Article:
Lopez-Oliva, I, Paropkari, AD, Saraswat, S et al. (10 more authors) (2018) Dysbiotic Subgingival Microbial Communities in Periodontally Healthy Patients With Rheumatoid Arthritis. Arthritis and Rheumatology, 70 (7). pp. 1008-1013. ISSN 2326-5191

https://doi.org/10.1002/art.40485

SUPPLEMENTARY INFORMATION

Ethical approval

Approval for this study was obtained from the South Birmingham Research Ethics Committee, UK (Approval codes: Inspired - 15/WM/0006, Opera - 11/WM/0235) and the study was conducted in accordance with the approved guidelines.

Recruitment of RA study population

Patients diagnosed with Rheumatoid Arthritis from three rheumatology departments ((University Hospitals Birmingham NHS Foundation Trust - Queen Elizabeth Hospital; Sandwell & West Birmingham Hospitals NHS Trust - Birmingham City Hospital, and Heart of England NHS Foundation Trust – Heartlands Hospital) from September 2014 to October 2015, in Birmingham, United Kingdom. All participants had a screening appointment at the Birmingham Dental Hospital, as part of the OPERA feasibility study (Outcomes of Periodontal Therapy in Rheumatoid Arthritis). During this visit, RA and periodontal status was assessed and biological samples were collected by a trained dentist. From 123 RA patients screened, 22 were considered for the study as periodontally healthy.

RA inclusion/exclusion criteria

Inclusion criteria: Patients with rheumatoid arthritis were classified according to the revised 1987 ACR criteria for RA; Disease Activity Score (DAS) 28 ≥3.2 (DAS28 score >5.1 only if patient on biologics or patient unwilling to take biologics); treatment with Disease Modifying Anti-Rheumatic Drugs (DMARDs) for ≥ 3 months and stable dose for ≥ 2 months.

Exclusion criteria: Rheumatic autoimmune disease other than RA, or significant systemic involvement secondary to RA; History of, or current, inflammatory joint disease other than RA; Diagnosis of juvenile idiopathic arthritis (JIA) or juvenile rheumatoid arthritis (JRA) and/or RA before age 16; a~~A~~ny surgical procedure or antibiotic exposure within 12 weeks prior to baseline.

Recruitment of systemically healthy study population

A total of 20 systemically healthy patients and periodontally healthy (NoRA group) where recruited as part of the INSPIRED feasibility study (Influence of Successful Periodontal Intervention on Renal and Vascular Systems in patients with Chronic Kidney Disease) from the oral surgery department of Birmingham dental hospital as well as staff, non-dentist/hygienist, from within the hospital (Approval codes: Inspired - 15/WM/0006). Exclusion criteria were any self-reported systemic illness including hypertension or diabetes.

Periodontal health criteria

Periodontal health for both the RA and non-RA groups was defined as a maximum 4 sites with 4mm probing depth and no probing depth of 5 mm or above, CAL <2mm on all teeth, BoP <20% (excluding wisdom teeth and distal of the second molars). Charting was performed using a UNC-15 probe on 4 interproximal sites per tooth in all the quadrants. Also, patients who had undergone periodontal treatment within 12 months prior to baseline were excluded for the study.

Sample collection

During the screening visit, a trained dentist investigated the patient general health and rheumatologic status such as the patient's height, weight and blood pressure along with a medical history questionnaire. Biological samples of blood, gingival crevicular fluid (GCF) and saliva were collected, followed by a periodontal pocket charting.

Subgingival plaque samples were collected, after removal of supragingival plaque with a cotton pellet, from the 6 sites (one site per sextant to include a molar, a premolar and an anterior tooth in each quadrant, ~~as~~ each tooth as distant as possible from ~~each~~ the other). In all subjects, the same teeth were sampled, as far as possible. In case of a missing tooth, a similar representative tooth was used (for example, a second premolar was sampled if the first was missing, and a canine was sampled if the central incisor was missing). The method of collection was using an endodontic paper point and the appropriate Gracey curette. Plaque samples from each patient were pooled together and stored in a cryotube containing Tris buffer and stored in a -80℃ freezer.

DNA isolation

Paper points were removed from RNA*later*, added to 180µl of phosphate buffered saline and agitated for 45 minutes; following which the supernatant was removed and used for analysis. Bacterial DNA was isolated using a Qiagen DNA MiniAmp kit (Qiagen, Valencia, CA, USA) according to instructions.

16s sequencing:

Two regions of the 16S rRNA genes were sequenced: V1–V3 (spanning *E.coli* 16S gene regions 8-27 and 519-536) and V7–V9 (spanning *E.coli* 16S gene regions 1099-1114 and 1528-1541). The primers used for sequencing have been previously described(1). The 16s amplicons were quantified using the Quant-iT PicoGreen dsDNA reagent and kit (Invitrogen). Equimolar concentrations of each amplicon were pooled and sequenced on the HiSeq 2500 system (Illumina). Two primers were used, since each primer is capable of detecting a range of genera that the other fails to recover. Together they allow the recovery of a wider range of the microbiome than is possible with a single primer alone. However, some genera are picked up by both primers. Thus, to prevent overcounting, the number of sequences assigned to an OTU by both primers was reduced by half. Primer averaging was carried out as previously described(1) using the implementation in the PhyloTOAST software suite(2). Analyses were conducted using the QIIME(3) and PhyloToAST.

Sequence analysis

The sequences were binned by sample and denoised using denoise_wrapper.py to reduce sequencing errors. All denoised sequences were aggregated and *de novo* operational taxonomic units (OTUs) were identified. Sequences were clustered into distinct OTUs at 97% similarity using the UCLUST[65] method. Chimeric sequences were depleted using ChimeraSlayer (v. 1.9.0, identify_chimeric_seqs.py)[66]. Sequences with an average quality score of 30 over a sliding window of 50bp and length >200 bp were assigned a taxonomic identity by alignment to the HOMD database(4) using the Blastn algorithm at 97% identity.

An OTU was included in the analysis if it was present in more than 5% of samples, at an abundance of at least 0.001%.

Alpha (within-group) and beta (between-group) diversity were computed. Since emergent evidence does not support rarefying the microbiome to compensate for sequencing effort(5), we used linear regression models to correct for sequencing depth (*phyloseq* in R). Shannon and ACE were used as estimators of alpha diversity. Both phylogenetic (UniFrac) and non-phylogenetic (Bray–Curtis, Jaccard) distance matrices were utilized to estimate beta diversity. Principal component analysis (PCoA) was performed on distance matrices, and significance of clustering was interrogated using Adonis with 999 permutations. PCoA plots and confidence ellipses were generated by the R package ggplot. Phylogenetic trees were created with iTOL (http://itol.embl.de/, version 3.4.1).

Probable gram staining characteristics and oxygen requirements were attributed to uncultivated species based on phylogenetic relatedness to the closest cultivated species.

Core species were identified using Qime's script (core_microbiome.py) when species were present in at least 80% of the patients in each group.

PCoA analysis of unifrac distances was performed to show group-wise clustering. Beta diversity was measured with Adonis and ANOSIM tests to estimate statistical difference between groups. Alpha diversity was analysed using Abundance Coverage Estimator (ACE), and differences between alpha diversities group-wise was measured using Wilcoxon test.

The Bioconductor package for R, *DESeq2,* was used to perform differential expression analysis of the annotated microbial transcripts(6). This function uses a negative binomial distribution of raw counts to estimate between-group differences, while accounting for sampling effort (library size) and dispersion of each category (taxon or functional gene). p-values were adjusted for multiple testing (FDR < 0.1, FDR-adjusted Wald Test).

Network correlations were determined by significant pairwise using Spearsman's correlation ($p<0.05$) and network graphs were calculated in Python (Networkx package) and visualized in Gephi. Network anchor OTUs were defined as significantly different in abundance between the 2 groups, high betweenness centrality (top 20%) and belonging to the core microbiome of each group.

To explore if the tight co-occurrence cluster in patients with Rheumatoid Arthritis could be attributed to common metabolic pathways, combinatorial analysis was conducted. PiCRUSt was used to estimate the functions encoded by the species involved in the cluster. The Ccore functional orthogs (present in ≥ 80% of Rheumatoid Arthritis patient cohort) were obtained using get_core_ids.py script(https://github.com/akshayparopkari/kadambari/blob/master/python/get_core_ids.py). FDR-corrected

significant and overlapping pairwise Spearman's ρ and Kendall-τ correlations were used to generate correlation matrix which was imported into Networkx[33] to create the graph structures, and Gephi[34] to visualize and label the network graphs ((using correlation.py script - https://github.com/akshayparopkari/kadambari/blob/master/python/correlation.py). A correlation value of ≥0.75 and significant p value <0.05 were used for visualization purposes, yielding 169 nodes and 5901 edges for overall KEGG pathway data and 115nodes and 2561 edges for metabolism related gene pathways only.

1. Kumar PS, Brooker MR, Dowd SE, Camerlengo T. Target region selection is a critical determinant of community fingerprints generated by 16S pyrosequencing. PLoS One. 2011;6(6):e20956.
2. Dabdoub SM, Fellows ML, Paropkari AD, Mason MR, Huja SS, Tsigarida AA, et al. PhyloToAST: Bioinformatics tools for species-level analysis and visualization of complex microbial datasets. Scientific reports. 2016;6:29123.
3. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335-6.
4. Chen T, Yu WH, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. Database : the journal of biological databases and curation. 2010;2010:baq013.
5. McMurdie PJ, Holmes S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLOS Computational Biology. 2014;10(4):e1003531.
6. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.