UCLA
UCLA Electronic Theses and Dissertations

Title
A Comparative Community Analysis of the Periodontitis Microbiome

Permalink
https://escholarship.org/uc/item/6sj6k714

Author
Chang, Michaela

Publication Date
2012

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
Los Angeles

A Comparative Community Analysis of the Periodontitis Microbiome

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science

in Oral Biology

by

Michaela H. Chang

2012

ABSTRACT OF THE THESIS

A Comparative Community Analysis of the Periodontitis Microbiome

by

Michaela H. Chang

Master of Science in Oral Biology
University of California, Los Angeles, 2012
Professor Huiying Li, Co-chair
Professor Renate Lux, Co-chair

Metagenomics using high-throughput DNA sequencing technologies is becoming a common approach for microbial community characterization. However, not many studies have compared the traditional 16S rRNA clone library approach with the new metagenomic approaches. In this study, we analyzed the community structure of the subgingival microbiome of periodontitis before and after initial therapy. Two approaches were employed in parallel - metagenomic shotgun sequencing and the traditional 16S rRNA clone library method using Sanger sequencing. This allowed us to compare the two approaches. DNA was extracted from subgingival plaque samples obtained from sites with chronic periodontitis in four systemically healthy subjects prior to and 4-6 weeks following initial therapy. For metagenomics sequencing, 100 bp paired-end reads were generated using Illumina sequencing platforms. Two analysis methods were used in analyzing the metagenomic sequencing data. First, the 16S rRNA reads

were extracted and aligned against SILVA rRNA database for taxonomic assignment. Second, the metagenomic reads were mapped against a database of microbial reference genomes for genome identification. For clone library analysis, the 16S rRNA gene was amplified and cloned. Plasmid inserts were sequenced bidirectionally and the taxa were assigned using the Microbiome Utilities Portal of Broad Institute. Statistical evaluation of the results demonstrated that the two sequencing approaches revealed consistent microbial community structure of the oral microbiome. The three analysis methods based on the metagenomic sequencing data and clone library data uncovered the presence of a diverse bacterial community in the subgingiva of periodontitis with little detection bias of the major oral microorganisms. The three analyses demonstrated consistent overall decreases in *Porphyromonas, Neisseria* and *Treponema*, and increases in *Fusobacterium, Veillonella, Prevotella* and *Streptococcus* after treatment. We observed microbial community shift from disease to resolution of periodontitis at genus level, which could allow the use of a few microbial markers in periodontitis diagnosis and monitoring.

The thesis of Michaela H. Chang is approved.

Ki-Hyuk Shin

Huiying Li, Committee Co-Chair

Renate Lux, Committee Co-Chair

University of California, Los Angeles

2012

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This thesis would not have been possible without the opportunity and direction given by Dr. Susan Kinder Haake, who has been my mentor, role model and greatest support. Words cannot express the profound impact she has made in my life, and I am truly blessed to have learned from such an amazing person.

I would like to express my deepest appreciation to my committee co-chair, Dr. Renate Lux for encouraging me to finish the thesis and for her time spent in reviewing my drafts. I would like to thank Dr. Huiying Li for her willingness to become a part of my committee and discussions during our microbiome meetings. I would also like to thank Dr. Ki-Hyuk Shin for his patience and warm advice during the preparation of this thesis.

I would like to extend my gratitude to Baochen Shi for his analysis of the metagenomic sequencing data, Shuta Tomida for his valuable inputs and Sorel Fitz-Gibbon for the clone library sequencing analysis. A special thank you to Jane Park for her assistance in preparing the numerous clones for sequencing. I would also like to thank everybody in the Haake and Lux/Shi labs for their moral support and heart-felt words of encouragement.

In addition, I would like to acknowledge The Genome Institute at Washington University for their scientific contribution. The studies described in this thesis were supported by NIH/NIDCR PHS Grants RC1DE020298 and 1R01DE021574, and NIH PHS Grant No. U54HG004968.

INTRODUCTION

Bacterial cells are estimated to exceed the number of body cells by a factor of ten in the human body (Zaura *et al.* 2009) and have a plethora of functions that are essential for the well being of the host. However, disruption of the healthy homeostatic condition, normally sustained by the indigenous bacteria, can lead to the emergence of pathogenic microorganisms and result in diseases. Bacterial communities of the oral cavity have been linked to a variety of systemic conditions (Beck *et al.* 1996; Berbari *et al.* 1997; Offenbacher *et al.* 1998; Dodman *et al.* 2000; Paster *et al.* 2001). Indigenous microorganisms compose distinct communities in various parts of the oral cavity such as the teeth, gingival sulcus, attached gingiva, tongue, cheek, lip, hard and soft palate (Dewhirst *et al.* 2010), and the balance between health and disease can be determined by the community of bacteria at each site. The oral microbiome is a complex population with over 600 unique taxa (Dewhirst *et al.* 2010), and a comprehensive understanding of microbial interactions can give essential knowledge about the contribution of the oral microbiota to health and disease (Jenkinson and Lamont 2005).

Recognizing the importance of the microbiota for the function of the human host, the Human Microbiome Project (HMP) was initiated by the National Institutes of Health (NIH) to characterize a core human microbiome in health, and to initiate studies on its variance in disease (Avila *et al.* 2009; Turnbaugh *et al.* 2009). This initiative aims to sequence a total of 3,000 reference genomes for submission to the public database (http://www.hmpdacc.org). Currently, 811 reference genome projects have been completed world-wide. 16S rRNA gene and shotgun metagenomics sequencing was used to survey the microbial communities at important body sites such as the oral and nasal cavities, skin, and urogenital and gastrointestinal tracts with half of the sites (nine sites) residing in the oralpharyngeal region. Though our ability to characterize

bacterial community profiles at a number of different body sites has evolved dramatically over the years, it has been estimated that 35% of bacterial species present in subgingival biofilms are still not yet cultivated (Teles *et al.* 2011).

Initially, microorganisms have been categorized by their phenotype: morphology, physiological activities, cell-component structure and ecological niches (Hugenholtz and Pace 1996). With the development of agar media, culturing bacteria for microscopic studies allowed many pathogens to be identified (Wade 2011). However, bacterial classification via phenotypic characteristics did not provide in-depth information on the evolutionary relationships among microorganisms. Another issue was that a large majority of host-associated community-dwelling bacteria resist cultivation because essential molecules for survival that are normally obtained from the host- and bacterial community environment cannot be supplemented in culture (Paster *et al.* 2001; Handelsman 2004).

Molecular approaches of characterizing microbial diversity became effective with the discovery of conserved 16S rRNA gene sequences within prokaryotic ribosomes (Saiki *et al.* 1985; Woese 1987). The 16S rRNA gene sequence became the gold standard of bacterial phylogeny because it is present in all bacteria and it is evolutionarily conserved due to its essential function. Sequence variations in the 16S rRNA gene very likely correlate to evolutionary distance of the organisms. Sequence-based bacterial phylogeny and taxonomy became easier to study with PCR-based studies compared to previous culture-dependent classifications (Belda-Ferre *et al.* 2012). Accumulation of 16S rRNA studies has amassed to sequence databases that can be used to compare and classify uncultivated bacteria. The more variable parts of the 16S rRNA sequences among organisms aid in distinguishing different taxa, while the conserved regions serve as universal sequences for PCR primer binding and grouping

of similar taxa (Frank *et al.* 2008). Another strength of PCR is that DNA from uncultivable bacteria and clinical isolates from endocarditis (Goldenberger *et al.* 1997), urinary tract infections (Domann *et al.* 2003), and various oral infectious diseases (Dewhirst *et al.* 2010) among others can be directly used in 16S rRNA sequencing for characterization (Woo *et al.* 2008). 2,049 genera comprising 10,929 species (http://www.bacterio.cict.fr/number.html# total) have been published to date, largely due to the use of 16S rRNA gene amplification and sequencing. The recent development of metagenomics, the genetic analysis of an entire population of microorganisms (Handelsman 2004) through direct isolation of genomic DNA from environmental samples, will aid in characterizing microbial function in addition to providing an inventory of species. An overall refinement of DNA amplification and bioinformatics has enabled researchers to obtain more information about taxonomy as well as bacterial functions.

The emergence of large-scale cultivation-independent approaches to microbial identification and community profiling largely relied on advances in DNA sequencing technologies. The majority of investigations utilizing DNA sequencing to date are based on Sanger sequencing, which was introduced in 1977 (Sanger *et al.* 1977) and expanded in the 1990's to a semi-automated capillary electrophoresis platform (Swerdlow and Gesteland 1990; Karger 1996). The Sanger method requires two steps: amplification and separation of DNA fragments. The modern version of this method uses fluorescently-labeled ddNTPs instead of radioactively-labeled ones for the chain termination DNA replication process to produce a mixture of labeled DNA extension products of varying length that correspond to the position of the specifically labeled nucleotide. Automated cycle sequencing allows reactions to amplify the terminated fragments, which are then assorted by capillary electrophoresis. A genetic analyzer

detects the emitted fluorescence at different wavelengths and translates the information into the corresponding DNA sequence. Though capillary electrophoresis has removed the need for manual separation of fragments via gel electrophoresis, the labeling and amplification of products are still a physical process (Shendure and Ji 2008). Sanger sequencing is currently able to generate read-lengths of approximately 1,000 base pairs with a per-base raw accuracies of 99.999% (Shendure and Ji 2008), features that are advantageous for first time (*de novo*) sequencing of genomes.

The DNA Analyzers from Applied Biosystems (ABI), which relies on the semi-automated, Sanger-based capillary electrophoresis technique, were introduced as a newer generation sequencing technology. This system is improved from the traditional Sanger sequencing-by-synthesis (chain termination) approach, incorporating ddNTPs chain terminators to the template and subsequently sorting DNA fragments with gel electrophoresis. Through an automated cycle sequencing process, DNA polymerase and random additions of ddNTP chain terminators generate amplified DNA fragments of different lengths from just one sample of template DNA (Mardis 1999). In standard Sanger sequencing, four separate solutions are required for each ddNTP, which demand more manual work and template DNA. Capillary electrophoresis removes the need for gel casting by injecting a separation matrix into the capillary tubes before loading a sample. Each sample is then injected into a capillary tube, where voltage is applied to separate the fragments as they travel through and simultaneously excited by a laser beam. Each fluorescently-labeled ddNTP produces light at different wavelengths, which is projected on a graph with the sequence. The Applied Biosystems 3730 DNA Analyzers are equipped with either a 48- or 96-capillary system to achieve parallel sequencing of multiple

samples. This machine is able to operate unmanned for 48 hours and generate reads of approximately 1000 base pairs within 3 hours (http://products.appliedbiosystems.com).

Key developments in sequencing involved miniaturization and pyrosequencing technology (also known as 454 sequencing) that allows direct detection of the incorporated nucleotide instead of requiring physical separation via electrophoresis to determine the next base in the sequence (Rothberg and Leamon 2008). For electrophoresis, DNA must be already amplified in order to produce lengths of sequences usable for an accurate separation process, which limits miniaturization. However, pyrosequencing could be condensed to any volume and concentration of DNA that produces detectable levels of light (Rothberg and Leamon 2008). 454 employs its own amplification process to compensate for low concentrations of template (Williams *et al.* 2006). Conventional PCR generates greater amounts of chimeras and relies on high DNA template concentrations (Williams *et al.* 2006). By utilizing compartmentalized beads, fragments are segregated from each other to further prevent the production of chimeras. While Sanger sequencing relies on fluorescent labeling in a chain termination process, 454 detects the light that is produced when a pyrophosphate is released as a nucleotide is added to the complementary unpaired base of the DNA template. 454 sequencing is able to generate simultaneously around 1 million reads of 400-700 base pair read lengths at a lesser cost than traditional methods (Shendure and Ji 2008). Major limitations of this new technology are the shorter read-lengths and less accurate base-calls compared to Sanger sequencing (Shendure and Ji 2008). However, the potential of using short-read sequencing to assemble whole genomes, increasing variety of molecular methods that need assessment by high-throughput sequencing and the overall advancement in technology (Shendure and Ji 2008), have led to the improvement of DNA sequencing.

Following the 454 platform, Illumina introduced a series of effective next-generation sequencing technologies. Illumina's Genome Analyzer technology is also becoming one of the prominent players in high-throughput sequencing. Similar to other "next-generation sequencing platforms" like the 454 technology, the Genome Analyzer involves parallel sequencing. First, fragments of DNA template are perpendicularly attached to a flow cell by short adapter sequences. The flow cell has a lawn of primers that match up to the sequences of the adapters. The amplification process begins with the template DNA bending over to find a complementary primer on the flow cell, where nucleotides are added along the template to form a complementary strand of DNA. The double stranded bridges are separated and amplification continues with the single strands until dense clusters of DNA are made (Figure 1). Fluorescently-labeled ddNTP chain terminators are added to each cluster, and the incorporated base is detected by laser excitation. The incorporated base is then cleaved and a new set of ddNTPs is added for detection of the next base. The sequencing concept is similar to the sequencing-by-synthesis (chain termination) method used in ABI sequencing technology. Illumina's Genome Analyzer provides more accurate base-by-base sequencing because the natural competition of all four ddNTPs at each sequencing cycle reduces incorporation bias (www.illumina.com). The HiSeq 2000 platform is able to produce sequence reads of 100 base pairs long and generate 200 million reads per lane. Paired-end sequencing is available too, aiming to compensate for the shorter read lengths by providing as much information possible about a genomic sequence. In paired-end sequencing, both ends of a DNA molecule are sequenced (currently up to 150 base pairs) resulting in two reads and an unknown region between them. Using computer programs, the shorter reads can be assembled into longer reads and contigs with defined sequence gaps between the paired-end reads. This can be done by mapping the longer reads or contigs to a reference

sequence (Chaisson *et al.* 2009). Despite the shorter base pair reads, compared to capillary-based methods, Illumina sequencing is able to generate much more sequences with shorter time and less cost.

While next-generation sequencing methods are necessary for large scale metagenomics projects, caution must be taken when relying on new technologies until results can be proven with repeated studies and the errors have been accounted for. Artifacts have been discovered in the 454 technology, first next-generation platform. 454 sequencing can lead to artificial amplification of more than 15% of the original DNA templates (Gomez-Alvarez *et al.* 2009). The amplification is random and unbiased, thus causing misinterpretation and overestimation of gene or taxon abundance. Though most of these duplicates are filtered, another problem is that the metagenomics libraries are also composed of 11-35% artificially replicated sequences. Comparing sequences to a database with replicates will lead to inaccurate identification and classification of biomarkers. Illumina DNA fragments reads are shorter, thus it is more difficult to sequence an organism for the first time, especially if a related genome is not already present in the databases. Error rates for Illumina sequencing are shown to increase after each sequencing cycle, which is suspected to occur during the substitution of bases at the 3' end (Dohm *et al.* 2008). When the terminator group on the base is incompletely removed, the next base cannot be added (phasing) and causes the bases to be shifted by one position (Dohm *et al.* 2008). As these shifts accumulate over the cycles, the sequence of the DNA molecule becomes increasingly inaccurate. In addition, mistakes in base-calling can occur during the detection of the bases. Green and red lasers are used for detection of G/T and A/C, respectively, and incorporation of a particular base is determined by intensities of light detected through optical filters. However, the filters display limited ability to clearly differentiate the bases leading to complications in

sequencing (Kircher *et al.* 2011). Errors are estimated to be 20.7 and 14.2 per megabase for 454 and Illumina-based contigs, demonstrating that longer reads may have greater error rates (Qin *et al.* 2010). Next-generation technology offers vaster opportunities of sequencing previously unidentified bacteria, and resolving these drawbacks will possibly uncover more new microorganisms.

These recent advances in sequencing technology have made the unveiling the oral microbiome a reality and will allow attaining critical knowledge about the maintenance of a healthy mouth as well as pointing out harmful bacteria that disturb the normal oral environment. Specific key pathogens have been identified in major oral diseases like dental caries and periodontitis (Marsh 2010), but our understanding of the consortia of species that constitute a "healthy" versus "diseased" microbiome still remain limited. Both caries and periodontitis manifests from polymicrobial communities that form on saliva-coated surfaces of the mouth (Jenkinson and Lamont 2005). Pioneer organisms such as *Streptococcus, Veillonella* and *Neisseria* initially adhere to the saliva-covered surfaces of the oral cavity. These early colonizers provide an adhesive surface for further bacterial colonization to form the microbial community known as dental plaque (Jenkinson and Lamont 2005). Dental plaque can create a beneficial or detrimental effect depending on the balance and imbalance of major microorganisms. The decalcification of the tooth surface in caries is mediated by acid-producing bacteria, the most common culprit being *Streptococcus mutans*. However, the effect of *S. mutans* may be moderated by the presence of alkali-producing bacteria (Nascimento *et al.* 2009), reflecting the impact of the bacterial consortia rather than specific pathogens on disease production. Similarly, it has also been recognized that periodontal diseases result from polymicrobial communities (Jenkinson and Lamont 2005; Dewhirst *et al.* 2010). Advanced periodontitis and alveolar bone

loss are commonly associated with a bacterial consortium comprised of *Porphyromonas gingivalis, Tannerella forsythia* (previously known as *Bacteroides forsythus*)*, Treponema denticola* and *Fusobacterium nucleatum* (Moore and Moore 1994; Ruby and Barbeau 2002; Haffajee *et al.* 2008). Several of the pathogens responsible for the progression of periodontitis are also implicated in systemic diseases (Paster *et al.* 2001) such as endocarditis (Berbari *et al.* 1997), osteomyelitis (Dodman *et al.* 2000), preterm low birth weight (Offenbacher *et al.* 1998; Han *et al.* 2010), cardiovascular and cerebrovascular disease (Beck *et al.* 1996; Wu *et al.* 2000). Thus, the oral microbiome remains a pertinent area of research for the diagnosis and treatment of oral and systemic diseases.

The purpose of this study was to examine the subgingival metagenome associated with chronic periodontitis using a longitudinal human model and to compare different sequencing approaches to profile the microbial communities. Specifically, this investigation had ***two specific aims***:

Specific Aim 1

Comparison of the 16S rRNA clone library, 16S rRNA metagenomics and reference genome alignment in the detection of the genera in subgingival plaque samples from patients with chronic periodontitis.

Specific Aim 2

To describe a preliminary characterization of the microbial community composition in the subgingival microbiome before and after initial periodontal treatment of patients with chronic periodontitis.

MATERIALS AND METHODS

Clinical Procedures

Healthy subjects with generalized moderate to severe chronic periodontitis who consented were recruited for participation in this study, which was approved by the UCLA Institutional Review Board (UCLA IRB#11-002319). Subjects with a history of antibiotic therapy in the past six months and any history of smoking or diabetes were excluded from the study. Following a 24 hour period of no oral hygiene, subgingival plaque samples were taken from four sites in four subjects. The area to be sampled was isolated and dried, the supragingival plaque removed, and the subgingival sample obtained with a sterile curette (Hu-Friedy Mfg. Co., Inc., Chicago, IL). The sampled plaque was suspended directly in 300 µl of ATL buffer (Qiagen, Inc., Valencia, CA) containing 0.25 ml of 0.1 mm glass beads (BioSpec Products, Inc., Bartlesville, OK) and immediately transported to the laboratory. Clinical parameters of gingival index, recession, pocket depth and bleeding on probing were recorded. Conventional initial periodontal therapy was subsequently conducted, including scaling and root planing as well as oral hygiene instructions. The subjects were sampled at the same sites 4 to 6 weeks after completion of initial therapy, again following a 24 hour period of no oral hygiene.

DNA Extraction

DNA was extracted from the samples with the QIAamp DNA micro kit (Qiagen, Inc.) using a modified "Isolation of Genomic DNA from Tissues" protocol with the addition of bead beating for maximal bacterial cell lysis. All procedures were completed in a laminar flow hood (NuAire Inc., Plymouth, MN) with RNase-free materials. Briefly, the bead beating was conducted on the sample after addition of Proteinase K (final concentration 0.62 µg/µl) in a total volume of 650 µl of ATL buffer. The bead beating protocol consisted of 90 seconds of bead beating (Biospec

Products, Inc.) with 30 seconds rest periods every 30 seconds. The vials were then horizontally incubated in a C24 Incubator/Shaker (New Brunswick Scientific, Co., Edison, NJ) at 56 °C with shaking at 200 rpm for 1 hour. The supernatant was transferred to a new 1.5 ml microcentrifuge tube, vortexed for 10 seconds after the addition of 85 µl of ethanol and 500 µl of AL buffer (Qiagen, Inc.), transferred to a chilled QIAamp MinElute Column (Qiagen, Inc.), centrifuged and the flow-through discarded. The columns were processed following the original protocol. The DNA was eluted with 25 µl of EB buffer (Qiagen, Inc.) and the extracted metagenomic DNA was stored at -20 °C.

16S rRNA Amplification

A 1501 base pair region of the 16S rRNA sequence was amplified from the metagenomic template with universal primers (Frank *et al.* 2008, F. Dewhirst, personal communication) as listed in Table 2. All PCR amplifications were performed in a reaction mixture consisting of 10X AccuPrime PCR Buffer II (Invitrogen Corp., Carlsbad, CA), 10 µM each of Sense and Anti-sense primer (Invitrogen Corp.), 1.0 U AccuPrime *Taq* DNA Polymerase High Fidelity (Invitrogen Corp.), 11-13 ng of metagenomic DNA template and adjusted to 50 µl with distilled water. The amplifications were carried out on an MJ Mini Gradient Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA) with an initial denaturation at 94 °C for 2 minutes followed by amplification under the following conditions: denaturation at 94 °C for 30 seconds, annealing at 48°C for 30 seconds, and elongation at 68°C for 1 minute. A total of 30 cycles were performed. The relative molecular mass of the amplified products were examined by electrophoresis in a 0.8% agarose gel and visually compared to molecular mass standards (Invitrogen Corp., Carlsbad, CA) under UV illumination (Gel Doc XR, Bio-Rad, Hercules, CA).

Cloning Procedures

The 16S amplicons were transformed into *E. coli* DH5α cells (Invitrogen Corp.) according to the manufacturer's protocol. The transformants were recovered on Luria-Bertani (LB) agar plates with 50 μg/ml of kanamycin and incubated overnight at 37 °C. For each of the eight samples, 200 white clones, to ensure that at least 100 clones would be successfully sequenced, as well as 2 blue colonies lacking insert DNA as controls, were subcultured on LB agar plates with 50 μg/ml of kanamycin and incubated overnight at 37 °C.

Screening and Purification of DNA

Subcultured colonies were lysed by resuspension in 20 μl of 1X lysis buffer (20% w/v sucrose, 200 mM NaOH, 120 mM KCl, 10 mM EDTA, 0.5% SDS, 1 mg bromophenol blue and 97 ml distilled water), incubated at 56°C for 7 minutes, then chilled on ice for 5 minutes. The cells were centrifuged at 18,000 x g for 10 minutes. 15 μl of the resulting supernatant was run on a 0.8% agarose gel and screened for the presence of the insert under UV illumination (Gel Doc XR, Bio-Rad). Positive clones were determined based on the relative molecular mass of the plasmids isolated from white colonies as compared to the blue colonies lacking insert DNA. The plasmid DNA from the positive clones was purified by the QIAprep Spin Miniprep Kit protocol (Qiagen, Inc., Valencia, CA) for sequencing and submitted to either Beckman Coulter Genomics (Beverly, MA) or the DNA Analysis Facility at Yale University (New Haven, CT) for sequencing.

Construction of a 16S rRNA Clone Library

Plasmid inserts were sequenced bidirectionally on an ABI 3730*xl* sequencer (Applied Biosystems, Inc., Carlsbad, CA). Bases were identified (base calling) and quality scores were determined with Phred (Ewing *et al*. 1998) default parameters. Bidirectional reads were assembled and aligned to a core set of NAST-formatted sequences (rRNA16S.gold) using

AmosCmp16Spipeline and NAST-ier, all from the Microbiome Utilities Portal (http://microbiomeutil.sourceforge.net/) of the Broad Institute (Cambridge, MA). Suspected chimeras were identified using ChimeraSlayer (Broad Institute). Sequences with at least 90% bootstrap support for a chimeric breakpoint were removed from further analysis. 16S sequences were compared to the Broad Institute's 16S rRNA GOLD database, as recommended (F. Dewhirst, personal communication), using BLAST (Altschul *et al*. 1990) and assigned to taxa based on the best match at greater than 97% nucleotide identity over at least 50 bases. Genera with at least one clone detected were included in abundance for comparison analyses.

Metagenomic Shotgun Sequence and Reference Genome Alignment Analyses

100 base paired-end reads were generated using Illumina GAIIx sequencing platform (Illumina, Inc., San Diego, CA) for each sample. Sequence reads were cleaned by filtering out any reads mapping to human DNA and by removing duplicate reads, which are artifacts of the sequencing process. In addition, reads with low compositional complexity or low quality were removed prior to analysis. Paired-end sequence reads were aligned against the SILVA rRNA database (SSU+LSU, release 104) using Bowtie (Langmead *et al.* 2009). Up to three mismatches were allowed in the alignment. We required that the paired-end reads align to the same reference sequence in different strand directions and the distance between the two is within 1Kb. In the case that the sequence read had the best hits to multiple organisms, we assigned evenly a fraction score to each aligned reference rRNA sequence with a sum of 1. The total number of genera found in each sample was calculated based on the rRNA sequences of the best hits and normalized by the total number of reads. Paired-end reads were additionally aligned to a concatenated reference genome scaffold database, which consists of 131 archaeal strains over 97 species, 326 lower eukaryotes over 326 species, and 1751 bacterial strains over 1253 species

(http://www.hmpdacc.org/HMREFG/). The depth of coverage for individual samples was normalized to 1 million reads and the relative abundance was calculated based on the total number of reads aligned per sample. Depth of coverage refers to the average number of times a base is represented in reads (https://ibi.uchicago.edu/ projects/sequencing_coverage_depth_ analysis/index.html). Genera with at least 1% abundance were included for comparison analyses.

Statistical Analysis

Abundance values of all methods were tested with non-parametric Kruskal-Wallis to determine significance of detected genera using Prism 5 (GraphPad Software, La Jolla, CA). In the case of significant difference ($p < 0.05$), subsequent pairwise Mann-Whitney U tests were performed between methods to determine if the difference exists between specific pairs of analyses or if the difference is a general one between all approaches.

RESULTS

Clinical resolution of sites after initial periodontal therapy

To compare different sequencing approaches in their detection of subgingival microbial community composition in periodontitis, subgingival plaque samples were collected from four systemically healthy patients with generalized moderate to severe chronic periodontitis before and after therapy for a total of eight samples. The gingival index, pocket depth (mm), recession of gums (mm) and the attachment level (mm) of the four pre-treatment (Pre-Tx) and four post-treatment (Post-Tx) sites were measured to assess the state of periodontal disease (Table 1). Average clinical probing depths in millimeters were 5.8 ± 0.5 and 3.8 ± 0.5 for the Pre- and Post-Tx sites, respectively. Average attachment level measurements in millimeters were 6.3 ± 1.3 for the Pre-Tx sites and 4.8 ± 1.0 for the Post-Tx sites. The average gingival index decreased from 1.3 ± 0.5 mm to 0 ± 0 mm in the sites after treatment. Average recession measurements in millimeters were 0.3 ± 0.7 and 1.6 ± 1.0 for Pre- and Post-Tx sites, respectively.

Summary of oral bacterial community analysis with different analysis methods

The bacterial community structure of the samples was analyzed using two types of sequencing technologies: the DNA Analyzer from ABI and the Genome Analyzer from Illumina. First, the 16S rRNA clone library (CL) from each sample was constructed using PCR and sequenced using the DNA Analyzer. The obtained 16S rRNA sequences were aligned to the 16S rRNA GOLD database (http://www.genomesonline.org) for taxon identification. Second, the DNA from each sample was subjected to whole genome shotgun sequencing using Illumina GAIIx. The Illumina reads obtained were analyzed in two ways: the 16S rRNA gene sequences were extracted and aligned against the 16S rRNA SILVA database (MG; http://www.arb-silve.de) for taxon identification; and the whole genome shotgun sequences were aligned against

a database of reference genomes (RG; http://www.hmpdacc. org/HMREFG/) for taxon and gene identification. All taxonomic identifications in this study were performed at the genus level to be more robust. The data obtained from these three analyses were examined for potential differences in identification of prevalent genera between the approaches (Table 2).

Genera with at least 1% relative abundance were included in the community analysis for all three data sets (Table 2). For the CL analysis, an average of 102 clones was analyzed and an average of 14 genera was detected for each sample. The ten most prevalent ("Top 10") genera (Table 6) were represented by 92% of the total number of clones. The metagenomics data using Illumina sequencing produced an average of 207 million original reads which were reduced to 38 million reads after applying the filtering procedures described in the materials and methods. These reads allowed for the identification of an average of 14 genera per sample in the MG data set, and separate alignment to the reference genomes (RG) resulted in the detection of an average of 13 genera in each sample. In both MG and RG analyses, approximately 84% of the filtered reads matched to the ten most common genera per sample. All genera identified in the CL data were present in both RG and MG data sets with the exception of *Anaeroglobus*, which was only detected in one sample of the CL set representing about 1% of the identified clones (Figure 3). In summary, the detection of prominent genera (Table 6) was comparable among all three approaches.

Statistical evaluation of the bacterial communities detected by the different methods

The significance of the apparent discrepancies in specific genera detection between the three methods evaluated in this study was assessed using non-parametric statistical analysis with the Kruskal-Wallis test. Genera that exhibited a difference in detection by the approaches (Figure 2; Table 7) were tested to assess if the differences were a result of significant detection biases

present among the methods. First, all eight Pre- and Post-Tx samples derived from the various patients were combined to investigate if the three detection methods exhibited a similar level of identification for the most prevalent genera. The null hypothesis was that there is no significant difference in the detection of genera between the three methods. The Kruskal-Wallis *p*-values of all tests, with the exception of *Actinomyces,* yielded a value greater than 0.1 to accept the null hypothesis (Table 3). Next, a pairwise Mann-Whitney U test between pairs of methodologies was performed to determine where the detection bias for *Actinomyces* occurs (Table 3). While the comparison of MG-RG complied with the null hypothesis, the analyses between CL-MG and CL-RG resulted each in a *p*-value less than 0.1. This rejected the null hypothesis and indicated a possible detection bias of *Actinomyces* between the CL and each of the metagenomics analyses (MG, RG).

Subsequently, the eight samples were separated into Pre- and Post-Tx groups for a more detailed examination of a possible detection bias between methods (Table 4). Therefore, Kruskal-Wallis tests were completed independently for the before and after treatment groups for the seven most prevalent genera along with *Actinomyces*. Again, the null hypotheses assumed that there was no significant difference in the proportion of genera before or after treatment for the approaches. With the exception of *Actinomyces* the *p*-values were greater than 0.1 for all genera in the Pre-Tx group and thus accepted the null hypothesis. Consistent with the initial findings for the combined samples, the pairwise Mann-Whitney U tests between pairs of methods indicated differential detection of *Actinomyces* for CL compared to MG and RG before therapy only (Table 4), while the two metagenomics-based approaches produced no significant difference for this genus. In the Post-Tx samples, the comparison of genera detection by the

different methods resulted in *p*-values greater than 0.1 for all to accept the null hypothesis that genus detection between the approaches was similar.

To further examine whether detection biases were present within individual approaches and where they occurred, Mann-Whitney U tests were carried out between the Pre- and Post-Tx samples of each method (Table 5). The null hypothesis was that there is no significant difference of genera detection between the Pre- and Post-Tx samples within a given methodology. With the exception of *Actinomyces* and *Porphyromonas* in the CL data, all genera yielded *p*-values greater than 0.1 to accept the null hypothesis. Since the samples obtained from subject C4 contained a very large proportion of *Porphyromonas*, an additional Mann-Whitney U test was completed with *Porphyromonas* in the CL method without considering this subject's data (Table 12). This was done to examine whether C4's contribution of the unusually high amount of *Porphyromonas* introduced the bias detected by the previous Mann-Whitney U test. The resulting *p*-value of 0.1000 indicated that the null hypothesis was valid in the detection of *Porphyromonas* among the rest of the samples. Thus, the only significant difference in detection was confirmed for *Actinomyces*.

Proportion of overall genera detected by the approaches

Since statistical analyses suggested that all methods compared in this study did not exhibit any significant difference in the detection of genera (with the exception of *Actinomyces* in the Pre-Tx group analyzed with CL), the obtained data sets were used for preliminary assessment of the effect of treatment on the bacterial community composition in the subgingiva. The proportion of common genera in the individual Pre- and Post-Tx samples was calculated from the CL, MG and RG datasets (Figure 3; Table 8). The diversity of genera detected in the samples was compared. In general, more variability was evident in the genera such as

Fusobacterium, Porphyromonas and *Prevotella* that were present in greater proportions. Even though statistical analysis confirmed that the sequencing approaches tested in this study do not significantly differ in the detection of most genera present in the tested samples, there are clearly differences among the three methods. The MG and RG approaches were, not surprisingly, most consistent with each other because they were derived from the same metagenomics sequence data sets. For example, *Actinomyces* and *Treponema* were detected more frequently using metagenomics sequence analyses; whereas *Fusobacterium* was detected more frequently in the CL analysis.

Relative abundance of genera indicate similar trends in all methods

The number of genera detected was comparable among individual samples and methods, and the relative abundance of genera in the Pre- and Post-Tx groups combined revealed that the same general trends of detection of community changes were apparent among the approaches (Figure 4; Table 7). For instance, decreases in the proportion of *Porphyromonas, Neisseria* and *Treponema* after periodontal therapy were consistently identified with each of the analyses used. Also, an overall increase in the proportion of *Fusobacterium, Prevotella* and *Veillonella* after treatment was identified with all of the approaches. These trends are supported by the statistical evaluation, which indicated that there was no significant bias in the detection of these genera among the methods used (Tables 3-5, 12). The increase in *Fusobacterium* appeared to be higher in the CL analysis than the MG and RG datasets; however, this difference is not statistically significant and could be due to the low sampling depth of CL method. Similarly, even though the increases observed in *Corynebacterium* in response to treatment were more pronounced in MG and RG approaches, statistics shows that this difference is not significant. In addition, *Actinomyces* was consistently detected in the MG and RG analyses but not CL method.

Statistical tests suggest that a detection bias in *Actinomyces* is present between the CL method and metagenomics sequence analyses.

The relative abundance of the genera in all the samples was calculated to examine the level of detection of individual genera among the three approaches (Figure 2; Table 9). *Porphyromonas, Streptococcus, Leptotrichia, Dialister, Gemella* and *Parvimonas* were uniformly detected in all the analyses. Genera that were detected in greater proportion such as *Fusobacterium, Prevotella, Treponema, Neisseria, Capnocytophaga* and *Veillonella* exhibited more apparent variation in detection between the three methods. However, these discrepancies are not influenced by a detection bias as indicated in the statistical evaluation (Table 3). Table 9 lists the relative abundance of common genera among the methods. These values are derived from the average amount of genera present in each sample rather than an indication of detection efficiency since statistical analyses infer that there is no detection bias among the methods.

Figure 5 and Table 10 show the relative abundance of phyla calculated for the Pre- and Post-Tx groups in all approaches. These phyla were identified in greater proportion than others. Notably, genera belonging to the phyla Bacteroidetes, Fusobacteria and Firmicutes were detected more frequently than other phyla (Table 10). The distribution of prevalent genera in each phylum is illustrated in Figure 6. The eight most common genera of each phylum are represented, while the rest of the genera were sorted into the 'other' category. For example, *Prevotella* and *Porphyromonas* make up at least 33% of the total Bacteroidetes in all methods. 57% of the Fusobacteria was composed of *Fusobacterium* in the CL data, whereas the detection of *Fusobacterium* comprised at least 35% of the phylum in the MG and RG analyses, respectively.

Effect of periodontal treatment on subgingival community profile

The changes in the relative abundance of genera between Pre- and Post-Tx groups were determined to examine the effect of treatment on the subgingival community profile. On the phylum level, there were consistent increases in the proportions of Firmicutes and Actinobacteria in the after treatment groups for all methods (Table 10). Fusobacteria was increased in the Post-Tx samples in the CL analysis, but did not differ greatly in the MG and RG analyses. *Prevotella, Streptococcus, Selenomonas, Capnocytophaga* and *Campylobacter* also showed general increase in the Post-Tx samples. The percentage of increase, and decrease, relative to the total abundance for individual samples was calculated for each method (Table 11). The genera showing the greatest changes are listed.

Decreases in the phyla Bacteroidetes, Proteobacteria and Spirochaetes were exhibited in the after treatment groups for all analyses (Table 10). These phyla include the genera *Porphyromonas, Neisseria* and *Treponema,* which decreased in proportion after treatment (Figure 7; Table 11). *Porphyromonas* showed the greatest decrease due to the profuse abundance of *Porphyromonas* in subject C4. This bias was confirmed by the previous statistical evaluation (Table 5, 12). Subsequent calculations, without inclusion of subject C4's data, indicated that *Porphyromonas* showed the second greatest decrease after *Neisseria* (Table 13). The main, if not the only, genus detected for the phylum Spirochaetes was *Treponema* (Figure 6), thus the relative change in Spirochaetes after treatment is a reflection of the loss of *Treponema*. Other genera that display a decrease in the Post-Tx groups albeit in lower proportions than the other common genera are *Gemella, Dialister* and *Leptotrichia* (Table 11). Though the changes observed by the different approaches were variable in proportion, statistical analysis confirmed that those apparent disparities between methods were not significant.

DISCUSSION

In this study, different analysis approaches were tested for their utility to identify the microbial oral communities present in diseased periodontal pockets and assess the potential community shift induced after successful periodontal therapy. A sizable number of studies have been dedicated to obtain a comprehensive inventory of the microbiota present in the oral cavity. A recent stringent analysis of existing 16S rRNA databases by Dewhirst *et al.* (2010) placed bacteria found in the oral cavity into 619 confirmed taxa and identified over 400 additional taxa that could be added to the oral microbiome upon confirmation. Accumulation of this information has become possible through advances in PCR and DNA sequencing technologies that have facilitated the characterization of microorganisms directly from environmental samples. These cultivation-independent approaches have allowed acquiring the type of data sets necessary for a comprehensive identification and analysis of microorganisms involved in the transition from health to disease. The advancement of sequencing technologies is also improving the power of uncovering formerly unclassified microbes in the human body. Rapid and cost-effective next-generation sequencing is answering the call for the rise of large-scale investigations of the human microbiome. Since these next generation sequencing methods are based on technologies that are very different from the more traditional clone library approaches that have been applied to generate the majority of analyses of oral microbial communities in health and disease (Aas *et al.* 2005; Paster *et al.* 2001; Kumar *et al.* 2006; Bik *et al.* 2010), it is necessary to perform an inclusive investigation of sequencing platforms to identify potential differences in the methods and the best approach for optimal results. In this study, a 16S rRNA clone library (CL) was compared to an Illumina sequencing platform based metagenomics approach in which 16S rRNA specific reads (MG) as well as reference genome alignment (RG) were used to identify the

genera in subgingival plaque samples from patients with periodontitis. Statistical evaluation of these different types of analyses for genera detection is necessary to assess the differences of the methods in characterization of the subgingival microbiome.

A comparison of genera detection between the three methods (Table 2) was conducted. Since there was an average of 102 clones per sample, matching one clone to a genus would represent 1% of that genus in the total sample and percentages less than one would essentially indicate zero to one clone. Thus, a cut-off value of 1% relative abundance was applied to the CL approach and the MG and RG methods, to be comparable. Though the methodologies are different from each other, they were able to detect a consistent number of prevalent genera in samples. The main difference among all methods was in the detection of *Actinomyces* (Figure 3), while all other apparent differences were found not to be statistically significant (Table 3-5) thus ruling out the possibility that these inconsistencies arose from detection biases within the methods. The sample sizes were too small to determine a normal distribution; normality tests with small sample populations would lead to inconclusive information about the distribution, thus, non-parametric tests were chosen for statistical analysis. The Kruskal-Wallis test was performed on all the samples for the three methods to see if any significant bias influenced how each genus was detected (Table 3). Generally, there was no significant bias in the detection of genera among methods with the exception of *Actinomyces*. Further Kruskal-Wallis and pairwise Mann-Whitney U tests indicated that the difference occurred between the Pre-Tx samples of CL-MG and CL-RG (Table 4). This could be due to the lack of detection of *Actinomyces* in the CL analysis and implies that the MG and RG approaches could be more adequate in detecting *Actinomyces* than the CL approach. Subsequent detailed statistical analysis of the Pre-and Post-Tx samples with the Mann-Whitney U test indicated that there was a significant bias in the

detection of *Porphyromonas* in the CL data. The significant bias present in the detection of *Porphyromonas* was due to subject C4's unusually high contribution to the genus (Figure 3). Because of random selection, the clones chosen for subject C4 may have happened to contain greater amounts of *Porphyromonas* than the other subjects' clones. This bias was accounted for, by repeating the Mann-Whitney U test without C4's Pre- and Post-Tx samples, when the statistics revealed that there was no significant bias of *Porphyromonas* detection between the remaining samples. A Mann-Whitney U comparison between the Pre- and Post-Tx samples for *Actinomyces* is not valid since it was not detected at all by the CL method. The low sampling depth and unequal number of clones in the CL analysis, may have introduced some degree of random variability (Kumar *et al.* 2006) in the outcome.

Since the statistical evaluation concluded that there is no bias in genera detection among the methods (with the exception of *Actinomyces* in the CL approach) and that the apparent differences are not significant, the data derived from this study could be further analyzed for the assessment of the core microbiome and shift of genus distribution after periodontal therapy. Furthermore, the ten most abundant genera identified were similar for all three approaches, even though the MG and RG analyses detection multiple-fold more genera per sample than the CL approach albeit in low proportions. While an obvious criticism could be that only 102 clones per samples were utilized to create the 16S rRNA clone library, other studies have demonstrated that an even lower number of clones were sufficient to identify the same core microbiome as a study using over a thousand clones per sample. Aas *et al.* (2005) and Bik *et al.* (2010) used an average of 57.5 and 1029 clones each sample, respectively, to investigate the normal bacterial diversity in oral health. Both studies found comparable major genera such as *Streptococcus, Gemella, Abiotrophia, Granulicatella, Rothia, Neisseria* and *Prevotella*, confirming that small clone

number does not necessarily introduce a bias in determining the abundant genera. Not surprisingly, and comparable to the differences that were observed in this study between the extensive metagenomics analyses and the relatively limited CL analysis, the study analyzing the larger clone number also produced a more extensive microbial spectrum. Similar to the study presented here, both these investigations also discovered inter-individual differences in microbial makeup.

Though the CL method seems sufficient in establishing a core oral microbiota consistent with the MG and RG methods, addressing issues with the CL approach may reveal additional information about the bacterial profile. In addition to effects of the relatively small clone number analyzed by the CL analysis on the assessment of the oral microbiome diversity discussed above, the amplification used in PCR in this approach may have introduced a bias in the detection of certain genera. The phylum Actinobacteria, which includes *Actinomyces,* is Gram-positive bacteria with a high G+C content that had low detection rates in previous studies (Munson *et al.* 2002; Munson *et al.* 2004; Wade 2011) as well as this study (Figure 3). The lower yield of Actinobacteria may be due to the combination of primer design and amplification conditions. The annealing temperatures of primers are important in the amplification of a comprehensive genera profile. An annealing temperature of 48 °C was used for the amplification of 16S rRNA gene according to Frank *et al.* (2008). This study utilized a lower temperature based on the assumption that it accommodates a diversity of primer-binding site sequences, while avoiding a high volume of chimeras. However, other studies have commonly used annealing temperatures of 55 °C or higher (Paster *et al.* 2001; Aas *et al.* 2005; Bik *et al.* 2010). Genera with high G+C content are thought to be better amplified with higher annealing temperature because sequences with high G+C have a low efficiency of dissociation from the template (Ishii and Fukui 2001).

An additional caveat in the amplification of genera with high G+C content is the observation that polymerases tend to prematurely terminate when they encounter regions of high G+C (Henke *et al.* 1997; Wade 2011). Overcoming these amplification biases against these types of bacteria is important in obtaining a more relevant picture of the microflora present in the mouth using methods like the CL approach in this study that require a PCR amplification step.

In the CL and other published 16S metagenomics studies using next-generation sequencing (Lazarevic *et al.* 2009; Kuczynski *et al.* 2012), the sequencing targets are amplicons of whole or certain hypervariable regions of 16S rRNA. The metagenomics approach in this study did not employ any amplification steps, but rather used the short reads generated by Illumina sequencing to align to the respective databases. This bypasses the ambiguities introduced by inefficient polymerase activities and the formation of PCR artifacts such as chimeras, allowing for a more comprehensive assessment of the bacterial diversity in a given sample on the genus-level and can play an essential role in identifying uncharacterized species that are less prevalent (Table 2). MG and RG analyses offer shorter sequencing reads, but the millions of reads per run are able to provide extensive sequencing information and more coverage by overlapping reads while reducing error rates (Lazarevic *et al.* 2009). For large scale metagenomics projects, the Illumina sequencing also provides a more time-efficient and cost-effective method for analyzing the microbial diversity in the samples compared to the capillary electrophoresis approach used for the CL analysis.

Since the statistical assessment indicated that there was no significant bias in genera detection among the three approaches, a preliminary characterization of the subgingival microbiome and treatment induced shifts in its composition were drawn using the CL, MG and RG data sets. Subgingival plaque was taken from four subjects with chronic periodontitis before

and after treatment for a total of eight samples. The clinical resolution of periodontitis was measured by the decrease of tissue inflammation and probing depths, from 5.8 $\pm$ 0.5 to 3.8 $\pm$ 0.5 mm. Despite this apparent resolution of disease, the microbiota in the subgingival pockets may never return to the state of naïve pockets. Thus, the oral microbiome in the after treatment samples does not represent absolute health because resolved sites have a high tendency for re-infection unless professional treatment is monitored frequently (Nyman *et al.* 1975; Melcher 1976; Cobb 2008). A previous study has assessed patients with periodontitis over a two year time period to observe the bacterial shift from disease to health, and reflected that subjects still may have been undergoing clinical change after the 2-year time point (Kumar *et al.* 2006). The samples taken after this 2-year resolution period contains the same common genera as the microbiome of the Post-Tx samples in this study, which were acquired after 3 months. Thus, the microflora detected after treatment in this study could likely be a description of the bacterial profile during the shift to resolved periodontitis rather than representing a healthy naïve oral microbial community. Though the changes in periodontium are evident of clinical resolution, the oral bacterial community may require more time to reach stability.

Overall, several genera including *Porphyromonas, Streptococcus, Leptotrichia, Dialister, Gemella* and *Parvimonas* were consistently detected in all the samples by the three analyses (Figure 2). Detection levels generally varied albeit not significantly for the more abundantly detected genera such as *Fusobacterium, Prevotella, Treponema, Neisseria* and *Veillonella*. Previous studies have recognized *Streptococcus, Gemella, Neisseria, Prevotella, Capnocytophaga, Treponema, Porphyromonas gingivalis* and species of *Fusobacterium* to be the most prominent microorganisms in the oral cavity associated with health and disease (Paster *et al.* 2001; Aas *et al.* 2005; Paster *et al.* 2006; Lazarevic *et al.* 2009; Bik *et al.* 2010). Many studies

have indicated that the subgingival microbiome differs between health and the various periodontitis-associated disease states (Sbordone and Bortolaia 2003; Aas *et al.* 2005; Kumar *et al.* 2006). In this study, the specific consortia of bacteria demonstrating increase and decrease, relative to treatment states, are similar to those formerly reported (Jenkinson and Lamont 2005; Paster *et al.* 2006; Bik *et al.* 2010).

The genera most prevalently detected and associated with periodontal disease were *Porphyromonas, Treponema, Prevotella* and *Fusobacterium* (Figure 3). *Porphyromonas gingivalis* (*P. gingivalis*) has been implicated in periodontitis, and increased levels of *P. gingivalis* have been found in disease active sites (Socransky and Haffajee 1992; Griffen *et al.* 1998). However, the current understanding is that one bacterial species alone cannot be attributed to cause disease; periodontal disease most likely results from the interactions of microbial consortia with one another (Jenkinson and Lamont 2005). The association of the Gram-negative anaerobes, *P. gingivalis, Treponema denticola* (*T. denticola*) and *Tannerella forsythia*, also known as the 'red-complex', has been linked with severe periodontal disease and was found to be present in decreased numbers in healthy subjects (Kinane 2001; Paster *et al.* 2006; Bik *et al.* 2010). Simonson *et al.* (1992) reported that the manifestation of *T. denticola* required the pre-existence of *P. gingivalis* in the subgingival plaque, and a physical interaction between the two species was evident since the heat treatment of *P. gingivalis* eliminated the coaggregation of the two (Onagawa *et al.* 1994). It is possible that *T. denticola* and *P. gingivalis* share a synergistic relationship, and their interaction influences pathogenesis in the periodontium (Simonson *et al.* 1992). Interestingly, *Porphyromonas* and *Treponema* were prominently decreased in the Post-Tx samples for all the approaches (Figure 7). This trend supports the finding above, in that therapy may have been effective in removing these bacteria that are linked

to disease. In addition, the reduction in *Porphyromonas* could have facilitated the simultaneous decline in the abundance of *Treponema.*

Both *Prevotella* and *Fusobacterium* are organisms of the orange complex, of which members were discovered to inhabit the subgingival plaque in greater proportions than other bacteria (Socransky *et al.* 1998), and found to increase in periodontitis. *Prevotella* is a suspected periodontopathogen, as it is commonly recovered in patient samples with periodontal disease (Dorn *et al.* 1998; Marcotte and Lavoie 1998). *Fusobacterium nucleatum* (*F. nucleatum*) plays a central role in periodontal disease due to its ability to bridge early Gram-positive and late Gram-negative bacteria in plaque biofilms (Kolenbrander and London 1993; Signat *et al.* 2011). Thus, it is speculated that elimination of *F. nucleatum* may reduce the adherence of pathogenic bacteria such as *P. gingivalis* and *T. denticola* (Socransky *et al.* 1998) that contribute to periodontitis. *Fusobacterium* frequently increase in sites of periodontal disease, but it is also one of the most common genera isolated from plaque isolated from healthy sites (Kolenbrander and London 1993). In this study, *Prevotella* and *Fusobacterium* were increased in their proportion after treatment in the three methods (Figure 7), which appears to contrast the results reported in prior findings (Socransky *et al.* 1998). Also, though they are especially detected in disease, they are considered moderate pathogens. When primary pathogens such as *Porphyromonas* and *Treponema* are mostly eliminated from periodontal pockets containing a myriad of bacteria, relative increase of moderate pathogens may seem to occur due to the reduction in total biomass. In actuality, because the resolved regions generally contain a much smaller total bacteria load than before treatment, a relatively low number of *Fusobacterium* and *Prevotella* can cause calculations of their proportions to appear far greater than in diseased sites. Thus, the actual

number of *Fusobacterium* and *Prevotella* is likely reduced in the Post-Tx samples compared to the Pre-Tx samples.

On the other hand, *Veillonella* and *Streptococcus*, which have been related to periodontal health, exhibited considerable increase after therapy (Figure 7) (Jenkinson and Lamont 2005; Kumar *et al.* 2005; Bik *et al.* 2010). These two genera are among the most commonly detected in the normal oral flora (Aas *et al.* 2005; Jenkinson and Lamont 2005). Kumar *et al.* (2006) discovered that *Veillonella* showed the greatest increase in improved health and decrease when disease conditions worsened. In this study, *Veillonella* also exhibited the highest increase in the Post-Tx samples. *Streptococci* are also generally linked to health, as they form the foundation of the microbiota in healthy gingiva (Stingu *et al.* 2008). It is plausible that some beneficial species of *Streptococcus* may act as barriers against pathogenic microbes inhabiting the subgingival pockets (Quirynen *et al.* 2001; Stingu *et al.* 2008). Particularly, *Streptococcus sanguinis* is a prominent organism that has been shown to display inhibitory effects on potential pathogens such as *Tannerella forsythia* and *Prevotella intermedia* (Stingu *et al.* 2008). Because they are located in the inner layers of the gingiva and the cell walls of Gram-positive bacteria are more difficult to lyse (Rogosa 1970; Chassy and Giuffrida 1980), it can explain why such a small number of them were detected in the samples in comparison to *Veillonella*; a Gram-negative bacteria (Table 11). Isolating bacteria, like *Veillonella,* on the superficial layers (Sbordone and Bortolaia 2003) are easier than obtaining *Streptococcus* from the base of deep pockets. It is also possible that the absence of pathogenic organisms after initial therapy allowed for more of these beneficial bacteria to colonize.

Besides the previously mentioned genera, several others displayed relative changes in relative abundance after therapy (Figure 7). Though less in proportion compared to the above genera, they were among the microorganisms that showed the greatest differences between Pre- and Post-Tx samples. Table 11 organizes the amount of increase and decrease of these genera. *Dialister* and *Leptotrichia* are suspected to play a role in periodontitis and other systemic diseases (Gundi *et al.* 2004; Kumar *et al.* 2005; Eribe and Olsen 2008), which is consistent with the overall decrease of the two genera in the Post-Tx samples. *Neisseria* and *Gemella* exist as commensals of the healthy periodontium (Aas *et al.* 2005; Kumar *et al.* 2005; Bik *et al.* 2010; Marri *et al.* 2010), yet decreases in the treated group. It is possible that these decreases can be attributed to the general loss of commensal bacteria after therapy (Marri *et al.* 2010), and studies in the future may reveal the exact interactions that are responsible for the decline of commensals in resolved states. The pathogenesis of *Campylobacter* (Savitt and Socransky 1984; Haffajee *et al.* 1998; Macuch and Tanner 2000), *Capnocytophaga* (Savitt and Socransky 1984; Macuch and Tanner 2000; Kumar *et al.* 2005), *Selenomonas* (Paster *et al.* 2001; Kumar *et al.* 2006) and *Corynebacterium* (Paster *et al.* 2001) are questionable because some studies identify some strains of these genera to be implicated in different stages of periodontitis and other strains to be health-associated. The changes in these low-abundant genera are more subjective to any fluctuations within subjects and samples, thus the information given by these genera should be further studied before accepting the conclusions. Prevalent genera, mentioned previously, can generally be categorized as health- or disease-associated because they embody species that are either predominantly commensal or pathogenic species. Thus, it is possible to monitor the prominent shifts associated with changes in periodontal health status with genus-level analysis.

It would be easy to explain a decrease in genera that are implicated in disease and an increase of bacteria that are health-associated. However, the oral microbiome of a disease-resolved cavity will have a different profile of bacteria than those that has never had disease. Once a periodontal pocket experiences disease, it will never fully recover to the state of a naïve periodontium (Melcher 1976; Cobb 2008). Other aspects to consider are the clinical criteria and amount of time for the resolution of disease. The Post-Tx samples may have been obtained while the periodontium was still undergoing a shift from disease to health and the Post-Tx microbiota presented may not be representative of the final profile of bacteria in resolved health. Thus, it is essential to assess the healing process of the gingiva and the bacterial stability at various time points after initial periodontal therapy. We plan to address this question in our future study. The microbial presence during the shift into health can offer knowledge about bacteria that play a role in mediating disease.

Recognizing the triggers of these bacterial shifts is essential for the development of diagnostic tests and therapeutic approaches for periodontal disease. Many factors, such as oral hygiene, host-associated biological modulators or introduction of new species of bacteria, may change a healthy microbiome into diseased (Kumar *et al.* 2006). A comparison of the oral microbiome between health and disease is an accessible method to evaluate those genera that determine the condition of the sub-gingiva. Though the CL approach is capable of species-level detection of bacteria, this study demonstrates that it is not required for assessment of the microbiota indicative of disease and resolved health because the comparison of the CL analysis with the metagenomics methods at the genus-level delivers reliable identification of the overall shift in genera at clinical resolution. Therefore, the use of next-generation sequencing will

become a great asset in determining the human oral microbiome in a time- and cost-efficient manner.

CONCLUSIONS

1. Statistical evaluation of the results demonstrated that there is little detection bias among the three analysis methods used in this study and that the differences in relative abundances of major genera are not significant among the methods except for *Actinomyces*, which was detected by metagenomics sequencing, but not by clone library method.

2. 16S rRNA clone library, 16S rRNA metagenomics and reference genome alignment sequencing analyses reveal the presence of a diverse bacterial community composition in the subgingiva of periodontitis.

3. Metagenomics sequencing based analyses were able to present a greater breadth of genera than the 16S rRNA clone library approach due to much greater sampling depth, but a very similar core oral microbiome was detected by all three methods.

4. There were consistent overall decreases in *Porphyromonas, Neisseria* and *Treponema*, and increases in *Fusobacterium, Veillonella* and *Prevotella* in the post-treatment groups detected by all approaches.

5. Genus level detection may be sufficient to monitor transition from disease to resolution of periodontitis thus allowing large scale analyses with the more cost-effective metagenomics approaches.

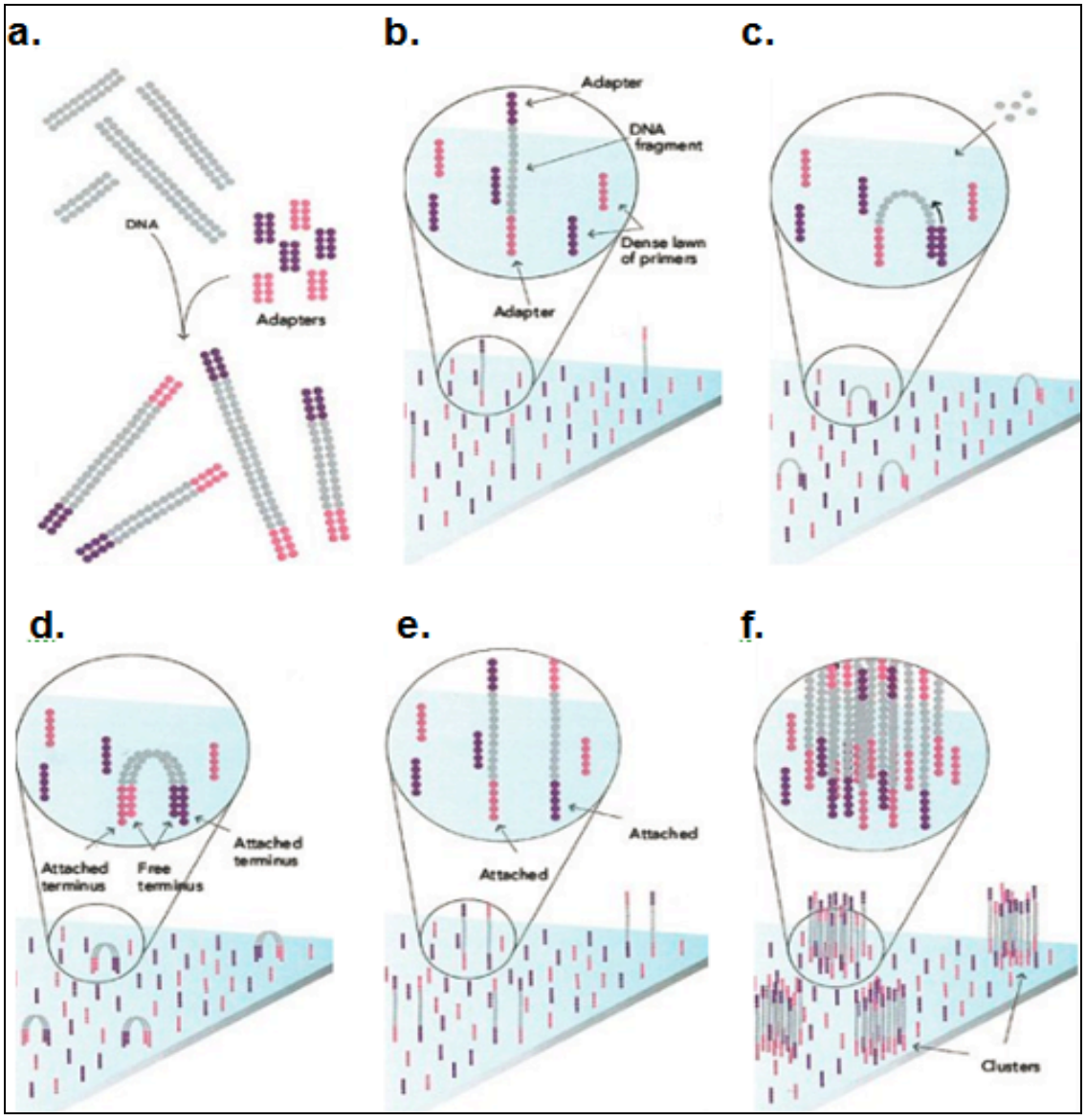

Figure 1. Illumina sample preparation and sequencing. (a,b) Short adapter sequences are ligated to DNA fragments and perpendicularly attached to a flow cell. The flow cell has a lawn of primers matching up to the adapter sequences. (c,d) Bridge amplification proceeds with the template DNA bending over to the complementary primer on the flow cell. Nucleotides are added along the template DNA to form a complementary strand of DNA. (e,f) Double stranded bridges are separated and amplification continues until dense DNA clusters are generated. Fluorescently-labeled ddNTP chain terminators are added to each cluster, and the incorporated base is detected by laser excitation. The incorporated base is then cleaved and a new set of ddNTPs is added for detection of the next base.

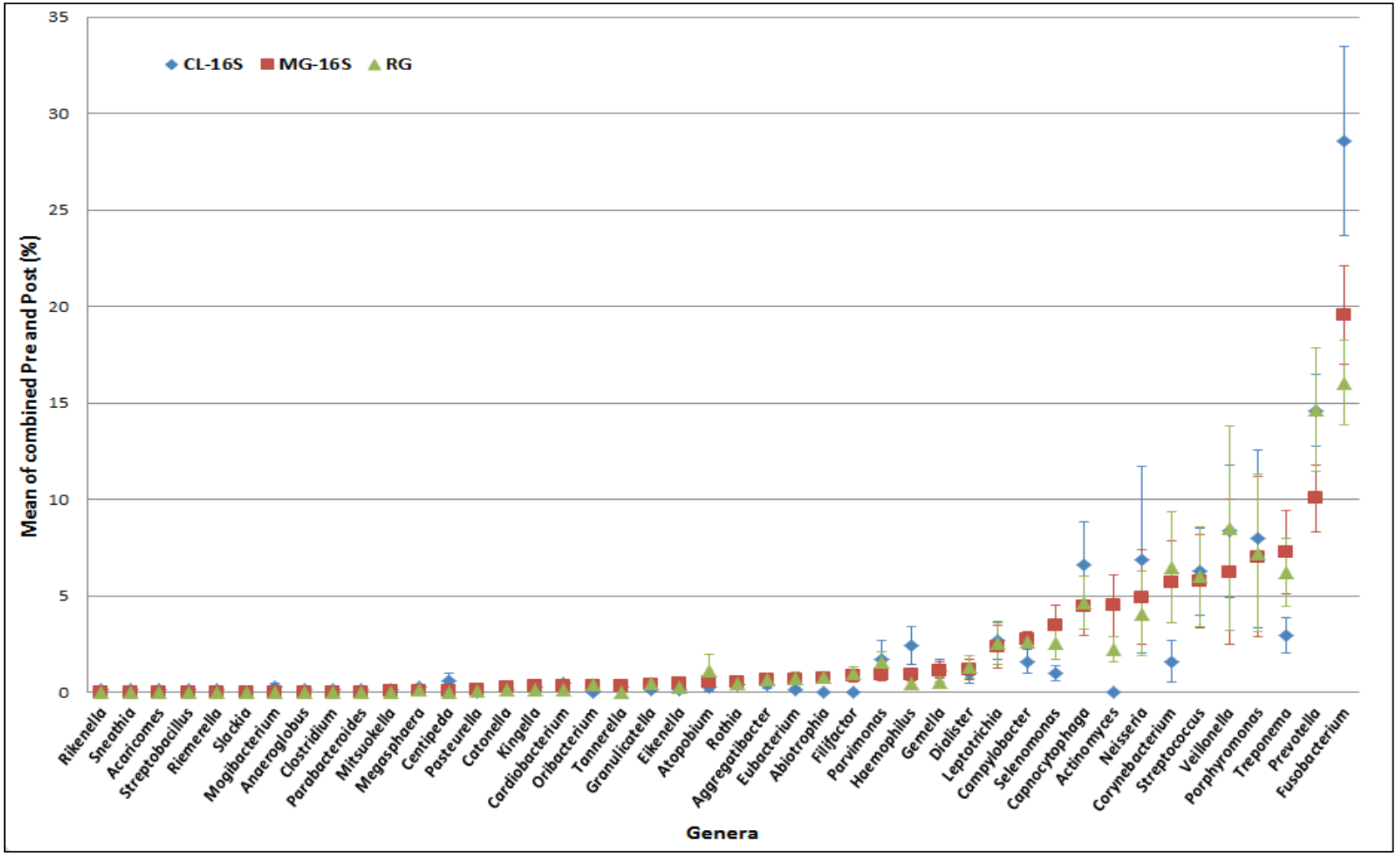

Figure 2. Relative abundance of genera detected by different methodologies. The relative abundances of individual genera were calculated by combining all samples for each of the analyses. Genera that had an abundance of 1% or greater in any of the methods are included. (CL-16S=16S rRNA Clone Library, MG-16S=16S rRNA Metagenomics, RG=Reference Genome Alignment)

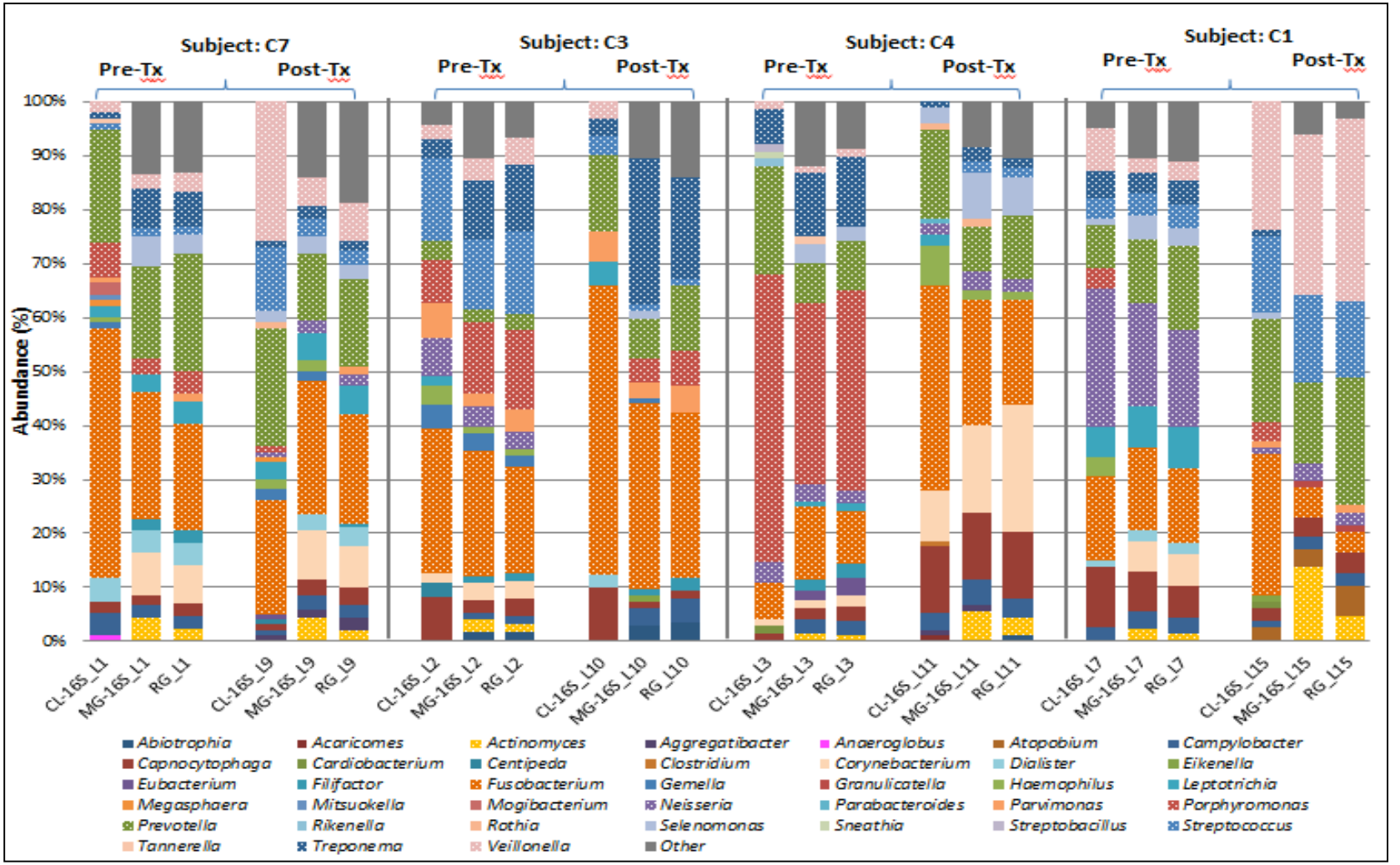

Figure 3. Proportion of Genera per sample for each method. The abundance and diversity of genera are compared in the Pre- and Post-Tx sample pairs. Genera with at least two clones detected were included in abundance graph for CL data. For metagenomics analyses, genera with at least 1% abundance, calculated from aligned reads, were included in the figure. Genera not meeting the cut off values were assigned to the 'Other' category (gray).

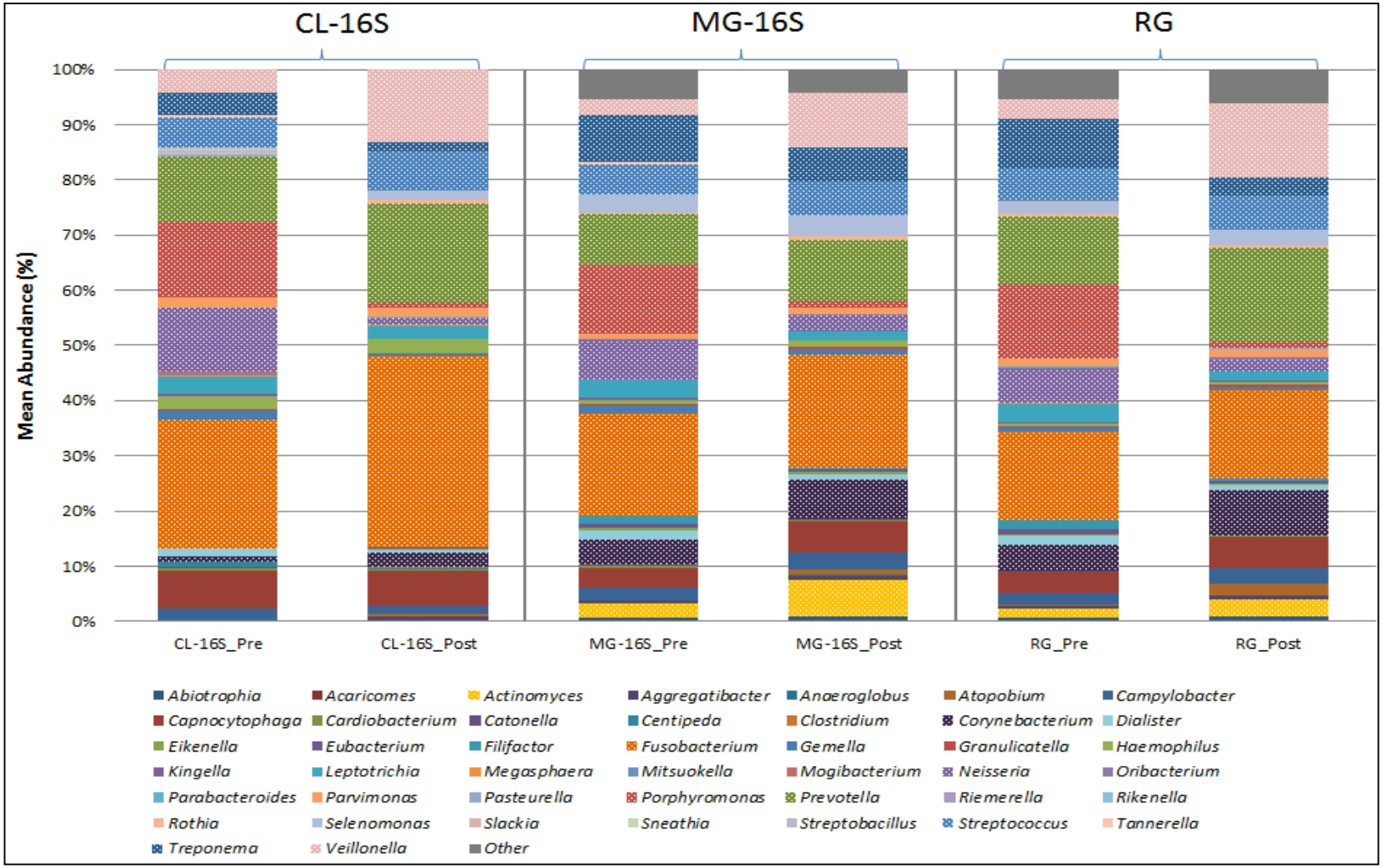

Figure 4. Relative abundances of genera in Pre- and Post-Tx groups for all three analyses. The mean abundance was calculated by combining all Pre-Tx samples for each method and also for all the Post-Tx samples. Genera inclusion was based on those chosen for Figure 3.

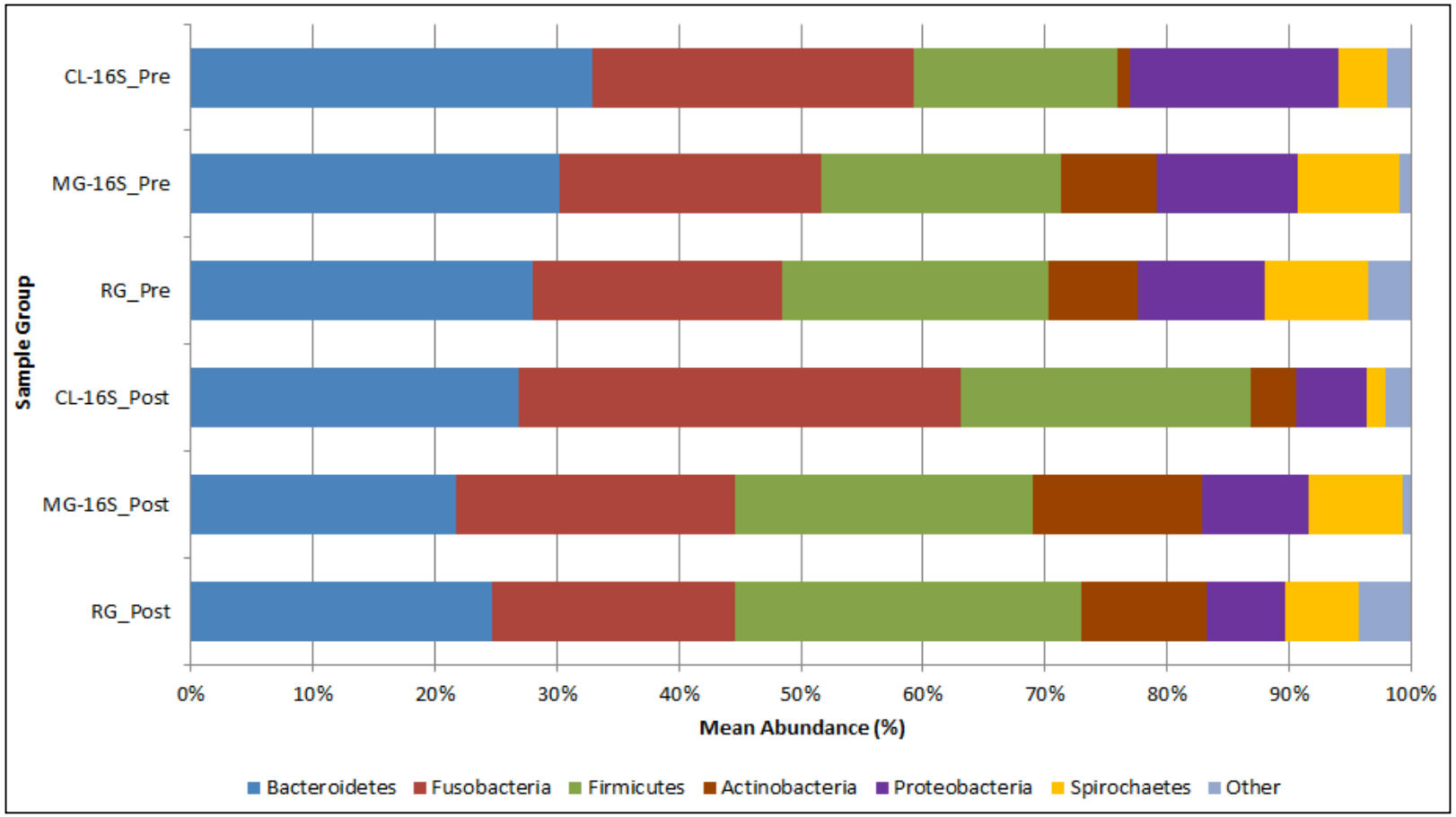

Figure 5. The relative abundances of phyla in Pre- and Post-TX groups for all three analyses. The mean abundances were calculated by combining all Pre-Tx samples of respective phyla for each method and also for all Post-Tx samples. Included in the analysis were the genera to six main phyla: Bacteroidetes (blue), Fusobacteria (red), Firmicutes (green), Proteobacteria (brown), Spirochaetes (purple) and Actinobacteria (yellow). Additional phyla were assigned to the 'Other' category (gray).

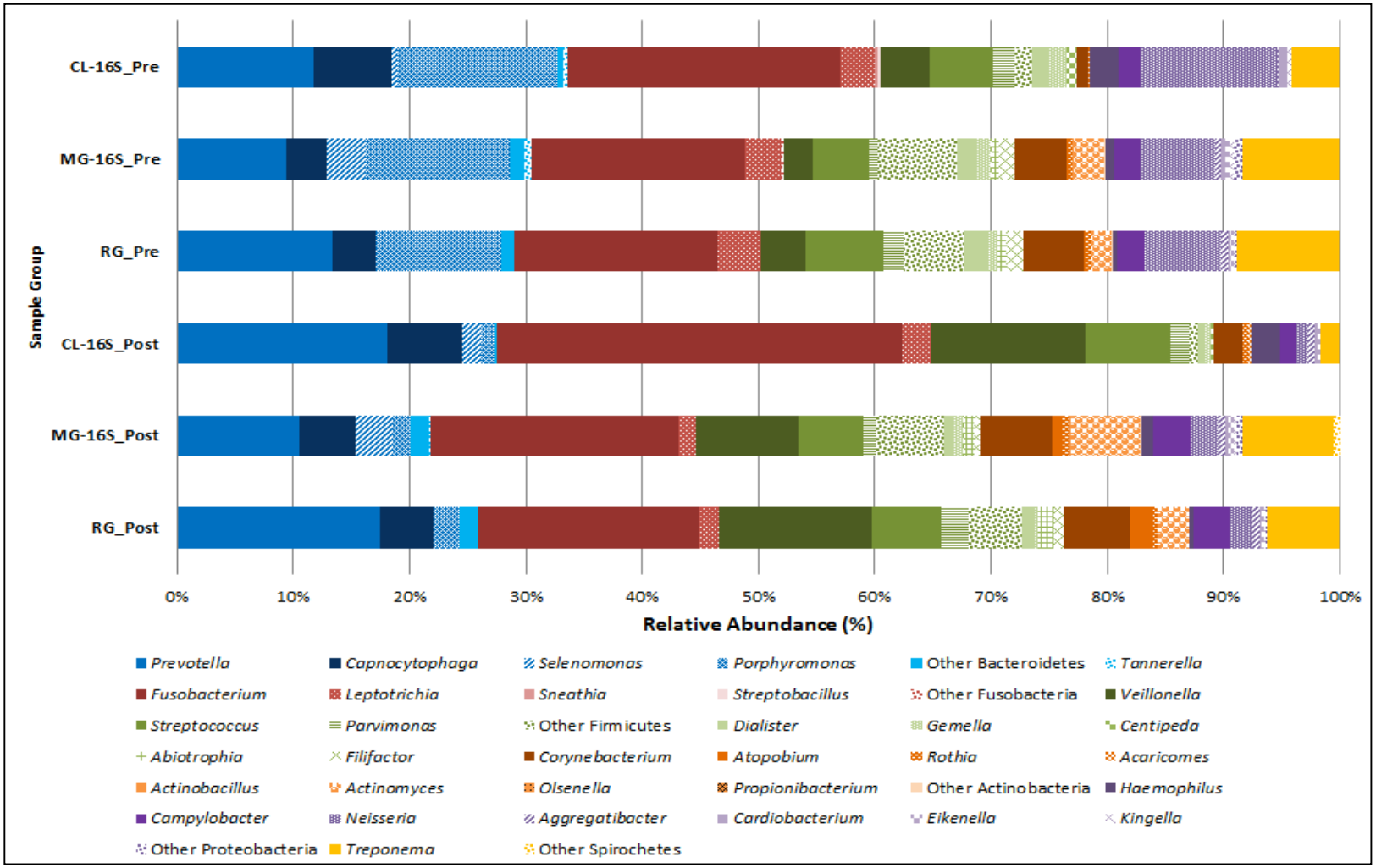

Figure 6. Distribution of genera in each phylum (Figure 4). The most abundant genera are displayed in the graph, whereas additional genera were assigned to the 'Other' category. Shades of blue=Genera in Bacteroidetes, red=Fusobacteria, green=Firmicutes, orange=Actinobacteria, purple=Proteobacteria and yellow=Spirochaetes.

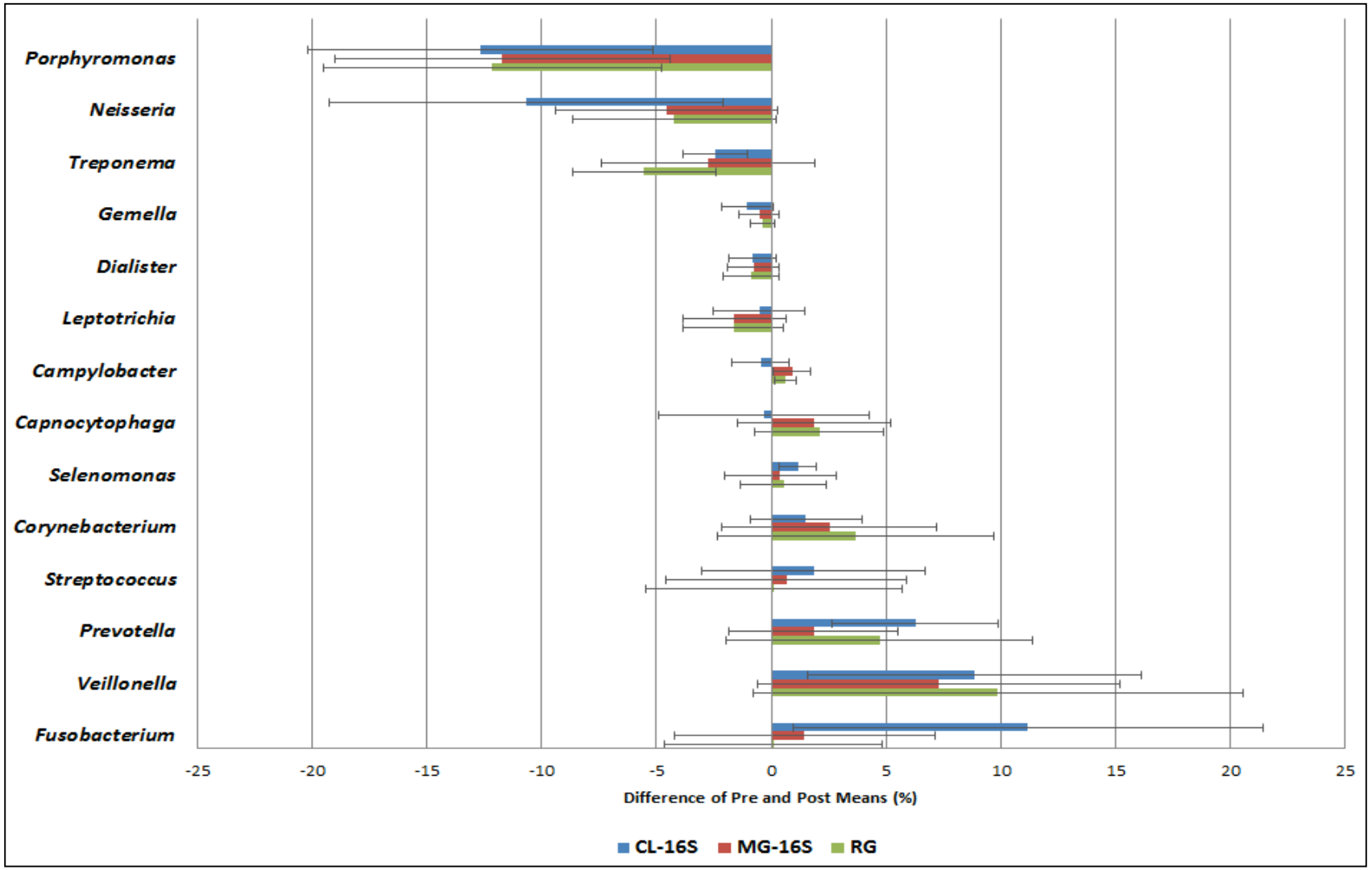

Figure 7. Effect of treatment on subgingival community profile. The difference between the mean abundance of all Pre-Tx and all Post-Tx samples of each genus, for all three analyses, were calculated. Change in relative abundance between Pre- and Post-Tx samples of the 14 most prevalent genera for each method are depicted.

	Gingival Index	Pocket Depth (mm)	Recession (mm)	Attachment Level (mm)
Pre-Tx	1.3 ± 0.5	5.8 ± 0.5	0.3 ± 0.7	6.3 ± 1.3
Post-Tx	0 ± 0	3.8 ± 0.5	1.6 ± 1.0	4.8 ± 1.0

Table 1. Clinical parameters of 8 samples (4 pairs) from 4 subjects Pre- and Post-Tx.
Sites were all interproximal and included 3 molars, 1 pre-molar, 1 cuspid and 2 incisors.

Data Summary										
Subject	C7		C3		C4		C1			
Sample Number	L1	L9	L2	L10	L3	L11	L7	L15	Average	Stdev
16S rRNA Clone Library Analysis:										
No. 16S rRNA clones analyzed	95	100	112	91	75	97	161	84	102	26
No. of Genera with >1% relative abundance	18	18	15	9	12	15	14	14	14	3
No. clones in Top 10 Genera	86	90	97	91	68	91	145	77	93	23
% relative abundance in Top 10 Genera	90%	90%	87%	100%	91%	94%	90%	92%	92%	4%
16S rRNA Metagenomics Analysis:										
# of Total original reads (million reads)	116	128	93	116	46	108	112	107	207	47
# of Total cleaned reads (million reads)	53	73	14	22	15	55	62	12	38	25
No. of Genera with >1% relative abundance	14	16	16	13	15	13	13	10	14	2
% relative abundance in Top 10 Genera	82%	77%	79%	86%	81%	87%	82%	94%	84%	5%
Reference Genome Analysis:										
# of Total original reads (million reads)	116	128	93	116	46	108	112	107	207	47
# of Total cleaned reads (million reads)	53	73	14	22	15	55	62	12	38	25
No. of Genera with >1% relative abundance	15	15	16	10	14	12	13	11	13	2
% relative abundance in Top 10 Genera	77%	71%	84%	86%	85%	87%	83%	96%	84%	7%
*16S Metagenomics and Reference Genome data produced by Baochen Shi										

Table 2. Summary of data statistics used for community analyses. About 100 clones were analyzed for the 16S rRNA clone library. 16S rRNA clone library sequences were compared to the Broad Institute's rRNA16S.gold database using BLAST. An average of 17.63 million cleaned reads was used to generate the data for metagenomics studies. 16S rRNA metagenomics sequences were aligned against the SILVA rRNA database. Paired end reads were additionally aligned to a concatenated reference genome scaffold. Subjects C# are listed with their respective Pre- and Post-Tx samples (L#). (Subject C7: Pre-Tx L1, Post-Tx L9; Subject C3: Pre-Tx L2, Post-Tx L10; Subject C4: Pre-Tx L3, Post-Tx L11; Subject C1: Pre-Tx L7, Post-Tx L15)

Genera	Methods	Statistical Test	*p* value	Significant Median Variation ($p < 0.05$)
Fusobacterium	ALL	Kruskal-Wallis	0.1291	No
Prevotella	ALL	Kruskal-Wallis	0.1795	No
Streptococcus	ALL	Kruskal-Wallis	0.994	No
Haemophilus	ALL	Kruskal-Wallis	0.1606	No
Selenomonas	ALL	Kruskal-Wallis	0.1153	No
Capnocytophaga	ALL	Kruskal-Wallis	0.7464	No
Neisseria	ALL	Kruskal-Wallis	0.743	No
Campylobacter	ALL	Kruskal-Wallis	0.1772	No
Corynebacterium	ALL	Kruskal-Wallis	0.1493	No
Porphyromonas	ALL	Kruskal-Wallis	0.8215	No
Treponema	ALL	Kruskal-Wallis	0.2541	No
Veillonella	ALL	Kruskal-Wallis	0.8531	No
Actinomyces	ALL	Kruskal-Wallis	0.0011	Yes
Genera	**Methods**	**Statistical Test**	***p* value**	**Significant Median Variation ($p < 0.05$)**
Actinomyces	CL and MG	Mann-Whitney U	0.0015	Yes
Actinomyces	CL and RG	Mann-Whitney U	0.0015	Yes
Actinomyces	MG and RG	Mann-Whitney U	0.1719	No

Table 3. **The statistical analysis of genera to evaluate the presence of detection bias between methods.** Kruskal-Wallis tests of the prevalent genera and *Actinomyces* were done on all approaches to examine whether the differences in detection are statistically significant. Null hypothesis is that there is no significant detection bias of genera between methods. Subsequent pairwise Mann-Whitney U tests, for *Actinomyces,* were done on the methodologies to assess the bias detected by the Kruskal-Wallis test.

Genera	Methods	Statistical Test	Pre_*p* value	Pre_Significant Median Variation ($p < 0.05$)	Post_*p* value	Post_Significant Median Variation ($p < 0.05$)
Veillonella	ALL	Kruskal-Wallis	0.6939	No	0.9206	No
Treponema	ALL	Kruskal-Wallis	0.1672	No	0.7031	No
Streptococcus	ALL	Kruskal-Wallis	0.925	No	0.7788	No
Prevotella	ALL	Kruskal-Wallis	0.6939	No	0.1183	No
Porphyromonas	ALL	Kruskal-Wallis	0.8126	No	0.9153	No
Neisseria	ALL	Kruskal-Wallis	0.6141	No	0.287	No
Fusobacterium	ALL	Kruskal-Wallis	0.6677	No	0.1738	No
Actinomyces	ALL	Kruskal-Wallis	0.0106	Yes	0.0717	No
Genera	**Methods**	**Statistical Test**	**Pre_*p* value**	**Pre_Significant Median Variation ($p < 0.05$)**		
Actinomyces	CL and MG	Mann-Whitney U	0.0211	Yes		
Actinomyces	CL and RG	Mann-Whitney U	0.0211	Yes		
Actinomyces	MG and RG	Mann-Whitney U	0.1143	No		

Table 4. **Assessment of detection bias between methods in a given clinical state.** Kruskal-Wallis tests of Pre-Tx and Post-Tx samples for all methods were performed to observe differences in genera detection before and after therapy. The null hypotheses are that there is no significant difference of genera abundance before treatment and after treatment between the methods. Subsequent pairwise Mann-Whitney U tests were completed to assess the bias detected in the Pre-Tx state for *Actinomyces.*

Genera	Statistical Test	CL_*p* value	CL_Signif.	MG_*p* value	MG_Signif.	RG_*p* value	RG_Signif.
Veillonella	Mann-Whitney U	0.4857	No	0.8846	No	0.8846	No
Treponema	Mann-Whitney U	0.1143	No	0.3429	No	0.3429	No
Streptococcus	Mann-Whitney U	1	No	0.8857	No	1	No
Prevotella	Mann-Whitney U	0.6857	No	0.6857	No	0.4857	No
Porphyromonas	Mann-Whitney U	0.0294	Yes	0.2186	No	0.2186	No
Neisseria	Mann-Whitney U	0.2454	No	0.5614	No	0.2454	No
Fusobacterium	Mann-Whitney U	0.4857	No	0.6857	No	0.6857	No
Actinomyces	Mann-Whitney U	< 0.0001	Yes	0.4857	No	0.4857	No

Table 5. **Statistical evaluation of genera detection within each method.** Pairwise Mann-Whitney U tests were performed with the four pre-treatment samples (L1-L7) and four post-treatment samples (L9-L15) to determine differences in detection before and after therapy within a method to further verify where the detection bias, found in Table 3 and 4, occurs. The null hypothesis is that there is no significant difference of genera between the pre and post-treatment samples within a given methodology.

CL-16S	MG-16S	RG
Capnocytophaga	*Actinomyces*	*Campylobacter*
Fusobacterium	*Capnocytophaga*	*Capnocytophaga*
Haemophilus	*Corynebacterium*	*Corynebacterium*
Leptotrichia	*Fusobacterium*	*Fusobacterium*
Neisseria	*Neisseria*	*Neisseria*
Porphyromonas	*Porphyromonas*	*Porphyromonas*
Prevotella	*Prevotella*	*Prevotella*
Streptococcus	*Streptococcus*	*Streptococcus*
Treponema	*Treponema*	*Treponema*
Veillonella	*Veillonella*	*Veillonella*

Table 6. Ten most prevalent genera for each methodology. For each method, the relative abundances (%) of each genus were calculated by combining all Pre- and Post-Tx samples. A minimum of 84% of the total clones/cleaned reads matched up to these genera for the individual approaches.

Genus	<CL-16S>	<MG-16S>	<RG>
Fusobacterium	28.6	19.6	16.0
Prevotella	14.6	10.0	14.7
Treponema	2.9	7.3	6.2
Veillonella	8.3	6.2	8.5
Corynebacterium	1.6	5.7	6.5
Neisseria	6.9	4.9	4.1
Actinomyces	0.0	4.5	2.2
Capnocytophaga	6.6	4.5	4.7
Selenomonas	1.0	3.4	2.6
Haemophilus	2.5	0.9	0.5
Campylobacter	1.6	2.8	2.6
Streptococcus	6.3	5.8	6.0
Porphyromonas	8.0	7.0	7.2

Table 7. Relative abundance (%) of prevalent genera based on methodology. The relative abundances were calculated by combining all Pre- and Post-Tx samples.

	C7 Pre-Tx			C7 Post-Tx		
Genus	**CL_L1**	**MG_L1**	**RG_L1**	**CL_L9**	**MG_L9**	**RG_L9**
Abiotrophia	0.0	0.3	0.4	0.0	0.4	0.5
Acaricomes	0.0	0.0	0.0	0.0	0.0	0.0
Actinomyces	0.0	4.3	2.3	0.0	4.2	1.9
Aggregatibacter	0.0	0.4	0.7	1.0	1.5	2.3
Anaeroglobus	1.1	0.2	0.0	0.0	0.0	0.0
Atopobium	0.0	0.2	0.4	0.0	0.3	0.6
Campylobacter	4.2	2.2	2.2	1.0	2.7	2.4
Capnocytophaga	2.1	1.8	2.5	1.0	3.1	3.3
Cardiobacterium	0.0	0.5	0.1	0.0	0.5	0.2
Catonella	0.0	0.1	0.0	0.0	0.1	0.0
Centipeda	0.0	0.1	0.0	1.0	0.1	0.0
Clostridium	0.0	0.0	0.1	0.0	0.1	0.1
Corynebacterium	0.0	8.1	7.0	0.0	9.0	7.8
Dialister	4.2	4.1	4.2	0.0	3.1	3.3
Eikenella	0.0	0.1	0.1	0.0	0.2	0.2
Eubacterium	0.0	0.8	0.6	1.0	0.6	0.4
Filifactor	0.0	2.1	2.3	0.0	0.5	0.6
Fusobacterium	46.3	23.4	19.9	21.0	24.8	20.4
Gemella	1.1	0.6	0.2	2.0	1.8	0.6
Granulicatella	0.0	0.2	0.2	0.0	0.5	0.5
Haemophilus	1.1	0.7	0.1	2.0	2.2	0.5
Kingella	0.0	0.1	0.0	0.0	0.2	0.1
Leptotrichia	2.1	3.4	4.1	3.0	4.9	5.3
Megasphaera	1.1	0.2	0.3	1.0	0.1	0.3
Mitsuokella	1.1	0.2	0.1	0.0	0.0	0.0
Neisseria	0.0	0.1	0.1	1.0	2.3	2.0
Parvimonas	1.1	0.7	1.3	0.0	0.9	1.5
Pasteurella	0.0	0.1	0.0	0.0	0.4	0.1
Porphyromonas	6.3	2.9	4.1	1.0	1.0	0.9
Prevotella	21.1	17.1	21.8	22.0	12.3	16.4
Rothia	0.0	0.3	0.3	1.0	0.4	0.4
Selenomonas	0.0	5.5	3.6	2.0	3.4	2.5
Sneathia	0.0	0.0	0.0	0.0	0.0	0.0
Streptococcus	1.1	1.5	1.5	12.0	3.2	2.8
Tannerella	1.1	0.8	0.0	0.0	0.4	0.0
Treponema	1.1	7.5	6.6	1.0	2.3	1.7
Veillonella	2.1	2.6	3.5	26.0	5.5	7.1
Other	2.1	6.9	9.7	0.0	7.4	13.6

Table 8. Proportion (%) of prevalent genera per sample for each subject. The percentage of each genus was calculated from total clones/cleaned reads aligned per sample and clones/reads that matched to individual genera per sample. CL, MG and RG denote each of the methods. (L1-L9=subject C7, L2-L10=subject C3, L3-L11=subject C4, L7-L15=subject C1)

	C3 Pre-Tx			C3 Post-Tx		
Genus	**CL_L2**	**MG_L2**	**RG_L2**	**CL_L10**	**MG_L10**	**RG_L10**
Abiotrophia	0.0	1.6	1.6	0.0	2.7	3.5
Acaricomes	0.0	0.0	0.0	0.0	0.0	0.0
Actinomyces	0.0	2.3	1.4	0.0	0.8	0.5
Aggregatibacter	0.0	0.9	0.5	0.0	0.2	0.2
Anaeroglobus	0.0	0.0	0.0	0.0	0.0	0.0
Atopobium	0.0	0.0	0.0	0.0	0.0	0.0
Campylobacter	0.0	1.4	1.6	0.0	3.3	4.2
Capnocytophaga	8.0	2.3	3.2	9.9	1.3	1.8
Cardiobacterium	0.9	0.0	0.0	0.0	0.0	0.0
Catonella	0.9	0.8	0.7	0.0	0.9	0.1
Centipeda	2.7	0.0	0.0	0.0	0.0	0.0
Clostridium	0.0	0.0	0.0	0.0	0.0	0.0
Corynebacterium	1.8	3.3	3.4	0.0	0.0	0.0
Dialister	0.0	0.1	0.2	2.2	0.0	0.1
Eikenella	0.0	0.9	0.6	0.0	1.2	0.6
Eubacterium	0.0	0.5	0.2	0.0	1.0	1.0
Filifactor	0.0	1.2	1.3	0.0	1.3	2.1
Fusobacterium	26.8	23.1	19.8	53.9	34.3	30.8
Gemella	4.5	3.2	2.1	0.0	1.0	0.5
Granulicatella	0.9	0.5	0.7	0.0	0.0	0.1
Haemophilus	3.6	1.2	1.2	0.0	0.0	0.1
Kingella	0.9	0.9	0.4	0.0	0.0	0.0
Leptotrichia	1.8	0.2	0.2	4.4	0.2	0.2
Megasphaera	0.0	0.0	0.0	0.0	0.0	0.0
Mitsuokella	0.0	0.0	0.0	0.0	0.0	0.0
Neisseria	7.1	3.8	3.0	0.0	0.5	0.3
Parvimonas	6.3	2.4	4.2	5.5	2.8	5.1
Pasteurella	0.0	0.1	0.2	0.0	0.0	0.0
Porphyromonas	8.0	13.3	14.7	0.0	4.5	6.5
Prevotella	3.6	2.3	3.1	14.3	7.4	11.9
Rothia	0.9	0.5	0.7	0.0	0.0	0.0
Selenomonas	0.0	0.1	0.1	0.0	1.3	0.9
Sneathia	0.0	0.0	0.0	0.0	0.0	0.0
Streptococcus	15.2	13.1	15.2	3.3	1.4	1.4
Tannerella	0.0	0.2	0.0	0.0	0.1	0.0
Treponema	3.6	11.0	12.3	3.3	27.0	18.7
Veillonella	2.7	3.3	5.2	3.3	0.7	0.9
Other	0.0	4.9	2.3	0.0	6.1	8.6

Table 8. (cont.) Proportion (%) of prevalent genera per sample for each subject.

	C4 Pre-Tx			C4 Post-Tx		
Genus	**CL_L3**	**MG_L3**	**RG_L3**	**CL_L11**	**MG_L11**	**RG_L11**
Abiotrophia	0.0	0.1	0.2	0.0	0.9	1.1
Acaricomes	0.0	0.0	0.0	1.0	0.0	0.0
Actinomyces	0.0	1.4	1.0	0.0	5.4	3.1
Aggregatibacter	0.0	0.2	0.1	1.0	1.3	0.9
Anaeroglobus	0.0	0.0	0.0	0.0	0.0	0.0
Atopobium	0.0	0.1	0.1	0.0	0.0	0.0
Campylobacter	0.0	2.4	2.6	3.2	4.7	3.8
Capnocytophaga	1.3	2.3	2.6	12.4	12.23	12.3
Cardiobacterium	1.3	0.4	0.1	0.0	0.5	0.3
Catonella	0.0	0.1	0.0	0.0	0.1	0.0
Centipeda	0.0	0.0	0.0	0.0	0.2	0.0
Clostridium	0.0	0.0	0.0	1.0	0.0	0.0
Corynebacterium	1.3	1.4	2.0	9.3	16.4	23.5
Dialister	0.0	0.5	0.6	0.0	0.0	0.0
Eikenella	0.0	0.3	0.2	0.0	0.3	0.2
Eubacterium	0.0	1.7	3.4	0.0	0.4	1.0
Filifactor	0.0	2.2	2.5	0.0	0.0	0.1
Fusobacterium	6.7	13.4	9.9	38.1	23.0	19.5
Gemella	0.0	0.3	0.1	0.0	0.2	0.1
Granulicatella	0.0	0.1	0.1	0.0	0.0	0.0
Haemophilus	0.0	0.1	0.1	7.2	2.0	1.4
Kingella	0.0	0.4	0.3	0.0	0.4	0.2
Leptotrichia	0.0	1.1	1.4	2.1	0.7	0.7
Megasphaera	0.0	0.0	0.1	0.0	0.0	0.0
Mitsuokella	0.0	0.0	0.0	0.0	0.1	0.0
Neisseria	4.0	3.2	2.5	2.1	3.6	2.3
Parvimonas	0.0	0.1	0.2	0.0	0.0	0.0
Pasteurella	0.0	0.0	0.0	0.0	0.3	0.3
Porphyromonas	53.3	33.5	37.0	0.0	0.2	0.2
Prevotella	20.0	7.4	9.3	16.5	8.2	11.7
Rothia	0.0	0.5	0.5	1.0	1.3	0.9
Selenomonas	0.0	3.7	2.4	3.1	8.6	7.1
Sneathia	1.3	0.0	0.0	0.0	0.0	0.0
Streptococcus	0.0	0.9	1.0	0.0	2.0	1.8
Tannerella	0.0	1.3	0.0	0.0	0.1	0.0
Treponema	6.7	11.7	13.2	1.0	2.7	1.8
Veillonella	1.3	1.3	1.5	0.0	0.6	0.7
Other	2.7	8.0	5.0	1.0	3.6	4.9

Table 8. (cont.) Proportion (%) of prevalent genera per sample for each subject.

	C1 Pre-Tx			C1 Post-Tx		
Genus	**CL_L7**	**MG_L7**	**RG_L7**	**CL_L15**	**MG_L15**	**RG_L15**
Abiotrophia	0.0	0.6	0.5	0.0	0.0	0.0
Acaricomes	0.0	0.0	0.0	0.0	0.0	0.0
Actinomyces	0.0	2.2	1.3	0.0	13.9	4.7
Aggregatibacter	0.6	0.4	0.7	0.0	0.0	0.0
Anaeroglobus	0.0	0.0	0.0	0.0	0.0	0.0
Atopobium	0.0	0.1	0.2	2.4	3.2	5.5
Campylobacter	2.5	3.2	2.9	1.2	2.4	2.2
Capnocytophaga	11.2	7.4	5.9	2.4	3.3	4.0
Cardiobacterium	0.6	0.6	0.2	1.2	0.0	0.0
Catonella	0.0	0.1	0.0	0.0	0.0	0.0
Centipeda	0.6	0.1	0.0	0.0	0.0	0.0
Clostridium	0.0	0.0	0.0	0.0	0.1	0.0
Corynebacterium	0.6	5.7	6.0	0.0	0.0	0.0
Dialister	1.2	2.1	2.2	0.0	0.0	0.0
Eikenella	0.0	0.5	0.3	1.2	0.3	0.1
Eubacterium	0.0	0.2	0.0	0.0	0.2	0.0
Filifactor	0.0	0.0	0.1	0.0	0.0	0.1
Fusobacterium	15.5	15.1	13.7	26.2	5.6	3.9
Gemella	0.6	0.9	0.3	0.0	0.7	0.3
Granulicatella	0.0	0.6	0.6	0.0	1.2	1.1
Haemophilus	3.7	0.8	0.2	0.0	0.0	0.0
Kingella	0.6	0.1	0.0	0.0	0.1	0.1
Leptotrichia	5.6	7.9	7.6	0.0	0.2	0.3
Megasphaera	0.0	0.0	0.0	0.0	0.0	0.1
Mitsuokella	0.0	0.0	0.0	0.0	0.0	0.0
Neisseria	25.5	19.2	17.9	1.2	3.3	2.4
Parvimonas	0.0	0.1	0.2	1.2	1.0	1.4
Pasteurella	0.0	0.1	0.0	0.0	0.0	0.0
Porphyromonas	3.7	0.8	0.5	3.6	0.0	0.0
Prevotella	8.1	11.7	15.7	19.1	15.1	23.8
Rothia	0.0	0.3	0.3	0.0	0.8	0.5
Selenomonas	1.2	4.4	3.3	1.2	0.4	0.4
Sneathia	0.0	0.0	0.0	0.0	0.0	0.0
Streptococcus	3.7	4.1	4.5	14.3	16.0	14.0
Tannerella	0.6	0.1	0.0	0.0	0.0	0.0
Treponema	5.0	3.9	4.3	1.2	0.0	0.0
Veillonella	8.1	2.8	3.6	23.8	29.7	33.8
Other	0.6	4.2	6.9	0.0	2.4	1.5

Table 8. (cont.) Proportion (%) of prevalent genera per sample for each subject.

Genus	<Pre-Tx _CL>	<Post-Tx _CL>	<Pre-Tx _MG>	<Post-Tx _MG>	<Pre-Tx _RG>	<Post-Tx _RG>
Abiotrophia	0.0	0.0	0.7	0.8	0.7	0.9
Acaricomes	0.0	0.3	0.0	0.0	0.0	0.0
Actinomyces	0.0	0.0	2.5	6.7	1.5	2.0
Aggregatibacter	0.2	0.5	0.5	0.8	0.5	0.9
Anaeroglobus	0.2	0.0	0.1	0.0	0.0	0.0
Atopobium	0.0	0.5	0.1	1.0	0.2	2.0
Campylobacter	1.8	1.3	2.3	3.3	2.3	3.0
Capnocytophaga	6.8	6.5	3.6	5.5	3.6	5.7
Cardiobacterium	0.7	0.3	0.3	0.3	0.1	0.1
Catonella	0.2	0.0	0.3	0.2	0.2	0.0
Centipeda	0.9	0.3	0.1	0.1	0.0	0.0
Clostridium	0.0	0.3	0.0	0.0	0.0	0.0
Corynebacterium	0.9	2.4	4.5	7.1	4.7	8.3
Dialister	1.4	0.5	1.6	0.8	1.7	0.9
Eikenella	0.0	0.3	0.5	0.4	0.3	0.2
Eubacterium	0.0	0.3	0.8	0.5	1.0	0.5
Filifactor	0.0	0.0	1.3	0.4	1.5	0.5
Fusobacterium	23.5	34.7	18.7	20.5	16.0	16.1
Gemella	1.6	0.5	1.4	0.9	0.7	0.3
Granulicatella	0.2	0.0	0.3	0.5	0.4	0.5
Haemophilus	2.5	2.4	0.7	1.1	0.4	0.5
Kingella	0.5	0.0	0.4	0.2	0.2	0.1
Leptotrichia	2.9	2.4	3.1	1.6	3.4	1.7
Megasphaera	0.2	0.3	0.1	0.0	0.1	0.1
Mitsuokella	0.2	0.0	0.1	0.0	0.0	0.0
Neisseria	11.7	1.1	7.1	2.6	6.2	2.0
Parvimonas	1.8	1.6	0.9	1.0	1.5	1.6
Pasteurella	0.0	0.0	0.1	0.2	0.1	1.0
Porphyromonas	13.8	1.1	12.5	1.1	13.3	1.2
Prevotella	11.7	18.0	9.1	11.1	12.3	17.0
Rothia	0.2	0.5	0.4	0.7	0.4	0.5
Selenomonas	0.5	1.6	3.2	3.7	2.3	2.8
Sneathia	0.2	0.0	0.0	0.0	0.0	0.0
Streptococcus	5.4	7.3	5.4	6.2	5.9	6.1
Tannerella	0.5	0.0	0.5	0.1	0.0	0.0
Treponema	4.1	1.6	8.5	5.9	9.0	3.5
Veillonella	4.3	13.2	2.7	10.1	3.5	13.4
Other	1.1	0.3	5.9	3.7	5.9	6.6

Table 9. Relative abundance (%) of prevalent genera by methodology. The mean abundance of each genus was calculated by combining all Pre-Tx samples and all Post-Tx samples for each method.

Phylum	CL_Pre	CL_Post	MG_Pre	MG_Post	RG_Pre	RG_Post
Bacteroidetes	33.0	26.8	30.2	21.8	28.0	24.8
Fusobacteria	26.3	36.3	21.45	22.8	20.5	19.8
Firmicutes	16.6	23.7	19.7	24.4	21.8	28.4
Proteobacteria	17.0	5.8	11.6	8.7	10.4	6.4
Spirochaetes	4.0	1.6	8.3	7.8	8.5	6.0
Actinobacteria	1.1	3.7	7.8	13.9	7.4	10.4
TM7	0.0	0.0	0.5	0.4	3.4	4.2
Other	2.0	2.1	0.5	0.2	0.1	0.2

Table 10. Relative abundance (%) of predominant phyla by methodology. The mean abundance was calculated by including all genera pertaining to each phylum for Pre-Tx as well as Post-Tx for each method.

Genus	CL-16S	MG-16S	RG
Fusobacterium	11.2	1.4	0.1
Veillonella	8.9	7.3	9.9
Prevotella	6.3	1.9	4.7
Streptococcus	1.8	0.7	0.1
Corynebacterium	1.5	2.5	3.7
Selenomonas	1.2	0.4	0.6
Capnocytophaga	-0.3	1.9	2.1
Campylobacter	-0.5	0.9	0.6
Leptotrichia	-0.5	-1.6	-1.6
Dialister	-0.8	-0.8	-0.9
Gemella	-1.0	-0.5	-0.4
Treponema	-2.5	-2.7	-5.5
Neisseria	-10.7	-4.6	-4.2
Porphyromonas	-12.7	-11.7	-12.2

Table 11. Relative change (%) in genera after treatment. The difference between the relative abundances of Pre-and Post-Tx samples for each genus was calculated for all methodologies to observe the effect of treatment on the subgingival community profile. Genera that exhibited the greatest changes, in proportion to other genera, are listed above.

Genera	Statistical Test	CL_*p* value	CL_Signif.
Porphyromonas	Mann-Whitney U	0.1000	No

Table 12. Mann-Whitney U test between the Pre- and Post-Tx samples of the CL method without subject C4's data. To test whether subject C4 introduced a bias into the original statistical analysis (Table 5) with its unusual amount of *Porphyromonas,* a Mann-Whitney U test was performed with only three subjects' (C7, C3, C1) data.

Genus	CL-16S	MG-16S	RG
Veillonella	12.9	10.7	14.0
Prevotella	8.5	2.4	5.8
Fusobacterium	6.6	-1.4	-2.9
Streptococcus	3.3	0.9	0.3
Selenomonas	0.6	-1.4	-1.0
Corynebacterium	-0.8	-2.3	-2.6
Dialister	-0.9	-0.8	-0.9
Leptotrichia	-1.0	-1.9	-1.9
Gemella	-1.2	-0.6	-0.5
Campylobacter	-1.5	0.4	0.4
Treponema	-1.7	-0.5	-3.7
Capnocytophaga	-3.5	-1.3	-0.6
Porphyromonas	-4.3	-4.8	-5.0
Neisseria	-12.6	-6.2	-5.4

Table 13. Relative change (%) in genera after treatment without subject C4's data. The difference between the mean abundance of all Pre-Tx and all Post-Tx samples of each genus, for all three analyses, were calculated with only C7, C3 and C1's data. Change in relative abundance between Pre- and Post-Tx samples of the 14 most prevalent genera for each method are depicted.

REFERENCES

Aas, J.A., B.J. Paster, L.N. Stokes, I. Olsen and F.E. Dewhirst (2005). "Defining the normal bacterial flora of the oral cavity." J Clin Microbiol **43**(11): 5721-5732.

Avila, M., D.M. Ojcius and O. Yilmaz (2009). "The oral microbiota: living with a permanent guest." DNA Cell Biol **28**(8): 405-411.

Beck, J., R. Garcia, G. Heiss, P.S. Vokonas and S. Offenbacher (1996). "Periodontal disease and cardiovascular disease." J Periodontol **67**(10 Suppl): 1123-1137.

Belda-Ferre, P., L.D. Alcaraz, R. Cabrera-Rubio, H. Romero, A. Simon-Soro, M. Pignatelli and A. Mira (2012). "The oral metagenome in health and disease." ISME J **6**(1): 46-56.

Berbari, E.F., F.R. Cockerill, 3rd and J.M. Steckelberg (1997). "Infective endocarditis due to unusual or fastidious microorganisms." Mayo Clin Proc **72**(6): 532-542.

Bik, E.M., C.D. Long, G.C. Armitage, P. Loomer, J. Emerson, E.F. Mongodin, K.E. Nelson, S.R. Gill, C.M. Fraser-Liggett and D.A. Relman (2010). "Bacterial diversity in the oral cavity of 10 healthy individuals." ISME J **4**(8): 962-974.

Chaisson, M.J., D. Brinza and P.A. Pevzner (2009). "De novo fragment assembly with short mate-paired reads: Does the read length matter?" Genome Res **19**(2): 336-346.

Chassy, B.M. and A. Giuffrida (1980). "Method for the Lysis of Gram-Positive, Asporogenous Bacteria with Lysozyme." Appl Environ Microbiol **39**(1): 153-158.

Cobb, C.M. (2008). "Microbes, inflammation, scaling and root planing, and the periodontal condition." J Dent Hyg **82 Suppl 3**: 4-9.

Dewhirst, F.E., T. Chen, J. Izard, B.J. Paster, A.C. Tanner, W.H. Yu, A. Lakshmanan and W.G. Wade (2010). "The human oral microbiome." J Bacteriol **192**(19): 5002-5017.

Dodman, T., J. Robson and D. Pincus (2000). "Kingella kingae infections in children." J Paediatr Child Health **36**(1): 87-90.

Dohm, J.C., C. Lottaz, T. Borodina and H. Himmelbauer (2008). "Substantial biases in ultra-short read data sets from high-throughput DNA sequencing." Nucleic Acids Res **36**(16): e105.

Domann, E., G. Hong, C. Imirzalioglu, S. Turschner, J. Kuhle, C. Watzel, T. Hain, H. Hossain and T. Chakraborty (2003). "Culture-independent identification of pathogenic bacteria and polymicrobial infections in the genitourinary tract of renal transplant recipients." J Clin Microbiol **41**(12): 5500-5510.

Dorn, B.R., K.L. Leung and A. Progulske-Fox (1998). "Invasion of human oral epithelial cells by Prevotella intermedia." Infect Immun **66**(12): 6054-6057.

Eribe, E.R. and I. Olsen (2008). "Leptotrichia species in human infections." Anaerobe **14**(3): 131-137.

Frank, J.A., C.I. Reich, S. Sharma, J.S. Weisbaum, B.A. Wilson and G.J. Olsen (2008). "Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes." Appl Environ Microbiol **74**(8): 2461-2470.

Goldenberger, D., A. Kunzli, P. Vogt, R. Zbinden and M. Altwegg (1997). "Molecular diagnosis of bacterial endocarditis by broad-range PCR amplification and direct sequencing." J Clin Microbiol **35**(11): 2733-2739.

Gomez-Alvarez, V., T.K. Teal and T.M. Schmidt (2009). "Systematic artifacts in metagenomes from complex microbial communities." ISME J **3**(11): 1314-1317.

Griffen, A.L., M.R. Becker, S.R. Lyons, M.L. Moeschberger and E.J. Leys (1998). "Prevalence of Porphyromonas gingivalis and periodontal health status." J Clin Microbiol **36**(11): 3239-3242.

Gundi, V.A., R. Desbriere and B. La Scola (2004). "Leptotrichia amnionii and the female reproductive tract." Emerg Infect Dis **10**(11): 2056-2057.

Haffajee, A.D., M.A. Cugini, A. Tanner, R.P. Pollack, C. Smith, R.L. Kent, Jr. and S.S. Socransky (1998). "Subgingival microbiota in healthy, well-maintained elder and periodontitis subjects." J Clin Periodontol **25**(5): 346-353.

Haffajee, A.D., S.S. Socransky, M.R. Patel and X. Song (2008). "Microbial complexes in supragingival plaque." Oral Microbiol Immunol **23**(3): 196-205.

Han, Y.W., Y. Fardini, C. Chen, K.G. Iacampo, V.A. Peraino, J.M. Shamonki and R.W. Redline (2010). "Term stillbirth caused by oral Fusobacterium nucleatum." Obstet Gynecol **115**(2 Pt 2): 442-445.

Handelsman, J. (2004). "Metagenomics: application of genomics to uncultured microorganisms." Microbiol Mol Biol Rev **68**(4): 669-685.

Henke, W., K. Herdel, K. Jung, D. Schnorr and S.A. Loening (1997). "Betaine improves the PCR amplification of GC-rich DNA sequences." Nucleic Acids Res **25**(19): 3957-3958.

Hugenholtz, P. and N.R. Pace (1996). "Identifying microbial diversity in the natural environment: a molecular phylogenetic approach." Trends Biotechnol **14**(6): 190-197.

Ishii, K. and M. Fukui (2001). "Optimization of annealing temperature to reduce bias caused by a primer mismatch in multitemplate PCR." Appl Environ Microbiol **67**(8): 3753-3755.

Jenkinson, H.F. and R.J. Lamont (2005). "Oral microbial communities in sickness and in health." Trends Microbiol **13**(12): 589-595.

Karger, A.E. (1996). "Separation of DNA sequencing fragments using an automated capillary electrophoresis instrument." Electrophoresis **17**(1): 144-151.

Kinane, D.F. (2001). "Causation and pathogenesis of periodontal disease." Periodontol 2000 **25**: 8-20.

Kircher, M., P. Heyn and J. Kelso (2011). "Addressing challenges in the production and analysis of illumina sequencing data." BMC Genomics **12**: 382.

Kolenbrander, P.E. and J. London (1993). "Adhere today, here tomorrow: oral bacterial adherence." J Bacteriol **175**(11): 3247-3252.

Kuczynski, J., C.L. Lauber, W.A. Walters, L.W. Parfrey, J.C. Clemente, D. Gevers and R. Knight (2012). "Experimental and analytical tools for studying the human microbiome." Nature Rev Genet **13**: 47-58.

Kumar, P.S., A.L. Griffen, M.L. Moeschberger and E.J. Leys (2005). "Identification of candidate periodontal pathogens and beneficial species by quantitative 16S clonal analysis." J Clin Microbiol **43**(8): 3944-3955.

Kumar, P.S., E.J. Leys, J.M. Bryk, F.J. Martinez, M.L. Moeschberger and A.L. Griffen (2006). "Changes in periodontal health status are associated with bacterial community shifts as assessed by quantitative 16S cloning and sequencing." J Clin Microbiol **44**(10): 3665-3673.

Langmead, B., C. Trapnell, M. Pop and S.L. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome Biol. **10**(3): R25.21-R25.10.

Lazarevic, V., K. Whiteson, S. Huse, D. Hernandez, L. Farinelli, M. Osteras, J. Schrenzel and P. Francois (2009). "Metagenomic study of the oral microbiota by Illumina high-throughput sequencing." J Microbiol Methods **79**(3): 266-271.

Macuch, P.J. and A.C. Tanner (2000). "Campylobacter species in health, gingivitis, and periodontitis." J Dent Res **79**(2): 785-792.

Marcotte, H. and M.C. Lavoie (1998). "Oral microbial ecology and the role of salivary immunoglobulin A." Microbiol Mol Biol Rev **62**(1): 71-109.

Mardis, E.R. (1999). "Capillary electrophoresis platforms for DNA sequence analysis." J Biomol Tech **10**(3): 137-143.

Marri, P.R., M. Paniscus, N.J. Weyand, M.A. Rendon, C.M. Calton, D.R. Hernandez, D.L. Higashi, E. Sodergren, G.M. Weinstock, S.D. Rounsley and M. So (2010). "Genome sequencing reveals widespread virulence gene exchange among human Neisseria species." PLoS One **5**(7): e11835.

Marsh, P.D. (2010). "Microbiology of dental plaque biofilms and their role in oral health and caries." Dent Clin North Am **54**(3): 441-454.
Melcher, A.H. (1976). "On the repair potential of periodontal tissues." J Periodontol **47**(5): 256-260.
Moore, W.E. and L.V. Moore (1994). "The bacteria of periodontal diseases." Periodontol 2000 **5**: 66-77.
Munson, M.A., A. Banerjee, T.F. Watson and W.G. Wade (2004). "Molecular analysis of the microflora associated with dental caries." J Clin Microbiol **42**(7): 3023-3029.
Munson, M.A., T. Pitt-Ford, B. Chong, A. Weightman and W.G. Wade (2002). "Molecular and cultural analysis of the microflora associated with endodontic infections." J Dent Res **81**(11): 761-766.
Nascimento, M.M., V.V. Gordan, C.W. Garvan, C.M. Browngardt and R.A. Burne (2009). "Correlations of oral bacterial arginine and urea catabolism with caries experience." Oral Microbiol Immunol **24**(2): 89-95.
Nyman, S., B. Rosling and J. Lindhe (1975). "Effect of professional tooth cleaning on healing after periodontal surgery." J Clin Periodontol **2**(2): 80-86.
Offenbacher, S., H.L. Jared, P.G. O'Reilly, S.R. Wells, G.E. Salvi, H.P. Lawrence, S.S. Socransky and J.D. Beck (1998). "Potential pathogenic mechanisms of periodontitis associated pregnancy complications." Ann Periodontol **3**(1): 233-250.
Onagawa, M., K. Ishihara and K. Okuda (1994). "Coaggregation between Porphyromonas gingivalis and Treponema denticola." Bull Tokyo Dent Coll **35**(4): 171-181.
Paster, B.J., S.K. Boches, J.L. Galvin, R.E. Ericson, C.N. Lau, V.A. Levanos, A. Sahasrabudhe and F.E. Dewhirst (2001). "Bacterial diversity in human subgingival plaque." J Bacteriol **183**(12): 3770-3783.
Paster, B.J., I. Olsen, J.A. Aas and F.E. Dewhirst (2006). "The breadth of bacterial diversity in the human periodontal pocket and other oral sites." Periodontol 2000 **42**: 80-87.
Qin, J., R. Li, J. Raes, M. Arumugam, K.S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D.R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage, M. Bertalan, J.M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H.B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, M. Jian, Y. Zhou, Y. Li, X. Zhang, N. Qin, H. Yang, J. Wang, S. Brunak, J. Dore, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, P. Bork and S.D. Ehrlich (2010). "A human gut microbial gene catalogue established by metagenomic sequencing." Nature **464**(7285): 59-65.
Quirynen, M., M. De Soete, K. Dierickx and D. van Steenberghe (2001). "The intra-oral translocation of periodontopathogens jeopardises the outcome of periodontal therapy. A review of the literature." J Clin Periodontol **28**(6): 499-507.
Rogosa, M. (1970). "Characters used in the classification of Lactobacilli." Int J Syst Bacteriol **20**(4): 519-533.
Rothberg, J.M. and J.H. Leamon (2008). "The development and impact of 454 sequencing." Nat Biotechnol **26**(10): 1117-1124.
Ruby, J. and J. Barbeau (2002). "The buccale puzzle: The symbiotic nature of endogenous infections of the oral cavity." Can J Infect Dis **13**(1): 34-41.
Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich and N. Arnheim (1985). "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia." Science **230**(4732): 1350-1354.
Sanger, F., S. Nicklen and A.R. Coulson (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-5467.
Savitt, E.D. and S.S. Socransky (1984). "Distribution of certain subgingival microbial species in selected periodontal conditions." J Periodontal Res **19**(2): 111-123.

Sbordone, L. and C. Bortolaia (2003). "Oral microbial biofilms and plaque-related diseases: microbial communities and their role in the shift from oral health to disease." Clin Oral Investig **7**(4): 181-188.

Shendure, J. and H. Ji (2008). "Next-generation DNA sequencing." Nat Biotechnol **26**(10): 1135-1145.

Signat, B., C. Roques, P. Poulet and D. Duffaut (2011). "Fusobacterium nucleatum in periodontal health and disease." Curr Issues Mol Biol **13**(2): 25-36.

Simonson, L.G., K.T. McMahon, D.W. Childers and H.E. Morton (1992). "Bacterial synergy of Treponema denticola and Porphyromonas gingivalis in a multinational population." Oral Microbiol Immunol **7**(2): 111-112.

Socransky, S.S. and A.D. Haffajee (1992). "The bacterial etiology of destructive periodontal disease: current concepts." J Periodontol **63**(4 Suppl): 322-331.

Socransky, S.S., A.D. Haffajee, M.A. Cugini, C. Smith and R.L. Kent, Jr. (1998). "Microbial complexes in subgingival plaque." J Clin Periodontol **25**: 134-144.

Stingu, C.S., K. Eschrich, A.C. Rodloff, R. Schaumann and H. Jentsch (2008). "Periodontitis is associated with a loss of colonization by Streptococcus sanguinis." J Med Microbiol **57**(Pt 4): 495-499.

Swerdlow, H. and R. Gesteland (1990). "Capillary gel electrophoresis for rapid, high resolution DNA sequencing." Nucleic Acids Res **18**(6): 1415-1419.

Teles, F.R., R.P. Teles, Y. Siegelin, B. Paster, A.D. Haffajee and S.S. Socransky (2011). "RNA-oligonucleotide quantification technique (ROQT) for the enumeration of uncultivated bacterial species in subgingival biofilms." Mol Oral Microbiol **26**(2): 127-139.

Turnbaugh, P.J., V.K. Ridaura, J.J. Faith, F.E. Rey, R. Knight and J.I. Gordon (2009). "The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice." Sci Transl Med **1**(6): 6ra14.

Wade, W.G. (2011). "Has the use of molecular methods for the characterization of the human oral microbiome changed our understanding of the role of bacteria in the pathogenesis of periodontal disease?" J Clin Periodontol **38 Suppl 11**: 7-16.

Williams, R., S.G. Peisajovich, O.J. Miller, S. Magdassi, D.S. Tawfik and A.D. Griffiths (2006). "Amplification of complex gene libraries by emulsion PCR." Nat Methods **3**(7): 545-550.

Woese, C.R. (1987). "Bacterial evolution." Microbiol Rev **51**(2): 221-271.

Woo, P.C., S.K. Lau, J.L. Teng, H. Tse and K.Y. Yuen (2008). "Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories." Clin Microbiol Infect **14**(10): 908-934.

Wu, T., M. Trevisan, R.J. Genco, J.P. Dorn, K.L. Falkner and C.T. Sempos (2000). "Periodontal disease and risk of cerebrovascular disease: the first national health and nutrition examination survey and its follow-up study." Arch Intern Med **160**(18): 2749-2755.

Zaura, E., B.J. Keijser, S.M. Huse and W. Crielaard (2009). "Defining the healthy "core microbiome" of oral microbial communities." BMC Microbiol **9**: 259.