

SCIENTIFIC REPORTS



OPEN

Variations in oral microbiome profiles in rheumatoid arthritis and osteoarthritis with potential biomarkers for arthritis screening

Bin Chen¹, Yan Zhao^{1,2}, Shufeng Li³, Lanxiu Yang⁴, Haiying Wang¹, Tao Wang², Bin Shi², Zhongtao Gai⁸, Xueyuan Heng⁹, Chunling Zhang⁷, Junjie Yang⁵ & Lei Zhang^{1,6,7,8,9,10}

The key to arthritis management is early diagnosis and treatment to prevent further joint destruction and maximize functional ability. Osteoarthritis (OA) and rheumatoid arthritis (RA) are two common types of arthritis that the primary care provider must differentiate, in terms of diagnosis and treatment. Effective and non-invasive strategies for early detection and disease identification are sorely needed. Growing evidence suggests that RA has a correlation with oral microbiome and may be affected by its dynamic variations. There is already a study comparing oral microbiome in patients with RA and OA, however, it did not screen for potential biomarkers for arthritis. In this study, we assessed the oral microbiome in saliva samples from 110 RA patients, 67 OA patients and 155 healthy subjects, using 16S rRNA gene amplicon sequencing. The structure and differences in oral microbiome between RA, OA and healthy subjects were analyzed. Eight oral bacterial biomarkers were identified to differentiate RA from OA. This report provides proof of oral microbiota as an informative source for discovering non-invasive biomarkers for arthritis screening.

Rheumatoid arthritis (RA) is a highly prevalent, chronic inflammatory and systemic autoimmune disease. Multiple clinical studies have suggested a potential link between RA and periodontopathic bacteria such as *Porphyromonas gingivalis*¹. Traditionally, osteoarthritis (OA) is viewed as a mechanically induced condition, and the role of inflammatory factors, such as those derived from microbiota, has been considered minor. Although there are significant differences in the pathogenesis of RA and OA, the clinical symptoms of rheumatoid arthritis and osteoarthritis are similar and sometimes difficult to distinguish, especially in the case of unilateral knee RA, which is more difficult to distinguish from OA. It may also affect prognosis in patients. Therefore, it is necessary to find a strategy for early detection of arthritis.

Recent molecular investigations have indicated that each individual carries over 200 species in their oral microbiome². The relationship between periodontopathic bacteria and systemic diseases has been explored for several years^{3–5}. Several studies have confirmed that oral bacteria have been associated with pancreatic cancer⁶, Parkinson's disease⁷, systemic lupus erythematosus⁸, colorectal cancer, lung inflammation⁹, type-2 diabetes^{10,11}. Oral bacteria can be mobile and affect sites in and around the oral cavity. It can enter the bloodstream easily and also during daily mastication, brushing, flossing and in diseased state. Oral bacteria may penetrate through

¹College of Life Science, Shandong Normal University, Jinan, 250014, China. ²Shandong Medicinal and Biotechnology Centre, Shandong Academy of Medical Sciences, Jinan, 250062, China. ³Department of Orthopedics, Qianfoshan Hospital Affiliated to Shandong University, Jinan, 250014, China. ⁴Guoyitang Hospital, Jinan, 250000, China. ⁵College of Life Science, Qilu Normal University, Jinan, 250200, China. ⁶Shandong Institutes for Food and Drug Control, Jinan, 250101, China. ⁷Qingdao Human Microbiome Center, The Affiliated Central Hospital of Qingdao University, Siliu South Road 127, Qingdao, Shandong Province, 266042, China. ⁸Shandong Children's Microbiome Center, Qilu Children's Hospital of Shandong University, Jinan, Shandong Province, 250022, China. ⁹Microbiological Laboratory, Department of Infection Management, Department of Neurosurgery, Lin Yi People's Hospital, Linyi, Shandong Province, 276003, China. ¹⁰Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Chemistry and Environment, Beihang University, Beijing, 100191, China. Bin Chen and Yan Zhao contributed equally. Correspondence and requests for materials should be addressed to L.Z. (email: microbiome@foxmail.com) or J.Y. (email: microbiota@foxmail.com) or C.Z. (email: qdzcl2011@163.com)

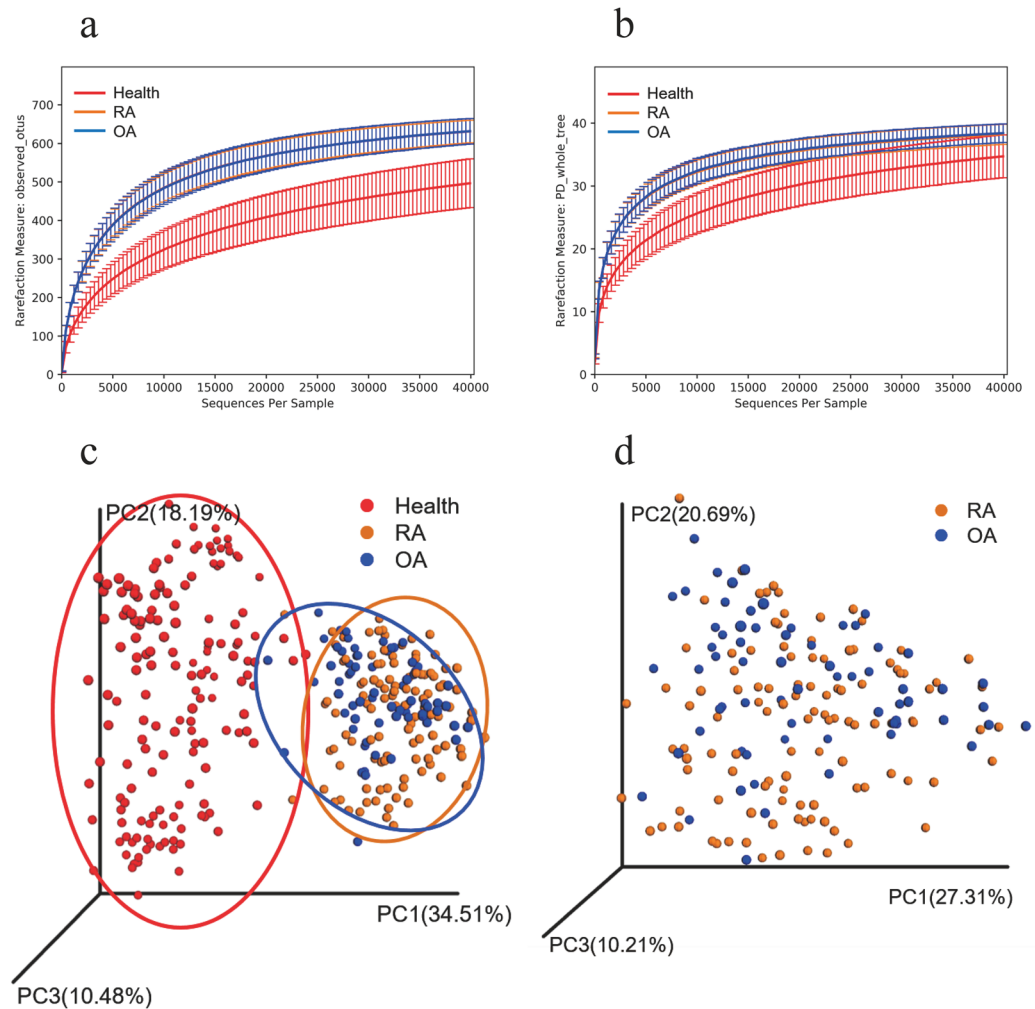


Figure 1. Alpha and beta diversity in RA, OA and healthy subjects. **(a)** Rarefaction analysis of observed OTUs. Using the mean of observed OTUs randomly sampled 150 times at 40000 sequencing depth. Error bars represent standard deviation. **(b)** Rarefaction analysis of Shannon index. Using the mean of observed OTUs randomly sampled 150 times at 40000 sequencing depth. Error bars represent standard deviation. **(c)** Weighted UniFrac principle coordinate analysis of RA, OA and healthy subjects. Ellipses are added to better visualize the cluster and separation between RA, OA and healthy control. **(d)** Weighted UniFrac principle coordinate analysis of RA and OA.

relatively permeable epithelial pockets to reach the underlying gingival connective tissue, enter the bloodstream and travel throughout the body^{12,13}.

Many clinical studies have shown that RA and OA may be associated with oral microbiota and are affected by it. Scher *et al.*¹⁴ used pyrosequencing to compare the composition of subgingival microbiota in patients with RA and controls, revealing that *Prevotella* and *Leptotrichia* species only existed in RA. Subsequently, a more intensive study performed metagenomic shotgun sequencing to detect the dental and salivary microbiome in patients with RA and controls. It was observed that *Haemophilus* spp. decreased in RA, and *Lactobacillus salivarius* numbers increased. Also, the functions of oral microbiome changed in RA¹⁵. However, recent findings suggest that OA may also be regulated by the microbiome, but there are very few studies describing the potential effects of oral microbiota in patients with OA.

In our study, we aim to describe the oral microbiome in patients with RA, OA and healthy subjects, and their potential association with RA and OA. We also attempted to distinguish RA and OA by identifying differences in oral microbiome and found a set of potential biomarkers for distinguishing RA and OA non-invasively.

Results

Variations in Oral Microbiota Profiles in RA and OA. We compared the oral microbiota profiles of 110 RA patients, 67 OA patients and 155 healthy subjects. Alpha diversity analysis reveals that there was no significant difference in observed operational taxonomic units (OTUs) ($P=0.213$, Welch's t-test, Fig. 1a) and PD whole tree index ($P=0.426$, Welch's t-test, Fig. 1b) between RA and OA, while the healthy group had significant differences compared to groups with RA ($P<0.001$) and OA ($P<0.001$). To visualize the overall differences in beta

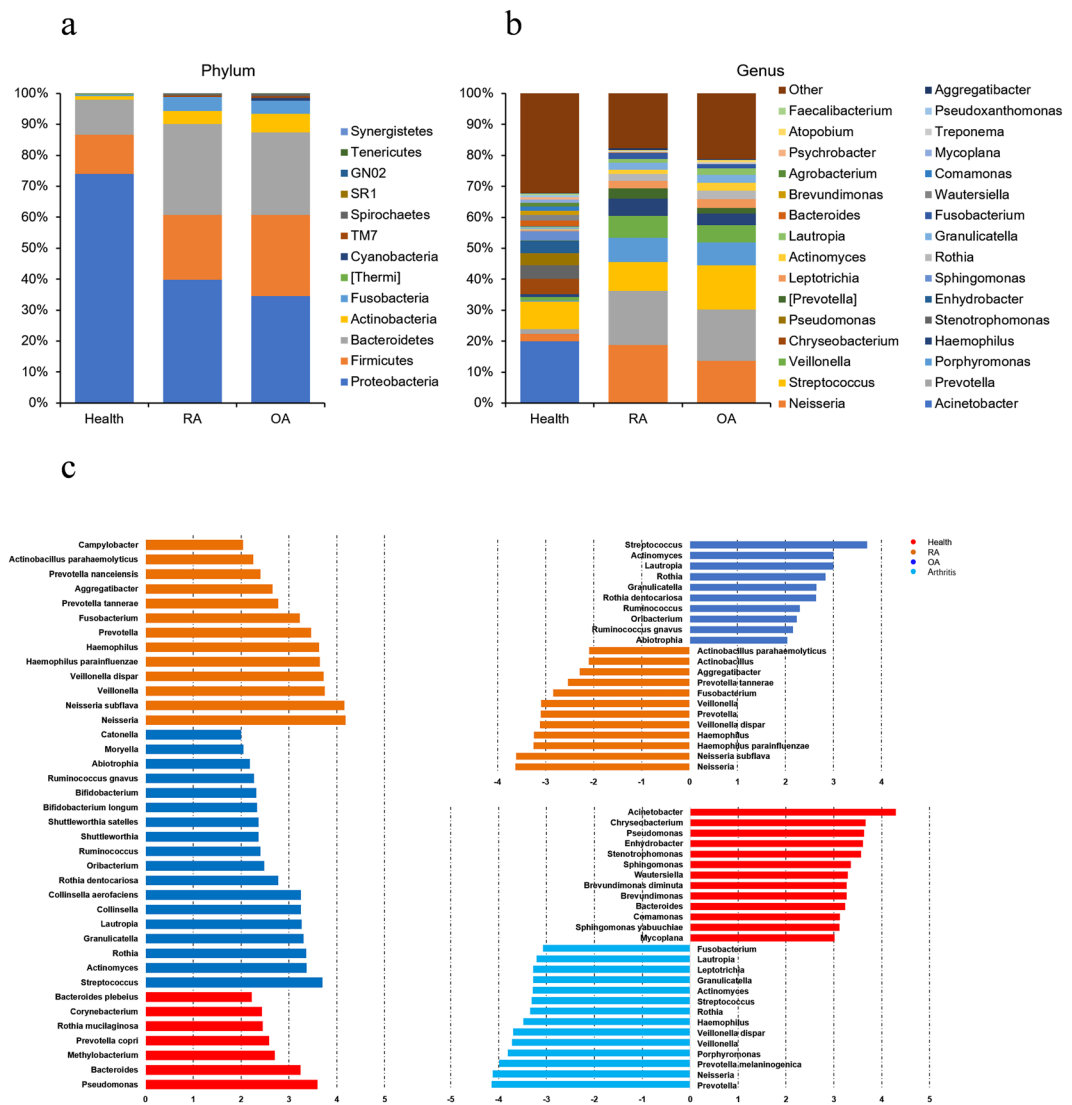


Figure 2. Taxonomic profiles and biomarkers of patients with RA, OA and healthy subjects. **(a)** Barplots of taxonomic profiles of patients with RA, OA and healthy subjects at the Phylum level. **(b)** Barplots of taxonomic profiles of patients with RA, OA and healthy subjects at the genus level. **(c)** Histogram of the LDA scores, where the LDA score indicates the effective size and ranking of each differentially abundant taxon (LDA > 2).

diversity between the microbiome profiles of these three groups, we conducted a Principal Coordinate Analysis (PCoA) of weighted and unweighted UniFrac distances. Beta diversity comparisons using Analysis of similarities (ANOSIM) for oral data suggested a minor, but statistically significant community difference between RA and OA (ANOSIM, $R = 0.07$, $P = 0.001$) (Fig. 1c,d). The data in healthy subjects suggested significant differences from that of RA and OA. The microbial composition of RA, OA and healthy subjects did differ at the phylum and genus levels (Fig. 2a,b). The most common phyla were Proteobacteria, Firmicutes, Bacteroidetes, Actinobacteria and Fusobacteria. The relative abundance of Proteobacteria in healthy subjects was significantly higher than in patients with RA and OA ($P < 0.001$, Welch's t-test), and the relative abundance of Firmicutes ($P < 0.029$, Welch's t-test) in patients with OA is significantly higher than those in patients with RA. To further evaluate microbiome differences among patients with RA, OA and healthy subjects, the Linear discriminant analysis (LDA) Effect Size (LEfSe) method was used to discover the difference in microbiota composition and identify significant biomarkers, thus revealing 24 genera and 14 species (LDA > 2) (Fig. 2c). Many of the taxa were differentially abundant in RA and OA and contained: *Neisseria subflava*, *Haemophilus parainfluenzae*, *Veillonella dispar*, *Prevotella tannerae*, *Actinobacillus paraahaemolyticus*, *Neisseria*, *Haemophilus*, *Prevotella*, *Veillonella*, *Fusobacterium*, *Aggregatibacter*, and *Actinobacillus*, which were more in RA, while *Rothia dentocariosa*, *Ruminococcus gnavus*, *Streptococcus*, *Actinomyces*, *Lautropia*, *Rothia*, *Granulicatella*, *Ruminococcus*, *Oribacterium*, and *Abiotrophia* numbers increased in OA (Fig. 2c). We also compared healthy subjects with the arthritis group (RA + OA), revealing potential opportunistic oral pathogens. *Prevotella melaninogenica*, *Veillonella dispar*, *Prevotella*, *Neisseria*, *Porphyromonas*, *Veillonella*, *Haemophilus*, *Rothia*, *Streptococcus*, *Actinomyces*, *Granulicatella*, *Leptotrichia*, *Lautropia*, and *Fusobacterium* were increased in the arthritis group (Fig. 2c). Since gender and age varied significantly among the

groups, it may confound the identified associations. We tested the age and gender effects on the oral microbiome using Spearman's rank correlation test (Supplemental Table S2). The results indicate that oral microbiome has no significant correlation with gender. There are only 5 taxonomies affected by the factor of age, however, they don't belong to the identified different abundant taxonomies among RA, OA and healthy subjects.

Imputed functions of Oral Microbiota in RA and OA. Given the structural changes within the oral microbiota in patients with RA and OA, we explored microbiota functions based on inferred metagenomes using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) algorithm. PICRUSt is a computational approach in which evolutionary modeling is used to predict the present gene families from 16S rRNA gene sequencing data and a reference genome database. In all, 6,909 predicted genes were assigned to 301 KEGG, level 3 modules. Forty-seven pathways were significantly different between RA and OA ($P < 0.05$; Welch's t-test, Supplemental Table S3), of which 28 KEGG pathways were related to metabolism. The differences between 22 pathways were extremely significant between RA and OA ($P < 0.01$; Welch's t-test, Fig. 3a). The most significantly different KEGG pathways included lipopolysaccharide biosynthesis (Fig. 2b), lipopolysaccharide biosynthesis proteins (Fig. 3c) and glycolysis/gluconeogenesis (Fig. 3d). For KEGG hierarchical level 1 functional modules, metabolism was the most abundant module, then genetic information processing and environmental information processing (Supplemental Fig. S1). The signaling and environmental information processes were significantly different between RA and OA ($P < 0.05$, Welch's t-test). For KEGG hierarchical level 3 functional modules, transporters were the most abundant module, followed by ABC transporters and DNA repair, then recombination proteins (Supplemental Fig. S2).

Potential biomarkers for distinguishing RA and OA. To evaluate the potential value of identified bacterial biomarkers for clinical differentiation of RA and OA, we constructed the receiver operating characteristic (ROC) curve and computed the area under the curve (AUC) values. We chose 10 OTUs with an LDA value greater than 3.0 in the LefSe analysis, as candidate biomarkers (*Streptococcus*, *Actinomyces*, *Lautropia*, *Neisseria*, *Neisseria subflava*, *Haemophilus parainfluenzae*, *Haemophilus*, *Veillonella dispar*, *Prevotella* and *Veillonella*). To verify the best prediction model, we constructed the ROC curve for each arrangement combination in these 10 candidate biomarkers, as the prediction model. The training and validation sets were randomly selected 5 times and the average AUC was calculated (Supplemental Table S4). The average AUCs of 3 biomarker models were greater than 0.85 (Fig. 4). Moreover, the biomarker model with 8 OTUs (*Actinomyces*, *Neisseria*, *Neisseria subflava*, *Haemophilus parainfluenzae*, *Haemophilus*, *Veillonella dispar*, *Prevotella* and *Veillonella*) generated the highest AUC value of 0.8756.

Discussions

We compared the oral microbiota in patients with RA, OA and healthy subjects. We showed that profiling the bacteria associated with the oral cavity may have value in the detection of RA and OA. We also found that patients with RA and OA had an oral microbiota with higher microbial diversity compared to healthy subjects, indicating that there could be more harmful bacteria or opportunistic pathogens in the oral cavity of patients with RA and OA. Interestingly, when compared to healthy subjects, the following potential opportunistic oral pathogens increased in the arthritis group (RA + OA): *Prevotella melaninogenica*, *Veillonella dispar*, *Prevotella*, *Neisseria*, *Porphyromonas*, *Veillonella*, *Haemophilus*, *Rothia*, *Streptococcus*, *Actinomyces*, *Granulicatella*, *Leptotrichia*, *Lautropia*, and *Fusobacterium*. Most of them are Gram-negative anaerobic bacteria. In keeping with this, *Porphyromonas*, *Prevotella melaninogenica*, *Actinomyces* and *Streptococcus* are considered to be directly responsible for periodontitis¹⁶. Previous studies also reported that they were associated with RA^{1,17,18}. *Porphyromonas gingivalis* has been demonstrated to be directly linked to RA through citrullination and induction of anti-peptidyl citrulline antibodies reacting to citrullinated human self-proteins¹⁹. *P. melaninogenica* causes a deficiency of terminal galactose residues by binding to the Fc region of the IgG molecule and metabolizing galactose with its enzymes²⁰. A lack of terminal galactose residues early in the course of RA is associated with a worsening prognosis²¹. In patients with chronic inflammatory diseases, oral microbiome dysbiosis is associated with the disease. Previous studies have found that alpha diversity of oral microbiome was significantly decreased in Sjögren syndrome²² and Behçet's disease¹¹, but our study found an increased alpha diversity in the patients, which could be due to the difference of disease. Consistent to these studies, our study also found an increase in the abundance of *Streptococcus* and *Haemophilus* in the diseased subjects.

Very few studies have been conducted on oral microbiome of patients with RA and OA using next-generation sequencing (NGS). Mikuls *et al.*²³ compared the subgingival microbiome of RA and OA patients. Compared to OA, 10 taxa were lower in numbers in patients with RA, and this study failed to identify a biomarker that could reliably distinguish RA from OA. The result from our study contrasted with the report from Mikuls *et al.*, and only *Streptococcus* was consistent. We found that *Prevotella* was more abundant in patients with RA instead of being lower. Scher *et al.*¹⁴ also revealed that *Prevotella* was found in greater abundance in patients with RA. Differences in the findings across studies may relate to the oral sampling site. Our study collected saliva samples instead of subgingival plaque like in Mikuls *et al.* The research indicated that the microbiome between supragingival and salivary communities displayed strong differences^{24,25}. We also used machine learning to conduct a classification analysis for selected biomarkers. We succeeded in identifying biomarkers that could reliably distinguish RA from OA in patients.

The translocation of microbiota-derived molecules into the systemic circulation is one route for the microbiome to mediate osteoarthritis. A recent study reported that systemic and synovial concentrations of bacterial Lipopolysaccharide (LPS) positively correlated with joint inflammatory response and the severity of joint degradation²⁶. LPS also plays a role in RA and can upregulate all of the cytokines²⁷. Furthermore, bacteria-derived LPS by itself is also capable of potentiating collagen-induced arthritis (CIA) in mice. In a nutshell, oral bacteria or

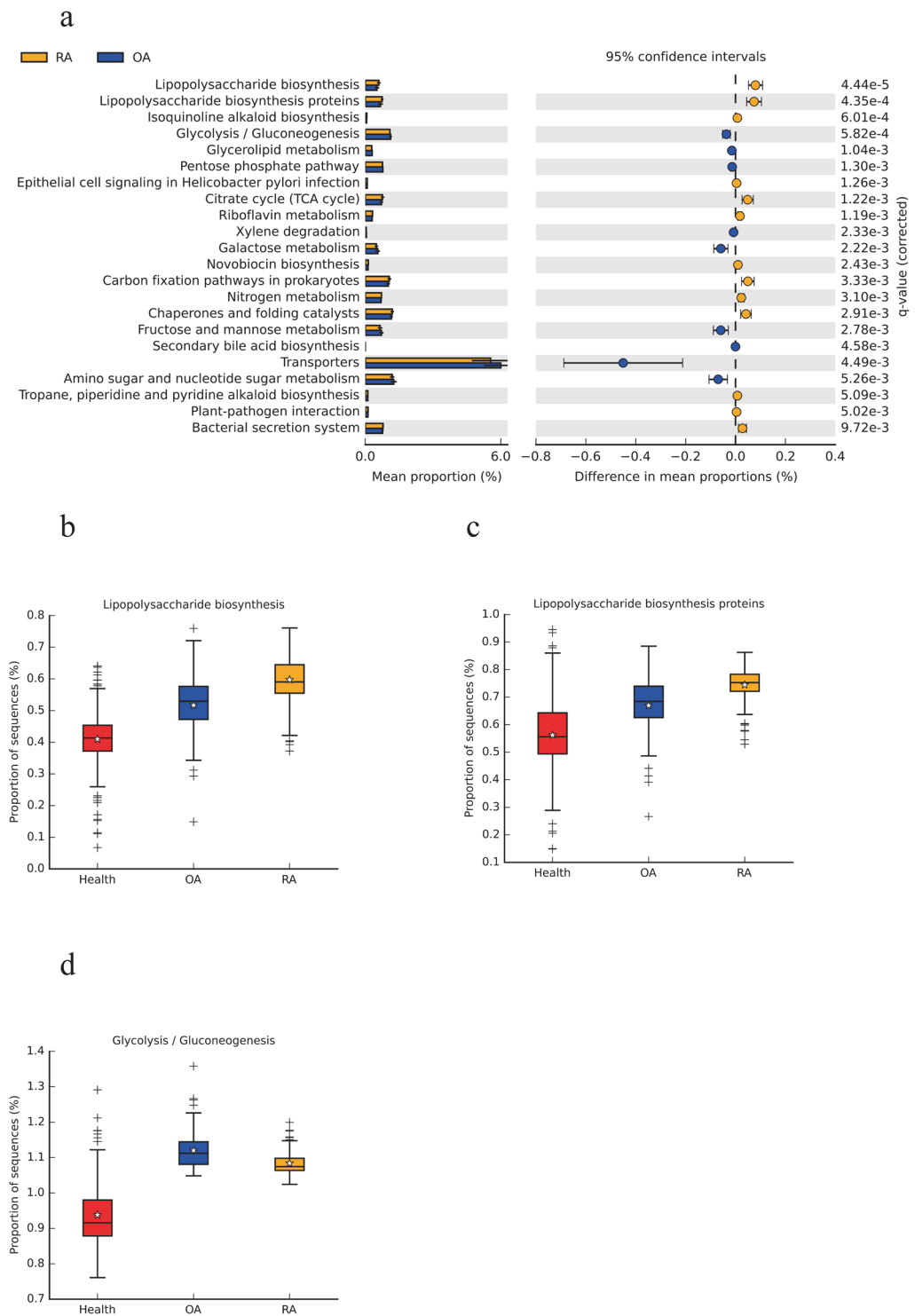


Figure 3. Functional analysis of oral microbiota in patients with RA, OA and healthy subjects. **(a)** Extended error barplot with 95% confidence intervals showing significantly different KEGG pathways between RA and OA. Corrected P-values are calculated using Benjamini-Hochberg FDR approach **(b)** Lipopolysaccharide biosynthesis **(c)** Lipopolysaccharide biosynthesis proteins **(d)** Glycolysis/Gluconeogenesis; *indicates the mean of the data, the data points outside of the whiskers are shown as crosses+.

bacteria-derived LPS could invade blood and be transported to joints, where they could stimulate innate immune receptors in bone, cartilage and synovium. Lyu²⁸ *et al.* detected the key drivers in oral microbial community related to dysbiosis in RA. They found that the development of biofilms is one of the drivers of persistent infections. LPS was found to play a role in biofilm formation²⁹, and previous study also found that *P. gingivalis* can grow in the biofilm, which can become destructive and may contribute to RA³⁰.

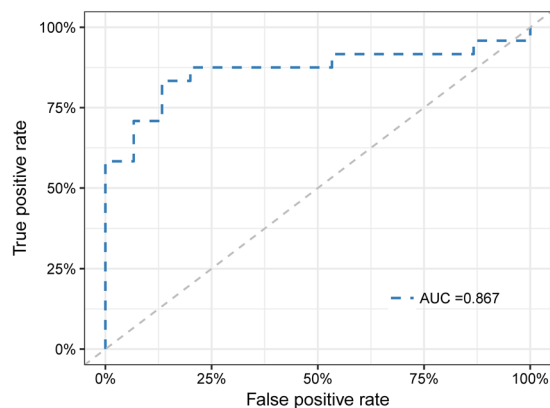


Figure 4. The ROC curve based on 8 most-distinctive OTUs. The maximum AUC value is then selected and drawn.

| | Healthy Subjects (n = 155) | RA (n = 110) | OA (n = 67) | P-values |
|---------------------|----------------------------|-------------------|-------------------|-------------|
| Age (mean \pm SD) | 49.96 \pm 11.17 | 56.65 \pm 11.36 | 57.79 \pm 9.712 | $P < 0.01$ |
| Gender (M/F) | 80/75 | 20/90 | 21/46 | $P < 0.001$ |

Table 1. Characteristics of all subjects.

Despite significant advances in understanding pathophysiology of RA and OA, early diagnosis and therapeutic intervention remain a challenge³¹. Rheumatoid factor (RF) and cyclic citrullinated peptides (CCPs) have usually been used as the biomarkers for RA, but have relatively low sensitivity³² and low specificity³³ for early RA. Magnetic resonance imaging (MRI) techniques have been developed for early-stage evaluation of cartilage damage in OA³⁴, however, MRI is expensive and not easily/widely accessible. There are a host of advantages in using oral microbiota as a source of biomarkers for predicting and screening disease. Saliva is a non-invasive collection method that does not cause discomfort and pain to subjects. Therefore, it is widely used in the study of biomarkers in diseases. We profiled the oral microbiome and discovered 8 biomarkers that offered an alternative strategy for the diagnosis of RA and OA. Further studies to validate these potential biomarkers using a larger clinical cohort are warranted.

Conclusions

We explored the differences in oral microbiome profiles of patients with RA, OA and healthy subjects using next-generation sequencing. Most importantly, we compared the oral microbiome and their potential functions in RA, OA and healthy subjects. Finally, 8 bacterial biomarkers were selected in the prediction model for distinguishing RA and OA in a non-invasive manner. This can be subjected to further clinical validation to develop new strategies for diagnosing RA and OA. Moreover, we need to further explore how bacteria can affect RA and OA, to help prevent and treat RA and OA.

Methods

Patients and Samples. This study was approved by the Institutional Review Board of Shandong Normal University and Shandong Academy of Medical Sciences. Written informed consents and questionnaire data sheets were obtained from all subjects who visited the Department of Orthopedics, Qianfoshan Hospital Affiliated to Shandong University, and Guoyitang Hospital. They agreed to serve as sample donors, in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki). All methods and experimental protocols in this study were performed in accordance with relevant guidelines and standard operating procedures. Saliva samples were collected using standard methods³⁵: 5–10 mL of unstimulated whole saliva was collected in a 50 mL sterile tube from each subject, between 9–10 a.m. (collection time not more than 30 minutes). All subjects were asked to refrain from drinking, eating and hygiene-related procedures for at least 2 hours prior to collection, and all patients with apparent oral problems (periodontal disease *et al.*) were excluded. Samples were immediately frozen and stored at -80°C . In all, 110 patients with RA who fulfilled the 2010 revised ACR/EULAR classification criteria were included³⁶. 67 patients with OA met the 1995 revised ACR classification criteria³⁷, and 155 healthy subjects served as controls (Table 1). The full list of sample information is available in the supplementary data (Table S1).

DNA Extraction and 16S rRNA Gene Sequencing. For each sample, 1.5 ml of saliva was used for DNA extraction. The saliva was centrifuged at $12,000 \times g$ for 15 min at 4°C , and then the precipitation was used for DNA extraction (Qiagen DNeasy Blood & Tissue Kit). Quantitation of DNA was measured by Nanodrop 2000 (Thermo Scientific). To generate 16S rRNA gene amplicons, in a 50 μl reaction, typically 50 ng of DNA was used as a template, with 0.4 μM of V1–V2 barcoded primers, targeting 27 F and 355 R of the bacterial 16S rRNA

gene (5' AGAGTTTGATCMTGGCTCAG3' and 5' GCTGCCTCCCGTAGGAGT3'). DNA was purified with QIAquick PCR Purification Kit (Qiagen) and PCR purification procedure. All amplicons were quantified and pooled to equalize concentrations for sequencing, using HiSeq 2500 (Illumina).

16S rRNA Gene Sequence Analysis. The 16S rRNA gene sequence paired-end data set was joined and quality filtered using the FLASH method, described by Magoč and Salzberg³⁸. All sequence analysis was provided in the Quantitative Insights Into Microbial Ecology (QIIME, version 1.9.1) software suite³⁹, according to the QIIME tutorial (<http://qiime.org/>) with some modifications. Chimeric sequences were removed using usearch61⁴⁰ with *de novo* models. Sequences were clustered against the 2013 Greengenes (13_8 release) ribosomal database's 97% reference data set (<http://greengenes.secondgenome.com/downloads>). Sequences that did not match any entries in this reference were subsequently clustered into *de novo* OTUs at 97% similarity with UCLUST. A taxonomy was assigned to all OTUs using RDP classifier within QIIME and the Greengenes reference data set. Rarefaction and rank abundance curves were calculated from OTU tables using alpha diversity and rank abundance scripts within the QIIME pipeline. The hierarchical clustering based on population profiles of most common and abundant taxa was performed using UPGMA clustering (Unweighted Pair Group Method with Arithmetic mean, also known as average linkage), on the distance matrix of OTU abundance. This resulted in a Newick-formatted tree, which was obtained using the QIIME package.

Statistical Analysis. To account for any bias caused by uneven sequencing depth, the least number of sequences present in any given sample from a sample category was selected randomly. Prior to calculating community-wide dissimilarity measures (α -diversity and β -diversity), we rarefied the OTU table to a sequencing depth of 40000 per sample for both diversity analyses. All PCoA were based on weighted UniFrac distances using evenly sampled OTU abundances. The prediction of the functional composition of a metagenome, using marker gene data and a database of reference genomes was done with PICRUSt as described by Langille *et al.*⁴¹. The graphical representation of the results was done with STAMP and the calculation of *P* values was done using Welch's t-test (*P* values were corrected for Benjamini-Hochberg FDR).

The ROC curves. The Support Vector Machine (SVM) classifier from R package e1071 was adopted for classification analysis of the selected biomarkers. Five-fold cross-validation was used to evaluate the performance of the prediction model. The ROC curves as well as the AUC value was calculated using the ROCR R package.

Data Availability

The data sets generated during and/or analyzed during the current study are available with the corresponding author, on reasonable request.

References

- Ogrendik, M. Rheumatoid arthritis is linked to oral bacteria: etiological association. *Mod Rheumatol* **19**, 453–456, <https://doi.org/10.1007/s10165-009-0194-9> (2009).
- Krishnan, K., Chen, T. & Paster, B. J. A practical guide to the oral microbiome and its relation to health and disease. *Oral Dis* **23**, 276–286, <https://doi.org/10.1111/odi.12509> (2017).
- Seymour, G. J., Ford, P. J., Cullinan, M. P., Leishman, S. & Yamazaki, K. Relationship between periodontal infections and systemic disease. *Clin Microbiol Infect* **13**(Suppl 4), 3–10, <https://doi.org/10.1111/j.1469-0691.2007.01798.x> (2007).
- Mawardi, H. H., Elbadawi, L. S. & Sonis, S. T. Current understanding of the relationship between periodontal and systemic diseases. *Saudi Med J* **36**, 150–158, <https://doi.org/10.15537/smj.2015.2.9424> (2015).
- Igari, K., Kudo, T., Toyofuku, T., Inoue, Y. & Iwai, T. Association between periodontitis and the development of systemic diseases. *Oral Biology & Dentistry* **2**, 4 (2014).
- Bracci, P. M. Oral Health and the Oral Microbiome in Pancreatic Cancer: An Overview of Epidemiological Studies. *Cancer J* **23**, 310–314, <https://doi.org/10.1097/PP0.0000000000000287> (2017).
- Pereira, P. A. B. *et al.* Oral and nasal microbiota in Parkinson's disease. *Parkinsonism Relat Disord* **38**, 61–67, <https://doi.org/10.1016/j.parkreldis.2017.02.026> (2017).
- Correa, J. D. *et al.* Subgingival microbiota dysbiosis in systemic lupus erythematosus: association with periodontal status. *Microbiome* **5**, 34, <https://doi.org/10.1186/s40168-017-0252-z> (2017).
- Segal, L. N. *et al.* Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat Microbiol* **1**, 16031, <https://doi.org/10.1038/nmicrobiol.2016.31> (2016).
- Casarin, R. C. *et al.* Subgingival biodiversity in subjects with uncontrolled type-2 diabetes and chronic periodontitis. *J Periodontal Res* **48**, 30–36, <https://doi.org/10.1111/j.1600-0765.2012.01498.x> (2013).
- Coit, P. *et al.* Sequencing of 16S rRNA reveals a distinct salivary microbiome signature in Behcet's disease. *Clin Immunol* **169**, 28–35, <https://doi.org/10.1016/j.clim.2016.06.002> (2016).
- Han, Y. W. *et al.* Interactions between periodontal bacteria and human oral epithelial cells: *Fusobacterium nucleatum* adheres to and invades epithelial cells. *Infect Immun* **68**, 3140–3146 (2000).
- Fish, E. W. The Teeth as a Source of Focal Sepsis. *Postgrad Med J* **16**, 26–33 (1940).
- Scher, J. U. *et al.* Periodontal disease and the oral microbiota in new-onset rheumatoid arthritis. *Arthritis Rheum* **64**, 3083–3094, <https://doi.org/10.1002/art.34539> (2012).
- Zhang, X. *et al.* The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat Med* **21**, 895–905, <https://doi.org/10.1038/nm.3914> (2015).
- Tatakis, D. N. & Kumar, P. S. Etiology and pathogenesis of periodontal diseases. *Dent Clin North Am* **49**(v), 491–516, <https://doi.org/10.1016/j.cden.2005.03.001> (2005).
- Koga, T. *et al.* Acute joint inflammation in mice after systemic injection of the cell wall, its peptidoglycan, and chemically defined peptidoglycan subunits from various bacteria. *Infect Immun* **50**, 27–34 (1985).
- Gomez, A. *et al.* Loss of sex and age driven differences in the gut microbiome characterize arthritis-susceptible 0401 mice but not arthritis-resistant 0402 mice. *PLoS one* **7**, e36095, <https://doi.org/10.1371/journal.pone.0036095> (2012).
- Rutger Persson, G. Rheumatoid arthritis and periodontitis - inflammatory and infectious connections. Review of the literature. *J Oral Microbiol* **4**, <https://doi.org/10.3402/jom.v4i0.11829> (2012).
- Haraldsson, G., Meurman, J. H., Kononen, E. & Holbrook, W. P. Properties of hemagglutination by *Prevotella melaninogenica*. *Anaerobe* **11**, 285–289, <https://doi.org/10.1016/j.anaerobe.2005.02.004> (2005).

21. Lacki, J. K., Porawska, W., Mackiewicz, U., Mackiewicz, S. & Muller, W. Changes in agalactosyl IgG levels correlate with radiological progression in early rheumatoid arthritis. *Ann Med* **28**, 265–269 (1996).
22. de Paiva, C. S. *et al.* Altered Mucosal Microbiome Diversity and Disease Severity in Sjogren Syndrome. *Sci Rep* **6**, 23561, <https://doi.org/10.1038/srep23561> (2016).
23. Mikuls, T. R. *et al.* The subgingival microbiome in patients with established rheumatoid arthritis. *Rheumatology (Oxford)*, <https://doi.org/10.1093/rheumatology/key052> (2018).
24. Hall, M. W. *et al.* Inter-personal diversity and temporal dynamics of dental, tongue, and salivary microbiota in the healthy oral cavity. *NPJ Biofilms Microbiomes* **3**, 2, <https://doi.org/10.1038/s41522-016-0011-0> (2017).
25. Segata, N. *et al.* Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome Biol* **13**, R42, <https://doi.org/10.1186/gb-2012-13-6-r42> (2012).
26. Huang, Z. Y., Stabler, T., Pei, F. X. & Kraus, V. B. Both systemic and local lipopolysaccharide (LPS) burden are associated with knee OA severity and inflammation. *Osteoarthritis Cartilage* **24**, 1769–1775, <https://doi.org/10.1016/j.joca.2016.05.008> (2016).
27. Pretorius, E., Akeredolu, O. O., Soma, P. & Kell, D. B. Major involvement of bacterial components in rheumatoid arthritis and its accompanying oxidative stress, systemic inflammation and hypercoagulability. *Exp Biol Med (Maywood)* **242**, 355–373, <https://doi.org/10.1177/1535370216681549> (2017).
28. Lyu, M. *et al.* KDiament: a package for detecting key drivers in a molecular ecological network of disease. *BMC Syst Biol* **12**, 5, <https://doi.org/10.1186/s12918-018-0531-8> (2018).
29. Van Houdt, R. & Michiels, C. W. Biofilm formation and the food industry, a focus on the bacterial outer surface. *J Appl Microbiol* **109**, 1117–1131, <https://doi.org/10.1111/j.1365-2672.2010.04756.x> (2010).
30. Marcinkiewicz, J., Strus, M. & Pasich, E. Antibiotic resistance: a “dark side” of biofilm-associated chronic infections. *Pol Arch Med Wewn* **123**, 309–313 (2013).
31. da Mota, L. M., dos Santos Neto, L. L. & de Carvalho, J. F. Autoantibodies and other serological markers in rheumatoid arthritis: predictors of disease activity? *Clin Rheumatol* **28**, 1127–1134, <https://doi.org/10.1007/s10067-009-1223-y> (2009).
32. van Venrooij, W. J., van Beers, J. J. & Pruijn, G. J. Anti-CCP antibodies: the past, the present and the future. *Nat Rev Rheumatol* **7**, 391–398, <https://doi.org/10.1038/nrrheum.2011.76> (2011).
33. van Boekel, M. A., Vossenaar, E. R., van den Hoogen, F. H. & van Venrooij, W. J. Autoantibody systems in rheumatoid arthritis: specificity, sensitivity and diagnostic value. *Arthritis Res* **4**, 87–93, <https://doi.org/10.1186/ar395> (2002).
34. Rousseau, J. & Garnero, P. Biological markers in osteoarthritis. *Bone* **51**, 265–277, <https://doi.org/10.1016/j.bone.2012.04.001> (2012).
35. Zhang, L. *et al.* Salivary transcriptomic biomarkers for detection of resectable pancreatic cancer. *Gastroenterology* **138**(949–957), e941–947, <https://doi.org/10.1053/j.gastro.2009.11.010> (2010).
36. Aletaha, D. *et al.* Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* **62**(2569–2581), 2010, <https://doi.org/10.1002/art.27584> (2010).
37. Hochberg, M. C. *et al.* Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology. *Arthritis Rheum* **38**, 1541–1546 (1995).
38. Magoc, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963, <https://doi.org/10.1093/bioinformatics/btr507> (2011).
39. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335–336, <https://doi.org/10.1038/nmeth.f.303> (2010).
40. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461, <https://doi.org/10.1093/bioinformatics/btq461> (2010).
41. Langille, M. G. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**, 814–821, <https://doi.org/10.1038/nbt.2676> (2013).

Acknowledgements

We thank all participants of this study, which was supported by the National Natural Science Foundation of China under contract No. 31471202 (Lei Zhang); the Shandong Provincial Key Research and Development Program under contract No. 2016YYSP009 (Lei Zhang); City of Weihai technique extension project under contract No. 2016GNS023 (Lei Zhang); Qingdao Key Research project No. 17–3–3–10-nsh (Chunling Zhang); the Innovation Project of Shandong Academy of Medical Sciences and the Natural Science Foundation of Shandong Province under contract No. ZR2014YL040 (to Yan Zhao). Lei Zhang is also supported by the Taishan Scholars Program of Shandong Province (No. tshw20120206).

Author Contributions

L.Z. and J.Y. designed the study. B.C., Y.Z. and H.W. did measurements and data analysis. S.L., L.Y., T.W., B.S., Z.G., X.H. and C.Z. obtained samples and clinical details. B.C. and L.Z. wrote the manuscript. All authors have read and critically revised the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-35473-6>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018