The objective of this study was to evaluate the most effective method of DNA extraction of oral mouthwash samples for use in microbiome studies that utilize next generation sequencing (NGS). Eight enzymatic and mechanical DNA extraction methods were tested. Extracted DNA was amplified using barcoded primers targeting the V6 variable region of the bacterial 16S rRNA gene and the ITS1 region of the fungal ribosomal gene cluster and sequenced using the Illumina NGS platform. Sequenced reads were analyzed using QIIME and R. The eight methods yielded significantly different quantities of DNA ($p < 0.001$), with the phenol-chloroform extraction method producing the highest total yield. There were no significant differences in observed bacterial or fungal Shannon diversity ($p = 0.64$, $p = 0.93$ respectively) by extraction method. Bray-Curtis beta-diversity did not demonstrate statistically significant differences between the eight extraction methods based on bacterial ($R^2 = 0.086$, $p = 1.00$) and fungal ($R^2 = 0.039$, $p = 1.00$) assays. No differences were seen between methods with or without bead-beating. These data indicate that choice of DNA extraction method affect total DNA recovery without significantly affecting the observed microbiome.
data processing, and statistical analyses\(^{21}\). Additionally, each of these steps has associated labor and cost factors that may influence a researcher's decision to use one method over another\(^{22}\). Previous research has shown that oral sampling techniques such as saliva, buccal swab, and oral rinse collection may influence overall DNA quantity and spectrum of microbes detected\(^{23-25}\). It has also been suggested that next-generation sequencing (NGS) may produce variable results particularly when analyzed using different classification algorithms\(^{26}\). Given that these processes can influence the understanding of microbial communities, investigating protocols for characterizing the biota of the oral cavity is important to allow inter-study comparisons.

Efficient and consistent methods of DNA extraction are central to accurately characterizing these communities. A number of studies have begun to examine the oral microorganisms using NGS with a variety of DNA extraction methods\(^{27-33}\). In addition, a large number of studies have collected and processed Scope mouthwash samples for genomic DNA that might be suitable for microbiome studies. The purpose of this investigation was to compare the most recent techniques to discern the most effective method of DNA extraction utilizing both enzymatic and mechanical lysis techniques across various human oral samples in order to determine the methods with the highest DNA yield and the most consistent results for characterization of both bacterial and fungal communities found in the oral cavity.

**Results**

For this study, eight DNA extraction methods, utilizing different combinations of enzymatic and mechanical lysis techniques, were compared across six oral samples (Table 1). The methods were evaluated for DNA yield and variation in the detected oral microbiome. There was a significant difference in DNA quantity among the eight extraction methods (\(p < 0.001\)). The phenol-chloroform extraction technique (Method 1) generated the highest DNA yield (Fig. 1) while the UltraClean Microbial DNA Isolation Kit (Method 7) and the UltraClean Microbial DNA Isolation Kit (Method 8) resulted in significantly lower DNA yields (\(p < 0.01\)) than the three non-bead-beating methods (Table 2).

DNA from the 48 DNA samples were amplified using 16S rRNA V6 barcoded primers and recently described primers for the ITS1 region and submitted for Illumina NGS. Raw sequences were processed for quality control and chimera removal, resulting in a total of 373,840 bacterial reads (average of 7,788 ± 1,579 reads per sample), and 363,881 fungal sequence reads (average of 5,965 ± 1,837 reads per sample). The bacterial community composition and normalized abundances in the oral cavity are displayed in the heat map (Fig. 2A). Dendrogram clustering based on the top 20 species shows a tendency of samples to cluster by original subject. DNA extraction method did not show clustering. Community clustering based on the top 20 fungi (Fig. 2B) displays a closer distance between samples than seen with the bacterial 16S data. However, the fungal heatmap also indicated that samples tended to cluster together based on subject and not extraction method.

Seven bacterial phyla were identified; Firmicutes, Bacteroidetes, Actinobacteria, Fusobacteria, Proteobacteria, Spirochaetes, and the candidate phylum TM7 (also known as Saccharibacteria), with the majority of OTUs assigned to Firmicutes and Bacteroidetes. At the genus/species level, Streptococcus dominated the oral cavity, consistent with published studies\(^{37}\), Rothia mucilaginosa\(^{34}\), an opportunistic pathogen in immunocompromised patients and Prevotella veroralis, a biofilm forming opportunistic pathogen\(^{35}\), were the second and third most abundant species, respectively (Fig. 2A).

The oral mycobiota was dominated by species from Ascomycota, Basidiomycota, an unidentified fungal phyla, and Zygomycota (order based on cumulative dominance across all samples). Constituents of the Candida genus were amongst the top identified OTUs consistent with previous reports on the oral mycobiome\(^{36}\). Several species

---

### Table 1. Methods of DNA extraction used in this study, including additional enzymatic and mechanical (bead-beating) cell disruption steps. Enzymes listed in this table are in addition to any lysis buffer included in each kit (either specified, such as proteinase K, or proprietary). All bead-beating was conducted on a FastPrep-24 Instrument (MP Biomedicals) at 6.0 m/s for 40 seconds.

| Method | Extraction Method | Commercially Available Kit | Enzymatic Lysis Step Added | Bead-beating |
|--------|-------------------|---------------------------|-----------------------------|--------------|
| M1     | Phenol/chloroform | No                        | No                          | No           |
| M2     | QIAamp DNA Mini Kit | Yes                      | No                          | No           |
| M3     | QIAamp DNA Mini Kit | Yes                      | Mutanolysin                 | No           |
|        |                   |                           | Lysozyme                    |              |
| M4     | QIAamp DNA Mini Kit | Yes                      | Mutanolysin                 | 0.1 mm–diameter zirconia/silica beads (BioSpec) |
|        |                   |                           | Lysozyme                    |              |
| M5     | PowerLyzer PowerSoil DNA Isolation Kit | Yes                      | Mutanolysin                 | 0.1 mm–diameter glass beads (MBio) |
|        |                   |                           | Lysozyme                    |              |
| M6     | PowerSoil DNA Isolation Kit | Yes                      | No                          | 0.7 mm–diameter garnet (MBio) |
| M7     | UltraClean Microbial DNA Isolation Kit | Yes                      | No                          | 0.7 mm–diameter garnet (MBio) |
| M8     | BioOstic Bacteremia DNA Isolation Kit | Yes                      | No                          | 0.15 mm–diameter garnet (MBio) |
of Malassezia were also identified in the oral cavity, including *Malassezia restricta* (Fig. 2B), a common lipid dependent human pathogen that is usually found on skin37.

Significant variation in sample evenness, based on the Shannon diversity index, was observed in the bacterial *p* < 0.001 and fungal *p* < 0.001 assays (Fig. 3A,B respectively). There was no significant difference in the Shannon diversity index among DNA extraction methods for either the bacterial *p* = 0.87 or fungal assays *p* = 0.93 (Fig. 3C,D, respectively). Similarly, β-diversity showed distinct clusters formed on the basis of subject in both the bacterial *p* < 0.001 and fungal *p* < 0.001 community analyses, which explained nearly all of the inter-sample community variance, R² = 0.80 and R² = 0.84, respectively (Fig. 4A,B). β-diversity analyses did not show significant sample clustering based on extracted method for either bacteria R² = 0.086, *p* = 0.996 or fungi R² = 0.039, *p* = 1.00 (Fig. 4C,D, respectively).

**Discussion**

In the current study, eight methods for DNA extraction from six oral cavity samples were used and DNA quantity and microbial community composition were compared. Our analysis revealed that DNA yield was significantly different among the eight DNA extraction methods with DNA recovery greatest after phenol-chloroform extraction (Fig. 1). The lower DNA yield of commercially available kits (Table 1) compared to the phenol-chloroform extraction method may be due to DNA loss during silica column purification. DNA yield tended to be greater with enzymatic digestion than using mechanical lysis (bead-beating) approaches. The lower DNA yield among bead-beaten samples is likely due to DNA degradation during mechanical disruption. Thus, for DNA yield, non-bead-beating methods, particularly phenol-chloroform extraction provides the greatest yield of total DNA.

Although DNA for human genetic studies has frequently been obtained using oral mouthwash and/or saliva collection methods38, compatibility of the DNA from these studies for future microbiome studies has not been examined in detail. Previous studies found differences in the oral bacterial microbiome based on DNA extraction methods32,33,39; whereas, other studies determined that DNA extraction methods did not seem to influence major differences in the oral microbiome32,34,40. Nevertheless, it is hard to do a direct comparison amongst studies in the literature, since many used saliva and/or plaque collection3,33,39,40, some compared crude DNA to purified DNA39,40, others used mock communities39, and one did not include NGS analysis of the microbiome32. Only one study examined both bacterial and fungal communities and surprisingly found no differences amongst 4 methods for bacterial communities, but found phenol-chloroform extraction best for fungal community diversity33.

Although we found that DNA extraction methods had an influence on DNA yield, we did not find a significant difference in oral microbiome composition across eight DNA extraction methods of oral rinse specimens. Shannon diversity measures for bacterial and fungal communities were similar across the employed extraction methods and did not achieve statistically significant differences. Similarly, PERMANOVA analysis on rank order Bray distances did not demonstrate differences in β-diversity for either assay. Our results instead demonstrated that individual subject differences drove diversity measures across both bacteria and fungi. Taken together, these data suggest that both α- and β-diversity measures were consistent for all eight-extraction measures, and that the choice of method does not have a major influence on the observed oral communities. The results of this study might have been influenced by the larger number of samples analyzed compared to previous studies.

All eight extraction methods were able to consistently recapitulate the original subject microbials as indicated by both alpha and beta diversity measures including Shannon diversity index and Bray-Curtis distances, respectively. These findings are consistent with previous studies that have demonstrated that each person’s oral microbiome is unique34,41. Additionally, all methods reported here detected hard to lyse gram-positive species,
such as Streptococcus\textsuperscript{43}, indicating sufficient lysis of cells. Moreover, the similarity of results for fungal community analyses across all methods is consistent with the one report that found phenol-chloroform extraction yielded the highest fungal diversity in saliva\textsuperscript{33}.

In summary, our study compared eight DNA extraction methods tested on oral rinse specimens that are commonly collected in large epidemiological studies and are used or may be used in the future to study the oral microbiome. While the eight methods tested in this study had significantly different DNA recovery, there was no difference in the observed oral microbiotas among methods. This study provides empiric evidence that research studies can select an appropriate DNA extraction method with or without bead-beating for characterization of the oral microbiota without influencing differences between the oral microbiome/mycobiome of individuals.

### Materials and Methods

#### Consent and Approval for Use of Human Participants.
Oral rinse specimens from six individuals were collected as part of a pilot study on sampling procedures for the Health and Nutrition Examination Survey in New York City 2013 (NYC HANES 2013), a collaborative project between the City University of New York (CUNY) Graduate School of Public Health and Health Policy and the NYC Department of Health and Mental Hygiene. IRB approval for analysis of pilot oral specimens was obtained from the Human Research Protection Program (HRPP) of CUNY. All methods performed in this study were conducted in accordance with Hunter College (CUNY) university integrated IRB approved protocol (PT: 346358-9). Informed consent was obtained from study participants prior to sample collection. Upon receipt all used human specimens received a lab Sample ID and no information regarding, age, race, gender or any other identifying information was used in the presented study.

#### Specimen Collection.
Consented study participants provided an oral sample by rinsing with 20 mL of Scope mouthwash for 20 seconds. The 20-second oral rinse was broken into two 5-second swish sessions and two 5-second gargle sessions. The oral rinse samples were frozen at \(-80\,^\circ\text{C}\) at the New York State Public Health Laboratory (NYPHL) office and were transported on dry ice to Albert Einstein College of Medicine, where they were immediately stored at \(-80\,^\circ\text{C}\).

#### DNA Extraction.
DNA was extracted from the oral rinse samples using eight DNA extraction methods based on physical and/or enzymatic lysis steps and isolation procedures (Table 1). Extraction methods with commercially available kits all used a silica-based column. One extraction method included a non-commercial method.

| Method Pairs  | Tukey HSD Q statistic | Tukey HSD p-value |
|---------------|-----------------------|-------------------|
| M1 vs M2      | 6.2                   | \(p < 0.01\)      |
| M1 vs M3      | 6.87                  | \(p < 0.01\)      |
| M1 vs M4      | 9.22                  | \(p < 0.01\)      |
| M1 vs M5      | 9.37                  | \(p < 0.01\)      |
| M1 vs M6      | 10.46                 | \(p < 0.01\)      |
| M1 vs M7      | 13.38                 | \(p < 0.01\)      |
| M1 vs M8      | 12.68                 | \(p < 0.01\)      |
| M2 vs M3      | 0.68                  | 0.9               |
| M2 vs M4      | 3.02                  | 0.41              |
| M2 vs M5      | 3.17                  | 0.35              |
| M2 vs M6      | 4.27                  | 0.08              |
| M2 vs M7      | 6.76                  | \(p < 0.01\)      |
| M2 vs M8      | 6.49                  | \(p < 0.01\)      |
| M3 vs M4      | 2.34                  | 0.69              |
| M3 vs M5      | 2.5                   | 0.63              |
| M3 vs M6      | 3.59                  | 0.21              |
| M3 vs M7      | 6.04                  | \(p < 0.01\)      |
| M3 vs M8      | 5.81                  | \(p < 0.01\)      |
| M4 vs M5      | 0.15                  | 0.9               |
| M4 vs M6      | 1.25                  | 0.9               |
| M4 vs M7      | 3.53                  | 0.22              |
| M4 vs M8      | 3.47                  | 0.24              |
| M5 vs M6      | 1.09                  | 0.9               |
| M5 vs M7      | 3.37                  | 0.28              |
| M5 vs M8      | 3.31                  | 0.3               |
| M6 vs M7      | 2.2                   | 0.75              |
| M6 vs M8      | 2.22                  | 0.74              |
| M7 vs M8      | 0.17                  | 0.9               |

Table 2. Tukey HSD post-hoc results of DNA yield between each DNA extraction method. Significant p-values are in bold.
using phenol-chloroform. All DNA isolation methods evaluated in this study are either commonly used in DNA extraction or have previously been used in microbial analysis studies. For each method, 1 mL from each oral rinse sample was centrifuged ($5,000 \times g$) for 5 minutes. The cell pellet was re-suspended in 100 $\mu$l TE buffer (10 mM Tris. Cl, pH 8.0, 1 mM EDTA) and used for DNA extraction. Upon completion of each extraction method, the purified DNA was eluted in 100 $\mu$l of elution buffer (pH 8.0) and DNA concentration was determined using a NanoDrop 2000 (Thermo Scientific, DE).

**Method 1 (Proteinase K/SDS/phenol chloroform extraction).** The cell pellet was directly processed in 200 $\mu$l cell lysis buffer (10 mmol/L Tris/HCl pH 8.0, 10 mmol/L EDTA, 0.1 mol/L NaCl, 2% SDS pH 8.0) and 10 $\mu$l proteinase K (20 mg/ml, Roche Diagnostics), and incubated overnight at 55 °C. The samples were treated with RNase A (100 mg/ml, Qiagen, Valencia, CA) for 20 minutes at 37 °C followed by phenol/chloroform extraction using Phase Lock Gel Tubes (PLG, 5 Prime Inc., Gaithersburg, MD) as described by the manufacturer.

**Method 2 (QIAamp DNA mini kit).** First, 20 $\mu$l of proteinase K (20 mg/ml) and 100 $\mu$l of Buffer AL were added to 100 $\mu$l of pelleted cells in TE. The samples were incubated at 56 °C for 10 minutes. After incubation, 100 $\mu$l of 100% ethanol was added to the samples and the DNA was purified following the manufacturer’s instructions.

**Method 3 (Enzymatic lysis followed by QIAamp DNA mini kit).** The pelleted cells in 100 $\mu$l TE were treated with lysozyme (0.84 mg/ml, Sigma Aldrich), mutanolysin (0.25 U/ml, Sigma Aldrich) and lysostaphin (21.10 U/ml, Sigma Aldrich) at 37 °C for 30 minutes. Subsequently, 20 $\mu$l proteinase K and 100 $\mu$l Buffer AL were added followed by incubation at 56 °C for 10 minutes. DNA was isolated using the QIAamp DNA Mini Kit as described above.

**Method 4 (Enzymatic and bead-beating lysis followed by QIAamp DNA mini kit).** Pelleted cells were digested using enzymes as in Method 3. After incubation, the mixture was treated with 15 $\mu$l proteinase K (10 mg/ml) and 150 $\mu$l Buffer AL (Qiagen) at 56 °C for 10 minutes. The samples were then transferred to a clean screw-cap tube containing 300 mg of 0.1 mm-diameter zirconia/silica beads (BioSpec, Bartlesville, OK) and mechanically lysed using a FastPrep-24 Instrument (MP Biomedicals, Santa Ana, CA) at 6.0 m/s for 40 seconds. Next, the samples were centrifuged (10,000 $\times g$) for 30 seconds and 200 $\mu$l of the supernatant was added to a clean microcentrifuge tube containing 100 $\mu$l of 100% ethanol. DNA was isolated using the QIAamp DNA Mini Kit (Qiagen) as described above.

**Method 5 (Enzymatic lysis followed by PowerLyzer PowerSoil DNA Isolation Kit).** The pelleted cells were incubated with the enzymes described in Method 3. After incubation, the mixture was transferred to a PowerLyzer Glass Bead Tube (0.1 mm) containing 650$\mu$l of Bead Solution. The remainder of the DNA isolation protocol was continued beginning with step 4 of the PowerLyzer PowerSoil DNA Isolation Kit instructions (MO
BIO laboratories, Inc., Carlsbad, CA). The bead-beating step used a FastPrep-24 Instrument (MP Biomedicals) set at 6.0 m/s for 40 seconds.

**Method 6 (PowerSoil DNA Isolation Kit).** DNA was extracted using the PowerSoil DNA Isolation Kit (MO BIO laboratories, Inc.) following the manufacturer's protocol without additional enzymatic lysis. The cells were mechanically lysed using manufacturer's provided bead tubes and a FastPrep-24 Instrument (MP Biomedicals) at 6.0 m/s for 40 seconds.

**Method 7 (UltraClean Microbial DNA Isolation Kit).** DNA was extracted using the UltraClean Microbial DNA Isolation Kit (MO BIO laboratories, Inc.) following the manufacturer’s protocol. The cells were mechanically lysed using manufacturer's provided bead tubes and a FastPrep-24 Instrument (MP Biomedicals) at 6.0 m/s for 40 seconds.

**Method 8 (BiOstic Bacteremia DNA Isolation Kit).** DNA was isolated using the BiOstic Bacteremia DNA Isolation Kit (MO BIO laboratories, Inc.) following the manufacturer's protocol. The cells were mechanically lysed using manufacturer's provided bead tubes and a FastPrep-24 Instrument (MP Biomedicals) at 6.0 m/s for 40 seconds.

---

**Figure 3.** Comparison of Fungal and Bacterial Shannon Alpha Diversity Measures. Shannon alpha diversity box plots of bacterial and fungal community composition based on variance in species evenness is shown for samples (panels A and B) and by methods (panels C and D). Significant variance is observed in bacterial sample evenness, $p < 0.001$ (panel A) as well as fungal community evenness, $p < 0.001$ (panel B). No significant differences are observed for Shannon diversity based on collection method for bacterial, $p = 0.87$ (panel C) or fungal diversity measures, $p = 0.93$ (panel D).
16S rRNA gene and ITS1 region amplification and massively parallel sequencing. To amplify the 16S rRNA gene region of bacterial species, an aliquot of 0.5 µl DNA from each sample and DNA isolation method was PCR amplified in a total reaction volume of 25 µl using barcoded primers spanning the V6 variable region of the 16S rRNA gene as previously described. In brief, an equal mixture of AmpliTaq Gold (Applied Biosystems, Carlsbad, CA) and HotStart-IT FidelityTaq DNA Polymerase (Affymetrix, Santa Clara, CA) was used. For all samples a unique 8-bp barcode was introduced to the PCR amplicons on the primers. Thermocycling conditions included an initial denaturation at 95 °C for 5 minutes, then 15 cycles at 95 °C for 1 minute, 55 °C for 1 minute, and 68 °C for 1 minute. This was followed by 15 cycles at 95 °C for 1 minute, 60 °C for 1 minute, and 68 °C for 1 minute; and a final extension for 10 minutes at 68 °C.

To amplify the ITS1 region of fungal species, 10 µl from each sample and DNA isolation method was PCR amplified in a total reaction volume of 25 µl using barcoded primers specific to the ITS1 region of the fungal ribosomal gene cluster. In brief, 9.25 µl of dd H2O, 2.5 µl of USB 10X buffer with MgCl2 (10 mM Tris-HCl, pH 8.6, 50 mM KCl, 1.5 Mn MgCl2, Affymetrix, Santa Clara, CA), 1 µl of USB MgCl2 (25 mM), 0.5 µl of dNTP mix (10 mM each, Roche Basel, Switzerland), 0.25 µl AmpliTaq Gold, polymerase (5 U/µl, Applied Biosystems, Carlsbad, CA), 0.5 µl of HotStart-IT DNA FidelityTaq Polymerase (2.5 U/µl, Affymetrix), and 1 µl (5 µM) of each primer (IDT, Coralville, IA). Thermocycling included an initial denaturation of 95 °C for 3 mins, followed by 35 cycles of 95 °C for 30 s, 55 °C for 30 s, 68 °C for 2 min, followed by a final extension of 68 °C for 10 min.
The 16S rRNA and ITS1 PCR products each were pooled at approximately equal molar DNA concentrations and purified using the QiAquick Gel Extraction Kit (Qiagen). Following library preparation using TruSeq DNA Sample Prep Kits (Illumina, San Diego, CA), the pooled 16S rRNA DNA was sequenced on an Illumina HiSeq. 2500 using paired-end 150 bp reads, while the pooled ITS1 DNA was sequenced on an Illumina MiSeq using paired-end 300 bp reads, by the Epigenomics and Genomics Core Facility, Albert Einstein College of Medicine (Bronx, NY).

Bioinformatics. MiSeq reads were demultiplexed using novocraft’s novobarc ode 1.0046 based on sample specific barcodes46. Reads were left and right trimmed with PrinSeq 0.20.457 to remove bases that fell below the PHRED score of 25. Paired end reads were merged with PANDASEQ 1.2048 using default settings.

For 16S RNA gene reads, OTUs were clustered using closed reference selection with USEARCH using a custom-in-house database that contains reference sequences from Green-Genes 13.848. Additionally, reference sequences of an oral microbiome specific database, Human Oral Microbiome Database (HOMD)49, were retrieved in order to account for bacteria specific to the human oral cavity. Representative sequences were aligned using PyNAST50 and phylogenetic analyses were performed using FastTree 2.052.

For fungal ITS1 reads, open reference_OTU picking was employed with QIIME 1.951, open-reference OTU picking protocol as previously described54. The protocol was modified to use VSEARCH version 1.4.053, which allowed for higher throughput. The OTU clustering threshold was changed from 97% to 99% sequence identity to account for fungal diversity. Sequence dereplication and chimera removal was performed as part of the QIIME’s usearch quality control protocol prior to OTU picking with VSEARCH. Representative sequences for each OTU cluster were chosen based on sequence abundance. BLAST was used to assign the taxonomy55.

All data were processed in R version 3.2.156. QIIME outputs were imported into R using the phyloseq57, package and further processed with vegan58, coin59, and resh ape60. Data visualization was performed using ggplot261. General community clustering was performed on the 20 most abundant OTUs (in terms of mean abundance across all samples) collapsed based on shared taxonomy at the species level using ward.D2 hierarchical clustering. β-diversity was assessed using Bray-Curtis distances and significance was calculated with PERMANOVA using the adonis function from the vegan package58. Statistical ellipses from the ggplot2 package were used to visualize the sample and method clusters on the NMDS plots. α-diversity was analyzed based on the Shannon’s alpha diversity and observed number of OTUs metrics and significance was determined using the Kruskal-Wallis test.

Data Availability

Data used in current study is available from the corresponding author upon reasonable request.

References

1. Benn, A. M., Heng, N. C., Broadbent, J. M. & Thomson, W. M. Studying the human oral microbiome: challenges and the evolution of solutions. Australasian dental journal 63, 14–24 (2018).
2. Yamashita, Y. & Takeshita, T. The oral microbiome and human health. Journal of oral science 59, 201–206, https://doi.org/10.2334/ jousanad16-0856 (2017).
3. Takahashi, N. & Nyvad, B. The role of bacteria in the caries process: ecological perspectives. J Dent Res 90, 294–303, https://doi. org/10.1177/0022034810379602 (2011).
4. Liu, B. et al. Deep Sequencing of the Oral Microbiome Reveals Signatures of Periodontal Disease. PLoS One 7 (2012).
5. Coventry, J., Griffiths, G., Scully, C. & Tonetti, M. Periodontal disease. BMJ 321, 36–39 (2000).
6. Hu, J., Polson, A. M. & Kang, C. Clinical and Molecular Applications of Periodontal Disease and Systemic Diseases: a Review of the Interrelations and Impacts of Microbes with Diabetes, Respiratory Diseases, Cardiovascular Diseases and Osteoporosis. Public Health 122, 417–433, https://doi.org/10.1016/j.puhe.2007.07.004 (2008).
7. Meurman, J. H., Sanz, M. & Janket, S. J. Oral health, atherosclerosis, and cardiovascular disease. Crit Rev Oral Biol Med 15, 403–413 (2004).
8. Huang, Y. J. & Lynch, S. V. The emerging relationship between the airway microbiota and chronic respiratory disease: clinical implications. Expert Rev Respir Med 5, 809–821 (2011).
9. de Pablo, P., Chapple, I. L., Buckley, C. D. & Dietrich, T. Periodontitis in Systemic Rheumatic Diseases. Nat Rev Rheumatol 5, 218–224, https://doi.org/10.1038/nrrheum.2009.28 (2009).
10. Tezel, M. et al. Chronic periodontitis and the incidence of head and neck squamous cell carcinoma. Cancer Epidemiol Biomarkers Prev 18, 2406–2412, https://doi.org/10.1158/1055-9965.epi-09-0334 (2009).
11. Michaud, D. S., Liu, Y., Meyer, M., Giovannucci, E. & Joshipura, K. Periodontal Disease, Tooth Loss and Cancer Risk in a Prospective Study of Male Health Professionals. Lancet Oncol 9, 550–558 (2008).
12. Saini, R., Saini, S. & Saini, S. R. Periodontitis: A risk for delivery of premature labor and low-birth-weight infants. J Nat Sci Biol Med I, 40–42 (2010).
13. Hayes, R. B. et al. Association of Oral Microbiome With Risk for Incident Head and Neck Squamous Cell Cancer. JAMA oncology, https://doi.org/10.1001/jamaoncol.2017.4777 (2018).
14. Whitmore, S. E. & Lamont, R. J. Oral bacteria and cancer. PLoS pathogens 10, e1003933, https://doi.org/10.1371/journal. ppat.1003933 (2014).
15. Lamont, R. J., Koo, H. & Hajishengallis, G. The oral microbiota: dynamic communities and host interactions. Nature Reviews Microbiology, 1 (2018).
16. Mukherjee, P. K. et al. Dysbiosis in the oral bacterial and fungal microbiome of HIV-infected subjects is associated with clinical and immunological variables of HIV infection. PloS one 13, e0200285 (2018).
17. Klimasova, J., Jiraskova Zakostelska, Z. & Tsakalova-Hogenova, H. Oral bacterial and fungal microbiome impacts colorectal carcinogenesis. Frontiers in microbiology 9, 774 (2018).
18. Delaney, C. et al. Fungi at the Scene of the Crime: Innocent Bystanders or Accomplices in Oral Infections? Current Clinical Microbiology Reports, 1–11 (2018).
19. Parker, B. J. et al. Bacterial Diversity in Human Subgingival Plaque. J Bacteriol 183, 3770–3783 (2001).
20. Consortium, H. M. P. A framework for human microbiome research. Nature 486, 215–221, https://doi.org/10.1038/nature11209 (2012).
21. von Wintzingerode, F., Gobel, U. B. & Stach, M. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol Rev 21, 213–229 (1997).
22. Wu, J., Lin, I.-H., Hayes, R. B. & Ahn, J. Comparison of DNA extraction methods for human oral microbiome research. *British Journal of Medicine and Medical Research* 4, 1980–1991 (2014).

23. Feigelson, H. S. et al. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. *Cancer Epidemiol Biomarkers Prev* 10, 1005–1008 (2001).

24. García-Closas, M. et al. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer Epidemiol Biomarkers Prev* 10, 687–696 (2001).

25. Lum, A. & Le Marchand, L. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. *Cancer Epidemiol Biomarkers Prev* 7, 719–724 (1998).

26. Smith, B. C. et al. The cervical microbiome over 7 years and a comparison of methodologies for its characterization. *PLoS One* 7, e40425, https://doi.org/10.1371/journal.pone.0040425 (2012).

27. Aas, J. A., Paster, B. J., Stokes, L. N., Olsen, I. & Dewhurst, F. E. Defining the normal bacterial flora of the oral cavity. *J Clin Microbiol* 43, 5721–5732, https://doi.org/10.1128/JCM.43.11.5721-5732.2005 (2005).

28. Dewhurst, F. E. et al. The human oral microbiome. *Journal of bacteriology* 192, 5002–5017, https://doi.org/10.1128/JB.00452-10 (2010).

29. Eren, A. M., Borisy, G. G., Huse, S. M. & Mark Welch, J. L. Oligotyping analysis of the human oral microbiome. *Proc Natl Acad Sci USA* 107, E875–E884, https://doi.org/10.1073/pnas.1409661111 (2014).

30. Melka, R., Fru, N., Morales, J. I. & Loewy, Z. G. Pharmacy: Oral Microbiome: Relationship of Stomatitis of Denture Wearers and Chronic Obstructive Pulmonary Disease (COPD) (2018).

31. Lim, Y., Totsika, M., Morrison, M. & Punyadeera, C. The saliva microbiome profiles are minimally affected by collection method or DNA extraction protocols. *Scientific reports* 7, 8523, https://doi.org/10.1038/s41598-017-07885-3 (2017).

32. Sohrabi, M. et al. The yield and quality of cellular and bacterial DNA extracts from human oral rinse samples are variably affected by the cell lysis methodology. *Journal of microbiological methods* 122, 64–72, https://doi.org/10.1016/j.mimet.2016.01.013 (2016).

33. Vesty, A., Biswas, K., Taylor, M. W., Gear, K. & Douglas, R. G. Evaluating the Impact of DNA Extraction Method on the Representation of Human Oral Bacterial and Fungal Communities. *PLoS One* 12, e0169677, https://doi.org/10.1371/journal.pone.0169677 (2017).

34. Cho, E. J., Sung, H., Park, S. J., Kim, M. N. & Lee, S. O. Rothia mucilaginosa Pneumonia Diagnosed by Quantitative Cultures and Intracellular Organisms of Bronchoalveolar Lavage in a Lymphoma Patient. *Ann Lab Med* 33, 145–149 (2013).

35. Binkley, C. J., Haugh, G. S., Kitchens, D. H., Wallace, D. L. & Sessler, D. I. Oral microbial and respiratory status of persons with mental retardation/intellectual and developmental disability: an observational cohort study. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 108, 722–731, https://doi.org/10.1016/j.tripleo.2009.06.027 (2009).

36. Ghannoum, M. A. et al. Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. *PLoS Pathogens* 6, e1000713 (2010).

37. Velegraki, A., Cafarchia, C., Gaitanis, G., Iatta, R. & Boekhout, T. Malassezia infections in humans and animals: pathophysiology, detection, and treatment. *PLoS pathogens* 11, e1004523 (2015).

38. Aguilu, I. et al. Associations of oral alpha-, beta-, and gamma-Human Papillomavirus Types With Risk of Incident Head and Neck Cancer. *JAMA oncology* https://doi.org/10.1001/jamaoncol.2015.5504 (2016).

39. Abusleme, L., Hong, B. Y., Dupuy, A. K., Straussbaugh, L. D. & Diaz, P. I. Influence of DNA extraction on oral microbial profiles obtained via 16S rRNA gene sequencing. *Journal of oral microbiology* 6, https://doi.org/10.3402/jom.v6i23990 (2014).

40. Lazarvic, V., Gaia, N., Girard, M., Francois, P. & Schrenzel, J. Comparison of DNA extraction methods in analysis of salivary bacterial communities. *PLoS One* 8, e67699, https://doi.org/10.1371/journal.pone.0067699 (2013).

41. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* 9, 357–359 (2012).

42. Ding, T. & Schloss, P. D. Dynamics and associations of microbial community types across the human body. *Nature* 509, 357–360, https://doi.org/10.1038/nature13178 (2014).

43. Biesbroek, G. et al. Deep Sequencing Analyses of Low Density Microbial Communities Working at the Boundary of Accurate Microbiota Detection. *PLoS One* 7 (2012).

44. Uyky, M., Zolnik, C. P., Patel, H., Levi, M. H. & Burk, R. D. Novel ITS1 Fungal Primers for Characterization of the Mycobiome. *mSphere* 2, e00488–e00417 (2017).

45. Hercules, C. Novocraf short read alignment package. http://www novocraft.com (2009).

46. Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J. & Knight, R. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature methods* 5, 235–237 (2008).

47. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 865–864 (2011).

48. Masella, A. R., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC bioinformatics* 13, 31 (2012).

49. DeSantis, T. Z. et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology* 72, 5069–5072 (2006).

50. Chen, T. et al. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* 2010 (2010).

51. Caporaso, J. G. et al. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26, 266–267 (2009).

52. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490 (2010).

53. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7, 335–336 (2010).

54. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Madé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584 (2016).

55. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* 215, 403–410 (1990).

56. R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria 2014).

57. McMurdie, P. J. & Holmes, S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8, e61217 (2013).

58. Olsen, J. et al. The vegan package. *Community ecology package* 10, 631–637 (2007).

59. Batdorf, C. S. (Google Patents, 1903).

60. Wickham, H. ggplot2: elegant graphics for data analysis. *Springer*, 2016.

Acknowledgements

This work was supported in part by the National Institute of Allergy and Infectious Diseases (AI072204), the NIDCR (DE026177), the Einstein Rockefeller-CUNY Center for AIDS Research funded by the NIH (AI24414) and the Einstein Cancer Research Center (P30CA013330) from the National Cancer Institute. Drs. Dowd and Waldron were funded in part by NIAID (AI121784).
Author Contributions
R.D.B. and C.P.Z. designed and conceived the study. J.R. drafted the manuscript and performed bench experiments. M.U. helped draft/edit the manuscript, performed bioinformatics analyses and prepared Figs 1–4. Z.C. helped edit/draft the manuscript and prepare Tables 1–2. C.P.Z. also helped edit the manuscript and performed bench experiments for ITS1 sequencing. H.E.J. was involved with study design and manuscript preparation. L.W. and J.B.D. were involved with study design and manuscript preparation. L.E.T. was involved with study design and manuscript preparation. In addition, R.D.B. helped draft/edit the manuscript. All authors reviewed the manuscript.

Additional Information
Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019