Distinguishing the Signals of Gingivitis and Periodontitis in Supragingival Plaque: A Cross-Sectional Cohort Study in Malawi

Authors:

Liam Shaw[1,2,*], Ulla Harjunmaa[3], Ronan Doyle[1], Simeon Mulewa[4], Davie Charlie[4], Ken Maleta[4], Robin Callard[1], A. Sarah Walker[5], Francois Balloux[6], Per Ashorn[3,7], Nigel Klein[1]

Affiliations:

1. Institute for Child Health, UCL, London, UK.
2. Centre for Mathematics and Physics in the Life Sciences and Experimental Biology, UCL, London, UK.
3. Center for Child Health Research, University of Tampere and Tampere University Hospital, Finland.
4. College of Medicine, University of Malawi, Malawi.
5. MRC Clinical Trials Unit at UCL, London.
6. Department of Genetics, Evolution and Environment, UCL, London, UK.
7. Department of Paediatrics, Tampere University Hospital, Tampere, Finland.

* Corresponding author: liam.shaw.13@ucl.ac.uk

Running Head: Gingivitis and periodontitis in supragingival plaque

Word count (abstract): 395 (abstract: 246, importance: 149)

Word count (main text): 4,440

Word count (funding information and acknowledgements): 192

Keywords: periodontitis, gingivitis, oral microbiome, microbial ecology, network analysis

Abstract

Periodontal disease ranges from gingival inflammation (gingivitis) to the inflammation and loss of tooth-supporting tissues (periodontitis). The bacterial composition of supragingival plaque across a range of periodontal severities has not previously been explored with high-throughput sequencing. Furthermore, quantitative modelling of bacterial abundances in supragingival plaque as a function of both gingivitis and periodontitis has not previously been attempted.

We assessed a cross-sectional cohort of 962 Malawian women for periodontal disease and used 16S rRNA gene amplicon sequencing (V5-V7 region) to characterise the bacterial composition of supragingival plaque samples. Associations between bacterial relative abundances and gingivitis/periodontitis were investigated by using negative binomial models, adjusting for epidemiological factors. We also examined bacterial co-occurrence networks to assess community structure.

The main differences in supragingival plaque composition were associated more with gingivitis than periodontitis, including higher bacterial diversity and greater abundance of particular species. However, even after controlling for gingivitis, the presence of subgingival periodontitis was associated with an altered supragingival plaque. A small number of species were associated with periodontitis but not gingivitis, including members of *Prevotella*, *Treponema,* and *Selemonas,* supporting a more complex disease model than linear progression following on from gingivitis. Co-occurrence networks of periodontitis-associated taxa clustered according to periodontitis across all

gingivitis severities. Species including *Filifactor alocis* and *Fusobacterium nucleatum* were central to this network, supporting their role in co-aggregation of periodontal biofilms during disease progression.

Our findings confirm that periodontitis cannot be considered simply an advanced stage of gingivitis, even when only considering supragingival plaque.

IMPORTANCE. Periodontal disease is a major public health problem associated with oral bacteria. While earlier studies focused on a small number of ‘periodontal pathogens’, it is now accepted that the whole bacterial community may be important. However, previous high-throughput marker gene sequencing studies of supragingival plaque have largely focused on high-income populations with good oral hygiene without including a range of periodontal disease severities. Our study includes a large number of low-income participants with poor oral hygiene and a wide range of severities. We are able to quantitatively model bacterial abundances as a function of both gingivitis and periodontitis, which has not previously been attempted. A signal associated with periodontitis remains after controlling for gingivitis severity, supporting the concept that even when only considering supragingival plaque, periodontitis is not simply an advanced stage of gingivitis. This suggests the future possibility of diagnosing periodontitis based on bacterial occurrences in supragingival plaque.

Introduction

Periodontal disease is a major public health problem, particularly in low-income settings like in sub-Saharan Africa (1). Aside from irreversible tooth loss, chronic periodontitis may also increase the risk of adverse systemic conditions (2) such as cardiovascular disease (3) and preterm birth, although for the latter, different studies have reported conflicting results (4). The association between periodontitis and systemic disease may be due both to increased systemic inflammation and to translocation of bacteria into the blood stream (5). Despite its importance, the microbial ecology of periodontal disease in different oral habitats remains incompletely understood. Studies of the oral microbiome in periodontal disease typically focus on small populations in developed countries with advanced dental healthcare systems, which may not be representative of the natural history of periodontal disease in the absence of treatment (6).

In periodontal disease, the immune system responds with inflammation to oral biofilms (7). After an initial focus on identifying particular periodontal ‘pathogens’ (8), it is now widely accepted that oral bacterial communities undergo a shift or dysbiosis (9) and that the presence of particular disease-associated species may exacerbate the inflammatory reaction to commensal bacteria (10). The two main features of periodontal disease are gingival inflammation (gingivitis) and the formation of periodontal pockets (periodontitis). While it is clear that gingivitis always precedes periodontitis (11), gingivitis does not always progress to periodontitis (12) suggesting that these may not simply represent different stages of a continuous spectrum of disease. While there is some evidence that

a steady continuous progression may be expected (13), most models involve acute bursts of exacerbation and longer periods of remission (14, 15).

Despite this knowledge, studies of oral bacteria in periodontal disease often fail to capture the full range of periodontal conditions: from health through gingivitis to periodontitis. In supragingival plaque in particular, comparing only healthy subjects with subjects suffering from periodontitis may lead to associations being attributed to periodontitis alone, despite the fact that they might also be present in subjects with gingivitis. To explain the progression of disease and identify factors uniquely attributable to periodontitis it is necessary to compare subjects across the full range of periodontal severities. In itself this is not a novel concept, with many previous studies investigating bacterial associations with disease using checkerboard DNA-DNA hybridization (16–18). Earlier studies were targeted at a small number of bacterial species, (typically around 40). The advent of high-throughput 16S rRNA gene amplicon sequencing has facilitated improved analysis of the total bacterial diversity in the oral cavity (19, 20) identifying around 1,000 species that may be present (10) and showing that samples from the mouth typically have higher alpha diversities than those from other body sites (21, 22). Recent studies have used such amplicon sequencing to characterize subgingival plaque across a range of periodontal conditions, finding differences between subjects with gingivitis and periodontitis (23, 24). Work on supragingival plaque has been less common due to the fact that it does not have a direct link to inflammation and subsequent loss of attachment in periodontitis. It therefore remains ambiguous whether, for supragingival plaque, periodontitis can be simply considered as an

advanced stage of gingivitis, or if there are detectable differences in bacterial composition.

To address this question, we investigated bacterial abundances in supragingival plaque using quantitative modelling that takes into account both gingivitis (quantified by bleeding-on-probing) and periodontitis (quantified by periodontal pocket depth) in a cross-sectional cohort of 962 Malawian women who had recently given birth (25). We used negative binomial models, originally developed for RNA-seq experiments (26), making use of absolute (i.e. un-normalized) read counts to avoid losing information – a downside of other statistical approaches applied to marker gene data like rarefying (27). After fitting a negative binomial distribution to count data for a given species, the mean of this distribution is then used as the output of a generalized linear model with a logarithmic link using experimental variables (e.g. disease severity) as inputs, allowing the identification of differentially abundant species. This approach considers bacterial species as independent, but in reality oral bacteria exist in complex polymicrobial biofilms (28). Therefore, we also applied co-occurrence analysis to periodontitis-associated bacteria to identify important members of the community.

In summary, we aimed to identify the effects of periodontitis on supragingival plaque after controlling for gingivitis severity, separating and distinguishing the signals of these two features of periodontal disease.

Materials and Methods

Study population

Women analyzed in this study were participants in the iLiNS-DYAD-M trial (registration ID: NCT01239693) (25). This was a randomized controlled trial into the effects of three nutritional supplements on birth outcomes: lipid-based nutrient supplement (LNS), multiple micronutrients (MMN) or iron folic-acid (IFA). Women were eligible for enrolment in the trial if they were pregnant <20 weeks, >14 years old, had no chronic illnesses requiring frequent medical care, no allergies, no evident pregnancy complications (edema, blood hemoglobin < 50 g / l, systolic blood pressure > 160 mmHg or diastolic > 100 mmHg), no earlier participation in the same trial and no concurrent participation in any other.

1,391 pregnant women were enrolled between February 2011 and August 2012 at antenatal clinics at two hospitals (Mangochi and Malindi) and two health centers (Lungwena and Namwera) in Mangochi district, Malawi. All women were self-reported non-smokers, and were given two courses of preventive malaria treatment with sulfadoxine–pyrimethamine (three tablets of 500 mg sulfadoxine and 25 mg pyrimethamine orally): one at enrolment and one between the 28^{th} and 34^{th} gestational week. After giving birth, 1229 women completed an oral health examination, consisting of a clinical examination and a panoramic x-ray of the jaws. 1024 women had this examination within six weeks of delivery of a single infant (mothers of twins were excluded) and were included in further analysis. After excluding women without a

152 supragingival sample (n=59), and those with unknown HIV status (n=3), 962 women
153 remained for our cross-sectional analysis.
154
155 **Classification of periodontal disease**
156 Gingivitis was measured by the number of dental arch sextants with bleeding-on-
157 probing (BoP) out of six, with three sextants on each jaw (left, middle, and right). For
158 periodontitis classification, each tooth was examined for evidence of deepened dental
159 pockets both clinically and radiologically. A tooth was defined as having periodontitis if
160 either a ≥4 mm pocket was measured in clinical examination or a vertical bony pocket
161 was identified at least at the cervical root level radiologically. A woman was defined as
162 having periodontitis if she had at least three teeth with periodontitis or at least one
163 dental arch sextant with horizontal bone loss (at least at cervical level). The examination
164 and classification methods are explained in detail elsewhere (29).
165
166 **Sample collection**
167 Supragingival dental plaque samples were collected by swabbing the gingival margin of
168 each tooth with a sterile plastic swab stick with a nylon fiber tip (microRheologics no.
169 552, Coban, Brescia, Italy). After transfer in a cold box with ice packs to a laboratory,
170 swabs were stored in cryovials at -20˚C before being transferred to -80˚C.
171
172 **DNA extraction and sequencing**
173 We used Illumina compatible primers (785F: GGATTAGATACCCBRGTAGTC

, 1175R: ACGTCRTCCCCDCCTTCCTC) (30) that amplify the V5-V7 region of the 16S rRNA gene to generate a sequencing library (31). Each sample was amplified with dual indexes on the forward and reverse primer. All barcodes and adapter sequences used have been previously published (32). Each reaction was set up with 1X Molzym PCR Buffer (Molzym), 200μM of dNTPs (Bioline), 0.4 μM of forward and reverse primer with barcode attached, 0.025μM of Moltaq (Molzym), 5μl of template DNA and PCR grade water (Bioline) to make a final reaction volume of 25μl. Cycling parameters were as follows: 94˚C x 3 min, 30 cycles of 94˚C x 30 sec, 60˚C x 40 sec, 72˚C x 90 sec and one final extension at 72˚C x 10 min.

Samples were purified and pooled into an equimolar solution using SequalPrep Normalization Plate Kit (Life technologies) and further cleaned using AMPure XP beads (Beckman Coulter), both as per manufacturer's instructions. After quantification using the Qubit 2.0 (Life technologies), the library was diluted and loaded into the MiSeq reagent cartridge at 10pM. MiSeq runs were set to generated 250bp paired-end reads and two 12bp index reads for each sample. Reads were deposited in the European Nucleotide Archive (accession number XXX, tbc before publication).

Taxonomic classification

Sequenced reads were merged, demultiplexed, and quality filtered (minimum average Phred score > 25) using QIIME v1.8.0 (33). Closed-reference operational taxonomic units (OTUs) were picked at 98.5% similarity against the Human Oral Microbiome Database (HOMD) v13.2 (20) using USEARCH v6.1.544 (34) in QIIME v1.8.0 (33) with

parallel_pick_otus_usearch61_ref.py. We used 98.5% sequence similarity because this is the threshold used to define taxa in HOMD, as it approximately corresponds to species level clusters for most oral bacteria (20). This approach identified 664 bacterial OTUs corresponding to 13,049,932 reads. The mean number of reads per sample was 13,565±6,833.

Closed-reference OTU picking suffers from a number of issues, including sensitivity to the order of reference sequences when sequences are identical over the region considered (35). This is a particular problem when sequences are similar; there exist oral bacteria that have >99% sequence similarity in given regions of the 16S rRNA gene but occupy separate oral habitats (36). For this reason, we also performed Minimum Entropy Decomposition (MED) on reads. MED is an unsupervised version of the oligotyping pipeline (37) which allows a greater resolution of microbial diversity by partitioning sequences based on sites with high positional entropy in a reference-free manner (36).

After the merging of overlapping reads, the average sequence length was 369 bases. We filtered sequences with an expected error greater than 1 using fastq_filter in VSEARCH v1.11.1 (38). We then discarded all sequences shorter than 350 or longer than 380 bases, but performed no other quality filtering (e.g. length truncation) because MED assumes that length variation is biologically meaningful. We ran MED v2.1 on 14,449,794 sequences (information on reads discarded at each stage is available in Supplementary Material). Because we wanted to be able to detect rare sequences, we

set the minimum substantive abundance parameter (M) to 1444 (0.1% of the total number of reads) and the maximum variation allowed within a node (V) to 3. All other parameters were set to their default values. We assigned taxonomy to MED phylotypes using GAST (39) with VSEARCH v1.11.1 replacing USEARCH.

Statistical analyses

Diversity. We fitted a multivariate linear regression model to predict species richness (observed number of species) and Shannon index (a measure of richness and evenness) using gingivitis, periodontitis, and the variables listed in Table 2 for 811/962 samples with complete data and >5,000 reads. Richness and Shannon index were averaged over 100 iterations of rarefying to 5,000 reads per sample. Backwards stepwise reduction by AIC (40) was used to select the final model.

Differential abundances. We used DESeq2 v1.6.3 (26) in Phyloseq to model abundances. DESeq2 uses negative binomial generalized linear models to compare the absolute number of reads for each taxa between categories (27). Gingivitis was included as a continuous variable (BoP ranging from 0 to 6) and periodontitis as a binary factor. The model also contained terms controlling for potential confounders (study site, nutritional intervention, HIV status, and sequencing run). P-values were corrected for multiple testing using the Benjamini-Hochberg procedure (41). Full DESeq results for gingivitis and periodontitis are available in Supplemental Material (Dataset S1).

Correlation networks. To facilitate higher resolution of the network of periodontitis-associated bacteria, we selected all MED phylotypes that had representative sequences with >98.5% sequence similarity to peridontitis-associated HOMD OTUs. We calculated pairwise Spearman correlation coefficients between these MED phylotypes across samples. We used the SparCC procedure for estimating correlations from compositional data using log-ratio transformed abundances (42) with default parameters (20 inference iterations and a correlation strength exclusion threshold of 0.1). To calculate pseudo p-values (two-sided t test) we shuffled the datasets for each group 100 times and repeated the procedure, removing correlations that were not significant ($p<0.05$, no multiple testing correction). Networks of strong correlations, defined as being outside the 95% CI for the mean correlation between nodes (mean + 1.96*s.d. e.g. 0.405 for the network in Figure 4a) were visualized as networks with qgraph v1.3.1 (43), using the Fruchterman-Reingold algorithm for node placement (44).

Results

Description of cohort

962 Malawian women were included in our analysis, with a mean age of 25.4 ± 6.2 years. 140 (14.6%) had no periodontal disease, 822 (85.4%) had gingivitis (bleeding-on-probing score [BoP] ≥1), and 307 (32.0%) had periodontitis (Table 1). Gingivitis and periodontitis were significantly correlated (Spearman's ρ=0.44) with the majority of women with periodontitis having high levels of gingivitis. Periodontitis and gingivitis were more common in women who were older, had lower socio-economic status, and fewer years of education (Table 2; for modelling see Supplemental Material, Table S1).

Plaque richness and diversity are higher in more severe gingivitis and periodontitis

Initial exploratory analysis with PCoA ordinations showed that although there was large variability in community composition across supragingival plaque samples, there was also a clear trend related to gingivitis severity that was robust to the analysis method used (HOMD OTUs or MED phylotypes, Figure 1). Stratifying by periodontitis in the same way did not indicate visually clear differences.

Quantitative analysis of diversity reflected this trend. Gingivitis was associated with higher microbial community richness (Figure 2a) and Shannon index (Figure 2b). Microbial communities did not markedly differ between healthy women and those with low levels of gingivitis. Both gingivitis and periodontitis were associated with higher supragingival plaque richness in a linear regression controlling for demographic

variables (Supplementary Table 3a). In the final model predicting Shannon index, periodontitis was not retained although gingivitis was (Supplementary Table 3b). Reversing the analysis, richness was retained in the final model for predicting gingivitis but not periodontitis (Supplementary Table S2).

Differences in bacterial abundances with gingivitis

Differential abundance analysis with DESeq2 (26, 27) found 118 OTUs that were significantly ($q<0.05$) associated with greater severity of gingivitis (Dataset S1), making up 16.6% of the dataset in terms of reads. Conversely, 47 OTUs were associated with lower severity (18.7% of the dataset), implying that gingivitis is not only related to bacterial load but also the nature of the microbial community.

Figure 3a and 3b show the cumulative abundances of health- and gingivitis-associated OTUs respectively, showing the progressive nature of changes with the degree of bleeding. Most of the pairwise comparisons of summed abundances of health- and gingivitis-associated OTUs were not significantly different between women with and without periodontitis (Kruskal-Wallis test, $p>0.05$). However, for women with periodontitis, severity of gingivitis was important, as there were microbial differences between women with and without periodontitis for both moderate gingivitis (BoP of 3; $p=0.014$) and severe gingivitis (BoP of 6; $p=0.011$).The most significantly gingivitis-associated OTU was *Peptostreptococcus stomatis*, which was present in over 75% of samples across severity categories and was an average of 1.45-fold more abundant (95% CI 1.37-1.54) with a unit increase in BoP.

Differences in bacterial abundances with periodontitis

While gingivitis had a stronger association with supragingival microbiota, there were also differences in microbial community composition with periodontitis (Figure 3c,3d). Seventy-one OTUs were significantly ($q<0.05$) more abundant in women with periodontitis (Dataset S1), making up 4.4% of the dataset in terms of reads. Thirteen OTUs were significantly more abundant in the absence of periodontitis, making up 3.6% of the dataset by reads. These health-associated OTUs were *Lautropia mirabilis, Rothia aeria, Streptococcus pyogenes, Streptococcus mutans* and seven members of *Actinomyces.*

At the genus level for periodontitis-associated OTUs, *Prevotella* (14 OTUs) and *Treponema* (10 OTUs) were the most represented. Only one member of the pathogenic red complex (8) was significantly associated with periodontitis: *Treponema denticola*. The other two members (*Porphyromonas gingivalis* and *Tannerella forsythia*) were additionally not identified as MED phylotypes in the dataset, possibly due to primer mismatch (see Supplementary Material). *Eubacterium nodatum,* previously identified as clustering with the red complex in supragingival plaque (45), was significantly associated with periodontitis.

Differences in bacterial abundances unique to periodontitis

Forty out of seventy-one periodontitis-associated OTUs (56%) were not associated with gingivitis (Supplementary Table S4). These taxa were rare: their mean cumulative

abundance was 2.2%, with only six OTUs having mean relative abundances >0.1%. The most represented genera were *Prevotella* (9 OTUs), *Treponema* (5 OTUs) and *Selemonas* (4 OTUs).

The presence or absence of periodontitis was not a significant determinant of cumulative abundances of these OTUs for women with the same levels of gingivitis (Kruskal-Wallis test, $p>0.05$), except for women with a BoP of 4 ($p=0.026$).

The co-occurrence network of periodontitis-associated taxa

The above analysis considers each OTU as independent, but in reality oral bacteria exist in complex polymicrobial biofilms where interactions are extremely important (28). Co-occurrence analysis can allow the identification of important members of microbial communities (46). We therefore analyzed the co-occurence networks of periodontitis-associated bacteria across all periodontal severities.

A preliminary network analysis of periodontitis-associated OTUs across periodontal severities indicated that the network was more connected in women with periodontitis across gingivitis severities (Supplementary Figure S1). However, we sought to confirm this co-occurrence pattern with a higher resolution analysis. We therefore selected all MED phylotypes that had >98.5% similarity to a periodontitis-associated OTU (see Methods). 81 MED phylotypes had representative sequences with >98.5% similarity to a periodontitis-associated OTU (see Dataset S2).

The strongly-connected co-occurrence network in women with severe gingivitis (BOP of 6) and periodontitis showed several genus-level clusters, including *Selenomonas*, *Peptostreptococcus*, and *Prevotella* (Figure 4a). Notably, these clusters were connected by a small group of central bacteria including *F. alocis* (phylotype 158) and several members of *Fusobacterium nucleatum* with phylotypes classified taxonomically as subspecies *vincentii* (phylotypes 3163 and 622) and *polymorphum* (phylotypes 618 and 619), suggesting their roles in co-aggregation of periodontal biofilms. Ranking phylotypes in the strongly-connected network according to their betweenness centrality, which measures potential for influence on information transfer in a network (47), the most connected phylotype was *F. nucleatum subsp. vincentii* (phylotype 3163) (see Supplementary Table 5). *T. denticola* was not present in this network, but when MED analysis was repeated with the minimum substantive abundance parameter reduced by a factor of 10 to 0.01% we found it was placed in the network in a central position.

To confirm that this altered community structure was a distinguishing feature of supragingival plaque between women with and without periodontitis, we clustered the correlation matrices based on Mantel distances for each category of periodontal disease (Figure 4b). Networks clustered by the periodontitis status of the women in the group, confirming that the altered community structure with periodontitis was detectable even in women with low levels of gingivitis. Within the periodontitis groupings, matrices clustered by gingivitis severity.

Discussion

In this study we investigated changes in the supragingival microbiome associated with periodontal disease severity in a large cross-sectional cohort in Malawi. Our main finding was that even though the composition of supragingival plaque is primarily associated with gingivitis, as quantified by bleeding-on-probing, rather than the presence or absence of periodontitis, the presence of periodontitis does have detectable associations with the supragingival microbiota that are unrelated to gingivitis. In particular the differences in co-occurrence patterns of taxa between women with and without periodontitis support a more complex etiology of disease than a simple progression from health through gingivitis to periodontitis.

Gingivitis and periodontitis were both associated with higher microbial community richness and Shannon index, and this association remained after adjustment for demographic factors including age, BMI, and socioeconomic status. This finding is consistent with previous research (48, 49), with higher diversity meaning that in periodontal disease the oral microbiota is added to rather than existing taxa undergoing replacement. This could correspond to primary ecological succession in a new environmental niche, as suggested by Abusleme *et al*. (50).

We found that many taxa were associated with gingivitis and periodontitis. The abundance of the majority of these taxa increased with gingivitis severity, and this pattern was not influenced by the presence of periodontitis. Furthermore, some women without gingivitis had similar summed percentage abundances of disease-associated

taxa to women with severe gingivitis. It would appear that relative bacterial abundances alone are insufficient to explain the presence of disease, consistent with a requirement for other factors such as the host inflammatory response to cause disease.

Periodontitis-associated OTUs were also identified including known periodontal pathogens like *F. alocis, T. denticola, F. nucleatum*, and *P. stomatis*, consistent with findings from other populations (28). OTUs including members of *Prevotella*, *Treponema,* and *Selemonas* were not significantly associated with gingivitis severity, supporting the idea that periodontitis is not just an advanced phase of gingivitis and involves additional bacteria. However, cumulative abundances of periodontitis-associated OTUs did not differ significantly between women with and without periodontitis who had the same levels of gingivitis, suggesting that abundances do not fully explain the disease.

What we did observe was different co-occurrence patterns across disease categories for periodontitis-associated bacteria, which indicated the presence of a consistent community structure in women with periodontitis across all gingivitis severities. Central nodes in this periodontitis-associated network included *F. alocis* and several subspecies of *F. nucleatum*, which acted as hubs connecting different clusters. Network analysis using betweenness centrality ranked *F. nucleatum subsp. vincentii* (phylotype 3163) as the most central phylotype in the strongly-connected co-occurrence network in women with severe gingivitis and periodontitis. These findings are consistent with the proposed roles as ‘bridging bacteria’ that contribute to the co-aggregation of periodontal biofilms

(51). *F. nucleatum* has been shown experimentally to 'facilitate the survival of obligate anaerobes in aerated environments' (52), and has been identified as one of the important precursors to attachment by later colonizers in periodontal disease (51). *F. alocis* has also been experimentally linked to the co-aggregation of periodontal biofilms (53, 54) and correlates with greater inflammation in periodontitis (24). Chen *et al.* also identified a similar *F. alocis*-centered co-occurrence group of taxa that was enriched in multiple oral habitats during periodontitis compared with healthy controls (49).

Limitations

The main strength of this study is that we were able to include women with different severities and combinations of periodontal disease, allowing us to distinguish signals from gingivitis and periodontitis. However, our observations about periodontitis only apply to supragingival plaque, as we did not sample from subgingival plaque due to the difficulty of collecting such a large number of samples from a cohort in a resource-limited setting. However, previous work has shown that sampling supragingival plaque still allows the detection of bacteria associated with periodontitis while being minimally invasive and simple to perform (55). Similarly, we were able to observe changes in abundances of rare taxa known to be associated with the subgingival plaque of periodontitis. For example, *Fretibacterium fastidiosum* (HOMD ID: 360BH017) which accounted for a mean of just 0.009% of reads was still significantly more abundant (2.5-fold) in women with periodontitis, consistent with the recent finding of a higher abundance in subgingival plaque when periodontitis was compared to gingivitis (23).

Another limitation was that samples were collected from across the mouth instead of localizing sampling to sites of specific interest. The distribution of bacterial species across the mouth is known to be heterogeneous, with supragingival plaque at sites adjacent to deepened periodontal pockets showing significantly higher counts of periodontitis-associated species (45). Due to the size of our cohort we used a single swab, which was probably responsible for the large amount of variability in our dataset when visualized in ordinations (Figure 1), and effectively pooled all supragingival sites. This precluded an investigation of heterogeneity between sites, but detectable associations with both gingivitis and periodontitis were still present even with this approach.

We treated gingivitis as a continuous variable but periodontitis as binary. In reality periodontitis is a complex disease with a problematic classification (15), and it is likely that our simple treatment of periodontitis obscures this complexity. This could cause bacterial co-occurrence patterns in women with periodontitis to appear stronger, as women with more severe disease may have greater abundances of associated bacterial species.

Our study is the largest to be conducted so far in a sub-Saharan population and our results appear consistent for the most part with previous work on bacterial associations with periodontal disease (16, 28, 45, 49, 56). However, it should be pointed out that our population was additionally notable in two respects. Firstly, all participants were women who had recently given birth. Pregnancy, particularly in its early to mid stages, is known

to be linked to periodontal disease and potential changes in the oral microbiome (57), with an increased susceptibility to gingivitis (58) although subgingival levels of known periodontal pathogens may remain unchanged (59). Qualitative differences between periodontal pathogens found during pregnancy and postpartum have also been observed (60). It is not clear for how long after pregnancy the oral microbiome remains altered, but evidence that significant changes are mainly detectable in early pregnancy (57) and the consistency of our results with other studies suggests that effects remaining after six weeks postpartum are small. Secondly, all women in the study were intermittently given sulfadoxine–pyrimethamine (SP) at enrolment and between the 28th and 34th gestational week for malaria prevention. Since systemic antibiotics can be given as a treatment for aggressive periodontitis (61), patients who have received antibiotic treatment in the previous 6 months are often excluded from studies of periodontitis. However, the salivary microbiome has been shown to be robust to disturbance by a week-long course of antibiotics (62). Given that SP treatment was intermittent, involved antibiotics not targeted at periodontal bacteria, and took place around two months before the oral sampling, we believe that it is unlikely to have played an important role, but have no direct evidence to support this claim.

Conclusion

This study represents the largest to date investigating associations between supragingival plaque composition and varying severities of periodontal disease, in a low-income sub-Saharan population with limited oral hygiene. We have identified distinct signals associated with gingivitis and periodontitis in supragingival plaque, with a dominant contribution from gingivitis. Future proposals for a diagnostic test for

484 periodontitis based on supragingival plaque sampling, which could be useful in low-
485 resource settings, will need to take this into account. Network analysis of observed co-
486 occurrence patterns and network analysis was consistent with the role of ‘bridging
487 bacteria’ like *F. nucleatum* and *F. alocis* in the co-aggregation of periodontal biofilms
488 prior to penetrance into subgingival regions. Although some periodontitis-associated
489 bacteria were also associated with gingivitis, the major change with periodontitis is in
490 the network of co-occurrences. Viewed this way, gingivitis sets the stage for
491 periodontitis to develop by providing an environment where periodontitis-associated
492 taxa can increase in abundance and co-aggregate into pathogenic biofilms that may
493 then penetrate to subgingival regions. More quantitative modelling of associations
494 between oral bacteria and various clinical features of disease will be necessary to
495 understand these complex relationships and explore the microbial ecology of
496 periodontitis.

Funding information

This publication is based on research funded in part by the Office of Health, Infectious Diseases, and Nutrition, Bureau for Global Health, U.S. Agency for International Development (USAID) under terms of Cooperative Agreement No. AID-OAA-A-12-00005, through the Food and Nutrition Technical Assistance III Project (FANTA), managed by FHI 360. Additional funding was provided by the Bill & Melinda Gates Foundation through a grant to the University of California, Davis. The findings and conclusions contained within the article are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation, USAID, the US government, or the other funders.

LS is funded by a EPSRC Centre for Doctoral Training studentship at UCL CoMPLEX (EP/F500351/1). Research was supported by the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and UCL.

Acknowledgements

We thank the study participants, the local communities, the research personnel at the study sites, and the iLiNS Project extended research team for their assistance with the study. Members of the iLiNS Project Steering Committee and Project Manager Mary Arimond (http://ilins.org) provided technical support.

References

1. **Petersen PE**, **Bourgeois D**, **Ogawa H**, **Estupinan-Day S**, **Ndiaye C**. 2005. The global burden of oral diseases and risks to oral health. Bull World Health Organ **83**:661–9.
2. **Li X**, **Kolltveit KM**, **Tronstad L**, **Olsen I**. 2000. Systemic diseases caused by oral infection. Clin Microbiol Rev **13**:547–58.
3. **Yu Y-H**, **Chasman DI**, **Buring JE**, **Rose L**, **Ridker PM**. 2015. Cardiovascular risks associated with incident and prevalent periodontal disease. J Clin Periodontol **42**:21–8.
4. **Ide M**, **Papapanou PN**. 2013. Epidemiology of association between maternal periodontal disease and adverse pregnancy outcomes--systematic review. J Periodontol **84**:S181–94.
5. **Hajishengallis G**. 2014. Periodontitis: from microbial immune subversion to systemic inflammation. Nat Rev Immunol **15**:30–44.
6. **Baelum V**, **Scheutz F**. 2002. Periodontal diseases in Africa. Periodontol 2000 **29**:79–103.
7. **Van Dyke TE**. 2008. The management of inflammation in periodontal disease. J Periodontol **79**:1601–8.
8. **Socransky SS**, **Haffajee AD**, **Cugini MA**, **Smith C**, **Kent RL**. 1998. Microbial complexes in subgingival plaque. J Clin Periodontol **25**:134–44.
9. **Jiao Y**, **Hasegawa M**, **Inohara N**. 2014. The Role of Oral Pathobionts in Dysbiosis during Periodontitis Development. J Dent Res **93**:539–546.
10. **Wade WG**. 2013. The oral microbiome in health and disease. Pharmacol Res

69:137–43.

11. **Schätzle M**, **Löe H**, **Bürgin W**, **Anerud A**, **Boysen H**, **Lang NP**. 2003. Clinical course of chronic periodontitis. I. Role of gingivitis. J Clin Periodontol **30**:887–901.

12. **Batchelor P**. 2014. Is periodontal disease a public health problem? Br Dent J **217**:405–9.

13. **Jeffcoat MK**, **Reddy MS**. 1991. Progression of probing attachment loss in adult periodontitis. J Periodontol **62**:185–9.

14. **Haffajee AD**, **Socransky SS**. 1986. Attachment level changes in destructive periodontal diseases. J Clin Periodontol **13**:461–75.

15. **Mdala I**, **Olsen I**, **Haffajee AD**, **Socransky SS**, **Thoresen M**, **de Blasio BF**. 2014. Comparing clinical attachment level and pocket depth for predicting periodontal disease progression in healthy sites of patients with chronic periodontitis using multi-state Markov models. J Clin Periodontol **41**:837–45.

16. **Ximénez-Fyvie LA**, **Haffajee AD**, **Socransky SS**. 2000. Microbial composition of supra- and subgingival plaque in subjects with adult periodontitis. J Clin Periodontol **27**:722–32.

17. **Ximénez-Fyvie LA**, **Haffajee AD**, **Socransky SS**. 2000. Comparison of the microbiota of supra- and subgingival plaque in health and periodontitis. J Clin Periodontol **27**:648–57.

18. **Haffajee AD**, **Teles RP**, **Patel MR**, **Song X**, **Veiga N**, **Socransky SS**. 2009. Factors affecting human supragingival biofilm composition. I. Plaque mass. J Periodontal Res **44**:511–9.

19. **Griffen AL**, **Beall CJ**, **Firestone ND**, **Gross EL**, **Difranco JM**, **Hardman JH**,

Vriesendorp B, **Faust RA**, **Janies DA**, **Leys EJ**. 2011. CORE: a phylogenetically-curated 16S rDNA database of the core oral microbiome. PLoS One **6**:e19051.

20. **Chen T**, **Yu W-H**, **Izard J**, **Baranova O V**, **Lakshmanan A**, **Dewhirst FE**. 2010. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. Database (Oxford) **2010**:baq013.

21. **Stearns JC**, **Lynch MDJ**, **Senadheera DB**, **Tenenbaum HC**, **Goldberg MB**, **Cvitkovitch DG**, **Croitoru K**, **Moreno-Hagelsieb G**, **Neufeld JD**. 2011. Bacterial biogeography of the human digestive tract. Sci Rep **1**:170.

22. **Huttenhower C**, **Gevers D**, **Knight R**, **Abubucker S**, **Badger JH**, **Chinwalla AT**, **Creasy HH**, **Earl AM**, **FitzGerald MG**, **Fulton RS**, **Giglio MG**, **Hallsworth-Pepin K**, **Lobos EA**, **Madupu R**, **Magrini V**, **Martin JC**, **Mitreva M**, **Muzny DM**, **Sodergren EJ**, **Versalovic J**, **Wollam AM**, **Worley KC**, **Wortman JR**, **Young SK**, **Zeng Q**, **Aagaard KM**, **Abolude OO**, **Allen-Vercoe E**, **Alm EJ**, **Alvarado L**, **Andersen GL**, **Anderson S**, **Appelbaum E**, **Arachchi HM**, **Armitage G**, **Arze CA**, **Ayyaz T**, **Baker CC**, **Begg L**, **Belachew T**, **Bhonagiri V**, **Bihan M**, **Blaser MJ**, **Bloom T**, **Bonazzi V**, **Paul Brooks J**, **Buck GA**, **Buhay CJ**, **Busam DA**, **Campbell JL**, **Canon SR**, **Cantarel BL**, **Chain PSG**, **Chen I-MA**, **Chen L**, **Chhibba S**, **Chu K**, **Ciulla DM**, **Clemente JC**, **Clifton SW**, **Conlan S**, **Crabtree J**, **Cutting MA**, **Davidovics NJ**, **Davis CC**, **DeSantis TZ**, **Deal C**, **Delehaunty KD**, **Dewhirst FE**, **Deych E**, **Ding Y**, **Dooling DJ**, **Dugan SP**, **Michael Dunne W**, **Scott Durkin A**, **Edgar RC**, **Erlich RL**, **Farmer CN**, **Farrell RM**, **Faust K**,

Feldgarden M, **Felix VM**, **Fisher S**, **Fodor AA**, **Forney LJ**, **Foster L**, **Di Francesco V**, **Friedman J**, **Friedrich DC**, **Fronick CC**, **Fulton LL**, **Gao H**, **Garcia N**, **Giannoukos G**, **Giblin C**, **Giovanni MY**, **Goldberg JM**, **Goll J**, **Gonzalez A**, **Griggs A**, **Gujja S**, **Kinder Haake S**, **Haas BJ**, **Hamilton HA**, **Harris EL**, **Hepburn TA**, **Herter B**, **Hoffmann DE**, **Holder ME**, **Howarth C**, **Huang KH**, **Huse SM**, **Izard J**, **Jansson JK**, **Jiang H**, **Jordan C**, **Joshi V**, **Katancik JA**, **Keitel WA**, **Kelley ST**, **Kells C**, **King NB**, **Knights D**, **Kong HH**, **Koren O**, **Koren S**, **Kota KC**, **Kovar CL**, **Kyrpides NC**, **La Rosa PS**, **Lee SL**, **Lemon KP**, **Lennon N**, **Lewis CM**, **Lewis L**, **Ley RE**, **Li K**, **Liolios K**, **Liu B**, **Liu Y**, **Lo C-C**, **Lozupone CA**, **Dwayne Lunsford R**, **Madden T**, **Mahurkar AA**, **Mannon PJ**, **Mardis ER**, **Markowitz VM**, **Mavromatis K**, **McCorrison JM**, **McDonald D**, **McEwen J**, **McGuire AL**, **McInnes P**, **Mehta T**, **Mihindukulasuriya KA**, **Miller JR**, **Minx PJ**, **Newsham I**, **Nusbaum C**, **O'Laughlin M**, **Orvis J**, **Pagani I**, **Palaniappan K**, **Patel SM**, **Pearson M**, **Peterson J**, **Podar M**, **Pohl C**, **Pollard KS**, **Pop M**, **Priest ME**, **Proctor LM**, **Qin X**, **Raes J**, **Ravel J**, **Reid JG**, **Rho M**, **Rhodes R**, **Riehle KP**, **Rivera MC**, **Rodriguez-Mueller B**, **Rogers Y-H**, **Ross MC**, **Russ C**, **Sanka RK**, **Sankar P**, **Fah Sathirapongsasuti J**, **Schloss JA**, **Schloss PD**, **Schmidt TM**, **Scholz M**, **Schriml L**, **Schubert AM**, **Segata N**, **Segre JA**, **Shannon WD**, **Sharp RR**, **Sharpton TJ**, **Shenoy N**, **Sheth NU**, **Simone GA**, **Singh I**, **Smillie CS**, **Sobel JD**, **Sommer DD**, **Spicer P**, **Sutton GG**, **Sykes SM**, **Tabbaa DG**, **Thiagarajan M**, **Tomlinson CM**, **Torralba M**, **Treangen TJ**, **Truty RM**, **Vishnivetskaya TA**, **Walker J**, **Wang L**, **Wang Z**, **Ward D V.**, **Warren W**, **Watson MA**, **Wellington C**,

608 **Wetterstrand KA**, **White JR**, **Wilczek-Boney K**, **Wu Y**, **Wylie KM**, **Wylie T**,
609 **Yandava C**, **Ye L**, **Ye Y**, **Yooseph S**, **Youmans BP**, **Zhang L**, **Zhou Y**, **Zhu Y**,
610 **Zoloth L**, **Zucker JD**, **Birren BW**, **Gibbs RA**, **Highlander SK**, **Methé BA**,
611 **Nelson KE**, **Petrosino JF**, **Weinstock GM**, **Wilson RK**, **White O**. 2012.
612 Structure, function and diversity of the healthy human microbiome. Nature
613 **486**:207–214.
614 23. **Park O-J**, **Yi H**, **Jeon JH**, **Kang S-S**, **Koo K-T**, **Kum K-Y**, **Chun J**, **Yun C-H**, **Han**
615 **SH**. 2015. Pyrosequencing Analysis of Subgingival Microbiota in Distinct
616 Periodontal Conditions. J Dent Res **94**:921–7.
617 24. **Camelo-Castillo A**, **Novoa L**, **Balsa-Castro C**, **Blanco J**, **Mira A**, **Tomás I**.
618 2015. Relationship between periodontitis-associated subgingival microbiota and
619 clinical inflammation by 16S pyrosequencing. J Clin Periodontol.
620 25. **Ashorn P**, **Alho L**, **Ashorn U**, **Yb C**, **Kg D**, **Harjunmaa U**, **Lartey A**, **Nkhoma M**,
621 **Phiri N**, **Phuka J**, **Sa V**, **Zeilani M**, **Maleta K**. 2015. The impact of lipid-based
622 nutrient supplement provision to pregnant women on newborn size in rural
623 Malawi : a randomized controlled trial . **101**:2014–2015.
624 26. **Love MI**, **Huber W**, **Anders S**. 2014. Moderated estimation of fold change and
625 dispersion for RNA-seq data with DESeq2. Genome Biol **15**:550.
626 27. **McMurdie PJ**, **Holmes S**. 2014. Waste not, want not: why rarefying microbiome
627 data is inadmissible. PLoS Comput Biol **10**:e1003531.
628 28. **Teles R**, **Teles F**, **Frias-Lopez J**, **Paster B**, **Haffajee A**. 2013. Lessons learned
629 and unlearned in periodontal microbiology. Periodontol 2000 **62**:95–162.
630 29. **Harjunmaa U**, **Järnstedt J**, **Alho L**, **Dewey KG**, **Cheung YB**, **Deitchler M**,

Ashorn U, **Maleta K**, **Klein NJ**, **Ashorn P**. 2015. Association between maternal dental periapical infections and pregnancy outcomes: results from a cross-sectional study in Malawi. Trop Med Int Health.

30. **Bonder MJ**, **Abeln S**, **Zaura E**, **Brandt BW**. 2012. Comparing clustering and pre-processing in taxonomy analysis. Bioinformatics **28**:2891–7.

31. **Doyle RM**, **Alber DG**, **Jones HE**, **Harris K**, **Fitzgerald F**, **Peebles D**, **Klein N**. 2014. Term and preterm labour are associated with distinct microbial community structures in placental membranes which are independent of mode of delivery. Placenta **35**:1099–101.

32. **Caporaso JG**, **Lauber CL**, **Walters WA**, **Berg-Lyons D**, **Huntley J**, **Fierer N**, **Owens SM**, **Betley J**, **Fraser L**, **Bauer M**, **Gormley N**, **Gilbert JA**, **Smith G**, **Knight R**. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J **6**:1621–4.

33. **Caporaso JG**, **Kuczynski J**, **Stombaugh J**, **Bittinger K**, **Bushman FD**, **Costello EK**, **Fierer N**, **Peña AG**, **Goodrich JK**, **Gordon JI**, **Huttley GA**, **Kelley ST**, **Knights D**, **Koenig JE**, **Ley RE**, **Lozupone CA**, **McDonald D**, **Muegge BD**, **Pirrung M**, **Reeder J**, **Sevinsky JR**, **Turnbaugh PJ**, **Walters WA**, **Widmann J**, **Yatsunenko T**, **Zaneveld J**, **Knight R**. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods **7**:335–6.

34. **Edgar RC**. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26**:2460–1.

35. **Westcott SL**, **Schloss PD**. 2015. De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational

taxonomic units. PeerJ **3**:e1487.

36. **Eren AM**, **Morrison HG**, **Lescault PJ**, **Reveillaud J**, **Vineis JH**, **Sogin ML**. 2014. Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. ISME J **9**:968–979.

37. **Eren AM**, **Maignien L**, **Sul WJ**, **Murphy LG**, **Grim SL**, **Morrison HG**, **Sogin ML**. 2013. Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. Methods Ecol Evol **4**.

38. **Rognes T**. 2016. Vsearch.

39. **Huse SM**, **Dethlefsen L**, **Huber JA**, **Welch DM**, **Relman DA**, **Sogin ML**. 2008. Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. PLoS Genet **4**:e1000255.

40. **Akaike H**. 1974. A new look at the statistical model identification. IEEE Trans Automat Contr **19**:716–23.

41. **Benjamini Y**, **Hochberg Y**. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B **57**:289 – 300.

42. **Friedman J**, **Alm EJ**. 2012. Inferring correlation networks from genomic survey data. PLoS Comput Biol **8**:e1002687.

43. **Epskamp S**, **Cramer AOJ**, **Waldorp LJ**, **Schmittmann VD**, **Borsboom D**. 2012. qgraph : Network Visualizations of Relationships in Psychometric Data. J Stat Softw **48**:1–18.

44. **Fruchterman TMJ**, **Reingold EM**. 1991. Graph drawing by force-directed placement. Softw Pract Exp **21**:1129–1164.

45. **Haffajee a. D**, **Socransky SS**, **Patel MR**, **Song X**. 2008. Microbial complexes in supragingival plaque. Oral Microbiol Immunol **23**:196–205.

46. **Faust K**, **Lahti L**, **Gonze D**, **de Vos WM**, **Raes J**. 2015. Metagenomics meets time series analysis: unraveling microbial community dynamics. Curr Opin Microbiol **25**:56–66.

47. **Freeman LC**. 1977. A Set of Measures of Centrality Based on Betweenness. Sociometry **40**:35.

48. **Kistler JO**, **Booth V**, **Bradshaw DJ**, **Wade WG**. 2013. Bacterial community development in experimental gingivitis. PLoS One **8**:e71227.

49. **Chen H**, **Liu Y**, **Zhang M**, **Wang G**, **Qi Z**, **Bridgewater L**, **Zhao L**, **Tang Z**, **Pang X**. 2015. A Filifactor alocis-centered co-occurrence group associates with periodontitis across different oral habitats. Sci Rep **5**:9053.

50. **Abusleme L**, **Dupuy AK**, **Dutzan N**, **Silva N**, **Burleson JA**, **Strausbaugh LD**, **Gamonal J**, **Diaz PI**. 2013. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. ISME J **7**:1016–25.

51. **Aruni AW**, **Dou Y**, **Mishra A**, **Fletcher HM**. 2015. The Biofilm Community-Rebels with a Cause. Curr oral Heal reports **2**:48–56.

52. **Bradshaw DJ**, **Marsh PD**, **Watson GK**, **Allison C**. 1998. Role of Fusobacterium nucleatum and coaggregation in anaerobe survival in planktonic and biofilm oral microbial communities during aeration. Infect Immun **66**:4729–32.

53. **Schlafer S**, **Riep B**, **Griffen AL**, **Petrich A**, **Hübner J**, **Berning M**, **Friedmann A**, **Göbel UB**, **Moter A**. 2010. Filifactor alocis--involvement in periodontal biofilms.

BMC Microbiol **10**:66.

54. **Fine DH**, **Markowitz K**, **Fairlie K**, **Tischio-Bereski D**, **Ferrendiz J**, **Furgang D**, **Paster BJ**, **Dewhirst FE**. 2013. A consortium of Aggregatibacter actinomycetemcomitans, Streptococcus parasanguinis, and Filifactor alocis is present in sites prior to bone loss in a longitudinal study of localized aggressive periodontitis. J Clin Microbiol **51**:2850–61.

55. **Galimanas V**, **Hall MW**, **Singh N**, **Lynch MDJ**, **Goldberg M**, **Tenenbaum H**, **Cvitkovitch DG**, **Neufeld JD**, **Senadheera DB**. 2014. Bacterial community composition of chronic periodontitis and novel oral sampling sites for detecting disease indicators. Microbiome **2**:32.

56. **Haffajee AD**, **Socransky SS**. 1994. Microbial etiological agents of destructive periodontal diseases. Periodontol 2000 **5**:78–111.

57. **Fujiwara N**, **Tsuruda K**, **Iwamoto Y**, **Kato F**, **Odaki T**, **Yamane N**, **Hori Y**, **Harashima Y**, **Sakoda A**, **Tagaya A**, **Komatsuzawa H**, **Sugai M**, **Noguchi M**. 2015. Significant increase of oral bacteria in the early pregnancy period in Japanese women. J Investig Clin Dent.

58. **Gürsoy M**, **Könönen E**, **Gürsoy UK**, **Tervahartiala T**, **Pajukanta R**, **Sorsa T**. 2010. Periodontal status and neutrophilic enzyme levels in gingival crevicular fluid during pregnancy and postpartum. J Periodontol **81**:1790–6.

59. **Adriaens LM**, **Alessandri R**, **Spörri S**, **Lang NP**, **Persson GR**. 2009. Does pregnancy have an impact on the subgingival microbiota? J Periodontol **80**:72–81.

60. **Carrillo-de-Albornoz A**, **Figuero E**, **Herrera D**, **Bascones-Martínez A**. 2010.

Gingival changes during pregnancy: II. Influence of hormonal variations on the subgingival biofilm. J Clin Periodontol **37**:230–40.

61. **Rabelo CC**, **Feres M**, **Gonçalves C**, **Figueiredo LC**, **Faveri M**, **Tu Y-K**, **Chambrone L**. 2015. Systemic antibiotics in the treatment of aggressive periodontitis. A systematic review and a Bayesian Network meta-analysis. J Clin Periodontol **42**:647–57.

62. **Zaura E**, **Brandt BW**, **Teixeira de Mattos MJ**, **Buijs MJ**, **Caspers MPM**, **Rashid M-U**, **Weintraub A**, **Nord CE**, **Savell A**, **Hu Y**, **Coates AR**, **Hubank M**, **Spratt DA**, **Wilson M**, **Keijser BJF**, **Crielaard W**. 2015. Same Exposure but Two Radically Different Responses to Antibiotics: Resilience of the Salivary Microbiome versus Long-Term Microbial Shifts in Feces. MBio **6**:e01693–15–e01693–15.

Figure legends and tables

Figure 1. PCoA ordination of supragingival plaque samples shows an approximate trend with gingivitis severity that is robust to analysis method.

PCoA ordinations based on Bray-Curtis dissimilarities between samples for (a, b) 626 HOMD OTUs and (c, d) 502 MED phylotypes. Filled ellipses show mean values for each gingivitis severity, ranging from 0 (yellow) to 6 (dark red). In both cases, an approximate trend is visible, despite the noisiness of the dataset. Before plotting, samples were rarefied to 5,000 reads to minimize the impact of sequencing depth.

Figure 2. Microbial community richness and Shannon index increase with gingivitis severity.

Both (a) richness (number of observed species) and (b) Shannon index (measure of diversity) of supragingival plaque increase with gingivitis severity. Estimates for each sample were calculated by sampling with replacement at a rarefaction depth of 5,000 sequences per sample and averaging over 100 iterations. The fitted line shows a local polynomial regression fit calculated using ‘loess’ in R, with the grey region indicating the 95% CI. 138/965 samples were excluded due to having fewer than 5000 sequences. Changing the rarefaction depth did not affect the conclusion that gingivitis severity was associated with an increase in both species richness and Shannon index.

Figure 3. Summed percentage abundances of OTUs associated with (a) decreased gingivitis, (b) increased gingivitis, (c) absence of periodontitis, and (d) presence of periodontitis for each periodontal disease category.

For plotting purposes, samples were rarefied to 10,000 reads per sample, resulting in the removal of 269/962 samples; this rarefaction was not used in the selection of the OTUs, which was performed using DESeq2 on the whole dataset. One outlier and two outliers in (c) and (d) respectively are not shown due to trimming the y-axis at a relative abundance of 30%.

Figure 4. The co-occurrence network of periodontitis-associated bacteria shows a distinct community structure with the presence of periodontitis across gingivitis severities.

(a) The strongly connected central co-occurrence network of periodontitis-associated bacteria across supragingival plaque samples from *n*=110 women with severe gingivitis (BOP=6) and periodontitis. Shown here are significant strong pairwise Spearman correlation coefficients ($p<0.01$, $\rho>0.405$) calculated with SparCC between MED phylotypes with >98.5% similarity to periodontitis-associated HOMD OTUs (see Methods). Node color indicates taxonomic genus, size is proportional to log-transformed mean relative abundance, and edge weight indicates the strength of the correlation. The red circle indicates the node with the highest betweenness centrality, classified taxonomically as *Fusobacterium nucleatum* ss. *vincentii*. Node layout was determined using the Fruchterman-Reingold algorithm in qgraph v1.3.1. 22 nodes without any

779 strong correlations connecting them to the rest of the network (i.e. no edges with
780 ρ>0.405) were removed during figure preparation.
781 (b) Clustering using hclust in R of the correlation matrices calculated in this way for all
782 severities of periodontal disease. The periodontitis-associated co-occurrence network is
783 more similar between women with periodontitis regardless of gingivitis severity.
784 Correlation matrices were not adjusted for significance due to the different numbers of
785 women between groups.

786 **Table 1. Breakdown of all women by severity of periodontal disease.**

	Number of dental arch sextants with bleeding-on-probing (BoP)						
	0	**1**	**2**	**3**	**4**	**5**	**6**
No periodontitis	137	72	95	111	72	63	102
Periodontitis	4	11	23	27	51	50	145

Table 2. Demographic characteristics broken down by severity of periodontal disease.

a. Malaria was diagnosed with a rapid diagnostic test obtained from a finger prick.

b. Anemia was defined as a haemoglobin count Hb < 110 g/l.

c. A proxy for socio-economic status was created from principal components analysis by combining information on the building material of the house, main source of water and electricity, sanitary facilities, and main type of cooking fuel used.

d. Women were enrolled at four sites: Lungwena / Malindi / Namwera / Mangochi.

e. Women received one of three nutritional interventions: IFA / MMN / LNS.

f .Supragingival samples were run on one of four sequencing runs on Illumina MiSeq.

BoP, bleeding-on-probing; BMI, body mass index; IFA, iron folate; MMN, mixed micro-nutrients; LNS, lipid-based nutritional supplement.

BoP	Periodontitis	N	Age (yrs)	Positive HIV test	Malaria [a]	BMI	Education (yrs)	Anemia [b]	Socio-economic status [c]	Site [d]	Nutritional intervention [e]	Sequencing run [f]
0	No	140	23.4 (5.8)	27 (19.3%)	37 (26.6%)	22.7 (3.2)	5.6 (3.6)	36 (25.7%)	0.38 (1.22)	36 / 37 / 18 / 49	43 / 53 / 44	47 / 49 / 41 / 3
1	No	72	23.9 (5.9)	7 (9.7%)	16 (22.2%)	22.6 (3.4)	5.1 (3.8)	12 (16.7%)	0.19 (1.11)	25 / 9 / 17 / 21	32 / 19 / 21	34 / 26 / 12 / 0
	Yes	11	31.6 (6.1)	1 (9.1%)	1 (9.1%)	22.7 (2.4)	4.4 (3.3)	3 (27.3%)	-0.35 (0.62)	6 / 2 / 1 / 2	8 / 0 / 3	3 / 5 / 3 / 0
2	No	95	24.7 (6.2)	11 (11.6%)	22 (23.2%)	22.1 (2.6)	4.4 (3.6)	19 (20.0%)	0.10 (1.10)	39 / 19 / 13 / 24	38 / 34 / 23	31 / 41 / 23 / 0
	Yes	23	27.5 (6.2)	5 (21.7%)	5 (21.7%)	21.7 (2.0)	2.7 (3.3)	4 (17.4%)	-0.16 (0.91)	13 / 1 / 4 / 5	5 / 11 / 7	9 / 7 / 7 / 0
3	No	111	24.4 (5.4)	11 (9.9%)	32 (28.8%)	21.7 (2.3)	4.3 (3.3)	21 (18.9%)	-0.12 (0.84)	41 / 22 / 22 / 26	40 / 34 / 37	36 / 34 / 39 / 2
	Yes	27	26.5 (5.7)	4 (14.8%)	3 (11.1%)	22.2 (2.7)	3.6 (3.0)	6 (22.2%)	-0.20 (0.91)	11 / 6 / 3 / 7	11 / 4 / 12	11 / 6 / 10 / 0
4	No	72	25.0 (6.4)	9 (12.5%)	16 (22.2%)	21.7 (2.2)	3.4 (3.0)	11 (15.3%)	-0.16 (0.80)	28 / 16 / 10 / 18	16 / 26 / 30	26 / 28 / 18 / 0
	Yes	51	26.9 (5.4)	8 (15.7%)	11 (21.6%)	21.8 (2.7)	3.3 (3.1)	7 (13.7%)	-0.17 (0.81)	27 / 3 / 7 / 14	14 / 19 / 18	23 / 7 / 21 / 0
5	No	63	24.9 (5.2)	7 (11.1%)	12 (19.0%)	21.6 (2.4)	4.0 (3.6)	15 (23.8%)	-0.16 (0.81)	22 / 11 / 9 / 21	22 / 23 / 18	26 / 13 / 24 / 0
	Yes	50	26.6 (5.9)	5 (10.0%)	7 (14.0%)	21.8 (3.1)	2.4 (2.8)	5 (10.0%)	-0.36 (0.61)	18 / 11 / 7 / 14	16 / 15 / 19	21 / 12 / 17 / 0
6	No	102	24.5 (5.5)	10 (9.8%)	18 (17.%)	21.9 (2.3)	3.5 (3.0)	26 (25.7%)	-0.20 (0.81)	36 / 24 / 16 / 26	33 / 41 / 28	18 / 46 / 36 / 2
	Yes	145	28.3 (7.0)	30 (20.7%)	28 (19.3%)	22.1 (2.5)	2.9 (3.0)	32 (22.1%)	-0.27 (0.74)	66 / 28 / 17 / 34	45 / 48 / 52	59 / 43 / 41 / 2