Original Article

A Comparative Study of the Oral Microbiome Compositions of Healthy Postmenopausal, Premenopausal, and Prepubertal Nigerian Females, Using 16S rRNA Metagenomics Methods

KC Anukam1,2,3, NR Agbakoba4

1Uzobiogene Genomics, London, Ontario, Canada, 2Department of Medical Laboratory Science, University of Benin, Edo State, 3Department of Medical Laboratory Sciences, Faculty of Health Sciences and Technology, College of Health Sciences, Nnamdi Azikiwe University, Nnewi Campus, Anambra State, Nigeria

Abstract

Introduction: There is a paucity of information on the oral microbiome compositions of Nigerians, mostly due to lack of appropriate molecular techniques. In this pilot study, we sought to determine and characterize the oral bacterial compositions of “healthy” females. Materials and Methods: Oral samples were collected from three randomly selected females aged 56, 28, and 8 years. DNA was extracted and 16S rRNA V4 region was amplified using custom-barcoded primers before sequencing with Illumina MiSeq platform. Quantitative Insights into Microbial Ecology pipeline was used for 16S rRNA recognition. Distribution of taxonomic categories at different levels of resolution was done using the ribosomal RNA similarities to entries in the RfSeq protein database. Diversity score was calculated based on the inverse Simpson’s index. Results: The inverse Simpson’s diversity index for the postmenopausal, premenopausal, and prepubertal was 7.74, 6.95, and 7.42 respectively. A total of 12 phyla, 70 genera, and 85 species were detected. Firmicutes followed by Proteobacteria, Actinobacteria, Bacteroidetes, and Fusobacteria dominated the oral microbiome of the subjects. Streptococcus thermophilus (33.19%) was the most abundance species in subject 1, while subject 2 was highly predominated by Haemophilus parainfluenzae (80.65%), and subject 3 was predominated by Haemophilus influenzae (23.05%). Conclusion: The study has revealed that bacteria with varying diversities colonized the subjects and it highlighted the importance of metagenomics in deciphering the oral bacterial compositions from females of different age groups. More studies are needed using metagenomics approach, to appreciate these bacterial organisms that are associated with health and disease in our environment.

Keywords: Metagenomics, Nigerian females, oral microbiome

Introduction

The human oral cavity is second to the gut in terms of diversity and relative abundance of microbes such as viruses, fungi, protozoa, archaea, and bacteria that inhabit the human body. It is a major gateway to the human body, and microorganisms colonizing the oral cavity have the propensity to spread to neighboring sites, influence the immune system, and alter gastrointestinal microbiome signatures.[1] Several authors have submitted that oral microbiota significantly influences health as some studies have previously linked oral microbes to systemic infections and other life-threatening maladies such as but not limited to preterm birth,[2] cardiovascular disorder,[3] stroke,[4] and pneumonia.[5]

However, the oral micro-ecology is influenced by external factors such as oral hygiene practices,[6] types

Address for correspondence: Dr. KC Anukam, Uzobiogene Genomics, London, Ontario, Canada. E-mail: kanukam@gmail.com

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

of food intake, oral fluids, temperature, and humidity and recently exposure to chemicals in the environment. The human microbiome project which includes the oral microbiome database (HOMD, www.homd.org) provides a comprehensive resource consisting of descriptions of oral bacterial taxa, a 16S rRNA identification tool, and a repository of oral bacterial genome sequences. These databases have little or no information from African populations. A recent study compared the salivary microbiome composition of different human populations living under very different climatic conditions and geographic locations, indicating that native Alaskans and Germans are more similar to each other than to Africans in their saliva microbiome composition at the genus level.

Microbiologists have in the past identified and characterized the microorganisms with the largest representation within the communities of healthy mouths to include the genera Streptococcus, Actinomyces, Veillonella, Fusobacterium, Porphyromonas, Prevotella, Treponema, Neisseria, Haemophilus, Eubacteria, Lactobacterium, Capnocytophaga, Eikenella, Leptotrichia, Peptostreptococcus, Staphylococcus, and Propionibacterium.

In recent time, half of oral bacteria are as yet uncultured, and culture-independent methods such as metagenomics have been successfully used to comprehensively describe the oral bacterial communities. The predominant taxa belonged to Firmicutes (Streptococcus, Veillonellaceae, Granulicatella), Proteobacteria (Neisseria, Haemophilus), Actinobacteria (Corynebacterium, Rothia, and Actinomyces), Bacteroidetes (Prevotella, Capnocytophaga, and Porphyromonas), and Fusobacteria (Fusobacterium).

While several studies have explored the use of metagenomics for oral microbiome compositions in health and diseases, including children at various developmental stages of their dentition in relation to health, little or no studies exists to our knowledge on the oral microbiome of Nigerian females and males. However, a recent study in Japan compared the abundance of seven common bacterial species in the oral cavity of nonpregnant women, early pregnancy, mid-pregnancy, and late pregnancy. The total viable microbial counts in all stages of pregnancy indicated higher than those of the nonpregnant women, especially in early pregnancy. The pathway that leads to changes in the oral microbiome of women remains unclear, and it is anticipated that the differences in the physiological repertoire of postmenopausal, premenopausal, and prepubertal females may have some roles to play. In the present metagenomics study, we sought to observe the differences in the oral microbiome compositions of healthy Nigerian females.

**Materials and Methods**

**Sample collection and data analysis**

Demographic data and oral health or disease history were randomly collected from the subjects by administering structured questionnaires mostly about the oral health and disease history. The subjects were recruited based on the following criteria: absence of recent antibiotic use in the last 2 months before sample collection, no steroid contraceptives, no recent periodontal treatment, no diabetes, human immunodeficiency virus infection, or pregnancy. Oral samples were collected from three female subjects (postmenopausal, premenopausal and prepubertal, hereby referred to as subject 1, 2, and 3, respectively) after informed verbal consent. Oral samples were self-collected following uBiome® sample collection instructions. Bacterial DNA was extracted and 16S rRNA V4 region amplified using custom-barcoded primers before sequencing with Illumina MiSeq platform.

**Sequence analysis**

The raw paired-end sequence FASTQ reads were imported into MG-RAST pipeline for quality check (QC). Quantitative Insights into Microbial Ecology (QIME) pipeline was used for 16S rRNA recognition. Sequences were prescreened using QIME-UCLUST algorithms for at least 70% identity to ribosomal sequences from the following RNA databases; Greengenes, Large Sub Unit (LSU), small sub unit (SSU), and Ribosomal Database Project (RDP). Operational taxonomic unit (OTU) picking was done at 97% identity against the RDP, LSU, and SSU databases. Microbial taxonomy was generated from the nonrarefied OTU table. Distribution of taxonomic categories at different levels of resolution was done using the ribosomal RNA similarities to entries in the REFseq protein database. Species diversity score was calculated based on the inverse Simpson’s index. Scores range from 0 to 10, with 10 being the most diverse.

**Results**

**Demographic data and oral history**

Subject 1 is a 56-year-old woman that had had three of her premolars extracted in 1981, 2004, and 2014. Before extraction, all the extracted teeth had holes in them, which were filled. The fillings lasted for a while and then wore off, and the pains started again, hence the need for the extractions. She had caries in one tooth (a 4th premolar), which eventually broke.
off on its own. The stump served as a reservoir for residual food items, which made her use toothpicks and dental floss frequently. She reported brushing her teeth two times daily with fluoride toothpaste. She last visited a dentist in 2014. As at the time the mouth swab was collected, she did not have any prevailing mouth problems. Subject 2 is a 28-year-old female who claimed to brush her teeth twice daily with a soft brush and oral B toothpaste. She has had tooth decay twice (when she was in primary and secondary school) that led to extraction of two premolars. She had scurvy several times and treated with Vitamin C. She ate lots

![Figure 1: Relative abundance of the phyla taxa represented as 100% stacked bar](image1.jpg)

![Figure 2: Two-D pie chart of the six phyla that are common in all the subjects, indicating higher abundance of *Firmicutes* and *Candidatus Saccharibacteria* in subject 3](image2.jpg)
of sugary confectionaries. There was no “remarkable” oral health issue at the time of sample collection. Subject 3 is an 8-year-old female with no oral maladies as at the time of sample collection.

**Sequence characteristics**

On average, the data sets generated 46,109 high-quality sequences per sample totaling a mean of 6,868,637 base-pairs (bp) with an average length of 149 bp (149 ± 13 bp). Post-QC produced an average base-pair count of 278,367 bp, postsequence count of 1972, mean sequence length of 141 ± 25 bp, and mean guanine-cytosine percent of 55% ± 4%. Subject 1 had 4400 sequences (9.54%) that failed to pass the QC pipeline, while subjects 2 and 3 had 4337 sequences (8.18%) and 5759 sequences (7.98%), respectively, that failed the QC pipeline.

**Oral bacterial compositions**

The diversity and bacterial community compositions of the three cases were explored by comparing the relative abundance at different taxonomic levels. The inverse Simpson’s diversity score for subjects 1, 2, and 3 was 7.74 (64th percentile), 6.95 (40th percentile), and 7.42 (54th percentile) respectively. The oral bacterial community structure was characterized by the relative abundances of the bacterial taxa. A total of 12 phyla, 70 genera, and 85 species were detected in the oral samples of the cases. Subject 1, as shown in Figure 1, had 10 phyla, mostly dominated by *Firmicutes* (53.92%), followed by...
**Proteobacteria** (24.48), **Actinobacteria** (15.10%), **Bacteroides** (5.11%), **Fusobacteria** (1.10%), **Acidobacteria** (0.25%), **Spirochaetes** (0.03%), **Cyanobacteria** (0.0029%), **Streptophyta** (0.00189%), and **Candidatus Saccharibacteria** (0.00189%). Subject 2 had 10 phyla, dominated by **Firmicutes** (48.70%), followed by **Proteobacteria** (45.99%), **Bacteroides** (2.26%), **Actinobacteria** (1.76%), **Fusobacteria** (0.80%), **Spirochaetes** (0.23%), **Tenericutes** (0.093%), **Synergistetes** (0.072%), **Streptophyta** (0.069%), and **Candidatus Saccharibacteria** (0.0047%). Subject 3 had only six phyla, mainly dominated by **Firmicutes** (74.40%), followed by **Proteobacteria** (10.55%), **Actinobacteria** (9.93%), **Bacteroides** (3.82%), **Fusobacteria** (1.21%), and **Candidatus Saccharibacteria** (0.072%). Figure 2 shows the six phyla that are common in all the subjects, indicating higher abundance of **Firmicutes** and **Candidatus Saccharibacteria** in subject 3 and more abundance of **Proteobacteria** in subject 2. Subject 1 had more abundance of **Actinobacteria** greater than subject 3 and subject 3 greater than subject 2. However, some sequence reads in the bacteria kingdom could not be classified.

A total of 70 genera were detected from the three cases. Subjects 1 and 3 had 52 genera, while subject 2 had 50 genera. In terms of absolute genera count, subject 1 had 966,411, subject 2 had 993,366, and subject 3 had 994,895 as shown in Figure 3. Subject 1 had eight exclusive genera representing 11.3%, which includes **Parascardovia**, **Weissella**, **Slackia**, **Comamonas**, **Shuttleworthia**, **Anoxybacillus**, **Howardella**, and **Enterobacter**. Subject 2 had nine exclusive genera (12.7%) comprising **Pasteurella**, **Mycoplasma**, **Odoribacter**, **Eggerthia**, **Peptococcus**, **Moraxella**, **Phocaecola**, **Desulfovibrio**, and **Bulleidia**. Subject 3 also had eight exclusive genera (11.3%); **Abiotrophia**, **Streptobacillus**, **Pelomonas**, **Enterococcus**, **Bacteroides**, **Aerococcus**, **Stenotrophomonas**, and **Crypotbacterium**. However,
Anukam and Agbakoba: Oral microbiome compositions of healthy Nigerian females

Subjects 1, 2, and 3 had 37 (52.1%) genera in common [Figure 4]. There was no significant difference in the relative abundance of the core genera among the subjects (subject 1 versus subject 2, $P = 0.5294$, subject 1 versus 3, $P = 0.3932$; subject 2 versus 3, $P = 0.4244$). The core genera that occurred with high relative abundance in subjects 1, 2, and 3 are shown in Figure 5. A total of 85 species were identified in the three subjects, showing 8 (9.1%) exclusive species in subject 1 comprising Aggregatibacter aphrophilus, Actinomyces gerencseriae, Parascardovia denticolens, Veillonella atypica, Slackia exigua, Actinomyces oris, Actinomyces naeslundii, and Gemella species 1754-94. Subject 2 had 14 (15.9%) exclusive bacterial species; Porphyromonas gingivalis, Pasteurella pneumotropica, Tannerella forsythia, Eggerthia catenaformis, Dialister pneumosintes, Actinomyces genome species C1, Moraxella species S12-08, Phocaeicola abscessus, Neisseria sicca, Desulfomicrobium orale, Johnsonella ignava, Bulleidia extructa, Prevotella micans, and Oribacterium species oral taxon 102. Subject 3 had 5 (5.7%) exclusive species (Granulicatella elegans, Gemella morbillorum, Vagococcus lutral, Bacteroides vulgatus, and Cryptobacterium curtum). The three subjects had 37 (42%) species in common [Figure 6]. The phylogenetic tree of the 37 common taxonomic species in the three subjects is represented in Figure 7.

At the species taxonomic level, Streptococcus thermophilus (33.19%) was the most abundant species in subject 1, followed by Haemophilus parainfluenzae (30.95%), Rothia dentocariosa (12.24%), Streptococcus gordii (8.58%), Streptococcus mitis (6.81%), and others [Figure 8]. Subject 2 was highly predominated by H. parainfluenzae (80.65%), followed by H. influenza (4.13%), Aggregatibacter segnis (2.96%), Actinobacillus porcinus (1.64%), Veillonella sp. oral

**Figure 8:** The pie chart of the most abundant species in subject 1

**Figure 9:** The pie chart of the most abundant species in subject 2

**Figure 10:** The pie chart of the most abundant species in subject 3

**Figure 11:** Some genera (100% stacked area) known to be associated with oral health
taxon 780 (1.52%), Lautropia sp. TeTO (1.36%), R. dentocariosa (0.95%), and others [Figure 9]. Similarly, subject 3 was predominated by H. influenza (23.05%), Rothia mucilaginosa (20.11%), S. thermophilus (9.64%), H. parainfluenzae (9.15%), R. dentocariosa (9.00%), and others as shown in Figure 10. Some genera known to be associated with oral health as they occurred in the subjects are indicated in Figure 11. At species taxonomic level, Figure 12 shows some taxa that may be associated with health and disease. The relative abundances of the 37 taxonomic genera that are common in the three subjects are exhibited in Figure 13. Figure 14 revealed some taxonomic species found to be associated with oral diseases from recently published literature.

**DISCUSSION**

To our knowledge, this is the first study from the Eastern part of Nigeria employing 16 S rRNA metagenomics platform targeting V-4 variable region for robust bacterial taxonomical classifications regarding the oral microbial compositions. The microbial diversity score indicates that subject 1’s oral microbiome was more diverse in terms of species richness than subject 2 and subject 3. Although the inverse Simpson’s index for subject 3 is higher than subject 2, it remains to be determined if age was any contributing factor for the differences observed in oral microbial diversity. The subject 3 diversity score was higher than subject 2 and one of the reasons that could be attributed to the difference may due to the oral history...
and bacterial species associated with oral disease present in subject 2.

The predominance of the phylum- *Firmicutes* in the subjects (53.92%, 48.70%, 74.40%) and other phyla such as *Proteobacteria*, *Actinobacteria*, *Bacteroides*, and *Fusobacteria* that contributed at least 0.1% are considered abundant in the oral microbiome. This is similar in a previous study that looked at the core microbiome of oral microbial communities by sequencing the microbiomes from several intraoral niches (dental surfaces, cheek, hard palate, tongue, and saliva) in three healthy individuals.\[^{11}\]

The preponderance of *Firmicutes* in subject 3 with 74.40% is very noteworthy when compared with subject 1 and subject 2. Other studies have compared the microbiome of adult saliva\[^{14}\] and child saliva appears to have a higher proportion of *Firmicutes* and *Actinobacteria* and a lower proportion of *Bacteroidetes*, *Fusobacteria*, and Spirochaetes at all stages of the dentition.\[^{12}\]

The presence of *Candidatus* Saccharibacteria in the subjects, less abundant but relatively higher in subject 3 needs more insight. This phylum was formally known as TM7 unculturable phyotype found in human mouth,\[^{15}\] but recently *Candidatus* Saccharibacteria phylum has been coined following complete genome sequencing of several Candidate Division TM7 members from wastewater.\[^{16}\] However, due to lack of cultured isolates and rareness of 16S rRNA gene sequences in depositories, knowledge on the biological insight of this group is still in the early stage.\[^{17}\]

The exclusive genera found in all the subjects indicate the individual variability of the oral microbiome as some previous studies have shown that these results point to the persistence of subject-specific taxa whose frequency fluctuates between the time points. For example, the genus *Gemella*, identified in subject 1 and subject 3 individuals, was not defined as a core-microbiome genus in previous studies of salivary bacterial communities.\[^{11}\] *Gemella* is depleted in the presence of tooth decay, and an increase is correlated with good oral hygiene and they are present in low abundance in subject 1 and 3 but not in subject 2. Bacterial genera associated with health are present in the three subjects with variable occurrence as shown in Figure 11, notably, *Streptococcus*, *Neisseria*, *Porphyromonas*, *Prevotella*, and *Veillonella*.

However, some species such as *P. gingivalis* and *T. forsythia* found exclusive in high numbers in subject 2 lend credence to the oral history. The high predominance by *H. parainfluenzae* (80.65%) in subject 2 is very interesting although the subject claimed to have “healthy” oral hygiene. High numbers are usually associated with asthma and chronic lung disease. The dysbiosis within oral microbiome is often associated with an increase in butyrate-producing pathogens such as *P. gingivalis*\[^{13}\] and *Filifactor alocis*, has also been recently implicated in diseased conditions such as periodontitis.\[^{19}\] The role of *Streptococcus mutans* as a primary caries pathogen appears more pronounced in subject 1 and recent study has found that populations with prevention programs compared to populations lacking caries treatment and prevention strategies have less *S. mutans*.\[^{16}\]

There is yet no consensus as to what constitutes normal oral microbiota, and it has been suggested that focus on functional rather than phylogenetic diversity may be required to fully understand host-microbiome interactions. The present study has revealed that age may be a factor in shaping the oral microbial diversity in women sampled, but larger population studies are needed to confirm this. However, other different biological interactions may affect the structural composition of the oral microbiome. Furthermore, the relationship of microorganisms with the oral environment may play a role in the composition of oral microbiota. For example, it has been suggested that oral environmental changes, such as high-sugar diet, low-pH, smoking, and fluoride use, may affect oral microbial diversity.\[^{20}\]

**Conclusion**

This study has demonstrated the importance of metagenomics in deciphering the oral bacterial compositions from females of different age groups. Most of the bacterial taxa have not been identified before in our environment with conventional culture methods, but with the use of metagenomics approach, we can now begin to appreciate an insight into these bacterial organisms that may be associated with health and disease.

**Financial support and sponsorship**

The study was funded by uBiome Inc., San Francisco, California, USA, under an academic consortium grant-in-kind award to Dr. Kingsley Anukam.

**Conflicts of interest**

There are no conflicts of interest.

**References**


