Florence Consensus Report. Gut 66(1):6-30.

2. David Bui, Heidi E. Brown, Robin B. Harris, Eyal Oren. 2016. Serologic Evidence for Fecal–Oral Transmission of Helicobacter pylori. Am J Trop Med Hyg Jan 6; 94(1): 82–88.

3. John KC Yee. 2016. Helicobacter pylori colonization of the oral cavity: A milestone discovery. World J Gastroenterol Jan 14; 22(2): 641–648.

4. Moodley,A., Wood,N.H., Shangase. 2013. The relationship between periodontitis and diabetes: a brief review. SADJ 68(6):260, 262.

5. Jianye Zhou, Nan Jiang, Zhenzhen Wang, Longqing Li, Jumei Zhang, Rui Ma, Hongbing Nie, and Zhiqiang Li. 2017. Influences of pH and Iron Concentration on the Salivary Microbiome in Individual Humans with and without Caries. Applied and Environmental Microbiology 83(1-4).

6. Ciric,L., Pratten,J., Wilson,M., Spratt,D. 2010. Development of a novel multi-triplex qPCR(q) PCR method for the assessment of bacterial community structure in oral populations. Environmental Microbiology Reports 2(6):770-774.

7. Chen, H., Jiang, W. 2014. Application of high-throughput sequencing in understanding human oral microbiome related with health and disease. Frontiers in Microbiology 5(508):508.

8. Sakiyama Y, Kato R, Inoue S, Suzuki K, Itabe H, Yamamoto M. 2010. Detection of oxidized low-density lipoproteins in gingival crevicular fluid from dental patients. Journal of Periodontal Research 45(2):216-222.

9. Kuczynski J, Stombaugh J, Walters WA, González A, Caporaso JG, Knight R. 2011. Using QIIME to

394 analyze 16S rRNA gene sequences from microbial communities. Current protocols in bioinformatics /
395 editoral board, Andreas D. Baxevanis Chapter 10(7):Unit 10.7.
396 10. Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics
397 26(19):2460-2461.
398 11. NIH/National Institute of Dental and Craniofacial Research. 2008. Scientists Launch first
399 comprehensive database of human oral microbiome. ScienceDaily.
400 12. Bokulich, N.A., Subramanian, S., Faith, J.J., Dirk Gevers, Jeffrey I Gordon, Rob Knight, David A
401 Mills, J Gregory Caporaso. 2012. Quality-filtering vastly improves diversity estimates from Illumina
402 amplicon sequencing. Nature Methods 10(1):57-59.
403 13. Paul D, Kumbhare SV, Mhatre SS, Chowdhury SP, Shetty SA, Marathe NP, Bhute S, Shouche YS.
404 2016. Exploration of microbial diversity and community structure of lonar lake: the only hypersaline
405 meteorite crater lake within basalt rock. Frontiers in Microbiology 6.
406 14. Ramette, A. 2010. Multivariate analyses in microbial ecology. Fems Microbiology Ecology
407 62(2):142-160.
408 15. Warton, D. I., S. T. Wright, Y. Wang. 2012. Distance-based multivariate analyses confound location
409 and dispersion effects. Methods in Ecology and Evolution 3(1): 89-101.
410 16. Francesco Asnicar, George Weingart, Timothy L. Tickle, Curtis Huttenhower, Nicola Segata. 2015.
411 Compact graphical representation of phylogenetic data and metadata with GraPhlAn. PeerJ 3: e1029.
412 17. Nicola Segata, Jacques Izard, Levi Waldron, Dirk Gevers, Larisa Miropolsky, Wendy S Garrett,
413 Curtis Huttenhower. 2011. Metagenomic biomarker discovery and explanation. Genome Biology
414 12(6):R60.

415 18. Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage,
416 Nada Amin, Benno Schwikowski, Trey Ideker. 2003. Cytoscape: a software environment for integrated
417 models of biomolecular interaction networks. Genome Res 13, 2498–2504.
418 19. Morgan G I Langille, Jesse Zaneveld, J Gregory Caporaso, Daniel McDonald, Dan Knights, Joshua
419 A Reyes, Jose C Clemente, Deron E Burkepile, Rebecca L Vega Thurber, Rob Knight, Robert G Beiko.
420 2013. Curtis Huttenhower Predictive functional profiling of microbial communities using 16S rRNA
421 marker gene sequences. Nature Biotechnology 31(9):814-821.
422 20. Christian Schulz, Kerstin Schütte, Nadine Koch, Ramiro Vilchez-Vargas, Melissa L Wos-Oxley,
423 Andrew P A Oxley, Marius Vital, Peter Malfertheiner, Dietmar H Piepe. 2018. The active bacterial
424 assemblages of the upper GI tract in individuals with and without Helicobacter infection. Gut
425 67:216–225.
426 21. Shira R. Abeles, Melissa Ly, Tasha M. Santiago-Rodriguez, David T. 2015. Effects of Long Term
427 Antibiotic Therapy on Human Oral and Fecal Viromes. PLoS One 10(8): e0134941.
428 22. Tsuyoshi Mishiro, Kentaro Oka, Yasutoshi Kuroki, Motomichi Takahashi, Kasumi Tatsumi, Tsukasa
429 Saitoh, Hiroshi Tobita, Norihisa Ishimura, Shuichi Sato, Shunji Ishihara, Joji Sekine, Koichiro Wada,
430 Yoshikazu Kinoshita. 2018. Oral microbiome alterations of healthy volunteers with proton pump
431 inhibitor. J Gastroenterol Hepatol May;33(5):1059-1066.
432 23. Niels Nørskov-Lauritsen. 2014. Classification, Identification, and Clinical Significance of
433 Haemophilus and Aggregatibacter Species with Host Specificity for Humans. Clin Microbiol Rev Apr;
434 27(2): 214–240.
435 24. Alyssa S. Shon, Rajinder P.S. Bajwa, Thomas A. Russo. 2013. Hypervirulent (hypermucoviscous)

436 Klebsiella pneumoniae: A new and dangerous breed. Virulence Feb 15; 4(2): 107–118.

437 25. Chaucer B, Smith N, Beatty D, Yadav M. 2018. Multiple Hepatic Abscess from Parvimonas micra:

438 An Emerging Gastrointestinal Microbe. ACG Case Rep J Sep 26;5:e70.

439 26. Darren Wong, Travis B. Nielsen, Robert A. Bonomo, Paul Pantapalangkoor, Brian Luna, Brad

440 Spellberg. 2017. Clinical and Pathophysiological Overview of Acinetobacter Infections: a Century of

441 Challenges. Clin Microbiol Rev Jan; 30(1): 409–447.

442 27. D. Belstrøm, M.L. Sembler-Møller, M.A. Grande, N. Kirkby, S.L. Cotton, B.J. Paster, S. Twetman,

443 P. Holmstrup. 2018. Impact of Oral Hygiene Discontinuation on Supragingival and Salivary

444 Microbiomes. JDR Clin Trans Res Jan;3(1):57-64.

445 28. Johnstone J, Nerenberg K, Loeb M. 2010. Meta-analysis: proton pump inhibitor use and the risk of

446 community-acquired pneumonia. Alimentary Pharmacology & Therapeutics 31(11):1165-1177.

447 29. JS Jin, M Touyama, S Yamada, T Yamazaki, Y Benno. 2014. Alteration of a human intestinal

448 microbiota under extreme life environment in the Antarctica. Biological & Pharmaceutical Bulletin

449 37(12):1899-906.

450 30. Kentaro Sugano, Jan, Tack, Ernst J Kuipers, David Y Graham, Emad M El-Omar, Soichiro Miura ,

451 Ken Haruma, Masahiro Asaka, Naomi Uemura, Peter Malfertheiner. 2015. Kyoto global consensus

452 report on Helicobacter pylori gastritis. Gut 64:1353–67.

453 31 Zou QH, Li RQ. 2011. Helicobacter pylori in the oral cavity and gastric mucosa: a meta-analysis. J

454 Oral Pathol Med 40: 317-324.

455 32. APA Kim, S. Y., Hyun, J. J., Jung, S. W., Koo, J. S., Yim, H. J., Lee,S.W.. 2014. Helicobacter pylori

456 recurrence after first- and second-line eradication therapy in Korea: the problem of recrudescence or

457 reinfection. Helicobacter Jun;19 (3):202-6.

458 33. Mitsushige Sugimoto, Jeng-Yih Wu, Suhaib Abudayyeh, Jill Hoffman, Hajer Brahem, Khaldun

459 Al-Khatib, Yoshio Yamaoka, and David Y. Graham. 2009. Unreliability of Results of PCR Detection of

460 Helicobacter pylori in Clinical or Environmental Samples. Journal of Clinical Microbiology

461 47(3):738-42.

Tables

Table 1: Alpha diversity indices for saliva bacteria in each group at 97% identity.

Group	Chao1		ACE		Shannon		Inverse Simpson		Coverage		Simpsoneven	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
N	1417.58	195.47	1491.22	210.41	7.94	0.36	1.02	0.01	0.98	0.01	0.05	0.02
P	1393.60	182.57	1465.97	196.06	7.89	0.42	1.02	0.01	0.98	0.00	0.05	0.01

SE, Standard Error. N=uninfected group, P=infected group.

No statistically significant difference was observed in all index among N group and P group ($p>0.05$, Student's t-test).

Table 2: Alpha diversity indices for saliva bacteria in each group at 97% identity.

Group	Chao1		ACE		Shannon		Inverse Simpson		Coverage		Simpsoneven	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
PE	1350.56*	200.94	1421.26*	215.43	7.83*	0.49	1.02	0.01	0.98	0.00	0.05	0.02
E	1143.95*	215.08	1206.85*	234.63	7.44*	0.53	1.03	0.02	0.98	0.00	0.05	0.02

SE, Standard Error. Asterisk indicates significant difference ($p < 0.05$, Student's t-test). PE=pre-eradicated group, E=eradicated group.

*Chao 1, *ACE, and *Shannon index between PE and E was statistically significant different ($p<0.05$).

Figure Legend

Figure 1: Distribution of the predominant bacteria at different taxonomic levels (phylum, class, order, family, and genus). The predominant taxa (>1% relative abundance) in each level are shown. N=uninfected group, P=infected group, E=eradicated group.

Figure 2: Principal coordinate analysis (PCoA) of unweighted UniFrac analysis. (A) PCoA analysis demonstrated that subjects of P group were significantly different from N group (PERMANOVAR, pseudo-F: 1.49, p=0.033). N=uninfected group, P=infected group. (B) PCoA analysis showed that the overall microbial composition showed significant difference between PE and E group (PERMANOVAR, pseudo-F: 3.34, p=0.001). E=eradicated group, N=uninfected group, PE=pre-eradicated group.

Figure 3: Comparison of microbial variations at the genus level, using the LEfSe online tool. (A) Histogram of the LDA scores for differentially abundant features among groups. The threshold on the logarithmic LDA score for discriminative features was set to 2.0. N=uninfected group, P=infected group. (B) Cladogram for taxonomic representation of significantly differences among groups. Differences are represented in the color of the most abundant taxa (red indicating P group, green indicating N group, and white indicating non-significant). N=uninfected group, P=infected group.

Figure 4: H. pylori in oral cavity of three groups. Red represent P group, yellow represent N group, and blue represent E group. N=uninfected group, P=infected group, E=eradicated group.

498

Supporting Information

Table S1 Clinical parameters of the 58 subjects. N=uninfected group, P=infected group, PE=pre-eradicated group, E=eradicated group.

Table S2: Modified OTU table at 97% identity. N=uninfected group, P=infected group, E=eradicated group.

Figure S1: Length distribution of sequences determined by 454 pyrosequencing.

Figure S2: A classification tree showing bacterial abundance by MEGAN. The taxonomy compositions and abundances were visualized by MEGAN (version 6.6.7). The larger the area of the colored pie chart, the greater the bacterial abundance. Different colors represent different groups, and the larger the colored sectorial area at a branch, the more the corresponding group contributed to the bacterial abundance. N=uninfected group, P=infected group, E=eradicated group.

Figure S3: Alpha diversity (observed species number) among groups. (A) N group and P group showed similar alpha diversity ($p>0.05$). N=uninfected group, P=infected group. (B) The observed species in E group were significantly lower than that of PE group and N group ($p<0.01$); One asterisk indicates significant differences ($p < 0.05$, Student's t-test), two asterisk indicates $p<0.01$, three asterisk indicates $p<0.001$. N=uninfected group, PE=pre-eradicated group, E=eradicated group.

Figure S4: Rank abundance curves for all OTUs. N=uninfected group, P=infected group, E=eradicated group.

Figure S5: Nonmetric Multidimensional Scaling (NMDS) based on unweighted UniFrac distances at the OUT level at 97% identity. Each sample is represented by a dot. （A）The samples formed well-separated clusters corresponding to the two groups, suggesting that the bacterial structures in N group and P group were different. N=uninfected group, P=infected group. Red squares represent the N samples. Blue triangles represent the P samples. (B) Blue triangles represent the N samples. Red circles represent the E samples. The samples formed well-separated clusters corresponding to the three groups, suggesting that the bacterial structures in E group, PE group, and N group were different. N=uninfected group, PE=pre-eradicated group, E=eradicated group.

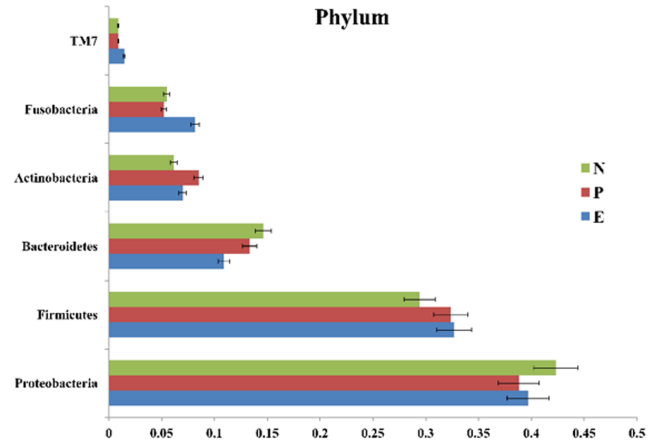
Figure S6: Comparison of microbial variations at the genus level, using the LEfSe online tool. (A) Histogram of the LDA scores for differentially abundant features among groups. The threshold on the logarithmic LDA score for discriminative features was set to 2.0. N=uninfected group, PE=pre-eradicated group, E=eradicated group. (B) Cladogram for taxonomic representation of significantly differences among groups. Differences are represented in the color of the most abundant taxa (red indicating N group, blue indicating PE group, green indicating E group, and white indicating non-significant). N=uninfected group, PE=pre-eradicated group, E=eradicated group.

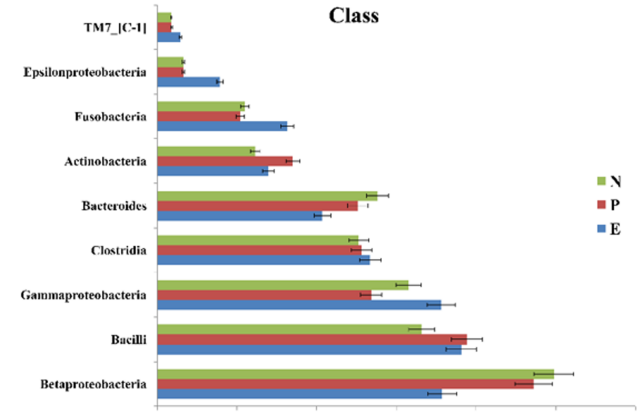
Figure S7: A network diagram of dominant genera showing associations among the 50 most

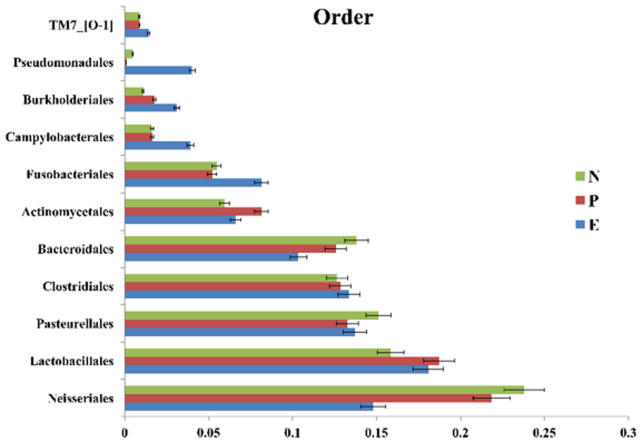
541 **abundant genera.** Red line represent positive associations, and green line represent negative
542 associations.

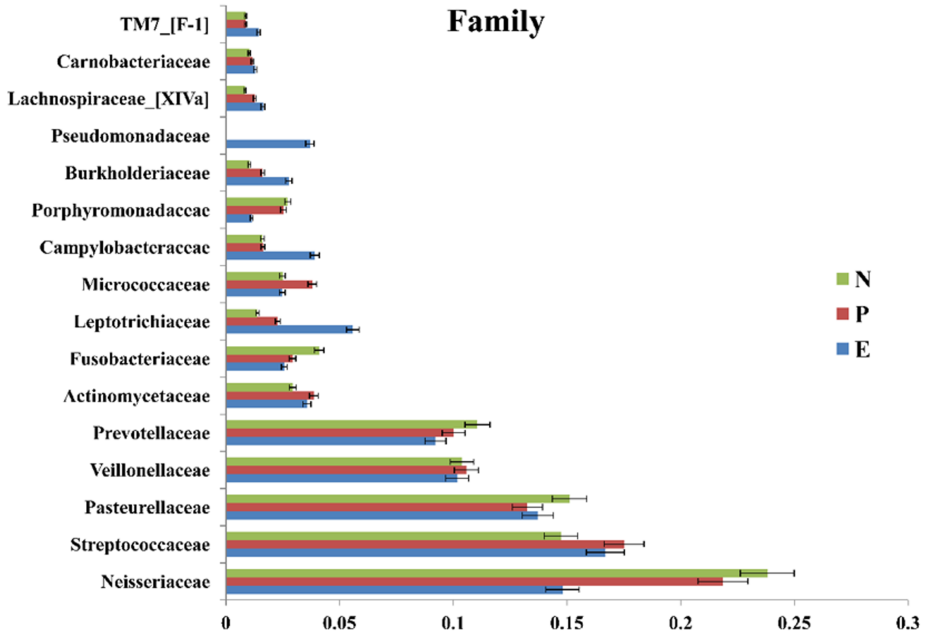
543

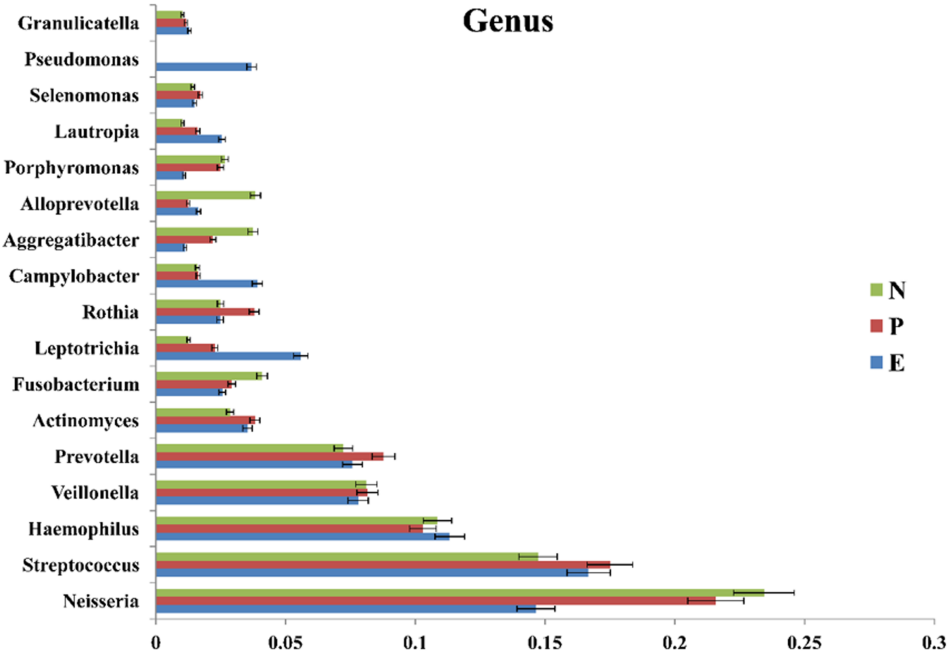
544 **Figures S8: Bacterial function prediction by PICRUSt analysis.** N=uninfected group, P=infected
545 group, E=eradicated group.

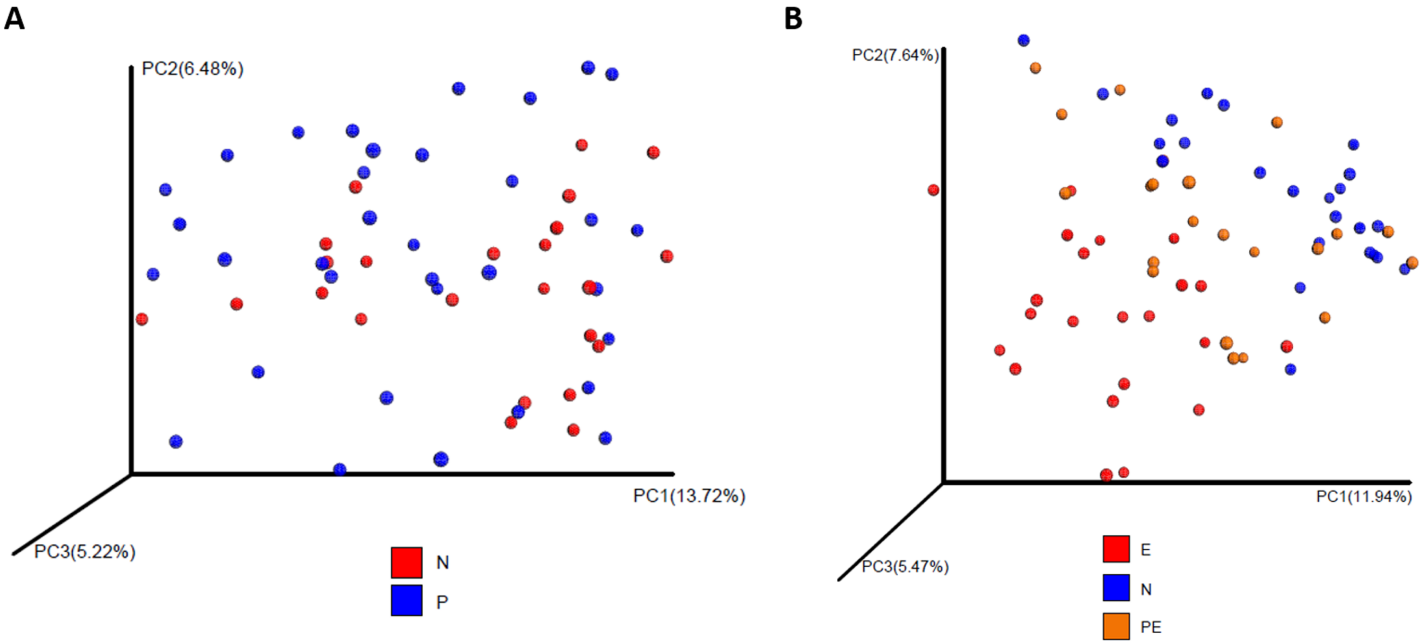


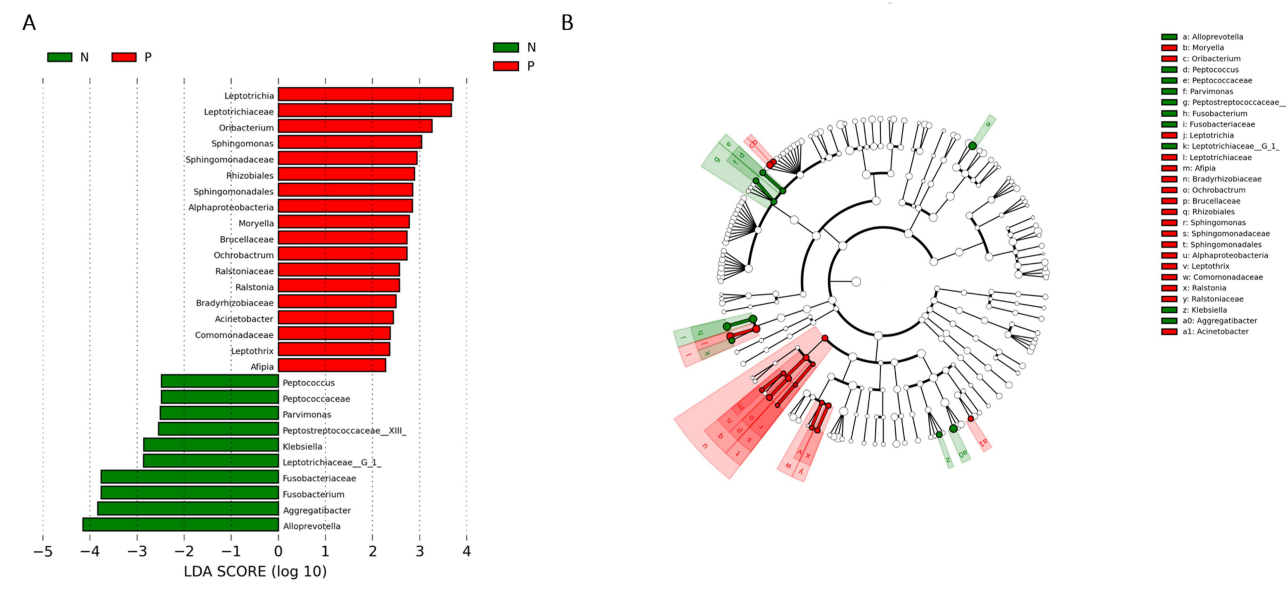











A
N
P
N
P
Leptotrichia
Leptotrichiaceae
Oribacterium
Sphingomonas
Sphingomonadaceae
Rhizobiales
Sphingomonadales
Alphaproteobacteria
Moryella
Brucellaceae
Ochrobactrum
Ralstoniaceae
Ralstonia
Bradyrhizobiaceae
Acinetobacter
Comomonadaceae
Leptothrix
Afipia
Peptococcus
Peptococcaceae
Parvimonas
Peptostreptococcaceae__XIII_
Klebsiella
Leptotrichiaceae__G_1_
Fusobacteriaceae
Fusobacterium
Aggregatibacter
Alloprevotella
−5 −4 −3 −2 −1 0 1 2 3 4
LDA SCORE (log 10)
B
a: Alloprevotella
b: Moryella
c: Oribacterium
d: Peptococcus
e: Peptococcaceae
f: Parvimonas
g: Peptostreptococcaceae__
h: Fusobacterium
i: Fusobacteriaceae
j: Leptotrichia
k: Leptotrichiaceae__G_1_
l: Leptotrichiaceae
m: Afipia
n: Bradyrhizobiaceae
o: Ochrobactrum
p: Brucellaceae
q: Rhizobiales
r: Sphingomonas
s: Sphingomonadaceae
t: Sphingomonadales
u: Alphaproteobacteria
v: Leptothrix
w: Comomonadaceae
x: Ralstonia
y: Ralstoniaceae
z: Klebsiella
a0: Aggregatibacter
a1: Acinetobacter

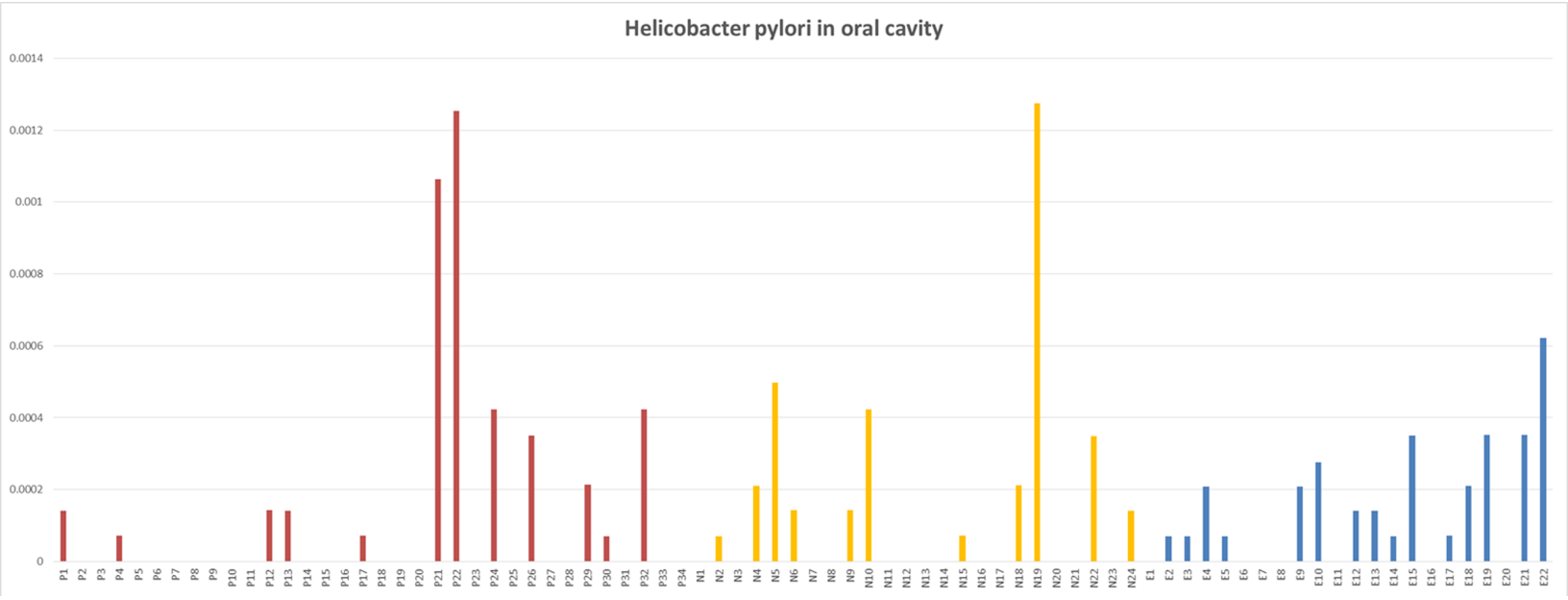


Table 1: Alpha diversity indices for saliva bacteria in each group at 97% identity.

Group	Chao1		ACE		Shannon		Inverse Simpson		Coverage		Simpsoneven	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
N	1417.58	195.47	1491.22	210.41	7.94	0.36	1.02	0.01	0.98	0.01	0.05	0.02
P	1393.60	182.57	1465.97	196.06	7.89	0.42	1.02	0.01	0.98	0.00	0.05	0.01

SE, Standard Error. N=uninfected group, P=infected group.

No statistically significant difference was observed in all index among N group and P group ($p>0.05$, Student's t-test).

Table 2: Alpha diversity indices for saliva bacteria in each group at 97% identity.

Group	Chao1		ACE		Shannon		Inverse Simpson		Coverage		Simpsoneven	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
PE	1350.56*	200.94	1421.26*	215.43	7.83*	0.49	1.02	0.01	0.98	0.00	0.05	0.02
E	1143.95*	215.08	1206.85*	234.63	7.44*	0.53	1.03	0.02	0.98	0.00	0.05	0.02

SE, Standard Error. Asterisk indicates significant difference ($p < 0.05$, Student's t-test).

PE=pre-eradicated group, E=eradicated group.

*Chao 1, *ACE, and *Shannon index between PE and E was statistically significant different ($p<0.05$).