



Clustering of Subgingival Microbiota Reveals Microbial Disease Ecotypes Associated with Clinical Stages of Periodontitis in a Cross-Sectional Study

Sébastien Boutin^{1,2*}, Daniel Hagenfeld^{3,4†}, Heiko Zimmermann⁵, Nihad El Sayed⁴, Tanja Höpker⁶, Halina K. Greiser⁶, Heiko Becher^{5,7}, Ti-Sun Kim⁴ and Alexander H. Dalpke^{1,2}

¹ Department of Infectious Diseases, Medical Microbiology and Hygiene, University Hospital Heidelberg, Heidelberg, Germany, ² Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), Heidelberg, Germany, ³ Department of Periodontology and Restorative Dentistry, University Hospital Münster, Münster, Germany, ⁴ Section of Periodontology, Department of Conservative Dentistry, University Hospital Heidelberg, Heidelberg, Germany, ⁵ Institute of Public Health, Epidemiology and Biostatistics, Heidelberg University, Heidelberg, Germany, ⁶ Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany, ⁷ Institute of Medical Biostatistics and Epidemiology, University Hospital Hamburg-Eppendorf, Hamburg, Germany

OPEN ACCESS

Edited by:

Steve Lindemann,
Purdue University, USA

Reviewed by:

J. Christopher Fenno,
University of Michigan, USA
Georgios Kotsakis,
University of Washington, USA

*Correspondence:

Sébastien Boutin
sebastien.boutin@med.uni-
heidelberg.de

†These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Systems Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 10 November 2016

Accepted: 17 February 2017

Published: 01 March 2017

Citation:

Boutin S, Hagenfeld D,
Zimmermann H, El Sayed N,
Höpker T, Greiser HK, Becher H,
Kim T-S and Dalpke AH (2017)
Clustering of Subgingival Microbiota
Reveals Microbial Disease Ecotypes
Associated with Clinical Stages
of Periodontitis in a Cross-Sectional
Study. *Front. Microbiol.* 8:340.
doi: 10.3389/fmicb.2017.00340

Periodontitis is characterized by chronic inflammation associated with alteration of the oral microbiota. In contrast to previous microbiome studies focusing *a priori* on comparison between extreme phenotypes, our study analyzed a random sample of 85 people. The aim of this study was to link microbial differences to disease's prevalence and severity. Using next generation sequencing of 16S rRNA amplicons and cluster analysis, we observed that the population can be divided into two major ecotypes: One mainly contained periodontal healthy/mild periodontitis individuals whereas the second ecotype showed a heterogeneous microbial distribution and clustered into three distinct sub-ecotypes. Those sub-ecotypes differed with respect to the frequency of diseased patients and displayed a gradual change in distinct subgingival microbiota that goes along with clinical disease symptoms. In ecotype 2, the subgroup with no clinical signs of disease was linked to an increase of *F. nucleatum vincentii* but also several other species, while only in "end-stage" dysbiosis classical red complex bacteria gained overweight. Therefore, the microbial disease ecotypes observed in our population can lead to an establishment of an early microbial risk profile for clinically healthy patients.

Keywords: periodontitis, microbiome, subgingival plaque, microbial ecology, polymicrobial infections

INTRODUCTION

Periodontitis is a chronic inflammatory disease caused by a dysbiosis of the oral microbiota (Darveau, 2010; Yost et al., 2015) and by the host response to these alterations (Costalonga and Herzberg, 2014). In the United States, 46% of adults aged ≥ 30 years suffer from periodontitis and approximately 10% of the population have deep periodontal pockets, worldwide

Abbreviations: AL, attachment loss; BOP, bleeding on probing; HOMD, human oral microbiome database; LDA, linear discriminant analysis; LEfSe, linear discriminant analysis effect size; NAKO, German National Cohort; NGS, next generation sequencing; OTU, operational taxonomic units; PCoA, principal coordinates analysis; PD, pocket depth; qPCR, quantitative PCR.

(Page and Eke, 2007; Petersen and Ogawa, 2012; Eke et al., 2015). The oral cavity and especially the periodontal pocket provides a unique eco-system for microbial organisms and harbors a diverse microbiota with up to 700 prokaryote species (Chen et al., 2010). One model for pathogenesis of periodontitis suggests that periodontal microbial communities can be clustered into complexes that are associated with disease severity (Socransky et al., 1998). Seminal work, at that time based on culture dependent techniques, of Socransky identified a “red complex” harboring *Porphyromonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola*, associated with the severe form of periodontitis. Further complexes included an intermediate orange complex with, e.g., *Fusobacterium nucleatum* and a yellow and green complex dominated by *Streptococcus* species, the latter being associated with health. Support for a classical role for the red complex as direct pathogens came from the observation by Holt showing induction of periodontitis upon oral implantation of these bacteria in non-human primates (Holt et al., 1988).

However, more recent concepts suggest that keystone pathogens can disrupt tissue homeostasis and change the composition of the commensal microbiota thereby generating host immune modulation and dysbiosis, that is responsible for periodontitis (Hajishengallis et al., 2011, 2012). Such a concept takes into account observations that periodontal pathogens often are low abundant and can be present in healthy people (Haffajee et al., 1998). Periodontitis resembles the process of microbial succession with an increase of periodontitis-associated taxa while health-associated species remain but decrease in number. In turn, the microbial community structure changes significantly, and biomass typically increases.

With the establishment of NGS, a nearly unbiased view of the bacterial composition is accessible. Conventional techniques are limited to cultivable organisms or a pre-selection of targeted pathogens is needed (Socransky et al., 1998, 2013). Using NGS, usually a part of the 16S rRNA gene is amplified and sequenced for taxonomic characterization of a bacterial community. Due to progress in sequencing efficiency and costs it is now possible to use NGS in population based studies to monitor larger cohorts.

Using NGS, comparisons between individuals with and without periodontitis revealed different microbiological compositions in periodontitis and health. *Fusobacteria*, *Porphyromonas*, *Tannerella*, and *Treponema* were increased in periodontitis. *Streptococcus*, *Firmicutes*, and *Proteobacteria* were associated with periodontal health (Griffen et al., 2012; Wichmann et al., 2012; Abusleme et al., 2013; Li et al., 2014). Communities in health and periodontitis differed, with higher diversity and biomass in periodontitis (Wichmann et al., 2012; Abusleme et al., 2013; Hong et al., 2015). Subgingival clusters in periodontally diseased patients were not associated with demographic, medical or disease-specific clinical parameters other than periodontitis extent (Li et al., 2014). However, most of the actual studies focus on extremes of the disease, mostly healthy versus severe disease. Cluster analysis with such a pre-selection could result in a bias, neglecting intermediate states of disease. Furthermore, population based studies with higher sample sizes are missing so far.

The objective of this study was primarily to identify the variation in the microbial complexes of subgingival samples from a population-based study and secondly to identify the relationship of oral health variables with the microbial composition within these complexes.

MATERIALS AND METHODS

Description of the Cohort

The NAKO aims at recruiting a representative sample from the general population in Germany (Wichmann et al., 2012). Recruitment takes place in 18 study centers distributed throughout Germany and will include 200,000 people aged between 20 and 69 years. Feasibility studies were conducted in all centers in 2012 to test specific aspects of the NAKO (Zimmermann et al., 2015). The aim of one feasibility study was a comparison of oral health and systemic parameters of Turks, Germans and Resettlers from the former Soviet Union (Resettlers) in the Rhine-Neckar metropolitan region. Since for this study both Germans and persons with a migration background were selected, different recruitment channels were chosen (Reiss et al., 2014). Briefly, on the one hand, people were drawn randomly based on the nationality from registration offices and on the other hand network recruitment strategies were used. For microbial analysis, participants ($n = 85$) from the study center Heidelberg were selected.

Periodontal Examination

Two calibrated dentists DH, NE (with minimal concordance of 90% in periodontal probing within an error interval of 1 mm) performed full-mouth periodontal examination at six sites per tooth, except third molars. AL was measured from the cemento-enamel junction or from the margin of the restoration to the bottom of the pocket. PD was measured from the gingival margin to the bottom of the pocket. BOP was measured as presence or absence of bleeding after probing (Lang et al., 1986). The periodontal examination was conducted using a UNC-PCP15 Color-Coded Probe (Hu-Friedy Europe, Rotterdam/Netherlands) with a black band for each millimeter up to 15 mm. Subjects with periodontitis were defined according to the CDC-AAP classification (Page and Eke, 2007). Briefly, participants with two or more approximal sites with AL of 4 mm or higher, but not on the same tooth or two or more sites with PD 5 mm or higher, but not on same tooth were classified as having moderate to severe periodontitis. All other participants were classified as mild or no periodontitis. One subgingival plaque sample per patient was taken from the mesio-buccal site of the first molar of a randomized quadrant. Randomization was performed with a randomization list. If the first molar was missing, the next mesial tooth in the same jaw and side was selected for plaque sampling. As a consequence of the randomized selection of the sampling site, the chosen site did not necessarily reflect the disease state of the host. In seven cases subjects had a full mouth diagnosis of moderate/severe periodontitis, but sampling occurred on teeth with no AL and presenting physiologic PD (defined as AL < 4 mm and

PD < 4mm). At these sites it is more likely to find an unchanged microbiome (Griffen et al., 2012; Shi et al., 2015). Thus, for these cases the samples were classified as having no or mild periodontitis.

Plaque Sampling

Subgingival plaque was sampled from the mesio-buccal site of the first molar of a randomized quadrant. Randomization was performed with a randomization list. If the molar was missing, the next mesial tooth in the same jaw and side was selected for plaque sampling. Firstly, visible supra-gingival plaque was removed from the sampling tooth with a sterile cotton pellet and secondly, aROEKO ISO25 sterile paper point (Coltene, Altstätten, Switzerland) was introduced to the bottom of pocket and remained there for 10 s. Plaque samples were placed in 1.5 ml safe-lock Eppendorf tubes and placed immediately at -20°C until DNA extraction.

DNA Extraction

DNA extractions were performed using the QIAamp Mini Kit (QIAGEN, Hilden, Germany). Protease solution (7.2 mAU) and 200 μL of Buffer AL were added to the sample followed by a 15 s vortex. Samples were incubated at 56°C for 10 min and then purified according to the manufacturer's protocol. DNA was eluted by adding 100 μL of buffer AE to the column, incubation for 1 min at room temperature and centrifugation at $6000 \times g$ for 1 min. Negative controls were performed by doing the extraction without clinical samples to control for absence of contamination.

Quantitative PCR

Quantitative PCRs were processed to evaluate the number of 16S rRNA gene copies using Unibac primer (forward: 5'-TGG AGC ATG TGG TTT AAT TCG A-3'; reverse: 5'-TGC GGG ACT TAA CCC AAC A-3') targeting the 16S rRNA gene. PCR reactions were performed in a final volume of 15 μL composed of 1X Sybr-green mastermix (Life technology, Darmstadt, Germany), 50 pmol of each primer and 2 μL of DNA (or plasmid DNA standards). The thermal cycler conditions were: a first denaturation at 95°C for 20 s, 40 amplification cycles (95°C for 3 s, 60°C for 30 s) and two final steps at 95°C for 15 s and 60°C for 1 min followed by a melt curve. All reactions were performed in duplicate in a StepOnePlus Real-time PCR system (Applied Biosystems, Foster City, CA, USA). Quantification of the 16S rRNA gene's number of copies was performed by comparison to the Cycle threshold value of a plasmid DNA standard, which had been quantified by spectrophotometry.

Library Preparation for Next Generation Sequencing (NGS)

DNA was amplified using universal bacterial primers flanking the V4 region [515F and 806R from (Caporaso et al., 2012)]. Each primer was tagged with an individual barcode to assign the sequences to the samples. PCR reactions were performed in 25 μL volumes composed of Q5 High-Fidelity 1X Master Mix (New England Biolabs GmbH, Germany), 0.5 μM of each primer and 2 μL of DNA. The thermal

cycler (Primus 25, Peqlab Biotechnologie GmbH, Germany or FlexCycler², Analytik Jena AG, Germany) conditions were: a first denaturation at 94°C for 3 min, 30 amplification cycles (94°C for 45 s, 50°C for 1 min, and 72°C for 1 min 30 s) and a final extension at 72°C for 10 min. Negative controls were performed using the negative control from the extraction step and using sterile water as template. PCR products were then checked by agarose gel electrophoresis (2%) for presence of amplicons. Both negative controls were negative on the gel. Amplicons were then purified by using Agencourt AMPure XP beads (Beckman Coulter, Germany) following the manufacturer's instructions. Purified products were checked for quality and concentration using a ND-1000 Nanodrop instrument (Nanodrop, Wilmington, DE, USA) and Bioanalyzer (Agilent Technologies Inc., Böblingen, Germany). Equimolar mix of all the PCR products was then sent to the Center for "Quantitative Analysis of Molecular and Cellular Biosystems" at Heidelberg University (BIOquant) which performed the ligation of the sequencing adapters to the library and the paired-end sequencing on an Illumina Miseq sequencing system with 250 cycles.

Analysis of Sequences

Paired sequences were assembled in contigs with stringent parameters: minimum overlap of 100 nt and maximum mismatch of 5 nt. Contigs were then filtered for quality; sequences with a quality score lower than 30 over 97% of the length were discarded. Each contig was assigned to the sample with the barcodes on the right and left ends (allowing no mismatch per barcode). The assignment and the pre-treatment of the sequences were done using MOTHUR software (Schloss et al., 2009). Sequences were screened for ambiguity in the sequences (maximum ambiguity allowed: 0) and for homopolymers (maximum homopolymer length allowed: 8 nt). Chimera detection was done by using the algorithm Uchime (Edgar et al., 2011) which allows for a fast, sensitive and accurate detection of chimera. Sequences were then clustered as OTU (using the threshold of 3% of divergence) and OTU representative sequences were then classified at the taxonomic levels by comparisons with sequences from a self-made unified database containing the SILVA, Ribosomal Database project and HOMD database (Bootstrap cutoff of 80). Species identification was only possible when the best match in the database was from the HOMD therefore other sequences were assigned as unknown species (sp.).

Statistical Analysis

An OTU table was used to calculate descriptive indices as alpha-diversity (non-parametric Shannon index), richness (Chao1 richness estimate, number of OTUs observed), evenness (non-parametric Shannon index-based measure of evenness) and dominance (Bergerparker-Index). Beta-diversity was assessed by calculating distance matrices based on Bray-Curtis distances. As inter-individual variation was important in our data, we assigned samples to ecotypes using a probabilistic modeling based on *k*-means algorithm (number of iterations 1,000,000) to cluster microbial communities into the major meta-communities. Clustering was confirmed and visualized by a heat-map and

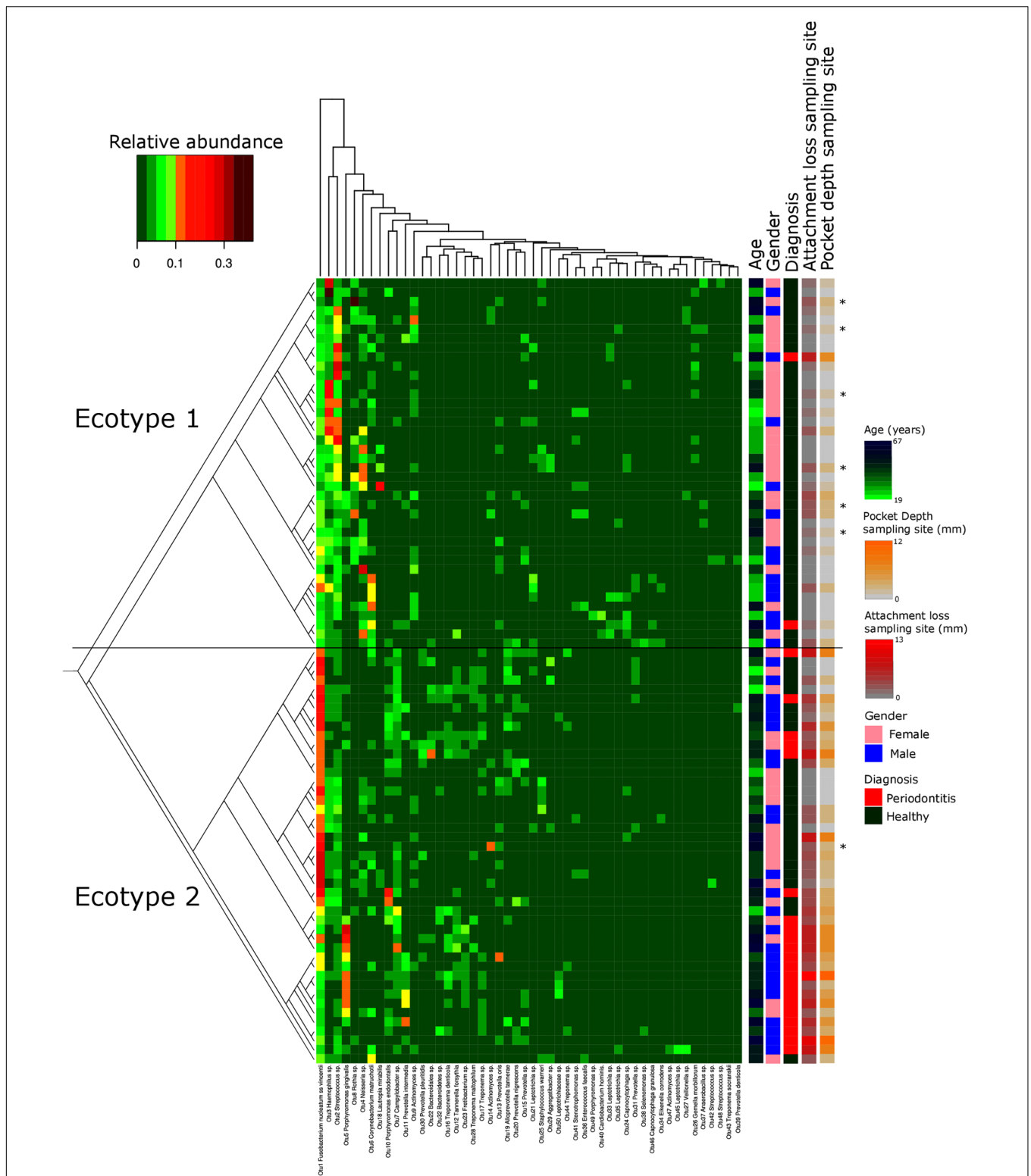
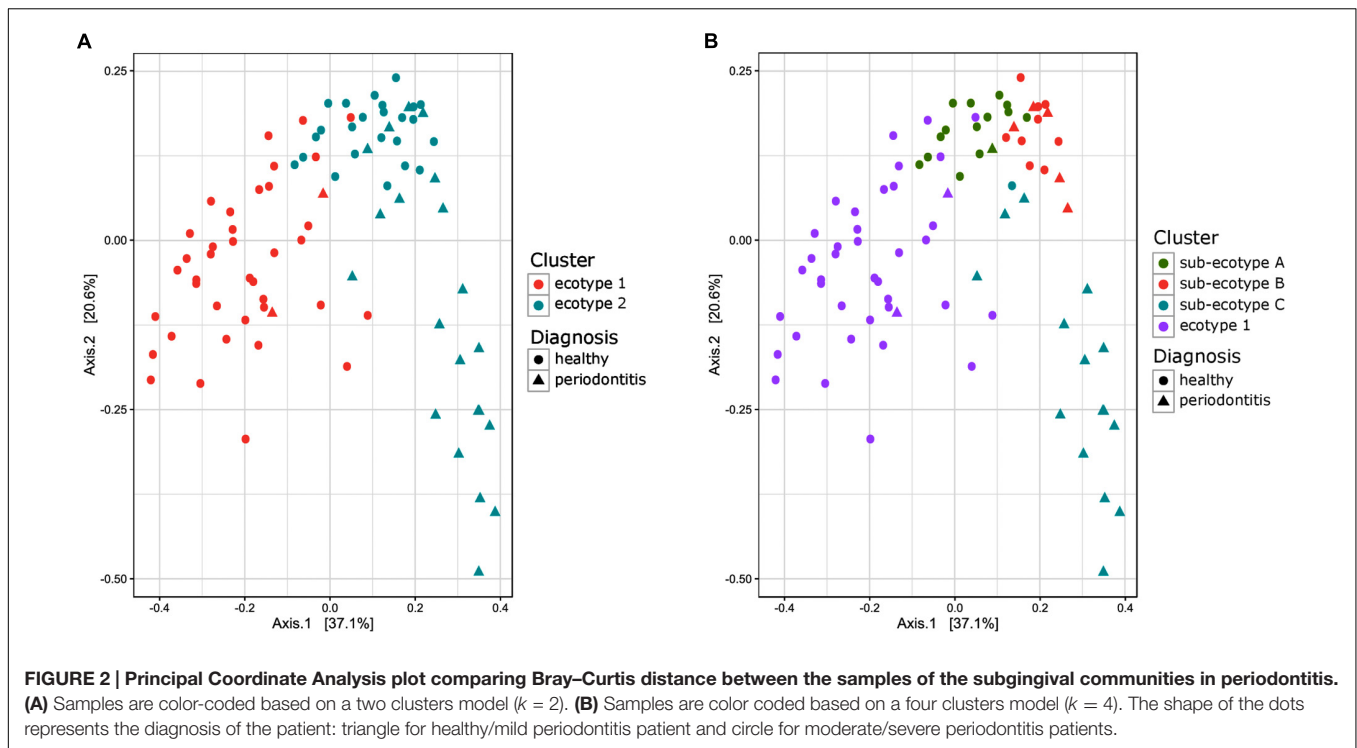


FIGURE 1 | Microbiota's structure analysis of the subgingival communities in periodontitis and correlation of microbial profiles with clinical and demographic measurements. Shown is a heatmap of the relative abundance for the top 50 OTUs found in the study cohort. Patients are in rows and OTUs are in columns. The color key is shown on the upper left. Patients are clustered based on the Bray–Curtis distance between their microbial communities validated by a *k*-mean algorithm probabilistic modeling (1000000 iterations). These distances are represented in the left tree showing two distinct clusters (Ecotypes 1 and 2). Relevant clinical and demographic characteristics are depicted in color codes at the right side. Patients diagnosed for periodontitis via the full mouth diagnosis but sampled at a non-diseased local site are symbolized by a *.



PCoA. The optimal number of clusters was defined by gap statistic (Tibshirani et al., 2001). A PERMANOVA adjusted for age and gender was performed to assess the statistical significance of differences in explanatory variables among samples or groups of samples. A LDA effect size (LEfSe) analysis was also performed to detect differentially abundant OTUs between groups. To compare the proportion of diseased patients between two or more groups, a multivariate analysis with a logistic regression model was performed adjusted for age and gender. For multiple groups' comparison, a *post hoc* Tukey test was done. All statistical analysis were performed with MOTHUR 1.33.0 (Schloss et al., 2009) and R 3.1.2 (R Development Core Team, 2014) [package: gplots (Warnes et al., 2009), vegan (Dixon, 2003), multcomp (Hothorn et al., 2008)].

RESULTS

Identification of Two Major Microbial Ecotypes in a Population-Wide Study

Thirty-eight males and 47 females were enrolled in this study. The mean age of all participants was 42.6 years (y) (range 19–67 years). Twenty-nine Resettlers from the former Soviet Union (34.1%), 17 Germans (20%), and 39 Turks (45.9%) participated in this study. Based on full mouth diagnosis according to the CDC AAP definition 30 (23+7) subjects (35.3%) were diagnosed with moderate to severe periodontitis, which is in the range of published representative studies (Eke et al., 2012). However, samples from diseased teeth were only obtained from 23 subjects whereas in 7 subjects with full mouth diagnosis “moderate/severe periodontitis” sampling was only done at a non-diseased site.

TABLE 1 | Microbiological and clinical differences between the ecotypes 1 and 2.

| | Ecotype 1 | Ecotype 2 | P-value |
|--|-----------|-----------|----------|
| Number of 16S rRNA gene copies | 116519 | 653924 | 3.57E-03 |
| log ₁₀ (number of 16S rRNA gene copies) | 5.07 | 5.81 | 9.76E-04 |
| Number of OTUs/samples | 229 | 199 | 0.00674 |
| Chao estimate | 410.1 | 359.2 | 0.07537 |
| Non-parametric Shannon index | 3.583 | 3.566 | 0.74244 |
| Evenness | 0.6461541 | 0.6621499 | 0.23617 |
| Bergerparker index | 0.1935 | 0.1979 | 0.75601 |
| Age (year) | 38.2 | 46.2 | 0.00525 |
| Pocket depth (mm) | 2.6 | 4.2 | 3.99E-03 |
| Attachment loss (mm) | 1.5 | 4.2 | 6.33E-04 |
| Gender (M/F) | 14/26 | 24/21 | 0.07435 |
| Periodontitis (Yes/No) | 2/38 | 21/24 | 0.00391 |
| Self reported periodontitis treatment (Yes/No/No answer) | 4/35/1 | 15/28/2 | 0.80236 |

This table summarizes the mean value of each ecotype for each parameter measured. Group comparisons were tested using a multivariate analysis with logistic regression with an adjustment for age and gender.

As these sites are probably containing a different microbiome (Griffen et al., 2012; Shi et al., 2015) they were reclassified as healthy/mild periodontitis according to the local disease status. From 2,431,538 reads that we obtained from all 85 subjects, we cleaned and removed chimera to end up with 2,248,918 clean

non-chimeric reads. We sub-sampled 4,218 reads per samples and validated sufficient coverage using rarefaction curves and calculating the Good's estimator of coverage [mean coverage: 97.8% (min–max: 95.6–99.0%)]. A total of 3,273 OTUs were found. The mean numbers of OTUs per sample was 213 (min–max: 139–332). To analyze the cohort we estimated major clustering of our samples using a k -means algorithm based on a Bray–Curtis distance matrix. This algorithm indicated that the study population clustered into two main clusters (ecotype 1: 47%; ecotype 2: 53%) (Figures 1, 2A), which were significantly divergent (PERMANOVA adjusted for age and gender: R^2 -value 0.26778, p -value < 0.001). Samples from the ecotype 1 showed a significantly richer microbiota (number of OTUs observed) and a lower biomass (estimated by the number of copies of the 16S rRNA gene by qPCR). No other differences in the alpha diversity were detected (Table 1).

Microbial Ecotypes Are Associated with Periodontal Disease Status

The strong divergence between the microbial structures of the two ecotypes was explained by the high number of OTUs differentially abundant (87 OTUs) and the weight of those OTUs; 34 OTUs of the top 50 most abundant OTUs (accounting for 83.2% of the microbiota) were divergent between the two ecotypes (Figure 3 and Supplementary Table S1). We noticed an increase in the ecotype 2 of four specific species (*Prevotella nigrescens*, *Fusobacterium nucleatum vincentii*, *Tannerella forsythia*, and *Treponema denticola*) known to be involved in periodontitis and a decrease of two species (*Eikenella corrodens* and *Streptococcus* sp.) known to belong to the healthy microbiota (Socransky et al., 1998). Of notice, the microbial clustering also reflected clinical measurements. Ecotype appears to be related to periodontitis: 2 out of 40 (5%) in ecotype 1 and 21 out of 45 (46.7%) in ecotype 2 showed diseased sampling sites. We used a generalized linear model (binomial family) to evaluate the accuracy of association between ecotype and disease status (accuracy: 72.9%, p -value < 0.001). The odds ratio for ecotype 2 estimated from a logistic regression model adjusted for age and sex is 11.0 (p -value = 0.00391) (Figure 1). This was reflected also by the current local inflammatory status at the sampling site, represented by 'BOP' (Supplementary Figure S1). Patients belonging to the ecotype 1 were younger and exhibited a lower PD and AL at the sample site (Table 1). No divergence of gender distribution was observed among the ecotypes and none of the other different clinical parameters showed a strong divergence.

Microbial Ecotype 2 Shows Further Sub-clustering That Associates with Distinct Clinical Parameters

Ecotype 2 was phenotypically more heterogeneous with an assembly of healthy participants, subjects with mild PD and AL as well as periodontitis patients, indicating a potential need of further clustering. This was also supported by the distribution on the PCoA plot and the dendrogram based on Bray–Curtis distances (Figures 2B, 4). Using gap statistic (Tibshirani et al., 2001) (Supplementary Figure S2), the

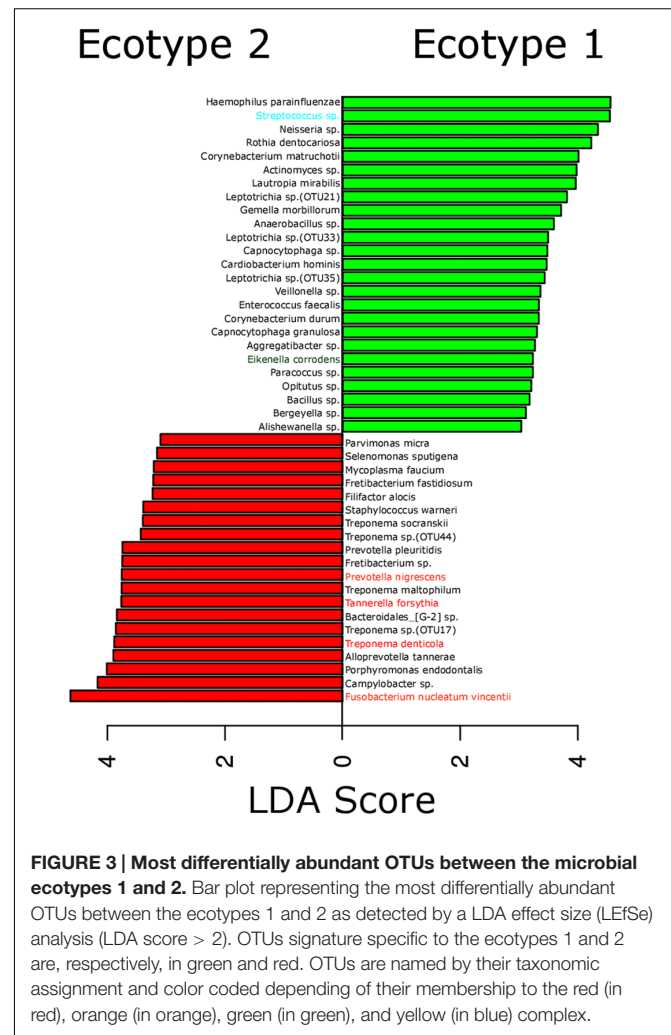


FIGURE 3 | Most differentially abundant OTUs between the microbial ecotypes 1 and 2. Bar plot representing the most differentially abundant OTUs between the ecotypes 1 and 2 as detected by a LDA effect size (LEfSe) analysis (LDA score > 2). OTUs signature specific to the ecotypes 1 and 2 are, respectively, in green and red. OTUs are named by their taxonomic assignment and color coded depending of their membership to the red (in red), orange (in orange), green (in green), and yellow (in blue) complex.

maximum number of reasonable clusters was estimated to be four clusters (corresponding to the ecotype 1 + three sub-ecotypes). According to this, three additional clusters (sub-ecotype A-B-C) within ecotype 2 could be identified. We again used a generalized linear model (binomial family) to evaluate the accuracy of association between ecotype and disease status (accuracy: 77.6%, p -value < 0.001). This finding was further confirmed by a PERMANOVA adjusted for age and gender, which showed an increase of the R^2 -value with the usage of four clusters (PERMANOVA Healthy vs. disease: R^2 -value = 0.44, p -value < 0.001). Furthermore the four ecotypes showed significantly different microbiota (Ecotype1 vs. sub-ecotype A: R^2 -value = 0.31, p -value < 0.01; Ecotype1 vs. sub-ecotype B: R^2 -value = 0.20, p -value < 0.01; Ecotype1 vs. sub-ecotype C: R^2 -value = 0.33, p -value < 0.01; sub-ecotype A vs. sub-ecotype B: R^2 -value = 0.19, p -value < 0.01; sub-ecotype A vs. sub-ecotype C: R^2 -value = 0.36, p -value < 0.01; sub-ecotype B vs. sub-ecotype C: R^2 -value = 0.42, p -value < 0.01).

Sub-ecotype A contain one moderate/severe periodontitis-patient (6.7%) and was significantly richer than the two other sub-ecotypes (Figure 5A), with a higher dominance index

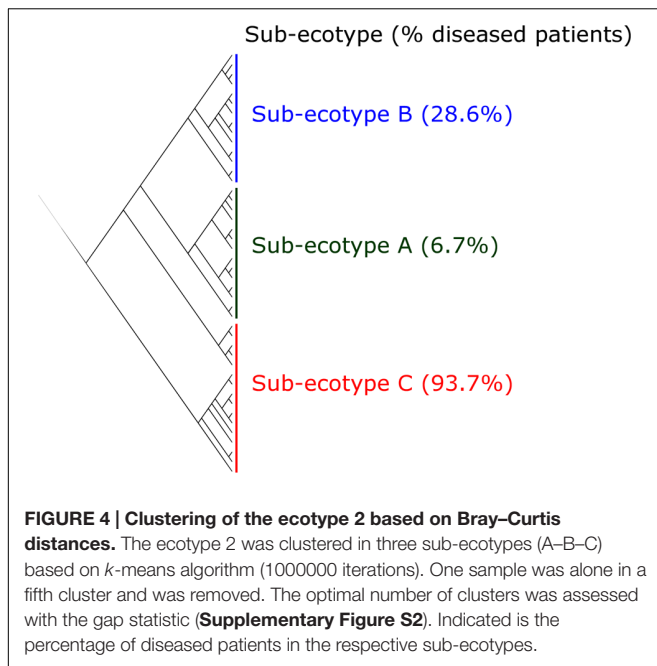


FIGURE 4 | Clustering of the ecotype 2 based on Bray-Curtis distances. The ecotype 2 was clustered in three sub-ecotypes (A–B–C) based on *k*-means algorithm (1000000 iterations). One sample was alone in a fifth cluster and was removed. The optimal number of clusters was assessed with the gap statistic (Supplementary Figure S2). Indicated is the percentage of diseased patients in the respective sub-ecotypes.

(Figure 5B) and a lower biomass of bacteria (Figure 5C). Sub-ecotype A also included subjects with the lowest PD (Figure 5D) and the lowest proportion of BOP positive sites (Supplementary Figure S1). Therefore, samples from this sub-ecotype A are phenotypically closer related to the healthy samples from ecotype 1.

On the opposite, sub-ecotype C showed the highest PD, the highest AL, a low richness and a high biomass (Figures 5D,E). Sub-ecotype C also had the lowest dominance measured by the Bergerparker index, probably because of a consortium of several different pathogens replacing otherwise dominating single OTUs (Figure 5B). This sub-ecotype contained most of the moderate/severe periodontitis patients with increased AL (median 5.5 mm), PDs (median 5 mm).

Yet, sub-ecotype B showed an intermediate phenotype with a low richness associated to a high biomass and dominance. In the clinical measurements, sub-ecotype B showed a mild PD (median: 4 mm), moderate gingival bleeding (proportions 50% positive) and a low AL (median: 3 mm). The data can indicate that those samples might be from mild periodontitis, gingivitis or show a controlled situation in treated patients. Despite the fact that none of the other clinical measurements was significantly different among the three clusters, we observed in patients from the ecotype 2 a division from healthy carrier (without clinical signs of periodontitis but mild signs of gingival bleeding) (sub-ecotype A) to severely diseased patients (sub-ecotype C) with an intermediate stage (sub-ecotype B).

Ecotype 2 Shows Heterogeneous Microbiota and a Clear Gradient of Distinct Pathogens

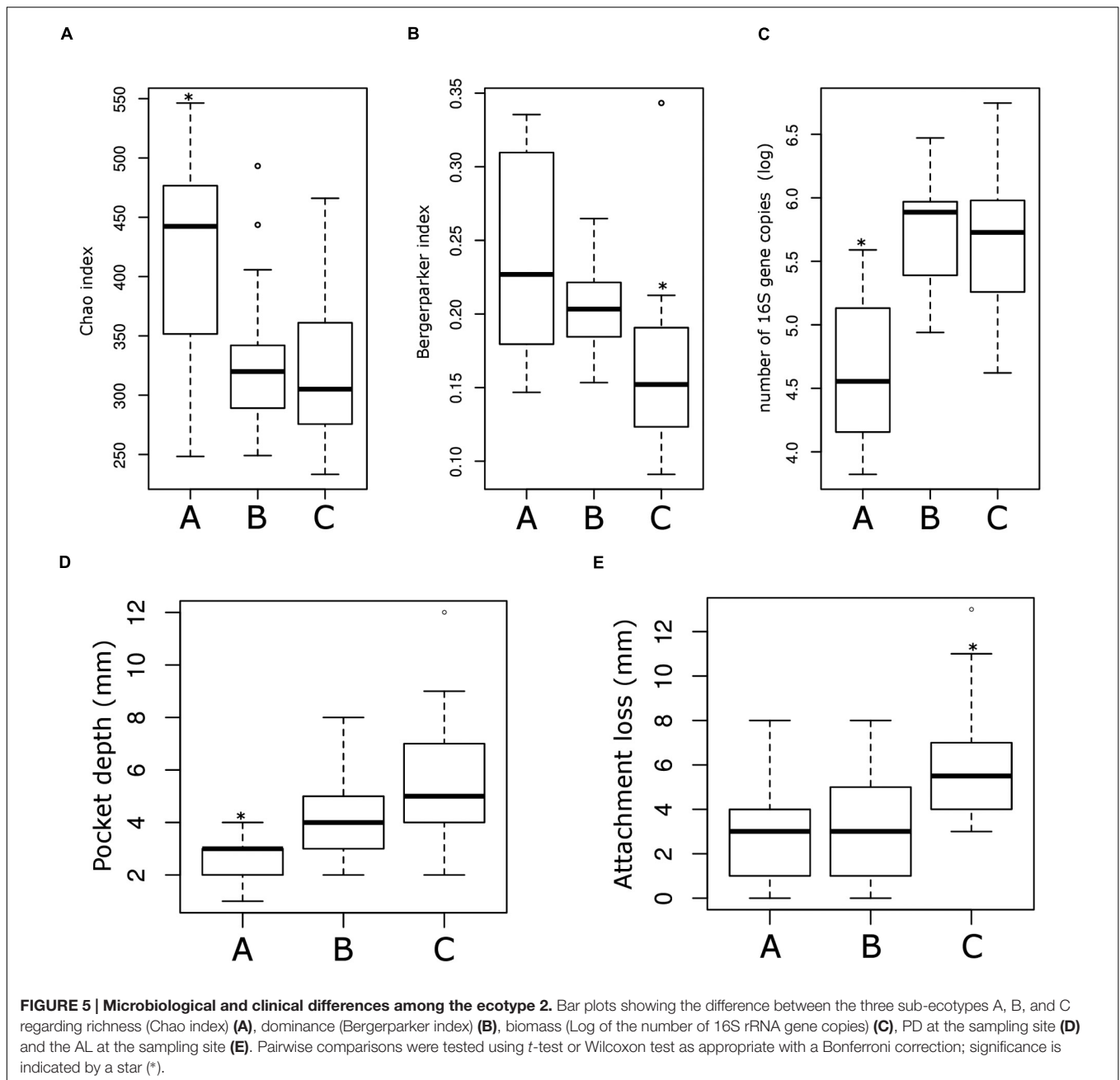
Differentiation of microbial ecotype 2 into three clusters was suggestive for a gradual division of microbial changes

between groups (Supplementary Table S2). We found 62 OTUs divergent between the sub-ecotype A and B (Bray-Curtis distance = 0.33), 29 OTUs between the sub-ecotype B and C (Bray-Curtis distance = 0.33) and 84 OTUs between the two extreme sub-ecotype A and C (Bray-Curtis distance = 0.51). Furthermore, most of the divergent OTUs showed a clear gradient pattern of abundance from sub-ecotype A to sub-ecotype B to sub-ecotype C (Figure 6). Based on Socranski definition of the key pathogens in periodontitis (Socransky et al., 1998), we observed that sub-ecotype A which mainly contained healthy people had higher relative abundances of the known “green complex” and “yellow complex” bacteria as well as one species from the “orange complex”; *F. nucleatum vincentii*, whereas the other “orange” and “red” complex bacteria showed up in sub-ecotype B and C, respectively (Figure 6 and for a close-up Supplementary Figure S3). Specifically, we analyzed abundance of *P. gingivalis*, *T. forsythia*, *T. denticola*, and *F. nucleatum vincentii* (Figure 7). Whereas *P. gingivalis* was not overall different between Ecotypes 1 and 2 (Figure 3) it turned out that within the ecotype 2 there was a clear and significant evolution: *P. gingivalis* increased significantly in abundance in sub-ecotype C where it became highly abundant (up to 35%). Of note, ecotype 1 and sub-ecotype A and B contained *P. gingivalis*, albeit at low abundance. A likewise trend (not significant) was seen for *T. denticola*. *T. forsythia* showed an increase in sub-ecotype B and then further in sub-ecotype C.

Interestingly, abundance of *F. nucleatum* was one of the decisive discriminators that differentiated healthy person in ecotype 1 from healthy subjects from the ecotype 2 (Sub-ecotype B). *F. nucleatum* declined from sub-ecotype A to B to C in abundance. Healthy patients from the ecotype 1 do not carry classical pathogens at high level of abundance. *F. nucleatum* increases greatly in early stages of the disease (Sub-ecotype A, minor PD and bone loss). In the second stage (Sub-ecotype B, mild PD and reduced bone loss), *T. forsythia* increases significantly and remains at this abundance while *P. gingivalis* becomes the dominant species during the severe stage (Sub-ecotype C, severe PD and bone loss). Therefore, we clearly observe an ecological clustering among those classical pathogens associated with the increased severity of the disease. Those classical key pathogens also co-occur with several other bacteria, which are not yet classified as associated with periodontitis (Figure 8). This global co-occurrence pattern highlights two major networks; one associated to the abundance of bacteria from the yellow and green complexes while the other is characterized by the abundance of bacteria from the orange and red complexes (Figure 8).

DISCUSSION

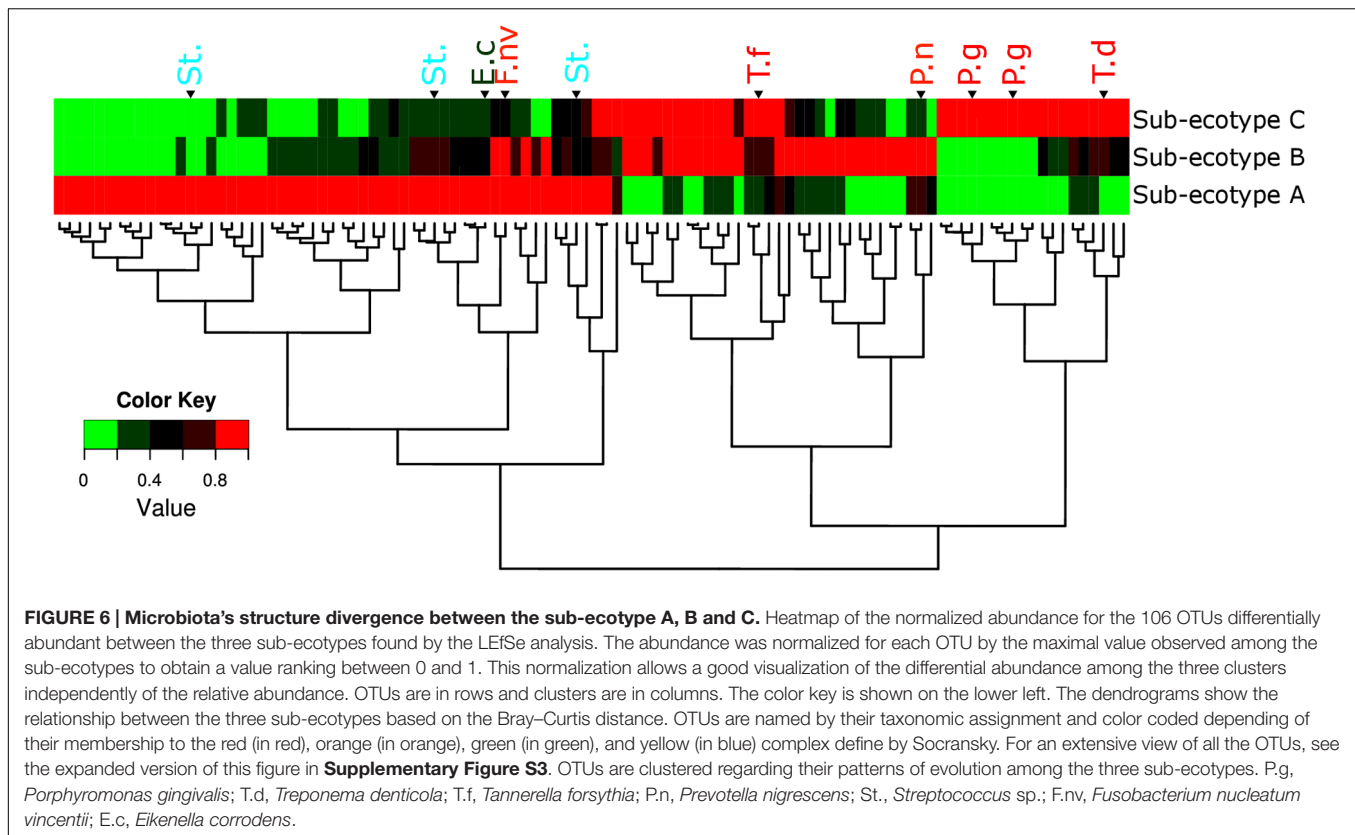
A recent concept in the field of periodontitis states that a succession of keystone pathogens can disrupt tissue homeostasis and changes the composition of the commensal microbiota thereby generating host immune modulation and dysbiosis that is responsible for periodontitis (Hajishengallis et al., 2011, 2012). In this study, we observed two different microbiological ecotypes



derived from sub-gingival samples within a population-based study; one major ecotype was homogeneously composed of healthy or only mild periodontitis subjects and the other more heterogeneous ecotype consisted of a mixture of three distinct clusters (sub-ecotypes A, B, and C). To our knowledge this is the first study of a population-based sample using NGS and showing a clear division in the microbiotas associated to the whole range of phenotypical variation in the disease from healthy to severe periodontitis.

In another study of 34 subjects with only periodontitis patients, two clusters could be found (Hong et al., 2015). In this study, one cluster was composed of bacteria known

as being keystone pathogens (*P. gingivalis*, *T. forsythia*, and *T. denticola*) while the other cluster was composed of bacteria known as first colonizer from the orange complex (*Fusobacterium* sp., *Prevotella* sp., and *Campylobacter* sp.) as well as bacteria from health-associated species. In a similar study by Park et al. (2015), 32 individuals have been allocated to three different groups (healthy, gingivitis, and periodontitis) according to clinical parameters. The healthy group dominated by *Streptococcus* sp., the gingivitis group was composed mostly by the genera *Streptococcus*, *Capnocytophaga*, *Haemophilus*, and *Leptotrichia* and the periodontitis group with a high abundance of *Porphyromonas*. Our findings match very well with those results

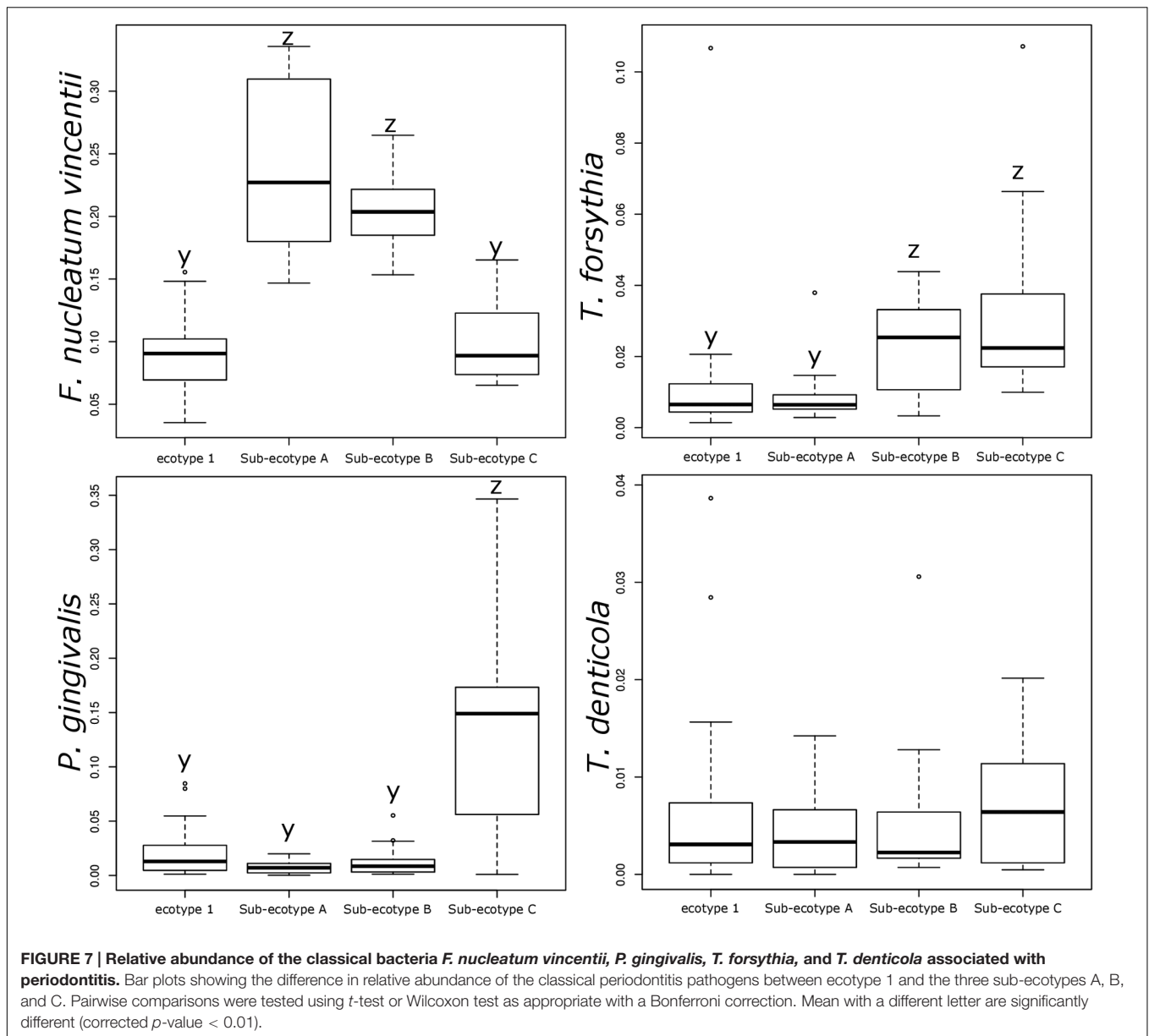


as our sub-ecotype A and B together compared to the sub-ecotype C display the same contrast as the one showed by Hong et al. (2015). Also, the three clusters depicted by Park et al. (2015) seem to be reflected by our cohorts from the ecotype 1 (healthy cohort), the sub-ecotypes A and B (mild disease, gingivitis) and the sub-ecotype C (periodontitis). However, in our study we were able to depict four distinct groups correlating with periodontally healthy or mild disease and moderate or severe periodontitis. This resolution is probably due to the fact that the clustering was based on a population design and only on the microbiota profiles of 85 patients without pre-selection based on the clinical parameters.

The four clusters (ecotype 1 + three sub-ecotypes A–C) observed in our study are nevertheless associated with distinct clinical conditions. Those clinical conditions could be ordered based on the known process of the disease from healthy to severely diseased. Patients from the ecotype 1 are mostly healthy and younger individuals with shallow pockets and nearly no AL at the sample site and low BOP. Those patients are also characterized by a low abundance of classical pathogens (Figures 1, 7) as shown in a previous study comparing the microbiota of periodontal compromised and matched healthy subjects (Griffen et al., 2012). Those results indicate the presence of a “healthy core microbiota” potentially linked to resistance against periodontitis. This concept of a “healthy core microbiota” was suggested in several studies before, which showed a high diversity and similarity to the different ecological niches within the oral cavity (Zaura et al., 2009; Bik et al., 2010). The

sub-ecotype A from the ecotype 2 also displays shallow pockets, slightly higher BOP and minimal AL at the sample site with only healthy or mild periodontitis patients in that cluster (Figures 3, 4). However, the microbiota undergo microbial colonization by *F. nucleatum vincentii* and first colonizers from the green and yellow complex indicating a new stage in the disease progression (Socransky et al., 1998; Kolenbrander et al., 2010). This result also supports the hypothesis that green complex is correlated to better health and that orange and yellow complexes are indicators of a first colonization leading to the establishment of the disease (Socransky et al., 1998; Kolenbrander et al., 2010). It also appears that yellow complex and green complex are co-occurring with species with yet unknown functions in the context of periodontitis during the early stage of the disease (Figure 7). On the other hand, *F. nucleatum* and *Prevotella* sp. seem to have a particular role as a bridging-organism between early and late colonizers (Socransky et al., 1998; Kolenbrander et al., 2010; Kirst et al., 2015). However, we identify additional bacteria that differ between those clusters and are directly co-occurring with this complex thus extending the classical complex concept (Figure 8).

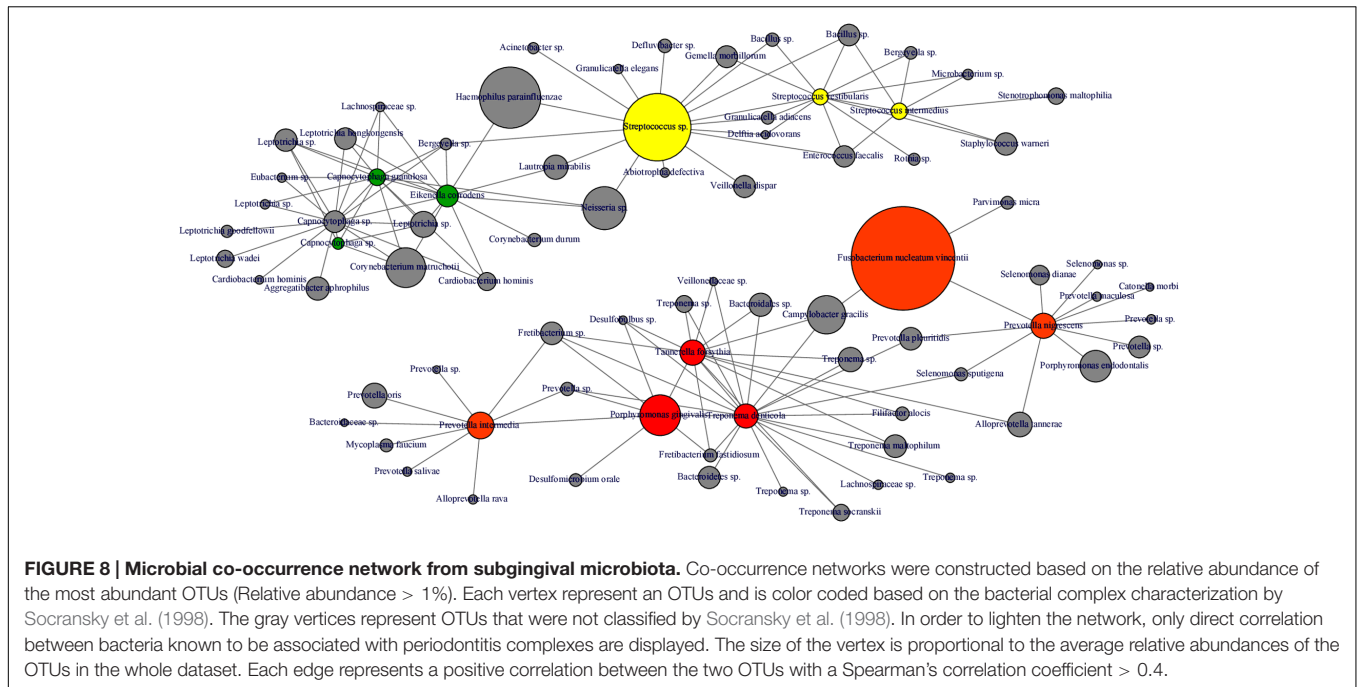
The sub-ecotype B showed an increased PD but still a low AL reflecting an increase in the severity and higher BOP. The microbiota from this cluster is characterized by a high abundance of *F. nucleatum vincentii* like in the sub-ecotype A but also an increase of bacteria from the orange and red complex (*T. forsythia*, *P. nigrescens*) (Socransky et al., 1998) as well as many other species with unknown function (Figure 6 and



Supplementary Figure S3, Table S2). Finally, the sub-ecotype C clinically reflects the most severe stage of the disease with deep pockets, frequent BOP and high AL; those changes go along with further alterations in the microbiota, notably an increase of species from the red complex (e.g., *P. gingivalis* and *T. denticola*) (Socransky et al., 1998) and numerous other species not known so far to be associated to the disease (Figure 8). The latter might indicate a possible pathogenic role or a hitch hiking effect (Figure 6 and Supplementary Figure S3, Table S2).

This study classified subjects primarily by the microbial ecotype, which was then associated to clinical signs of periodontitis. Periodontitis was classified by the CDC-AAP classification, which takes AL as well as PD into account. Thus

it is a more specific definition as the widely used CPITN-index, which only uses maximal PD (Cutress et al., 1987). A periodontitis patient can have different stages of inflammation on different teeth, which might also be represented by intra-individual microbiome differences (Shi et al., 2015). We took only one sample from each patient from one randomized sampling site. In many other studies different sites were pooled, which might lead to mixing different microbial stages together. Our approach also has its drawbacks, because sites can be randomly selected from a patient classified as periodontitis, but with no apparent signs of inflammation or AL. Future studies should consider different inflammatory conditions by taking multiple samples at different stages of disease in one subject to account for these intra-individual differences.



CONCLUSION

In this study, we observed the emergence of four discrete microbial communities within our population associated with clinical outcome of periodontitis ranging from healthy/mild periodontitis to moderate/severe periodontitis. The evolution of the abundance of the known pathogens involved in the disease as well as the increase of the severity parameters (AL, PD, and BOP) along the four ecotypes tend to prove a gradual change of microbiota during disease. Therefore, all together, our findings are supporting the concept of induced dysbiosis or polymicrobial synergy (Hajishengallis and Lamont, 2012). We propose a scenario of gradual change in subgingival microbiota, which goes along with clinical disease symptoms. Only in “end-stage” dysbiosis classical red complex bacteria gain overweight. Such a concept of induced dysbiosis including early, low abundant keystone pathogens has been suggested and supported by others before (Hajishengallis and Lamont, 2012; Hajishengallis et al., 2012; Park et al., 2015) but is now underlined by our findings in a population-based study. This scenario needs to be confirmed by future studies based on longitudinal design to validate the microbial variation over the different clusters in time. However, the findings in this study can already lead to an establishment of a microbial risk profile for clinically healthy patients.

ETHICS STATEMENT

The ethical board of the medical faculty of Heidelberg (S-108/2011) had approved this study and all participants gave their written informed consent.

AVAILABILITY OF DATA AND MATERIALS

Clean sequences, metadata files and R script used for analysis have been uploaded, and are publicly available in Figshare (<https://figshare.com/s/8da5157b93740ae58b8e>).

AUTHOR CONTRIBUTIONS

SB, DH, HZ, NE, TH, HG, HB, T-SK, and AD made substantial contributions to conception and design. SB, DH, HZ, HB, NE, T-SK, and AD contribute to the acquisition of data, analysis and interpretation of data. SB, DH, HZ, NE, TH, HG, HB, T-SK, and AD have been involved in drafting the manuscript or revising it critically for important intellectual content. SB, DH, HZ, NE, TH, HG, HB, T-SK, and AD gave a final approval of the version to be published.

FUNDING

This project was conducted in the context of the pretest studies of the German National Cohort (www.nationale-kohorte.de) and funded by the Federal Ministry of Education and Research (BMBF: 01ER1001A-I) and supported by the Helmholtz Association. DH received a Post-Doc fellowship from the Medical Faculty of Heidelberg.

ACKNOWLEDGMENT

We acknowledge the excellent technical support by Jessica Panitz, Selina Hassel, and David Ibberson as well as the technical team from the Bioquant.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.00340/full#supplementary-material>

FIGURE S1 | Proportions of bleeding on probing (BOP) in each sub-cohort. BOP was measured as presence or absence of bleeding on probing of the

periodontal pocket. Statistical difference in BOP was calculated by a Tukey test based on a logistic regression. Column with a different letter are significantly different (corrected p -value < 0.01).

FIGURE S2 | Estimation of the optimal number of clusters in the population. The number of clusters in our dataset was estimated via the gap statistic as described by Tibshirani et al. (2001).

FIGURE S3 | Microbiota's structure divergence between the sub-ecotypes A, B, and C (expanded version).

REFERENCES

- Abusleme, L., Dupuy, A. K., Dutzan, N., Silva, N., Burleson, J. A., Strausbaugh, L. D., et al. (2013). The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *ISME J.* 7, 1016–1025. doi: 10.1038/ismej.2012.174
- Bik, E. M., Long, C. D., Armitage, G. C., Loomer, P., Emerson, J., Mongodin, E. F., et al. (2010). Bacterial diversity in the oral cavity of 10 healthy individuals. *ISME J.* 4, 962–974. doi: 10.1038/ismej.2010.30
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624. doi: 10.1038/ismej.2012.8
- Chen, T., Yu, W.-H., Izard, J., Baranova, O. V., Lakshmanan, A., and Dewhirst, F. E. (2010). The human oral microbiome database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* 2010:baq013. doi: 10.1093/database/baq013
- Costalonga, M., and Herzberg, M. C. (2014). The oral microbiome and the immunobiology of periodontal disease and caries. *Immunol. Lett.* 162, 22–38. doi: 10.1016/j.imlet.2014.08.017
- Cutress, T. W., Ainamo, J., and Sardo-Infirri, J. (1987). The community periodontal index of treatment needs (CPITN) procedure for population groups and individuals. *Int. Dent. J.* 37, 222–233.
- Darveau, R. P. (2010). Periodontitis: a polymicrobial disruption of host homeostasis. *Nat. Rev. Microbiol.* 8, 481–490. doi: 10.1038/nrmicro2337
- Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927–930.
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinforma.* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381
- Eke, P. I., Dye, B. A., Wei, L., Slade, G. D., Thornton-Evans, G. O., Borgnakke, W. S., et al. (2015). Update on prevalence of periodontitis in adults in the United States: NHANES 2009 to 2012. *J. Periodontol.* 86, 611–622. doi: 10.1902/jop.2015.140520
- Eke, P. I., Dye, B. A., Wei, L., Thornton-Evans, G. O., Genco, R. J., and CDC Periodontal Disease Surveillance workgroup: James Beck (University of North Carolina, Chapel Hill, USA), Gordon Douglass (Past President, American Academy of Periodontology), Roy Page (University of Washington) (2012). Prevalence of periodontitis in adults in the United States: 2009 and 2010. *J. Dent. Res.* 91, 914–920. doi: 10.1177/0022034512457373
- Griffen, A. L., Beall, C. J., Campbell, J. H., Firestone, N. D., Kumar, P. S., Yang, Z. K., et al. (2012). Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *ISME J.* 6, 1176–1185. doi: 10.1038/ismej.2011.191
- Haffajee, A. D., Cugini, M. A., Tanner, A., Pollack, R. P., Smith, C., Kent, R. L., et al. (1998). Subgingival microbiota in healthy, well-maintained elder and periodontitis subjects. *J. Clin. Periodontol.* 25, 346–353. doi: 10.1111/j.1600-051X.1998.tb02454.x
- Hajishengallis, G., Darveau, R. P., and Curtis, M. A. (2012). The keystone-pathogen hypothesis. *Nat. Rev. Microbiol.* 10, 717–725. doi: 10.1038/nrmicro.02873
- Hajishengallis, G., and Lamont, R. J. (2012). Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology. *Mol. Oral Microbiol.* 27, 409–419. doi: 10.1111/j.2041-1014.2012.00663.x
- Hajishengallis, G., Liang, S., Payne, M. A., Hashim, A., Jotwani, R., Eskan, M. A., et al. (2011). Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. *Cell Host Microbe* 10, 497–506. doi: 10.1016/j.chom.2011.10.006
- Holt, S., Ebersole, J., Felton, J., Brunsvold, M., and Kornman, M. (1988). Implantation of *Bacteroides gingivalis* in nonhuman primates initiates progression of periodontitis. *Science* 239, 55–57.
- Hong, B.-Y., Furtado Araujo, M. V., Strausbaugh, L. D., Terzi, E., Ioannidou, E., and Diaz, P. I. (2015). Microbiome profiles in periodontitis in relation to host and disease characteristics. *PLoS ONE* 10:e0127077. doi: 10.1371/journal.pone.0127077
- Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom. J.* 50, 346–363. doi: 10.1002/bimj.200810425
- Kirst, M. E., Li, E. C., Alfant, B., Chi, Y.-Y., Walker, C., Magnusson, I., et al. (2015). Dysbiosis and alterations in predicted functions of the subgingival microbiome in chronic periodontitis. *Appl. Environ. Microbiol.* 81, 783–793. doi: 10.1128/AEM.02712-14
- Kolenbrander, P. E., Palmer, R. J., Periasamy, S., and Jakubovics, N. S. (2010). Oral multispecies biofilm development and the key role of cell-cell distance. *Nat. Rev. Microbiol.* 8, 471–480. doi: 10.1038/nrmicro2381
- Lang, N. P., Joss, A., Orsanic, T., Gusberty, F. A., and Siegrist, B. E. (1986). Bleeding on probing. A predictor for the progression of periodontal disease? *J. Clin. Periodontol.* 13, 590–596.
- Li, Y., He, J., He, Z., Zhou, Y., Yuan, M., Xu, X., et al. (2014). Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. *ISME J.* 8, 1879–1891. doi: 10.1038/ismej.2014.28
- Page, R. C., and Eke, P. I. (2007). Case definitions for use in population-based surveillance of periodontitis. *J. Periodontol.* 78, 1387–1399. doi: 10.1902/jop.2007.060264
- Park, O.-J., Yi, H., Jeon, J. H., Kang, S.-S., Koo, K.-T., Kum, K.-Y., et al. (2015). Pyrosequencing analysis of subgingival microbiota in distinct periodontal conditions. *J. Dent. Res.* 94, 1–7. doi: 10.1177/0022034515583531
- Petersen, P. E., and Ogawa, H. (2012). The global burden of periodontal disease: towards integration with chronic disease prevention and control. *Periodontol.* 2000, 15–39. doi: 10.1111/j.1600-0757.2011.00425.x
- R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Reiss, K., Dragano, N., Ellert, U., Fricke, J., Greiser, K. H., Keil, T., et al. (2014). Comparing sampling strategies to recruit migrants for an epidemiological study. Results from a German feasibility study. *Eur. J. Public Health* 24, 721–726. doi: 10.1093/eurpub/cku046
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09
- Shi, B., Chang, M., Martin, J., Mitreva, M., Lux, R., Klokkevold, P., et al. (2015). Dynamic changes in the subgingival microbiome and their potential for diagnosis and prognosis of periodontitis. *mBio* 6:e01926-14. doi: 10.1128/mBio.01926-14
- Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C., and Kent, R. L. (1998). Microbial complexes in subgingival plaque. *J. Clin. Periodontol.* 25, 134–144.
- Socransky, S. S., Haffajee, A. D., Teles, R., Wennstrom, J. L., Lindhe, J., Bogren, A., et al. (2013). Effect of periodontal therapy on the subgingival microbiota over a 2-year monitoring period. I. Overall effect and kinetics of change. *J. Clin. Periodontol.* 40, 771–780. doi: 10.1111/jcpe.12117
- Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 63, 411–423. doi: 10.1111/1467-9868.00293

- Warnes, G., Bolker, B., and Bonebakker, L. (2009). gplots: various R programming tools for plotting data. *R Package Version 2*.
- Wichmann, H.-E., Kaaks, R., Hoffmann, W., Jöckel, K.-H., Greiser, K. H., and Linseisen, J. (2012). The German National Cohort. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 55, 781–787. doi: 10.1007/s00103-012-1499-y
- Yost, S., Duran-Pinedo, A. E., Teles, R., Krishnan, K., and Frias-Lopez, J. (2015). Functional signatures of oral dysbiosis during periodontitis progression revealed by microbial metatranscriptome analysis. *Genome Med.* 7, 27. doi: 10.1186/s13073-015-0153-3
- Zaura, E., Keijser, B. J. F., Huse, S. M., and Crielaard, W. (2009). Defining the healthy “core microbiome” of oral microbial communities. *BMC Microbiol.* 9:259. doi: 10.1186/1471-2180-9-259
- Zimmermann, H., Hagenfeld, D., Diercke, K., El-Sayed, N., Fricke, J., Greiser, K. H., et al. (2015). Pocket depth and bleeding on probing and their associations with dental, lifestyle, socioeconomic and blood variables: a cross-sectional, multicenter feasibility study of the German National Cohort. *BMC Oral Health* 15:7. doi: 10.1186/1472-6831-15-7
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Boutin, Hagenfeld, Zimmermann, El Sayed, Höpker, Greiser, Becher, Kim and Dalpke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.