Journal of Oral Microbiology

COACTION PUBLISHING

ORIGINAL ARTICLE

The hidden 'mycobacteriome' of the human healthy oral cavity and upper respiratory tract

Lilia Macovei[1]*, Jon McCafferty[1], Tsute Chen[1], Flavia Teles[1,2], Hatice Hasturk[1], Bruce J. Paster[1,2] and Antonio Campos-Neto[1]*

[1]The Forsyth Institute, Cambridge, MA, USA; [2]Harvard School of Dental Medicine, Boston, MA, USA

The incidence of opportunistic non-tuberculous mycobacteria (NTM) infections has increased considerably in the past decades causing an array of infections, including respiratory and soft-tissue infections. NTM are ubiquitous and can be found in numerous environments, including households and water plants. However, NTM have not been reported to be associated with the healthy human oral microbiome. Since the oral cavity and upper respiratory track are the main ports of entry of microorganisms into the human body, elucidating NTM diversity and prevalence will assist in the assessment of the potential risks of infection elicited by these opportunistic pathogens. Here, we report the identification of a '*non-tuberculous mycobacteriome*' in healthy individuals. We employed a modified DNA extraction procedure in conjunction with mycobacterial-specific primers to screen niches in the oral cavity (buccal mucosa and dental plaque) and upper respiratory tract (nostrils and oropharynx) of 10 healthy subjects. A total of 50 prevalent operational taxonomic units sequenced on MiSeq (Illumina) using 16S rRNA V3–V4 region were detected across all screened niches, showing the presence of diverse NTM communities. NTM DNA was detected in the nostrils of all 10 subjects, in buccal mucosa of 8 subjects, in the oropharynx of 7 subjects, and in the dental plaques of 5 subjects. Results from quantitative PCR showed each individual harbored 10^3–10^4 predicted NTM per each screened niche. The modification of standard DNA isolation methods to increase sensitivity toward mycobacterial species represents an important step to advance the knowledge of the oral as well as the overall human microbiome. These findings clearly reveal for the first time that healthy individuals harbor a 'non-tuberculous mycobacteriome' in their oral cavity and upper respiratory tract and may have important implications in our understanding of infections caused by NTM.

Keywords: *oral microbiome*; *Mycobacterium*; *non-tuberculous-mycobacteria*; *periodontal disease*

*Correspondence to: Lilia Macovei, Bioanalytical and Biomarker Development, Shire 300 Shire Way, Lexington, MA 02421, USA, Email: lmacovei@shire.com; Antonio Campos-Neto, The Forsyth Institute, 245 First Street, Cambridge, MA 02142, USA, Email: acampos@forsyth.org

To access the supplementary material for this article, please see Supplementary files under 'Article Tools'

Received: 22 September 2014; Accepted: 19 January 2015; Published: 13 February 2015

Non-tuberculous mycobacteria (NTM) are natural inhabitants of various environments and are opportunistic pathogens particularly in HIV-positive individuals, children, and elderly populations (1, 2). In the past decades, the incidence of NTM infections increased considerably in the United States (3–7). Infections with NTM were shown to lead to mycobacteriosis; chronic pulmonary infections, such as non-cystic fibrosis bronchiectasis; and chronic obstructive pulmonary diseases in both immunocompromised and immunocompetent individuals with no definable risk factors (8, 9). NTM mimic diseases caused by the intracellular pathogen *Mycobacterium tuberculosis* (Mtb), involving any organ or part of the body (10). Often radiological and clinical manifestations of NTM infections are difficult to distinguish from tuberculosis (TB), especially in advanced HIV co-infected patients. Most importantly, these infections are often misdiagnosed as TB or wrongly labeled as non-responsive or drug-resistant TB cases, as most of NTM species are resistant to conventional anti-TB drugs (11, 12).

Although over 180 species of *Mycobacterium* are widely found in the environment, including soils, natural water sources, tap water, shower cap biofilms, insects, aerosols, vertebrates, dust, and sawdust (13–16) (www.dmsz.de), NTM have not been described as components of the human microbiome. Therefore, efforts to identify NTM in human-associated niches are highly significant for the

Citation: Journal of Oral Microbiology 2015, **7**: 26094 - http://dx.doi.org/10.3402/jom.v7.26094

understanding of the potential etiology and physiopathology of opportunistic diseases caused by these organisms.

The 16S ribosomal RNA currently represents the most accepted marker used to characterize the human-associated bacterial communities (17, 18). To facilitate inter-study comparisons in the Human Microbiome Project, conventional lysis with bead beating using the PowerSoil DNA extraction kit (MOBIO) has been used (19, 20). This DNA extraction protocol has been used in numerous studies to characterize human-associated communities, including oral and upper respiratory tract microbiota (17, 20, 21). Although these studies have identified organisms representative of most bacterial taxons for each human-associated niche, NTM were rarely detected. Few issues that can cause biases toward hard-to-lyse bacteria as the NTM may be a consequence of the inefficient extraction methods, leading to low genomic DNA yields (22). The low genomic DNA coupled with relative polymerase chain reaction (PCR) amplification and the low copy number of the 16S rRNA genes (1, 2) on NTM genomes compared to multi-copies of 16S rRNA genes (5–15) of other bacterial species represent potential biases in detection of this segment of the microbiota (23–25).

In this study, we used a modified DNA extraction procedure to detect the presence of NTM in healthy oral and upper respiratory tract microbiome. We used 16S rRNA gene sequencing approach and quantitative PCR to determine NTM relative abundance and diversity. Our results demonstrate that NTM species are clearly present in oral and upper respiratory tract communities, representing a potential reservoir for infection. We posit that efforts to identify NTM in human-associated niches are needed and will be highly significant for the understanding of the potential etiology and physiopathology of the diseases they cause.

Materials and Methods

Participant enrollment

Ten healthy individuals, four males and six females aged between 21 and 65 were recruited at The Forsyth Institute Center for Clinical and Translational Research (CCTR) (Supplementary Table 1). The study protocol was approved by the Institutional Review Board at The Forsyth Institute (IRB#00000037). A written informed consent was obtained from all subjects after providing an explanation of the study and details about sample collection, prior to participation. Exclusion criteria for the study were as follows: 1) use of any antimicrobials within the past 2 months, 2) pregnancy, and 3) any chronic pulmonary and other systemic diseases or chronic use of medications that could potentially affect the oral or upper respiratory tract.

Sample collection and DNA extraction

Individual samples were collected from various niches in the oral cavity and upper respiratory tract. Oral cavity samples were collected from buccal mucosa (pooled from right and left sides), posterior wall of oropharynx using sterile swabs (BBL CultureSwab; Becton, Dickinson and Co.), and teeth (supra and subgingival samples) using sterile Gracey curettes (Hu-Friedy Mfg. Co., LLC; Chicago, IL) from sites with probing pocket depth of <4 mm. Upper respiratory tract samples included the samples from the nostrils (pooled from right and left sides) collected using sterile swabs. The posterior wall of the oropharynx was swabbed without touching the tongue, tonsils, uvula, or other oral parts. Dental samples collected from four subgingival and four supragingival sites were pooled separately into Eppendorf tubes containing 150 µl of Tris-EDTA buffer and processed for DNA extraction. Samples were rapidly frozen at -80°C until analysis. The samples, both the swabs and dental plaque, were diluted into a sterile 2-ml tube with 750 µl beads and 60 ul C1 buffers (PowerLyser Soil kit; MP Biomedicals) for immediate nucleic acid extraction. To enhance genomic DNA extraction, samples were incubated for 10 min in a boiling water bath. The tubes were run on a bead-beating machine (MoBio) for 60 s at 6.0 m/s. DNA was purified from the supernatants using the manufacturer's protocol (PowerLyser Soil kit; MP Biomedicals).

16S rRNA gene PCR amplification and mycobacterial cloning

To characterize mycobacterial communities, we amplified 16S rRNA genes from extracted gDNA using two sets of primers: mycobacterial-specific and universal primers. The strategy for obtaining a mycobacterial community profile was a two-step approach due to the 16S rRNA sequence homology with *Corynebacteria* spp., inducing a detection bias in NTM detection. Therefore, the first step in amplification of the mycobacteria DNA was using a set of universal F24 (5′-GAG TTT GAT YMT GGC TCA G-3′) primer combined with mycobacterial-specific reverse primer R990 (5′-CGT CCT GTG CAT GTC AAA-3′) (yielding a product of 981 bp). This primer set amplified NTM and Corynebacteria spp., as shown by the sequenced PCR products, and to select for NTM a nested PCR using 1 ul of purified product from step 1 with mycobacterial-specific primers MycF121 (5′- CGT GGG TGA TCT GCC CT-3′) and Myc858R (5′-CGG CAC GGA TCC CAA GG-3′) (yielding a product of 737 bp) (15). For each sample, three amplified products were pooled, purified using PCR purification kit (Invitrogen) and quantified by gel electrophoresis using a 1% agarose DNA quantification ladder (Invitrogen Corp.), which was also used for determination of the product size. As a control for the PCR amplification, the same amount (5 ng) of gDNA was amplified with primers for *rpoB* gene coding

Citation: Journal of Oral Microbiology 2015, **7**: 26094 - http://dx.doi.org/10.3402/jom.v7.26094

the β subunit of the RNA polymerase, which is known to be distinctively conserved in members of the genus *Mycobacterium* (26). All samples with products from mycobacterial-specific primers were positive with *rpoB* gene (not shown).

To confirm that the lysis of the total community DNA was not affected by our method, we characterized total bacterial communities from each sample using PCR amplification with universal primers (9F: 5′-GRG TTY GAT YMT GGC TCA G-3′ and 1541R: 5′-RAA GGA GGT GWT CCA DCC-3′) (27) and the purified product was used for MiSeq library preparation and sequencing as previously described using MiSeq (Illumina) (17). Primers for MiSeq sequencing were chosen to amplify in V3–V4 region, as these were considered better at detecting individual organisms within V3–V5 region generating two libraries (28). Mycobacterial species homology using the 16S rRNA gene has previously been reported to be 96.6% (29). However, this percent similarity is based on the full length 16S rRNA and may not be applied to the V3–V4 region. We chose the conventional 97% cutoff during the operational taxonomic unit (OTU) clustering stage and 98.5% BLAST cutoff for taxonomy assignment, both higher than 96.6%, to ensure better classification.

To confirm the amplification of mycobacterial-specific nucleic acid sequences, samples from two different niches were cloned and sequenced (10 clones per site). Amplicon pools were ligated and cloned using the standard protocol from the TOPO TA cloning kit for sequencing (Invitrogen). Individual cloned 16S rRNA gene sequences were first amplified using the M13F and M13R primers (TOPO TA cloning kit for sequencing manual) and then sequenced from the 5'end with the Myc121 primer using an ABI3700 (Applied Biosystems, Inc.). After primer and vector sequences were removed, the 16S rRNA gene sequences were trimmed by removing any leading and trailing bases that contained ambiguities and for which confidence was less than 25%. Each sequence of 400–500 bp was manually inspected for any remaining base caller errors by using Sequencher (Gene Codes Corp.). Sequences were examined by using a BLAST analysis (30).

16S rRNA profiling

We used an open reference-based OTU approach as previously described (17, 18, 31) using QIIME 1.7.0 (32). Briefly, after sequences were demultiplexed, quality filtered, and the paired-end reads were merged together (33). The resulting 7,120,989 sequences across all samples with OTUs at the 97% sequence similarity level were picked using UCLUST (34) against the Greengenes database pre-clustered at 97% identity (35). Chimera screening was conducted through UCHIME (36) and OTUs were assigned taxonomy by BLASTn v2.2.22 (30) mapping to the RDP Release 11 database (37) (http://rdp.cme.msu.edu/) and a 98.5% alignment threshold. Before BLAST analysis of the OTUs to the database, we filtered out any short reads ($<1{,}000$ bases), reads without species level annotation, and reads that contained an unclassified annotation. This reduced the overall size of the RDP Release 11 database to 124,464 sequences. Alpha and Beta diversity was calculated through packages provided in the QIIME 1.7.0 software environment (32). Clustering sequences into OTU space is a way of representing taxonomic similarity in the dataset. QIIME's default setting for OTU clustering sequence similarity as representation of species is 97% (32). However, we increased the similarity threshold for the BLAST algorithm to 98.5% to ensure that the OTUs would get the best possible assignment from the reference database.

Library comparison

Oral and upper respiratory samples from 10 healthy subjects (Table 1) were employed for this study. Samples from buccal mucosa, nostril, dental plaque (pooled subgingival and supragingival), and oropharynx were collected for the profiling of microbial communities. Ten minutes boiling step and an enzymatic digestion step were introduced before mechanically processing the samples. The mycobacterial community was targeted using the nested PCR approach. The first PCR amplification was performed with universal primer F24 and a reverse mycobacterial-specific primer. The amplified product was used in a second PCR with mycobacterial-specific primers only. PCR products were attached with sequencing barcodes unique for each sample and pooled together for a single MiSeq sequencing run. A total number of 17,051,480 raw reads (paired-end) were obtained from MiSeq and assembled with FLASH (33) into 13,581,978 input sequences for QIIME (version 1.7.0) run. After quality checking of sequence length (250–500 bp) and quality filter (qscore ≥ 20) a total number of 7,120,989 sequences were used with UCLUST for OTU clustering. For mycobacterial community, the final sequence number was 4,860,872 and for the total microbial community

Table 1. Characterization of the human subjects enrolled in the study

Subject	Gender	Age	Ethnicity
1	M	48	Asian
2	M	49	Asian
3	M	46	Caucasian
4	F	38	Asian
5	F	44	African American
6	F	42	Hispanic
7	F	43	Asian
8	M	51	Asian
9	F	46	Hispanic
10	F	49	Caucasian

2,260,117 sequences were obtained. The sequences were deposited and publicly available at the HOMD database (ftp://www.homd.org/publication_data/20140722/).

Statistical analysis

All detected taxa by the MiSeq from all 10 healthy subjects and were added to the local database containing downloaded V3–V4 region of the NCBI sequences of Mycobacteria spp. and phylogenetic relationships were compared using neighbor-joining method in MEGA 6 with 500 replicates (38). To compare the numerically dominant phyla from each body site, correlation coefficients (Pearson and Spearman correlation indices) and linear regression were performed using Prism 6 (GraphPad Software, Inc.) using relative bacterial prevalence calculated based on percentages of the total community. To determine if the community compositions within and between groups were different, we used one-way analysis of variance (ANOVA). We performed weighted and unweighted UniFrac analyses using the neighbor-joining tree of all taxa represented on the libraries that have >250–500 bp 16S rRNA sequences. A *t*-test on the UniFrac distance matrix was used to determine if the UniFrac distances were on average significantly different for the bacterial communities detected in the two body sites. To determine how bacterial community compositions varied across samples, we also compared unweighted UniFrac profiles for each sample using principal coordinate analysis (PCoA) in QIIME (32).

Quantification of NTM community using q-PCR

The content of *Mycobacterium* rDNA operons was quantified with q-PCR to have an estimation in each community. FastSYBR® Green Master Mix (Applied Biosciences) was used in a final volume of 20 µl with 5 ng of total DNA per reaction for each sample with two sets of primers. The first set of forward primer MycoARB210 (5′-TTT GCG GTG TGG GAT GGGC-3′) and reverse MycoARB585: (5′-CGA ACA ACG CGA CAA ACCA-3′) amplifying mycobacterial 16S rRNA gene (39) and the 65 kDa heat-shock protein (*hsp*65), previously identified as mycobacterial-specific (40), with the forward Hsp65F: (5′-ACC AAC GAT GGT GTG TCC AT-3′) and reverse primer Hsp65R: (5′-CTT GTC GAA CCG CAT ACC-3′) were used at 0.5 µM each. Thermal program was 95°C for 15 min, 45 cycles of denaturing 15 s at 94°C, annealing 30 s at 61°C for MycoARB primer set and 60°C for Hsp65, extension 30 s at 72°C, plate read at 72°C for 1 s and plate read at 80°C for 1 s using Applied Biosystems StepOnePlus™ Instrument. The cycling was followed by a final extension at 72°C for 7 min, and a melting curve analysis from 65–95°C with a plate read every 0.5°C. Two MycoARB585 PCR products were sequenced and confirmed the correct target amplification product. To compare with the total bacterial gDNA in each sample, we used universal 16S rRNA gene primers Univ F: (5′-TCC TAC GGG AGG CAG CAGT-3′) and reverse primer Univ R: (5′-GGA CTA CCA GGG TAT CTA ATC CTG TT-3′) (41).

The universal primers were used with 95°C for 20 s; 40 cycles: 95°C for 1 s and 58°C for 20 s and extension 30 s at 72°C. For melting curve analysis, PCR products were incubated for 15 s at 95°C, annealing at 60°C for 1 min, ramp up to 95°C with plate read every 0.3°C increase, and a final 95°C for 15 s.

Samples were run in duplicate and 10-fold serial dilutions of genomic *M. fortuitum* subsp. *fortuitum* DSM 46621 genome (NCBI RefSeq Assembly ID: GCF_000295855.1) DNA (starting with 10 ng/µl/reaction) were used to generate a standard curve. The standard curve for the bacterial q-PCR reaction had an R^2 value of 0.98.

Results

Diversity of mycobacteria and total bacterial community present in the oral and upper respiratory tract

Correlations between the relative abundance of taxa within each tested niche among 10 healthy subjects were initially examined. For the purpose of these analyses, relative abundance refers to the evenness of distribution of individuals among species in a community. Two communities may be equally rich in species but differ in relative abundance. In this study, relative abundance was calculated for each library as the number of sequences identified for each genus (99% similarity) divided by total number of sequenced reads. The same principle was applied for the sequenced libraries with mycobacterial-specific primers. Figure 1 shows that the total bacterial community of nostrils was characterized by significantly reduced bacterial richness (number of bacterial types detected; $P < 0.0001$, one-way ANOVA test; Fig. 1a) with a detected taxa mean of 55.3 ± 9.6 ($n = 10$) compared to buccal mucosa, dental plaque, and oropharynx [average of 175 ± 8.73 (SEM)]. Evenness (relative distribution of bacterial types; $P < 0.0001$, one-way ANOVA test; Fig. 1b) and Shannon's diversity (metric calculated using richness and evenness indices; $P < 0.0001$, one-way ANOVA test; Fig. 1c) confirmed reduced diversity observed in nostrils compared to the other tested body sites.

Sequenced mycobacterial genomes from NCBI contain only one or two copies of rRNA genes; therefore, the total number of sequences detected can only reflect the frequency of the 16S rRNA genes, not the cellular abundance, due to different gene/cell ratio. To evaluate diversity of mycobacterial community in oral and upper respiratory tract, we used targeted nested PCR approach. Positive PCR amplification products obtained with mycobacterial primers per each subject were sequenced and were used to estimate the relative abundance of the mycobacterial community. Three libraries were cloned and sequenced to

Citation: Journal of Oral Microbiology 2015, **7**: 26094 - http://dx.doi.org/10.3402/jom.v7.26094

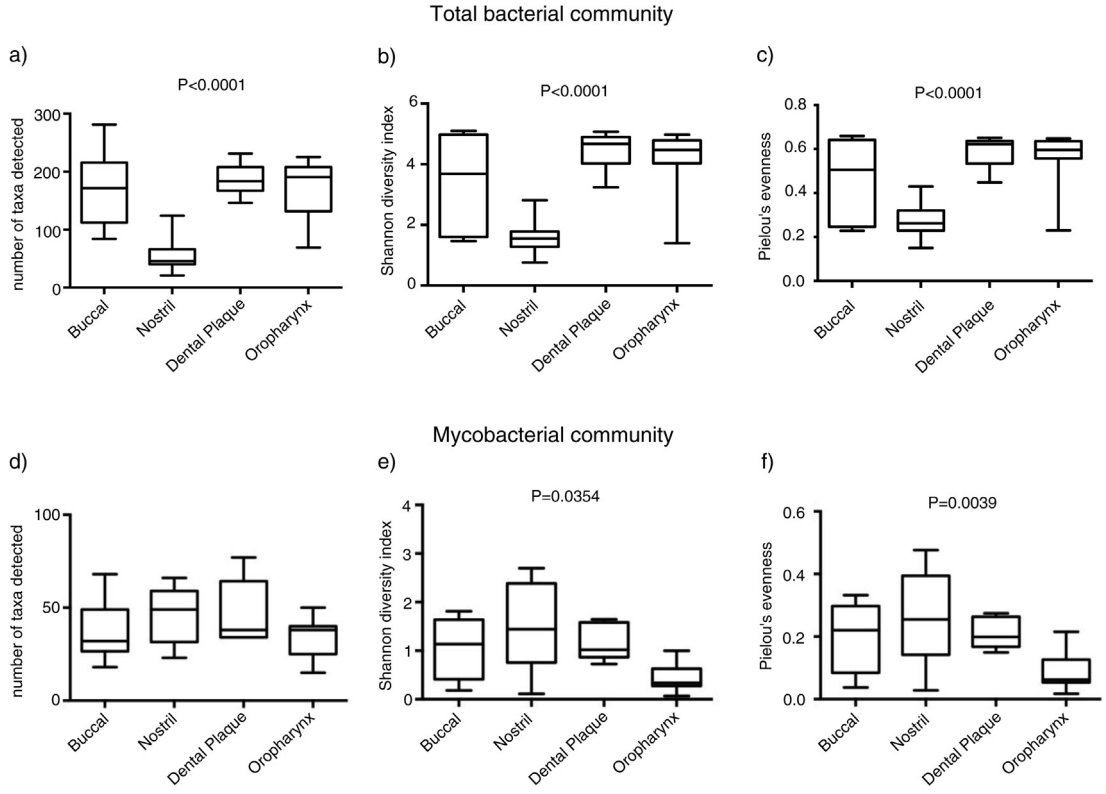

Fig. 1. Comparative analyses of bacterial community in oral and upper respiratory tract. (a–c) Richness (a), Shannon diversity (b), and evenness (c) indices of total bacterial community grouped by niche from 10 individuals. (d–f) Richness (d), Shannon diversity (e), and evenness (f) indices of mycobacterial-specific community based on total number of positive PCR products: Buccal ($n=8$), Nostril ($n=10$), Dental Plaque ($n=5$), and Oropharynx ($n=7$). Values represent means $\pm$ SEM. Shown P-value was calculated using one-way ANOVA.

confirm the presence of the mycobacteria. Detected diversity of mycobacterial community ranged from 34.0 ± 4.3 (SEM) in oropharynx to 46.8 ± 7.2 (SEM) detected taxa in dental plaque ($P > 0.005$, one-way ANOVA; Fig. 1d). The low mycobacterial diversity in oropharynx is confirmed by the evenness ($P = 0.0354$, one-way ANOVA test; Fig. 1b) and Shannon diversity indices ($P = 0.039$, one-way ANOVA test; Fig. 1c). The inter-subject variability tested for total bacteria and mycobacterial communities were insignificant (see Supplementary Fig. 1).

Phylogenetic distribution of mycobacterial community

A total of 600 OTUs were detected across all screened niches in healthy individuals and 61 prevalent OTUs were detected when a cutoff of >200 sequences per library was applied. We chose to classify OTU assignment using the BLAST algorithm because this method is currently and frequently being used in metagenomic studies. It provides high performance and accuracy, and it has been tested and demonstrated to be comparable to the RDP classifier (42, 43). Eleven taxa clustered with *Corynebacterium* genus and 50 OTUs belong to *Mycobacterium* genus. These taxa cluster with 85 out of 191 species sequences as depicted by Unweighted Pair Group Method with Arithmetic Mean (UPGMA) hierarchical clustering is shown in Fig. 2. None of the detected taxa clustered with the intracellular pathogens' clade of *M. tuberculosis* complex and other pathogens (*M. leprae*). This branch contained 39 other mycobacterial species due to limited nucleotide variability in the V3-V4 region of the 16S rRNA sequences. The majority of the Mycobacterial taxa were found in all four niches with few exceptions. The nostril and dental plaque environments harbored five and two unique mycobacterial OTUs, respectively, as shown by stacked colored bars (Figs. 2 and 3).

Relative quantitative distribution of NTM communities present in oral and upper respiratory tract

To confirm the NTM sequence numbers present in each sample, we used quantitative PCR (qPCR) with two sets of mycobacterial-specific primers for 16S rRNA and *hsp*65 genes and a set of universal 16S rRNA primers to screen DNAs from 40 libraries (Fig. 4a–c). qPCR showed good sensitivity in detecting NTM ($r^2 = 0.98$) in each run with all primer sets. Buccal mucosa and nostril communities had the higher numbers of NTM based on *hsp*65

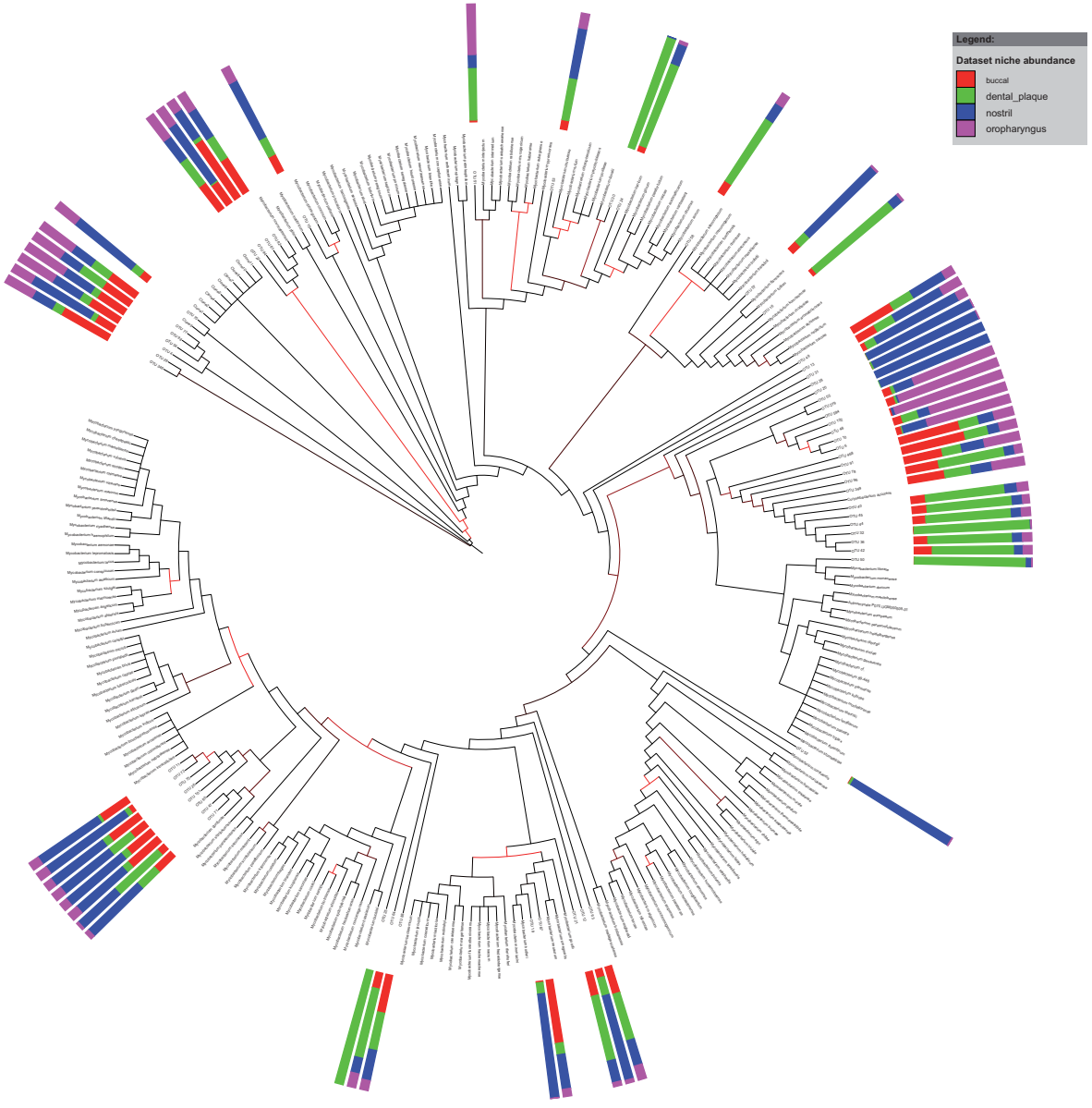

Fig. 2. Phylogenetic distribution of Mycobacterial community from oral and upper respiratory tract. Phylogenetic tree based on UPGMA clustering of the sequences representing V3–V4 region of the 16S gene sequences (see Materials and Methods). Stacked bar plots show the relative abundance of taxonomic assignments to each niche within a clade of potential species.

qPCRs (Fig. 4b), while fewer NTM were detected in dental plaque and oropharynx. Of the detected NTM species, some species seem to contain two copies of 16S rRNA, since the detected mycobacterial-specific 16S rRNA qPCR is showing higher numbers (Fig. 4c).

qPCR confirmed nested PCR amplification. NTM DNA was detected in every screened nostril library, in only 80% of buccal communities, 70% of oropharynx NTM, and 70% of the screened dental plaque compared to 50% of sequenced libraries that showed NTMs by nested PCR (five out of ten individuals) (Supplementary Table 1). Two positive samples with *hsp*65 qPCR in dental plaque showed no product with nested PCR approach, which could be due to a series of factors including suboptimal number of gDNA for detection, primer miss-amplification, or dimerization. The individual numbers of NTM were found to be variable 10^3–10^4 copies per each screened niche, which is a representative proportion of the total bacterial community.

Phylum-level comparison of total bacterial communities present in oral and upper respiratory tract of healthy subjects

Although this evaluation was not intended to validate the overall oral microbiome, the next experiments were designed to verify the impact that the boiling step to lysis

Citation: Journal of Oral Microbiology 2015, **7**: 26094 - http://dx.doi.org/10.3402/jom.v7.26094

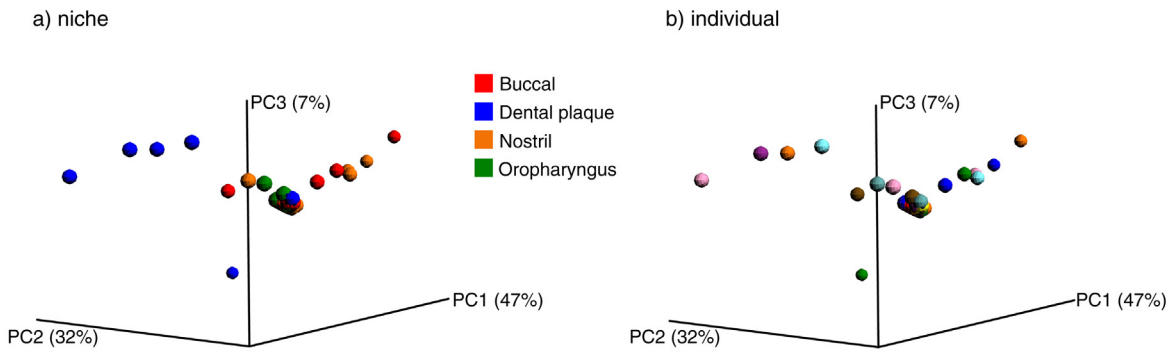

Fig. 3. Mycobacterial communities diversity grouped by site (a) and healthy individual or subject (b). Weighted UniFrac-based PCoA, where each symbol represents the value for a sample, with the shape of the symbol indicating the sites. The percentages of variation explained by the plotted principal coordinates (PCo2 and PCo3) are indicated on the axes.

mycobacterial protocol had on the general bacterial community. For this assessment, we used universal 16S rRNA primers to amplify 5 ng per each sample of the same DNA used for mycobacterial-specific primers. To classify sequence taxonomy, we used 99% similarity match to RDP database (32). The analyses detected overall nine phyla in four tested sites, with seven from the nostrils (least diverse) and nine for the rest (Fig. 5). On average, the most abundant were the *Firmicutes* (61.9%), *Proteobacteria* (11.3%), *Fusobacteria* (10.6%), *Bacteroidetes* (8.8%), *Actinobacteria* (5.4%), and *Spirochaetes* (0.34%). Rare phyla (those with average abundances of $<0.01\%$) included the *Chloroflexi*, *Tenericutes*, and *Synergistetes*. Few phyla accounted for the majority of the sequenced 16S rRNA bacterial community, with similar phylogenetic distribution patterns as previously described for these colonization sites, suggesting good concordance between DNA isolation method used in this study compared to other profiling approaches (44–46).

Of all sequenced niches, nostrils exhibited the lowest diversity (Fig. 1a–c and 5). In nostril samples, two main phyla were detected, *Firmicutes* and *Actinobacteria* (light blue and dark blue, respectively, in Fig. 5), or *Firmicutes* and *Proteobacteria* (purple). An averaged phylum-level distribution pattern for each site demonstrated that each has a phylum-level distribution distinct from each other. Inter-individual variability in each site is observed.

Dental plaque and buccal communities showed inverse correlation between *Firmicutes* and *Fusobacterium* phyla (Fig. 6a, b and d). Relative prevalence of *Firmicutes* and *Fusobacteria* in dental plaque communities has inverse Pearson correlation coefficient $r=-0.945$ $P<0.001$, $r^2=0.894$) (Fig. 6a, b). The buccal community relationship between the two phyla did not appear linear and was the reason for using the Spearman correlation coefficient rather than the Pearson correlation coefficient (Spearman $r=-0.648$; $P=0.049$; $P<0.05$) (Fig. 6c). A similar trend was observed in the data from the nostrils

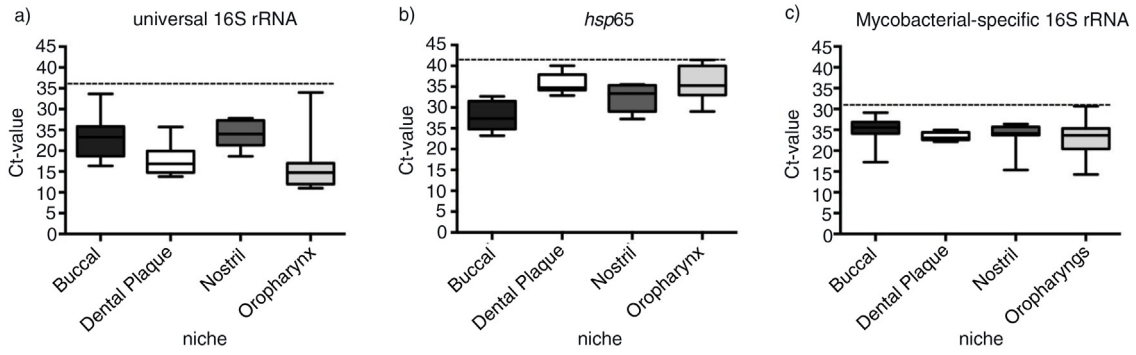

Fig. 4. NTM abundance measured by qPCR from oral and upper respiratory tract communities. (a) Universal 16S rRNA; (b) hsp65 gene (1 copy/genome); (c) Mycobacterial-specific 16S rRNA gene (1–2 copies/genome). 5 ng/ul of gDNA was used in all reactions of 20 ul. A duplicate 10-fold dilution series of genomic DNA of *M. fortuitum* ranging from 10–105 CFUs was used to generate a standard curve for each primer set (both with R^2 for each set of 0.98); q-PCR was performed using the Fast SYBRT® Green Master Mix (Life Technologies). Dotted line shows the qPCR cutoff determined by the reactions with no template run.

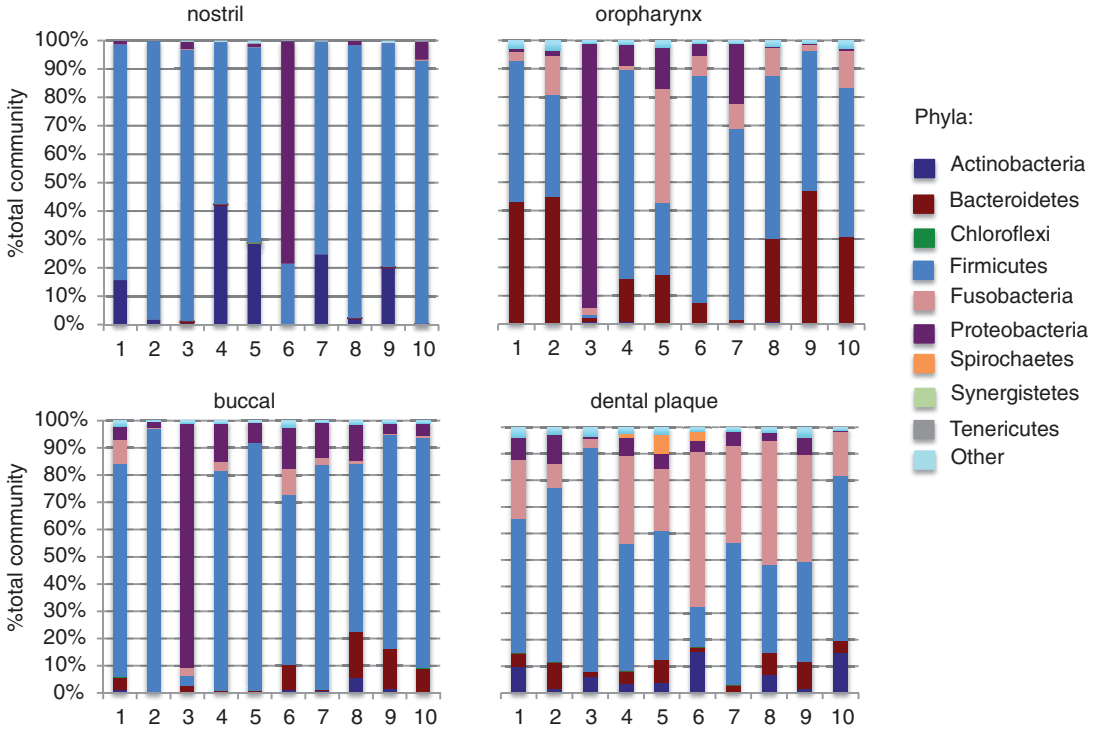

Fig. 5. Phylum-level comparison of oral and upper respiratory tract total bacterial communities amplified with universal 16S rRNA primers. Stacked bar plots show relative abundance of main phyla in each niche of 10 healthy individuals and were generated by BLAST analysis of the V3–V4 region sequences with 99% similarity against RDP11 database.

between *Actinobacteria* and *Proteobacteria* phyla (Fig. 6d; Spearman $r = -0.697$; $P = 0.0306$; $P < 0.05$).

The *Actinobacteria* in total bacterial communities were mainly represented by *Corynebacteria*, *Actinomycetaceae*, *Micrococcaceae* families, and *Mycobacteriacea* (Supplementary Table 2). Compared with previous published microbiome libraries using the same universal primers, the current lysis protocol improved detection of mycobacterial spp. from 1–2 sequences to average of 46.1 sequences (0.02–0.05% relative abundance) per individual microbiome. Therefore, when universal primers are used, the modified DNA lysis protocol showed that detection of the most common bacteria present in the human oral microbiome was not impaired and indeed improved detection of actinobacterial taxa.

Discussion

In this study, we detected NTM species in human oral and upper respiratory microbiomes in healthy subjects. NTM were not detected as a part of these communities in other studies, likely due to inadequate bacterial cell lysis. We used a modified community DNA extraction protocol (10 min boiling before bead beating) and surveyed these communities using mycobacterial-specific 16S rRNA gene, qPCR and cloning. Notably, when total microbial community was investigated with universal 16S rRNA primers after improved lysis, the sequence number for NTM increased (<46.1 seq. per library), while amplification with mycobacterial-specific primers allowed detection of 10^3–10^4 NTM the oral and upper respiratory tract. It is important to note that these two different results come from separate experiments, that is, Illumina sequencing and qPCR, respectively. Nonetheless, both showed improved/positive detection of NTM after the use of the improved lysis method. Using this approach, we found diverse NTM taxa (OTUs) present in each niche, branching with opportunistic pathogens such as *M. fortuitum*, *M. mucogenicum* and *M. neoaurum*, and saprophytes found in soil.

These results are supported by other studies, in which the 16S rRNA oligo-based microarray (PhyloChip) was employed and mycobacterial DNA was detected in the nostril and upper respiratory tract (44, 45) and the gut (47). However, when the hybridization datasets were compared to data generated by direct 16S rRNA sequencing, low abundant taxa such as Mycobacteria were not found (45). Hard-to-lyse bacteria, such as species of *Mycobacterium* genus, require a more complex DNA extraction (22). Another interfering issue in appreciating existing mycobacterial numbers is apparently the 16S rRNA gene copy number per genome. Species with multiple copies will have a prevalent abundance in the sequencing library of the community, while those with one or two copies per genome (as in the case of

Citation: Journal of Oral Microbiology 2015, **7**: 26094 - http://dx.doi.org/10.3402/jom.v7.26094

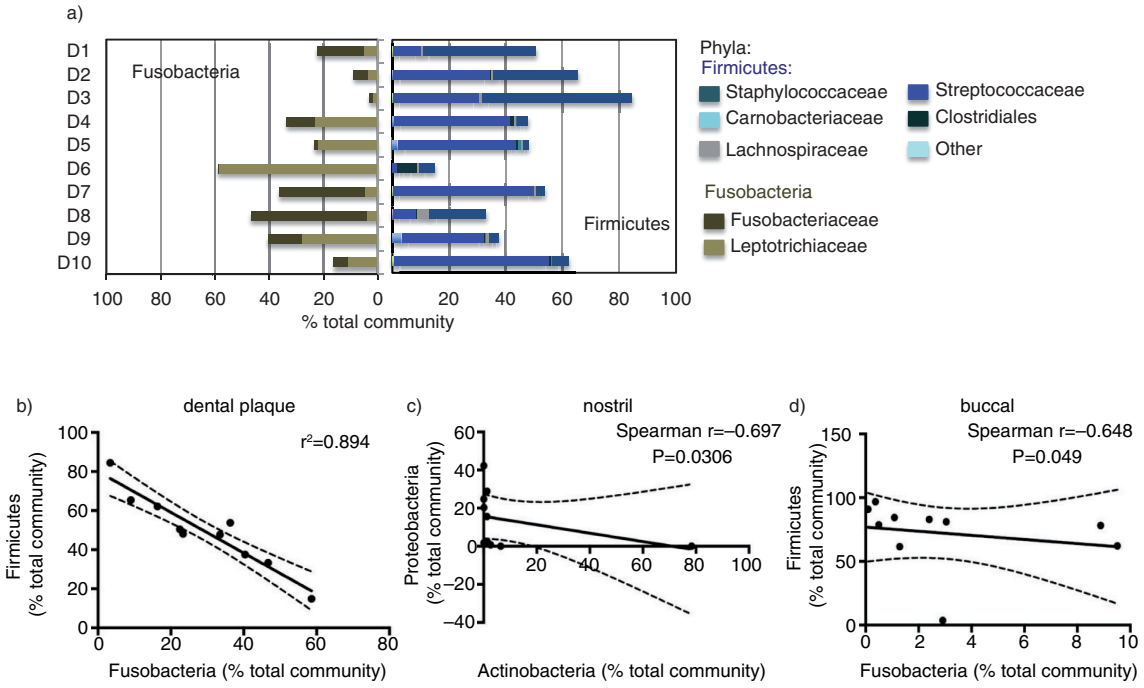

Fig. 6. Correlations of phyla in total bacterial communities. (a) Stacked bar plots show relative abundance of main phyla in dental plaque which were generated by BLAST analysis of the V3–V4 region with 99% similarity against RDP11 database; (b) Inverse correlation between the relative prevalence (percentages of the total community) in dental plaque of members of the phylum *Firmicutes* and *Fusobacterium*, where dashed lines indicate 95% confidence intervals and Pearson correlation coefficient = −0.945; $P < 0.0001$; (c) Inverse correlation between phylum Proteobacteria and the phylum Actinobacteria in the nostril communities (Spearman correlation coefficient $r = -0.0.697$; $P < 0.05$); (d) Inverse correlation between phylum *Firmicutes* and the phylum *Fusobacteria* in the buccal mucosae communities (Spearman correlation coefficient $r = -0.648$; $P < 0.05$).

mycobacteria), although equivalent in numbers, will remain underrepresented. The relative PCR amplification efficiency leaves the low-proportion taxa underrepresented so that a larger number of sequences are needed to detect the 'rare biosphere' bacterial taxa (24, 25).

To determine relative numbers of NTM community in healthy individuals, we used q-PCR with *Mycobacterium*-specific primers to screen DNAs from all 40 sources. q-PCR demonstrated that different niches exhibited substantial NTM microbiota variation in healthy individuals. NTM are known to abound in diverse environments, including municipal water systems and showerheads (16, 48). Consequently, nostrils and buccal mucosa are the most common niches for NTM detection, compared to dental plaque and oropharynx. Our oral and upper respiratory tract encounters many incoming bacteria originating from our daily habits, social interaction, food, beverages, and breathing. However, little is known about how trespassing bacteria colonize or affect these niches. In fact, only about 35% of oral bacteria are known species, the others are unknown and/or unculturable (46).

The presence of NTM in the microbiome of the oral cavity has important medical implications. First, it should help us to better understand and track the original source of the NTM that are associated with increased infections particularly in immunocompromised patients. Second, NTM present in the oral microbiome could be associated with the high variability that is observed in humans vaccinated with the attenuated *Mycobacterium bovis* bacillus Calmette–Guérin (BCG). Experimental evidence obtained from animal studies suggest that this failure could be related to preexisting immune responses to antigens that are shared between BCG and environmental NTM (49–51). Thus, prior exposure of guinea pigs to *M. fortuitum*, *M. avium*, or *M. kansasii* impaired the anti-tuberculosis protection induced by vaccination of the animals with BCG in 15, 50, or 85%, respectively (52). However, when the combination of BCG plus an NTM was used, the immune responses were no more protective than BCG alone. These studies lead to the hypothesis that NTM exposure may influence BCG protection probably by providing a degree of partial protection on which BCG cannot improve, or by 'immunizing' the animals against the upcoming vaccine, thus preventing the 'take' of the vaccine (53–55). Therefore, it is possible that presence of NTM in the oral microbiome may indeed influence the fate BCG vaccination in humans.

In addition, because of the unique glycolipid content that *Mycobacterium* species have in their cell walls, NTM may be important players in the development of

periodontal disease. Mycobacterial lipids are potent immunomodulators and stimulants of pro-inflammatory cytokines and of Toll-like receptors (56–58). Therefore, under circumstances that favor the replication of the NTM cells in the dental plaque, it is possible that the inflammatory signals induced by the NTM lipids could contribute to the development of periodontal disease including bone loss.

Our study is the first culture-independent molecular survey for NTM in healthy oral and upper respiratory tract community. The finding of NTM taxa in these niches prompts more questions and further investigations of a careful characterization at the species level of these community members. Oral cavity and upper respiratory tract have long been recognized niches for pathogen carriage with critical role in both, invasive diseases as well as respiratory tract infections. The dynamics of pathogen colonization results in various outcomes, ranging from persistence with disease, asymptomatic states, or complete clearance. Little is known about mechanisms of NTM colonization and persistence in human body sites. However, the increase of NTM infections in developed countries motivates further studies to better understand the role of these bacteria in health and disease.

Acknowledgements

This work was supported by Pilot Grant from The Forsyth Institute (LM) and R01 AI076425 (A.C-N.).

Conflict of interest and funding

There is no conflict of interest in the present study for any of the authors.

References

1. Danielides V, Patrikakos G, Moerman M, Bonte K, Dhooge C, Vermeersch H. Diagnosis, management and surgical treatment of non-tuberculous mycobacterial head and neck infection in children. ORL J Otorhinolaryngol Relat Spec 2002; 64: 284–9.
2. Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, Gordin F, et al. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med 2007; 175: 367–416.
3. Kendall BA, Winthrop KL. Update on the epidemiology of pulmonary nontuberculous mycobacterial infections. Semin Respir Crit Care Med 2013; 34: 87–94.
4. Khan K, Wang J, Marras TK. Nontuberculous mycobacterial sensitization in the United States: national trends over three decades. Am J Respir Crit Care Med 2007; 176: 306–13.
5. Cassidy PM, Hedberg K, Saulson A, McNelly E, Winthrop KL. Nontuberculous mycobacterial disease prevalence and risk factors: a changing epidemiology. Clin Infect Dis 2009; 49: e124–9.
6. Adjemian J, Olivier KN, Seitz AE, Holland SM, Prevots DR. Prevalence of nontuberculous mycobacterial lung disease in U.S. Medicare beneficiaries. Am J Respir Crit Care Med 2012; 185: 881–6.
7. Prevots DR, Shaw PA, Strickland D, Jackson LA, Raebel MA, Blosky MA, et al. Nontuberculous mycobacterial lung disease prevalence at four integrated health care delivery systems. Am J Respir Crit Care Med 2010; 182: 970–6.
8. Iseman MD, Marras TK. The importance of nontuberculous mycobacterial lung disease. Am J Respir Crit Care Med 2008; 178: 999–1000.
9. Wallace RJ, Jr, Brown BA, Griffith DE. Nosocomial outbreaks/pseudo-outbreaks caused by nontuberculous mycobacteria. Annu Rev Microbiol 1998; 52: 453–90.
10. Gopinath K, Singh S. Non-tuberculous mycobacteria in TB-endemic countries: are we neglecting the danger? PLoS Negl Trop Dis 2010; 4: e615.
11. Munawwar A, Singh S. AIDS associated tuberculosis: a catastrophic collision to evade the host immune system. Tuberculosis (Edinb) 2012; 92: 384–7.
12. Singh S, Gopinath K, Shahdad S, Kaur M, Singh B, Sharma P. Nontuberculous mycobacterial infections in Indian AIDS patients detected by a novel set of ESAT-6 polymerase chain reaction primers. Jpn J Infect Dis 2007; 60: 14–18.
13. van Ingen J, Boeree MJ, Dekhuijzen PN, van Soolingen D. Environmental sources of rapid growing nontuberculous mycobacteria causing disease in humans. Clin Microbiol Infect 2009; 15: 888–93.
14. Matlova L, Dvorska L, Palecek K, Maurenc L, Bartos M, Pavlik I. Impact of sawdust and wood shavings in bedding on pig tuberculous lesions in lymph nodes, and IS1245 RFLP analysis of *Mycobacterium avium* subsp. hominissuis of serotypes 6 and 8 isolated from pigs and environment. Vet Microbiol 2004; 102: 227–36.
15. Cheung PY, Kinkle BK. Mycobacterium diversity and pyrene mineralization in petroleum-contaminated soils. Appl Environ Microbiol 2001; 67: 2222–9.
16. Feazel LM, Baumgartner LK, Peterson KL, Frank DN, Harris JK, Pace NR. Opportunistic pathogens enriched in showerhead biofilms. Proc Natl Acad Sci U S A 2009; 106: 16393–9.
17. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 2012; 6: 1621–4.
18. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A 2011; 108: 4516–22.
19. Human Microbiome Project Consortium H. A framework for human microbiome research. Nature 2012; 486: 215–21.
20. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012; 486: 207–14.
21. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. Science 2009; 326: 1694–7.
22. Hosek J, Svastova P, Moravkova M, Pavlik I, Bartos M. Methods of mycobacterial DNA isolation from different biological material: a review. Vet Med 2006; 51: 180–92.
23. Vetrovsky T, Baldrian P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PLoS One 2013; 8: e57923.
24. Gonzalez JM, Portillo MC, Belda-Ferre P, Mira A. Amplification by PCR artificially reduces the proportion of the rare biosphere in microbial communities. PLoS One 2012; 7: e29973.
25. Reeder J, Knight R. The 'rare biosphere': a reality check. Nat Methods 2009; 6: 636–37.
26. Kim BJ, Lee SH, Lyu MA, Kim SJ, Bai GH, Chae GT, et al. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (rpoB). J Clin Microbiol 1999; 37: 1714–20.

27. Bahrani-Mougeot FK, Paster BJ, Coleman S, Ashar J, Barbuto S, et al. Diverse and novel oral bacterial species in blood following dental procedures. J Clin Microbiol 2008; 46: 2129–32.
28. Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, et al. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med 2013; 187: 1067–75.
29. Yamada-Noda M, Ohkusu K, Hata H, Shah MM, Nhung PH, Sun XS, et al. Mycobacterium species identification--a new approach via dnaJ gene sequencing. Syst Appl Microbiol 2007; 30: 453–62.
30. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389–402.
31. Hamady M, Lozupone C, Knight R. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. ISME J 2010; 4: 17–27.
32. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010; 7: 335–6.
33. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 2011; 27: 2957–63.
34. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010; 26: 2460–1.
35. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 2006; 72: 5069–72.
36. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 2011; 27: 2194–200.
37. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 2013; 42: D633–42.
38. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 2013; 30: 2725–9.
39. Tirkkonen T, Nieminen T, Ali-Vehmas T, Peltoniemi OA, Wellenberg GJ, Pakarinen J. Quantification of *Mycobacterium avium* subspecies in pig tissues by real-time quantitative PCR. Acta Vet Scand 2013; 55: 26.
40. Zelazny AM, Root JM, Shea YR, Colombo RE, Shamputa IC, Stock F, et al. Cohort study of molecular identification and typing of *Mycobacterium abscessus*, *Mycobacterium massiliense*, and *Mycobacterium bolletii*. J Clin Microbiol 2009; 47: 1985–95.
41. Kirakodu SS, Govindaswami M, Novak MJ, Ebersole JL, Novak KF. Optimizing qPCR for the quantification of periodontal pathogens in a complex plaque biofilm. Open Dent J 2008; 2: 49–55.
42. Liu Z, DeSantis TZ, Andersen GL, Knight R. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. Nucleic Acids Res 2008; 36: e120.
43. Liu KL, Porras-Alfaro A, Kuske CR, Eichorst SA, Xie G. Accurate, rapid taxonomic classification of fungal large-subunit rRNA genes. Appl Environ Microbiol 2012; 78: 1523–33.
44. Iwai S, Fei M, Huang D, Fong S, Subramanian A, Grieco K, et al. Oral and airway microbiota in HIV-infected pneumonia patients. J Clin Microbiol 2012; 50: 2995–3002.
45. Lemon KP, Klepac-Ceraj V, Schiffer HK, Brodie EL, Lynch SV, Kolter R. Comparative analyses of the bacterial microbiota of the human nostril and oropharynx. MBio 2010; 1: e00129–10.
46. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC, Yu WH, et al. The human oral microbiome. J Bacteriol 2010; 192: 5002–17.
47. Vujkovic-Cvijin I, Dunham RM, Iwai S, Maher MC, Albright RG, Broadhurst MJ, et al. Dysbiosis of the gut microbiota is associated with hiv disease progression and tryptophan catabolism. Sci Transl Med 2013; 5: 193ra91.
48. Falkinham JO, 3rd. Surrounded by mycobacteria: nontuberculous mycobacteria in the human environment. J Appl Microbiol 2009; 107: 356–67.
49. Colditz GA, Berkey CS, Mosteller F, Brewer TF, Wilson ME, Burdick E, et al. The efficacy of bacillus Calmette-Guerin vaccination of newborns and infants in the prevention of tuberculosis: meta-analyses of the published literature. Pediatrics 1995; 96: 29–35.
50. Beverley P, Ronan E, Lee L, Arnold I, Bolinger B, Powrie F, et al. Environmental effects on protection against Mycobacterium tuberculosis after immunization with Ad85A. Vaccine 2013; 31: 1086–93.
51. Brandt L, Feino Cunha J, Weinreich Olsen A, Chilima B, Hirsch P, Appelberg R, et al. Failure of the *Mycobacterium bovis* BCG vaccine: some species of environmental mycobacteria block multiplication of BCG and induction of protective immunity to tuberculosis. Infect Immun 2002; 70: 672–8.
52. Palmer CE, Edwards LB. Identifying the tuberculous infected. The dual-test technique. JAMA 1968; 205: 167–9.
53. Fine PE. Variation in protection by BCG: implications of and for heterologous immunity. Lancet 1995; 346: 1339–45.
54. Hernandez-Pando R, Pavon L, Arriaga K, Orozco H, Madrid-Marina V, Rook G. Pathogenesis of tuberculosis in mice exposed to low and high doses of an environmental mycobacterial saprophyte before infection. Infect Immun 1997; 65: 3317–27.
55. Fine PE, Vynnycky E. The effect of heterologous immunity upon the apparent efficacy of (e.g. BCG) vaccines. Vaccine 1998; 16: 1923–8.
56. Geisel RE, Sakamoto K, Russell DG, Rhoades ER. *In vivo* activity of released cell wall lipids of *Mycobacterium bovis* bacillus Calmette-Guerin is due principally to trehalose mycolates. J Immunol 2005; 174: 5007–15.
57. Lima VM, Bonato VL, Lima KM, Dos Santos SA, Dos Santos RR, Goncalves ED, et al. Role of trehalose dimycolate in recruitment of cells and modulation of production of cytokines and NO in tuberculosis. Infect Immun 2001; 69: 5305–12.
58. Medzhitov R. Toll-like receptors and innate immunity. Nat Rev Immunol 2001; 1: 135–45.