Altered oral microbiota composition associated with recurrent aphthous stomatitis in young females

Zhengyang Zhu, MSc*a, Zhixing He, PhDb, Guanqun Xie, MDb, Yongsheng Fan, MDb, Tiejuan Shao, PhDb,*

Abstract
Oral microbiota has been implicated in pathogenesis of recurrent aphthous stomatitis (RAS), which is a common mucosal disorder with unclear etiology. This study has explored the association between oral microbiota disorder and RAS in high-risk young female population.

Forty-five young females were enrolled, including 24 RAS patients and 21 healthy individuals. Oral microbiome was analyzed by illumina miseq sequencing.

Oral microbiota associated with RAS was characterized by the lower alpha-diversity indices (Chao1 and ACE). Several infectious pathogens increased in RAS, such as genera Actinobacillus, Haemophilus, Prevotella and Vibrio. The PICRUSt analysis indicated that the oral microbiota might be related with the up-regulation of genes involving infectious and neurodegenerative diseases, environmental adaptation, the down-regulation of genes involving basal metabolism, such as carbohydrate, energy, and amino acid metabolism.

This study indicated that oral microbiota may play a significant role in RAS development.

Abbreviations: BMI = Body Mass index, CRP = C-reactive protein, HCs = healthy controls, KEGG = kyoto encyclopedia of genes and genomes, LDA = linear discriminant analysis, OTUs = Operational Taxonomic Units, PCoA = principal coordinates analysis, QIIME = quantitative insights into microbial ecology, RAS = recurrent aphthous stomatitis.

Keywords: illumina miseq, oral microbiota, recurrent aphthous stomatitis, young females

1. Introduction
Recurrent aphthous stomatitis (RAS) is a common clinical condition producing painful ulcerations in the oral cavity.[1] Epidemiological studies have shown that the prevalence of RAS was approximately 20% in the general population, and it was especially higher (male, 48.3%; female, 57.2%) in professional-school students aged between 10 and 19 years.[2,3] RAS is characterized by chronic inflammation of the oral mucosa with multiple recurrent small, rounds, ovoid ulcers with circumscribed margins, erythematous haloes, and yellow or gray floors.[4] The potential trigger factors of RAS include genetic predisposition, immunologic disturbances, viral and bacterial infections, food allergies, systemic diseases, and stress; however, the etiology of RAS remains unknown.[5]

To date, it is widely accepted that genetic, immunological and environmental factors are involved with RAS.[6] Microbial infection and disturbance of the oral microbiota have been identified as the important environmental factors. Streptococcal species and Helicobacter pylori were the two main infectious bacteria that may contribute to the pathogenesis.[7,8] The infectious bacteria may act as either pathogens or as antigens that induce the production of antibodies, which can cross-react with oral mucosal keratinocytes.[9] However, a complex commensal microbiota exists in the oral cavity, which is characterized by the oral microbiota. Previous studies have reported the association between oral microbiota and RAS,[9–11] however, among various RAS populations, the profiles of the oral microbiota differed resulting from the intervening factors, such as diet, age, and host gene.[12–14] Therefore, the relationship between RAS and the oral microbiota in young adults may be different from the previous reported literature.
To weaken the intervening effects of other factors on the association between RAS and oral microbiota, forty-five females, aged 19–22 years, were enrolled in this study. Illumina MiSeq sequencing was used to reveal the oral microbiome differences between RAS and healthy students. The aim of this study was to ascertain the association between oral microbiota and RAS in young females.

2. Materials and methods

2.1. Study subjects

A cross-sectional collection of saliva samples was performed for female students from Zhejiang Chinese Medical University who were suffering from ulcer in the oral cavity and diagnosed with RAS. In addition, another group of female from the same university, who were matched by gender and age, were also enrolled for saliva sample collection as healthy controls (HCs). All participants were aged between 19 and 22 years old. Subjects who have received antibiotics or antibacterial mouth rinses or drugs within one month were excluded. Additional exclusion criteria included other types of oral mucosal lesions, smoking, and consumption of alcohol within one month were also reordered. Twenty RAS females were summarized in Table 1. No difference in age and BMI index collected from each subject and stored at Chinese Medical University, and written informed consent was collected from healthy females. A minimum of 2ml saliva was day after ulcer appearance. Unstimulated whole saliva was collected from the diluted DNA using bacterial primers 338F (5′- ACTCTACGGGAGGCAGCA-3′) and 806R (5′-GGACTACHVGGGTWTCTAA T-3′), with the reverse primer containing a 6-bp barcode used to tag each sample. PCR amplification was performed in a 30μl mixture containing 1.0μl of forward primer (10μM), 1.0μl of reverse primer (10μM), 5.0μl of dNTP (100mM), 2μl of DNA template (20ng/μl), 5.0μl of 5× reaction buffer, 5.0μl of 5×GC buffer, 9.0μl of ddH2O and 2.0μl of QS High-Fidelity DNA Polymerase (NEB). The reactions were hot-started at 98°C for 2min, followed by 25 cycles of 98°C for 15s, 55°C for 30s, and 72°C for 30s, with a final extension step at 72°C for 5min.

Amplicons were extracted from 1.5% agarose gel and purified using the QIAquick Gel Extraction kit (Qiaegen, Valencia, CA, USA). Purified amplicons were pooled in equimolar quantities and pair-end sequenced (2 x 250) on an Illumina MiSeq platform (Personal Biotechnology Co., Ltd, Shanghai) according to the standard protocols. Libraries were constructed according to the TruSeq™ DNA Sample Prep Kit protocol (Illumina, USA).

2.2. Sample collection and DNA extraction

For sampling, the saliva of RAS females was collected on the third day after ulcer appearance. Unstimulated whole saliva was collected from healthy females. A minimum of 2ml saliva was collected from each subject and stored at −80°C until DNA extraction.

Before DNA extraction, the frozen saliva samples underwent a freeze-drying process for removing moisture. The concentrated samples were used to extract genomic DNA using the QIAamp DNA Micro Kit (Qiaegen, Hilden, Germany) according to the manufacturer’s protocols. The DNA concentration was determined using NanoDrop (Thermo Scientific) and DNA was visualized using agarose gel electrophoresis.

2.3. 16S rRNA gene amplification and sequencing

The hyper-variable V3-V4 regions of the bacterial 16S rRNA gene sequences were amplified from the diluted DNA using bacterial primers 338F (5′- ACTCTACGGGAGGCAGCA-3′) and 806R (5′-GGACTACHVGGGTWTCTAA T-3′), with the reverse primer containing a 6-bp barcode used to tag each sample. PCR amplification was performed in a 30μl mixture containing 1.0μl of forward primer (10μM), 1.0μl of reverse primer (10μM), 5.0μl of dNTP (100mM), 2μl of DNA template (20ng/μl), 5.0μl of 5× reaction buffer, 5.0μl of 5×GC buffer, 9.0μl of ddH2O and 2.0μl of QS High-Fidelity DNA Polymerase (NEB). The reactions were hot-started at 98°C for 2min, followed by 25 cycles of 98°C for 15s, 55°C for 30s, and 72°C for 30s, with a final extension step at 72°C for 5min.

Amplicons were extracted from 1.5% agarose gel and purified using the QIAquick Gel Extraction kit (Qiaegen, Valencia, CA, USA). Purified amplicons were pooled in equimolar quantities and pair-end sequenced (2 x 250) on an Illumina MiSeq platform (Personal Biotechnology Co., Ltd, Shanghai) according to the standard protocols. Libraries were constructed according to the TruSeq™ DNA Sample Prep Kit protocol (Illumina, USA).

2.4. Processing of sequencing data

After sequencing, raw FASTQ files were demultiplexed and quality-filtered using Quantitative Insights Into Microbial Ecology (QIIME, version 1.17) under the following criteria: (i) reads were truncated at any site receiving an average quality score of <20bp over a 50-bp sliding window, and truncated reads shorter than 50bp were discarded; (ii) exact barcode matching, two nucleotide mismatch in primer matching, and reads containing ambiguous characters were removed; (iii) only sequences that overlapped for more than 20bp were merged according to their overlap sequence, reads that could not be merged were discarded.

The high-quality sequences were clustered into the 16S rRNA Operational Taxonomic Units (OTUs) with 97% similarity cutoff using UPARSE (version 7.1). Chimera detection and removal were assessed using the GOLD reference database and UCHIME. The taxonomic affiliation of each OTU was analyzed using Ribosomal Database Project (RDP) classifier against the SILVA 16S rRNA database.

2.5. Sequence analysis

Bacterial diversity was determined by sampling based analysis of OTUs, and shown by Chao1, ACE, Shannon and Simpson indices that were calculated by using R program package “vegan”. Principal coordinates analysis (PCoA) based on OTUs distribution was conducted by the R package to visualize interaction among bacterial communities.

The specific characterization of oral microbiota was analyzed by linear discriminant analysis (LDA) effect size (LEfSe) method (http://huttenhower.sph.harvard.edu/lefse/). LEfSe performs the Kruskall-Wallis test to determine the features with significantly different abundances between RAS and HCs and LDA is used to assess the effect size of each feature.

### Table 1

|                         | Healthy females (n = 21) | RAS females (n = 24) |
|-------------------------|--------------------------|----------------------|
| Age                     | 19–22                    | 19–22                |
| Ulcer numbers           | NA                       | Single: 20           |
|                         |                          | Multiple: 4          |
| BMI                     | 21.98 ± 0.85             | 22.48 ± 1.20         |
| CRP (mg/L)              | 0.48 ± 0.14              | 1.98 ± 0.73          |
| Medical history         |                          |                      |
| Chronic hypertension    | NA                       | NA                   |
| Diabetes mellitus       | NA                       | NA                   |
| Autoimmune disease      | NA                       | NA                   |
Microbial function was inferred using PICRUSt that predict the molecular functions of each sample based on the 16S rRNA maker gene sequences. These predictions were pre-calculated for genes in databases including KEGG and COGs. Briefly, the PICRUSt that consisted of two steps: gene content inference and metagenome inference was performed as described previously.[18]

Multiple group differences in diversity indices and predicted metabolic function were analyzed via Mann-Whitney non-parametric test in SPSS software 16.0.

3. Results
3.1. Structure and composition of the oral microbiome in RAS

Alpha and beta diversity was calculated to assess overall differences in microbial community structure of saliva between RAS and HCs. As shown in Figure 1, the two alpha-diversity indices (Chao1 and ACE) of RAS were significantly lower than HCs ($p < 0.05$), however, the other two alpha-diversity indices (Shannon and Simpson) were not statistically different. The above results indicated that lower diversity of oral microbiome was accompanied with RAS in young females.

The beta-diversity difference using the unweighted UniFrac distance was shown by PCoA analysis (Fig. 2). On the PCoA plot, oral microbial communities separated in RAS and HC samples, suggesting a unique oral microbiota in RAS.

3.2. RAS-associated microbiota changes in saliva

LEfSe method was used to display the greatest differences of microbial structure between HCs and RAS. Seventeen different microbial phylotypes with LDA score $>2$ and $p$ value $<0.05$ were displayed in Figure 3. Compared with HCs, 10 phylotypes increased while 7 phylotypes decreased in RAS.

At the phylum level, Proteobacteria was significantly more abundant in the oral microbiota of the RAS than that in the HCs (Fig. 3). The class Gammaproteobacteria and its derivates (Order Pasteurellales, genus Actinobacillus and Haemophilus, and order Vibrionales, Family Pseudomonadaceae, genus Vibrio) contributed to the significant increase of the phylum Proteobacteria in RAS. In addition, Family Prevotellaceae and its genus Prevotella were also significantly more enriched in the oral microbiota of RAS.

Conversely, Firmicutes were significantly reduced in the oral microbiota of RAS (Fig. 3). Compared with the HCs, the class Bacilli and its derivates (Order Gemellales, Family Gemellaceae and Order Bacillales, Family Paenibacillaceae) decreased in the RAS. Meanwhile, the decreased phylotypes of RAS also included the genus Anaerovorax and Xanthomonadaceae.

3.3. Predicted functional profile of oral microbiota

Phylogeny and function of oral microbial community were connected. Hence, the PICRUSt approach was used to analyze the KEGG pathways compositions in bacterial population based on 16S rRNA maker gene sequences. 12 second-level KEGG pathways differed in abundance between RAS and HCs (Fig. 4, $p < 0.05$).

Firstly, the pathways of infectious diseases and neurodegenerative diseases significantly increased in oral microbiota of RAS (Fig. 4). Secondly, the PICRUSt analyses also indicated that oral microbiota in RAS had the significantly more abundant genes involving in cell motility, cellular processes and signaling, as well as environment adaptation. Thirdly, oral microbiota of RAS had the significantly lower basal metabolism, such as carbohydrate metabolism, biosynthesis of other secondary metabolites, cell growth and death, energy metabolism, amino acid metabolism, and nucleotide metabolism.

4. Discussion

It has been well known that microbial factor is a causative agent of RAS.[19] Therefore, imbalances of oral microbiota are frequently accompanied with RAS.[9–11] The present study also characterized the oral microbiome of RAS. The study cohort was selected in young females aged from 19 to 22, since the onset of
RAS seems to peak in young females according to epidemiologic studies.\(^\text{[20]}\)

Using Illumina Miseq sequencing, the present study showed a significantly decreased alpha diversity in the oral microbiota of RAS. The previous literature also studied alpha diversity of oral microbiota in RAS, but there is no obvious difference in alpha diversity between healthy controls and RAS in study population with the mean age of 44.\(^\text{[11]}\) It has been widely accepted that the lower alpha diversity represented more severe imbalances of microbial community.\(^\text{[21,22]}\) Therefore, the young adults suffered a sharper imbalance of oral microbiota than the middle-aged adults with RAS.

In young females, several associations between RAS and oral microbiota was consistent with the former studies. It was well known that Streptococcal species and *Helicobacter pylori* were the main RAS-related pathogenic bacteria. Due to the limit of 16S rRNA Miseq sequencing, *Helicobacter pylori* was not identified in this study. In addition, Streptococcal species were not found being the significantly altered microbial phylotypes. Indeed, Streptococcal species and *Helicobacter pylori* have been proposed to provoke development of RAS,\(^\text{[7,23]}\) but there is no evidence to prove that they are related with RAS in oral microbiome.\(^\text{[21,22]}\) The reason might be that Streptococcal species were initial factors perturbing the global oral microbiota at the early stage of inflammation in RAS but their proportion were lower to normal levels after oral microbiota imbalances. In addition, the genus *Haemophilus* and *Prevotella*, which were increased in RAS, were also reported to be positive with RAS in the previous literature.\(^\text{[11]}\) It has been well-known that genus *Haemophilus* and *Prevotella* were two infectious pathogen through intervening immunity, energy metabolism and amino acid metabolism.\(^\text{[24,25]}\)

Furthermore, there are some alterations of oral microbiota specifically being found in this study. In oral cavity of young females with RAS, the family *Pseudoalteromonadaceae* which was known as the infectious bacteria also significantly increased. The family *Pseudoalteromonadaceae* was always considered as the exogenous invasion in the oral cavity, and its genus *Vibrio* was a common type of infectious bacterium in human oral cavity.\(^\text{[26]}\) In addition, the increased genus *Actinobacillus* of RAS has been strongly implicated in the development of periodontitis.\(^\text{[27,28]}\) The above specific findings may be the underlying reasons leading to the high risk of RAS in young adults.

Phylogeny and functions of gut microbiota are connected.\(^\text{[18]}\) The altered oral microbiota could determine the potential functional of oral microbiota in the samples. RAS is a systemic disease with pathogenesis of infection, immunity and metabolism disorders.\(^\text{[29]}\) The study inferred that oral microbiota could contribute to RAS pathogenesis due to the different function pathways of oral microbiota in the infectious diseases and basal metabolisms. The appearance of RAS could change the micro-environment of the oral mucosal with inflammation and immune activity.\(^\text{[30]}\) The above stress would have adverse effects on oral microbiota and decline general metabolic functions (such as amino acid metabolism, carbohydrate metabolism and nucleotide metabolism) of oral microbiota in RAS. This is consistent with effects of microbiota on other diseases.\(^\text{[31,32]}\) Meanwhile, the adverse environment in oral mucosal of RAS could induce some microbiota increasing, which could perform functional pathway (environmental adaptation) in decreasing negative effects.

### 5. Conclusions

In conclusion, the present study explored association between oral microbiota and RAS in young females for the first time. Some specific oral microbiota significantly altered in young females suffered from RAS, such as *Haemophilus*, *Prevotella*, *Vibrio* and *Actinobacillus*. The predicted function pathways using PICRUSt also verified the correlation between oral microbiota and RAS. Thus, our study may facilitate understanding the role of oral microbiota in pathogenesis of RAS.
microbiota in RAS development in young females. However, this study also had several limitations. Firstly, the sample population come from a small group of people, and both sample source and size need to be expanded to verify our results. Secondly, there is little basic clinical information of sample population, and more clinical information needs to be collected to exclude their influence on oral microbiota.

Author contributions
Conceptualization: Tiejuan Shao, Yongsheng Fan, Zhixing He.
Data curation: Zhengyang Zhu, Zhixing He, Tiejuan Shao.
Formal analysis: Guanqun Xie.
Funding acquisition: Yongsheng Fan.
Methodology: Zhengyang Zhu, Guanqun Xie.
Software: Zhixing He.
Supervision: Guanqun Xie.
Validation: Guanqun Xie.
Writing - original draft: Zhengyang Zhu, Tiejuan Shao.
Writing-review & editing: Zhengyang Zhu, Tiejuan Shao, Yongsheng Fan, Zhixing He.

References
[1] Edgar NR, Saleh D, Miller RA. Recurrent Aphthous Stomatitis: A Review. The Journal of clinical and aesthetic dermatology 2017;10:26–36.
[2] Gallo Cde B, Mimura MA, Sugaya NN. Psychological stress and recurrent aphthous stomatitis. Clinics (Sao Paulo) 2009;64:645–8.
[3] Chavan M, Jain H, Diwan N, et al. Recurrent aphthous stomatitis: a review. Journal of oral pathology & medicine: official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology 2012;41:577–83.
[4] Jurge S, Kuffer R, Scully C, et al. Muco pathogenesis of recurrent aphthous stomatitis: literature review. Arch Immunol Ther Exp (Warsz) 2014;62:203–15.
Zhu et al. Medicine (2021) 100:10

[6] Akintoye SO, Greenberg MS. Recurrent aphthous stomatitis. Dent Clin North Am 2014;58:281–97.

[7] Di Pierro F, Adami T, Rapaccoli G, et al. Clinical evaluation of the oral probiotic Streptococcus salivarius K12 in the prevention of recurrent pharyngitis and/or tonsillitis caused by Streptococcus pyogenes in adults. Expert Opin Biol Th 2013;13:339–43.

[8] Guiseren D, Karaduman A, Kutsal D, et al. The relationship between recurrent aphthous stomatitis, and periodontal disease and Helicobacter Pylori infection. Clin Oral Invest 2016;20:2035–60.

[9] Bankvall M, Spéberg F, Gale G, et al. The oral microbiota of patients with recurrent aphthous stomatitis. J Oral Microbiol 2014;6:25739.

[10] Hijazi K, Lowe T, Meharg C, et al. Mucosal Microbiome in Patients with Recurrent Aphthous Stomatitis. J Dent Res 2015;94:87s.

[11] Kim YJ, Choi YS, Baek KJ, et al. Mucosal and salivary microbiota associated with recurrent aphthous stomatitis. Bmc Microbiol 2016;16 Suppl 1:57. doi: 10.1186/s12866-016-0673-z.

[12] Adler CJ, Dobney K, Weyrich LS, et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet 2013;45:450–5.

[13] Mason MR, Nagaraja HN, Camerlengo T, et al. Deep Sequencing Identifies Ethnicity-Specific Bacterial Signatures in the Oral Microbiome. Plos One 2013;8:e77287.

[14] Shoemark DK, Allen SJ. The Microbiome and Disease: Reviewing the Neolithic and Industrial revolutions. Quintessence Int 2000;31:95–112.