Original Paper

Microbiome Diversity in Sputum of Nontuberculous Mycobacteria Infected Women with a History of Breast Cancer

Julie V. Philley[a]   Anbarasu Kannan[b]   Patti Olusola[c]   Paul McGaha[d]
Karan P. Singh[e]   Buka Samten[f]   David E. Griffith[a]   Santanu Dasgupta[b]

[a]Department of Medicine, The University of Texas Health Science Center at Tyler, Tyler, TX, USA, [b]Department of Cellular and Molecular Biology, The University of Texas Health Science Center at Tyler, Tyler, TX, USA, [c]Department of Family Medicine, The University of Texas Health Science Center at Tyler, Tyler, TX, USA, [d]Department of Community Health and Preventative Medicine, The University of Texas Health Science Center at Tyler, Tyler, TX, USA, [e]Department of Epidemiology and Biostatistics, The University of Texas Health Science Center at Tyler, Tyler, TX, USA, [f]Department of Pulmonary Immunology, The University of Texas Health Science Center at Tyler, Tyler, TX, USA

Abstract

Background/Aims: The nontuberculous mycobacterial lung disease (NTM), caused by *Mycobacterium avium* complex (MAC) is an increasing health problem in the USA and worldwide. The NTM disease is prevalent in Caucasian women with a current diagnosis or history of breast cancer (BCa), posing a significant challenge towards treatment. We hypothesize that NTM affected women with considerable therapeutic resistance may harbor pathogenic microbes other than nontuberculous mycobacterium, aiding in disease progression and therapeutic resistance. ***Methods:*** We assessed microbiome diversity in sputa from healthy women, women with nontuberculous mycobacterial lung disease (NTM) and women with both nontuberculous mycobacterial lung disease and breast cancer (NTM-BCa). First, we collected sputa and isolated DNA from sputa of these healthy women and women with NTM and NTM-BCa. We also isolated DNA from sera derived extracellular vesicles from women with NTM-BCa. To identify diverse pathogenic microbes in various groups of subjects, we then performed 16S rDNA sequencing. Data analysis was performed utilizing the analytical pipelines at the Center for Metagenomic and Microbiome Research (CMMR), Baylor College of Medicine. ***Results:*** A large community of resident microbes, including bacteria, virus, Archeas and Fungi live in the human body are being increasingly recognized as the key components of human health and disease. We identified a diverse microbiome community in the sputa and the extracellular vesicles dominated by *Streptococcus*, *Haemophillus*, *Veillonella*, *Neisseria*, *Prevotella*, *Fusobacterium*, *Bacteroides*, *Allistipes*, *Faecalibacterium* and *Staphylococcus* in women with nontuberculous

Santanu Dasgupta   Department of Cellular and Molecular Biology, The University of Texas Health Science Center at Tyler
Tyler, Texas (USA)
Tel. 903-877-7007, Fax 903-877-7558; E-Mail santanu.dasgupta@uthct.edu

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

264

Philley et al.: Microbiome of Sputa

mycobacterial lung disease as well as women with both nontuberculous mycobacterial lung disease and breast cancer. Some of these genera, including *Fusobacterium*, *Bacteroides*, *and Allistipes* have estrobolome activity and associated with breast and other neoplasms. **Conclusion:** This work confirms the presence of a distinct pathogenic microbiome other than nontuberculous mycobacteria in the sputa and the circulating extracellular vesicles of these patients. This information could be useful for better therapeutic design to treat the NTM patients.

Introduction

The incidence of nontuberculous mycobacterial (NTM) lung disease has increased considerably in the United States during the past decades [1]. Being natural inhabitants under different environmental conditions, NTM can infect diverse populations. This could lead to mycobacteriosis, chronic pulmonary infection, including non-cystic fibrosis bronchiectasis and chronic obstructive pulmonary diseases in both immunocompetent and immunocompromised individuals such as cancer patients [1-2]. The majority of the NTM infection involved *mycobacterium avium* complex (MAC) [2-3]. Notably, there is an increasing rate of NTM lung disease in Caucasian women with a current diagnosis or history of breast cancer [2].

Diverse populations of microorganisms, including bacteria, fungi and viruses living in a human body are collectively known as the "microbiota" [4-5]. Being an integral part of the human body and various organ systems, they are involved in a variety of cellular processes, including metabolism and influence development of different diseases. Emerging studies indicate that residence of diverse populations of microbiota in various cellular compartments influence immune system function, metabolism leading to chronic inflammation and cellular transformation [4-9]. Many of these studies focusing on gut microbiota identified that infection with a gram negative bacterium *Helicobacter pylori*, residing in the gastric epithelium induces gastric inflammation and correlates with increased rate of gastric cancer [10]. In this light, it is likely that other than MAC, population and functional disequilibrium in other microbes residing in the lung could potentially create a microenvironment to increase susceptibility and therapeutic resistance towards NTM lung disease. Remarkably, certain opportunistic microbes have the ability to metabolize Estrogen or Estrogen like compounds (Estrobolome), thereby enhancing their activities, which could potentially promote tumorigenic transformation [11-12]. Estrogen is an abundantly expressed protein and an attractive therapeutic target for BCa and Estrogen/Estrogen like compounds when modified through the estrobolomes, could potentially contribute to BCa tumorigenesis [11-12].

Extracellular vesicles (EVs) are small (50-200 nm), secreted vesicles present in all cell types, circulation and body fluids [13-18]. The EVs contain specific nucleic acids, proteins, metabolites and have the ability to transport these cellular factors to the neighboring cells [13-19]. As evident in many studies, including a study from our laboratory, the EVs could influence immune system function, host-pathogen interaction during microbial infection and carcinogenesis [3, 13-19].

In this study, our goal was to identify the population of pathogenic microbes residing with the MAC species in the NTM infected women. We examined the microbiota of sputa and sera derived EVs in women with NTM, NTM-BCa (women with both nontuberculous Mycobacterial lung disease and breast cancer) along with the healthy controls (HC). We identified variable abundance of different genera such as *Streptococcus*, *Haemophillus*, *Veillonella*, *Neisseria*, *Rothia*, *Prevotella, Fusobacterium*, and *Leptotrichia* among various groups. Compared to the sputa, the circulating EVs from some NTM-BCa subjects harbor unique microbiota including *Bacteroides*, *Escherichia/Shegella, Streptococcus*, *Allistipes*, *Cloacbacterium*, *Faecalibacterium*, *Lachnoclostridium, Staphylococcus, Nisseria*,

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

265

Philley et al.: Microbiome of Sputa

Fusicatenibacter, *Fusobacterium*, *Lachnospira* and *Tepidimonus*. Many of the identified genera possess estrobolome activity and are associated with the development of various lung and oral cavity diseases including malignancies.

Materials and Methods

Human samples and ethical statement

Sputa were collected from 15 women with both NTM and breast cancer (NTM-BCa), 5 NTM and 5 Healthy controls (HC) at The University of Texas Health Science Center at Tyler under an IRB approved protocol (#974). Informed consent was obtained from all the patients. Sera EVs from 4/15 NTM-BCa subjects were collected. All patients were de-identified and only relevant clinical information such as age, grade, stage, diagnosis, etc. was collected (Table 1 and Table 2). All methods were performed in accordance with the relevant guidelines and regulations.

Sputum collection

Routine expectorated sputa were collected and cultured as necessary for acid-fast bacilli (AFB) [20-21]. For patients unable to produce sputum by spontaneous expectoration, sputum induction was performed with nebulized hypertonic saline. Samples were processed using standard decontamination procedures, fluorochrome microscopy and cultured on solid media on a plate of Middlebrook 7H10 agar with and without antibiotics and a broth culture (ESP, Thermofisher, formerly Trek Diagnostic Systems, and Cleveland, OH) as described [20-23]. MAC isolates were identified using AccuProbe (Hologic Gen-Probe Inc.) [21]. For decontamination, N-acetyl-L-cysteine-sodium hydroxide (NALC-NaOH) method was used [24]. Alternatively, a combination of N-acetyl-L-cysteine-sodium hydroxide (NALC-NaOH) and Oxalic acid was also used as necessary for decontamination [24]. All methods were performed in accordance with the relevant guidelines and regulations. As the environmental control, we utilized a 15 ml falcon tube containing PBS and kept it open for the time period when the sputa were processed. We were unable to isolate bacteria from this culture and as a reason, the sample was not processed further.

Isolation of extracellular vesicles

The EVs were isolated from human sera and their purity was confirmed by Nanosight Tracking and Western blotting with EV markers Syntenin and CD63 [2]. DNA was isolated from the EVs using a standardized protocol [25].

16S rRNA gene sequencing for microbiome detection

For the microbiome sequencing, sputa from 15 NTM-BCa subjects, 5 NTM, 5 healthy controls and exosomes from 4 NTM-BCa subjects (Table 1) were examined (Total 29 samples). The 16S rRNA gene sequencing methods were adapted from the NIH-Human Microbiome

Table 1. Demographic information of the NTM-BCa subjects. [1]NTM-BCa: Women with both NTM lung disease and breast cancer. [2]Birth control measures taken by the subjects. N: No; Y: Yes; UK: Unknown; YR: Year. [3]Hormore therapy received after the menopause. N: No; Y: Yes; UK: Unknown. [4]Types of hormone therapy received. HMP: Homeopathy; YR: Year; M: Month; UK: Unknown

[1]NTM-BCa cases	Age at NTM Diagnosis	Age at BCa Diagnosis	[2]Hormonal Contraceptive/Starting Age	Types/Duration	Age of Menopause	[3]Hormone for Menopause	[4]Kind of Hormones/Age started/Duration
1	75	38	N/UK	-	45	N	-
2	60	65	Y/21YR	Birth control/5YR	45	Y	Premarin/45YR/19YR
3	66	45	Y/32YR	Implant/4YR	45	Y	Premarin/HMP/45YR/1YR
4	57	61	N/UK	-	50	Y	Estrogen pills/45YR/10YR
5	71	75	N/UK	-	48	Y	Ogen/45YR/27YR
6	73	68	Y/19YR	Pills/4YR	52	Y	Estrogen/52YR/10YR
7	54	52	Y/UK	UK/UK	-	N	-
8	63	58	Y/32YR	Pills/UK	51	N	-
9	61	45	Y/20YR	Birth control/2 Weeks	-	Y	Premarin/50YR/current
10	55	51	Y/UK	Birth control/2YR	50	N	-
11	68	68	Y/23YR	Pills/10YR	55	Y	Estrogen/UK/9M
12	55	61	Y/21YR	Pills/5YR	42	N	-
13	66	56	N/UK	-	-	Y	Premarin/45YR/12YR
14	55	48	Y/22YR	Pills/1.5YR	48	Y	UK/46Y/6M
15	82	76	N/UK	-	UK	UK	-

266

Philley et al.: Microbiome of Sputa

Project [26-27] and followed appropriate guidelines for sample processing and data analysis [28]. Briefly, bacterial genomic DNA was extracted using MO BIO PowerSoil DNA Isolation Kit (MO BIO Laboratories, (MoBIO PowerSoil v3.4). The 16S rDNA V4 region was amplified (Illumina 16Sv4 v1.2) by PCR using specific primer sets (Table 3) and sequenced on the MiSeq platform (Illumina MiSeq v2 2x250 v1.8) using the 2x250 bp paired-end protocol yielding pair-end reads that overlap almost completely. The primers used for amplification contained adapters for MiSeq sequencing and single-end barcodes (Table 3) allowing pooling and direct sequencing of PCR products [29].

Quality control and compositional analysis

The 16S rRNA gene pipeline data incorporates phylogenetic and alignment based approaches to maximize data resolution (Analytical Pipeline CMMR 16Sv4 v1.0) [30]. The read pairs were de-multiplexed based on the unique molecular barcodes and reads were merged using USEARCH v7.0.1090, which has a *de novo* built-in chimera filter and all singleton OTUs were discarded [31-34]. We have also applied contig filtering to remove as much sequencing error utilizing the count files as necessary [31-34]. For the sequencing and data analysis, we used the analytical packages and

Table 2. Demographic information of the Healthy control and NTM subjects. [1] Healthy women (HC) and women with nontuberculous Mycobacterial (NTM) lung disease. [2]Birth control measures taken by the subjects. N: No; Y: Yes; UK: Unknown; YR: Year. [3]Hormore therapy received after the menopause. N: No; Y: Yes; UK: Unknown. [4]Types of hormone therapy received. YR: Year; M: Month UK: Unknown

[1]Patient ID	Age at NTM Diagnosis	[2]Hormonal Contraceptives/Starting Age	Types/ Duration	Age of Menopause	[3]Hormone for Menopause	[4]Kind of Hormones/Age started/Duration
HC1	-	N	-	UK	Y	Lotion/UK/1YR
HC2	-	Y/20	UK/10YR	42YR	Y	Pemerin/58/1M
HC3	-	Y/Early 20	UK/UK	UK	UK	UK
HC4	-	UK	UK	UK	UK	UK
HC5	-	UK	UK	UK	UK	UK
NTM1	49	Y/17YR	R	55YR	N	-
NTM2	63	Y/22YR	UK/UK	35YR	Y	UK/UK/2YR
NTM3	58	Y/17YR	UK/6M	UK	N	-
NTM4	69	Y/24YR	Pills/31YR	45YR	Y	UK/46YR/10YR
NTM5	62	Y/25YR	Pills/5YR	45YR	Y	UK/46YR/5YR

Table 3. Sequences of the primers used for the microbiome analysis in sputa and extracellular vesicles (E) from healthy control (HC), NTM (BC) and NTM-BCa (BR) subjects

# Sample ID	Barcode sequence	Linker primer sequence	Reverse primer
BR07	GTATGACTGGCT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR06	CGCAGACAGACT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
HC04	GCCTATACTACA	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR12E	GGCTCTGCTCTT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BC03	GAGAAAGACGGT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR03	TAAGCCTCCAAG	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR03E	TCAACTAATGCC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR11	GATCAACTACCG	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR13E	ATGGGCTCTGAC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
HC01	CGCAATCGGTAG	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
HC05	ATATGCCAGTGC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR08	TGTCCAGGTTTA	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
HC02	AGCAACGGACTC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR12	ATGGGAATGGGA	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR13	GTCCGCTGAATA	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BC02	TACGCGCTGAGA	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR09	GCAGCCGAGTAT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR04	CTTCTAAGGAGC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR01	AGGGTTCCAGTT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BC01	ACTGATCCTAGT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR14	ACTTGAGTGCGG	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
HC03	TATGACGGTTCT	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BC05	CGTACAGTTATC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR10	ATTTCACGGGAC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BC04	TACCTTGGGTGA	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR15	TAGCCGGCATAG	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR05	AGGTCTAGGGTC	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR02	GCATTGCGTGAG	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA
BR11E	TTCTCCGTGTCG	GGACTACHVGGGTWTCTAAT	GTGCCAGCMGCCGCGGTAA

pipelines at the Center for Metagenomic and Microbiome Research (CMMR), Baylor College of Medicine as utilized by other groups [30]. We obtained summary statistics and quality control measurements for each sequencing run, as well as multi-run reports and data-merging capabilities for validating built-in controls and characterizing microbial communities. The 16S rRNA gene sequences were clustered into Operational Taxonomic Units (OTUs) [30]. OTUs are collapsed based into phylotypes defined by taxonomic similarity. Operational Taxonomic Units (OTUs) identified were mapped to the SILVA database at 97% identity to obtain phylum to genus taxonomies. The OTUs were mapped to an optimized version of the SILVA Database [35-36], containing only the 16S v4 region to determine taxonomies. Unclassified OTUs at any level are kept in the analysis and are represented by the next lower level of taxonomic resolution. Abundances were recovered by mapping the de-multiplexed reads to the UPARSE OTUs. A custom script constructed an OTU table from the output files generated in the previous two steps for downstream analyses of alpha-diversity, beta-diversity and phylogenetic trends [37].

Data analysis

The Agile Toolkit for Incisive Microbial Analyses (AITMA) from CMMR was used for analyzing and visualizing microbiome data sets. We examined the microbiome populations in alpha diversity (AD), beta diversity (BD), and taxa abundances (TA) through customizable box/hierarchical plots, Principal Coordinate Analysis (PCoA) ordinations, and heatmaps. Appropriate statistical annotations were added to infer significant correlations with the metadata. We performed Mann-Whitney and Kruskal-Wallis tests for all the statistical data analysis with appropriate FDR application utilizing the Bonferroni correction method (CMMR 16Sv4 v1.0). The Heatmaps and the hierarchical clustering were performed utilizing Weighted Unifrac/PCoA methods based on taxa abundance. Rarefaction analysis was employed to normalize all samples to the same total amount of reads. Samples with total read counts below the designated threshold were excluded. An optimal rarefaction cutoff is represented by a complete or almost complete saturation of OTU richness demonstrated by a clear asymptote of the curve. Alpha diversity was used to measure the diversity within a sample. The different alpha diversity metrics use the counts (richness) and distribution (evenness) of Operational Taxonomic Units (OTUs) within a sample as the basic values for these calculations. Beta diversity ordinations were used to compare the samples to each other and to measure the distance or dissimilarity between samples in the datasets. UniFrac is a phylogenetic distance metric used for comparing microbial communities, which incorporates phylogenetic distances between observed organisms in the computation. Both weighted (quantitative, structure) and unweighted (qualitative, composition) variants of UniFrac were used in this study as necessary. Weighted UniFrac accounts for the abundance of observed organisms, while unweighted UniFrac only considers their presence or absence. Using this phylogenetic metric, data was plotted using the PCoA ordination method.
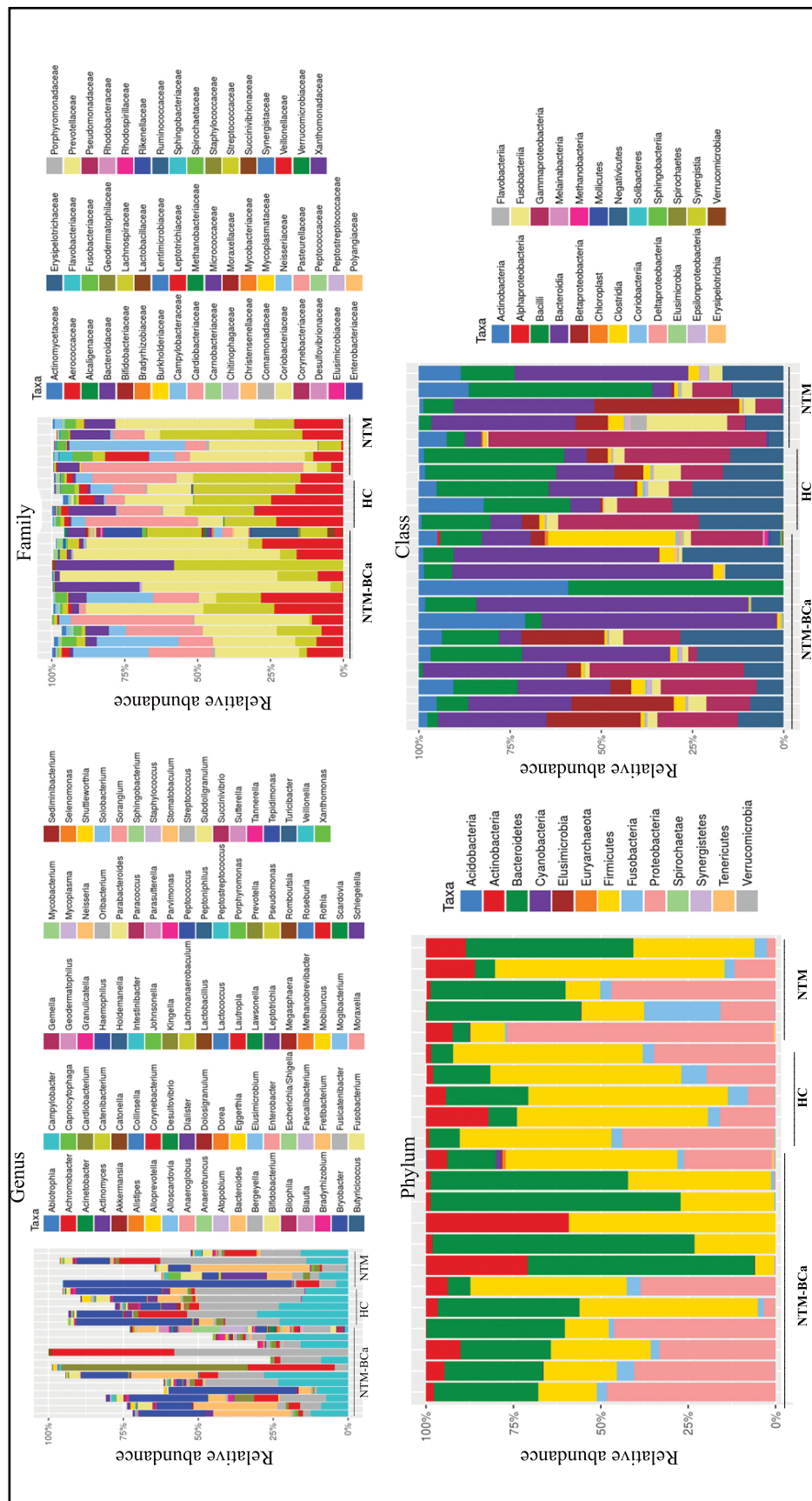
Results

Pre and postmenopausal hormone intake survey of the study population

Based on the available clinical history, 67% (10/15) of the NTM-BCa subjects had taken hormonal contraceptives (Table 1). Also, 64% (9/14, no information available for 1 case) of these patients had undergone postmenopausal hormone therapy (Table 1). On the other hand, 67% (2/3, no information available for 2 cases) of the HC and 100% (5/5) of the NTM
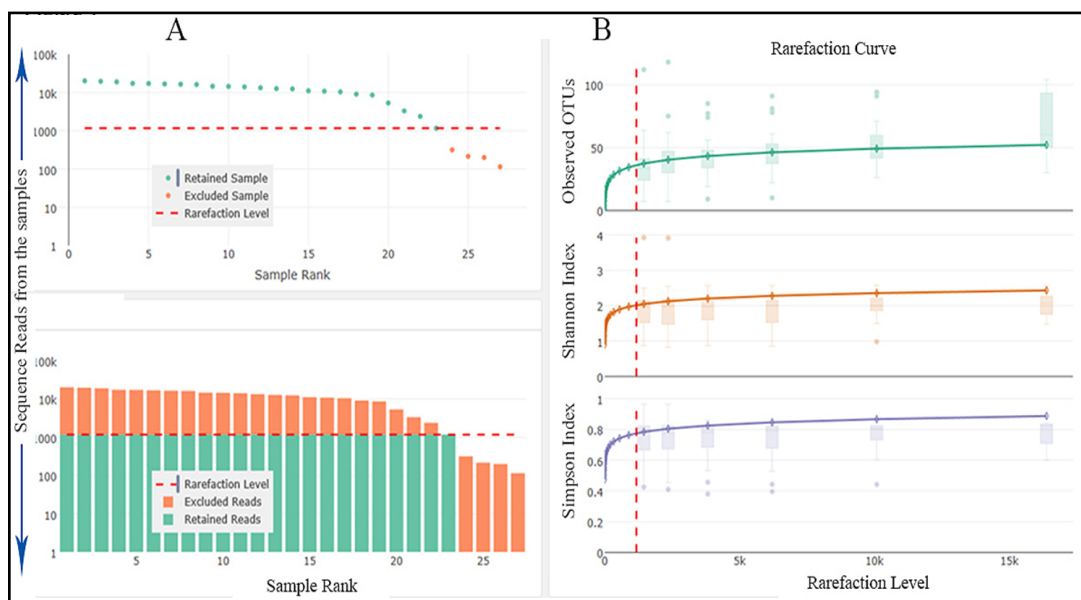
Fig. 1. Rarefaction analysis of the samples. (A) Validated sequence reads from all the samples. (B) Rarefaction curves of the various samples indicating sample richness and evenness.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

268

Philley et al.: Microbiome of Sputa

Fig. 2. Overall microbiome of sputum. The relative abundance of bacterial taxa identified in different sputum samples was represented as stacked bar plots. Each bar represents an individual and each colored box, a bacterial taxon, family and phylum. The height of a colored box represents the relative abundance at various taxonomic levels as indicated. HC: Healthy controls; NTM: Women with NTM lung disease only; NTM-BCa: Women with both NTM lung disease and breast cancer. A minimum average abundance of 0.05% was considered for calculating the overall taxa abundance.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

269

Philley et al.: Microbiome of Sputa

Fig. 3. Taxa distribution in the sputa of various groups. (A) Distribution pattern of the Operational Taxonomic Units (OTUs) across various groups. Compared to the HC group, the OTUs, Unc02v3h and UncB2490 were significantly lower in the NTM-BCa group (p=0.007-0.017). No significant difference was observed in the OTU distribution between HC vs. NTM (p=0.06-0.08) or NTM vs. NTM-BCa groups (p=0.9-0.1). (B) Significantly lower abundance of Leptotrichia (0.4% vs. 2.8%, p=0.006) and Streptococcus (13% vs. 28%, p=0.01) in the BCa-NTM group compared to the HC subjects. Abundance of genus Veilonella was significantly lower in the NTM group compared to the HC group (9% vs. 21%, p=0.03). No considerable difference in genera distribution between the NTM and NTM-BCa groups (p=0.16-0.95). A minimum average abundance of 1% was considered. (C) At the phyla level, the majority of the taxa in the HC group belonged to Firmicutes (52%) Proteobacteria (24%) and Bacteriods (13%). Compared to the HC, the top phyla in the NTM group were Proteobacteria (28%), Firmicutes (27%) and Bacteriods (13%). In the NTM-BCa group the top phyla were Bacteriods (36%), Firmicutes (33%) and Proteobacteria (19%). The phyla Firmicutes and Fusobacteria were dominant (p=0.04) in the HC compared to the NTM-BCa group. No considerable difference in the distribution of phyla was observed between the HC vs. NTM or the NTM vs. NTM-BCa groups (=0.29-0.80). HC: Healthy controls; NTM: Women with NTM lung disease only; NTM-BCa: Women with both NTM lung disease and breast cancer.
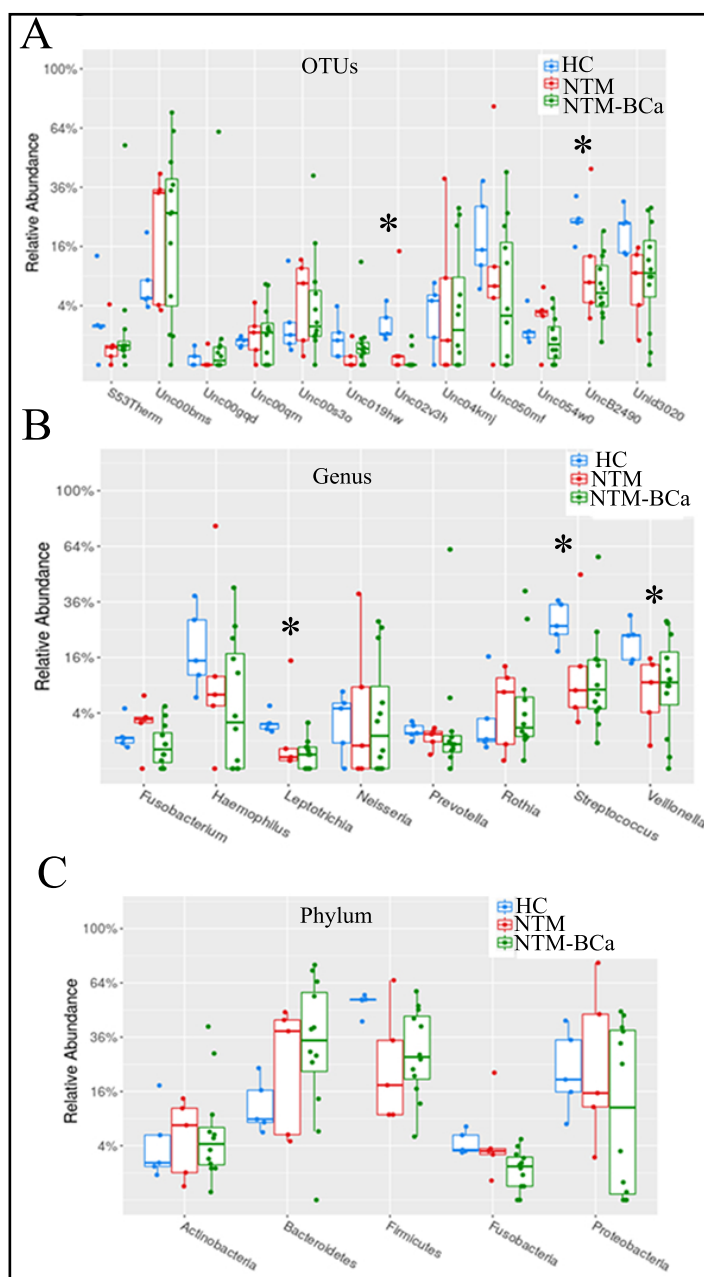

subjects had taken hormonal contraceptives. Hundred percent (2/2, no information available for 3 cases) of the HC and 60% (3/5) of the NTM patients had undergone postmenopausal hormone therapy (Table 2).

Overall sequence reads and data filtration

In this study, sputum samples from 15 NTM-BCa, 5 NTM and 5 HC women along with sera EVs from 4/15 NTM-BCa cases (Total 29 samples) were subjected to 16S rDNA V4 region sequencing. The sequence reads were first clustered into unique OTUs (Operational Taxonomic Units). In the next step, the OTUs were classified to various taxonomic levels,

Philley et al.: Microbiome of Sputa
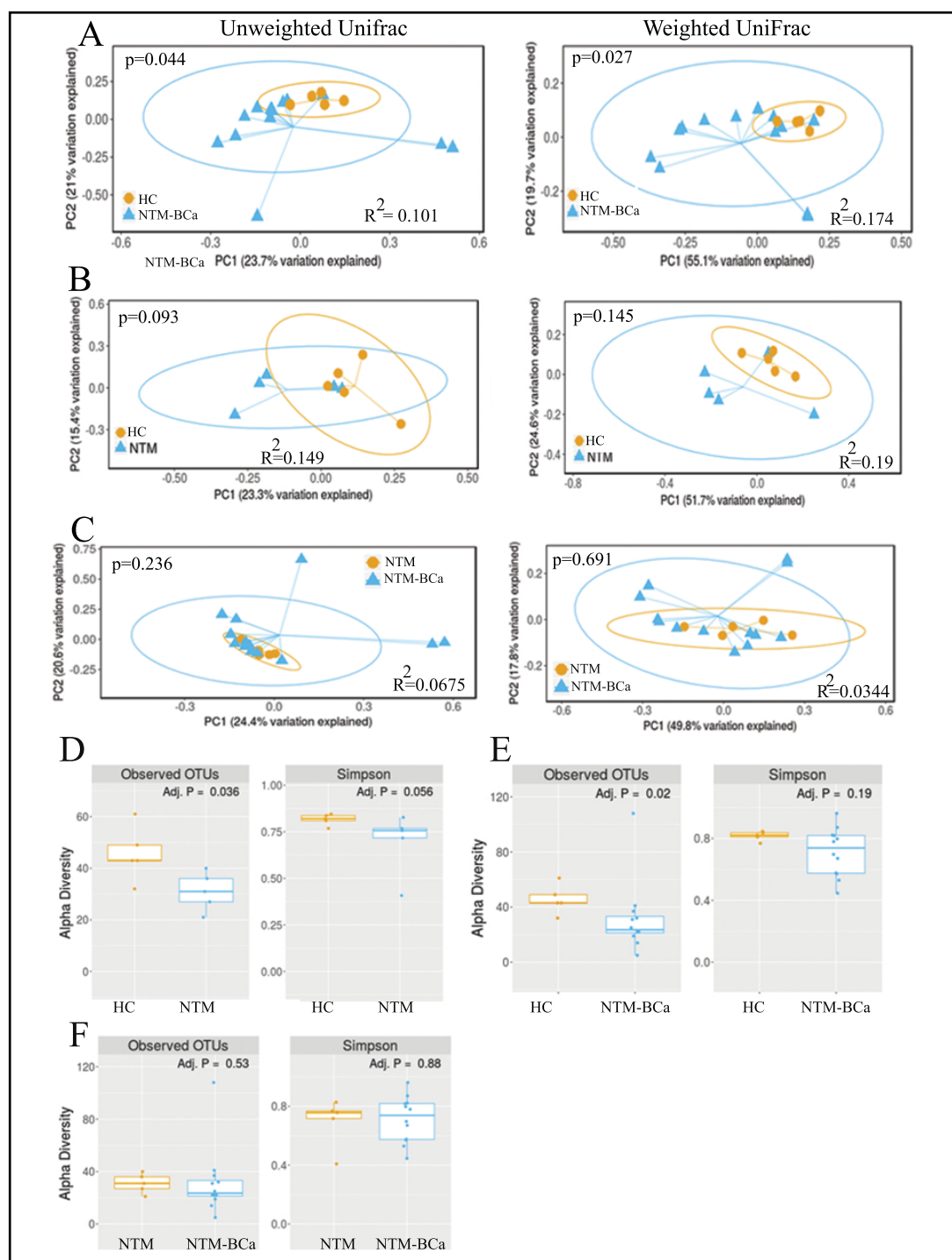
Fig. 4. Beta and alpha diversity ordination plots. (A-C) Measurement and comparison of beta diversity in various groups using both unweighted and weighted UniFrac (phylogenetic) distance matrices. Data were plotted using PCoA ordination method utilizing the OTUs as the basic units. The percentage of variability of each principal coordinates (PC) has been explained on the axis of each plot. Each colored circle or triangle represents a sample from a group, as indicated. The probability distribution has been represented by the ellipses. (D-F) Measurement of the intra-sample variation by alpha diversity analysis. The alpha diversity metrics use the counts (richness) and distribution (evenness, Simpson Index) of the OTUs within a sample as the basic values for these calculations. HC: Healthy controls; NTM: Women with NTM lung disease only; NTM-BCa: Women with both NTM lung disease and breast cancer.

such as genus, phylum, class, order, kingdom, etc. Three sputum samples from the NTM-BCa group were excluded due to a low number of sequence reads after the rarefaction analysis (Fig. 1). A total of 288385 reads after data filtration was obtained from all the samples. We have detected a total of 220 OTUs across all the samples after the rarefaction analysis. The average number of observed OTUs in all samples was 32.

Nature and distribution pattern of microbiota in sputum

While comparing the bacterial taxa in the sputum samples of the various groups, a total of 93 genera belonging to 13 phyla, 52 families, and 24 classes was detected (Fig. 2). Considering a minimum average of 1% abundance, the most frequently occurring OTUs observed across all the samples were *unc00bms* (22%), *unc050mf* (15%), *unc04kmj* (7%), *unc00s30* (6%), *S53therm* (4%), *unc054w0* (2%), *unc00qrm* (2%), *unc02v3h* (1%) and *unc019hw* (1%) (Fig. 3A). The major genera in the sputum samples from various groups were *Streptococcus* (19%), *Haemophillus* (17%), *Veillonella* (14%), *Neisseria* (7%), *Rothia* (6%), *Prevotella (3%) Fusobacterium* (2%), and *Leptotrichia* (2%) (Fig. 3B). The majority of the microbiota population observed in the sputa across all the groups belonged to the phyla *Firmicutes* (38%), *Bacteriodetes* (26%), *Proteobacteria* (25%), *Actinobacteria* (4%) and *Fusobacteria* (4%) (Fig. 3C).

Microbiome in sputum of healthy women and women with NTM and NTM-BCa

To determine the distribution pattern and richness of the microbiota in various groups, we performed both beta and alpha diversity analyses. Beta diversity analysis using both unweighted and weighted UniFrac PCoA (Principal Coordinate Analysis) methods demonstrated a significant difference (p=0.044-0.027) in taxa distribution among the NTM-BCa subjects compared to the HC (Fig. 4A). Distribution of taxa among the healthy and the NTM group varied considerably, however, it did not reach a significant level (Fig. 4B, p=0.09-0.14). No appreciable difference in the taxa distribution was observed between the NTM and NTM-BCa group by the beta diversity analysis (Fig. 4C, p=0.23-0.69). On the other hand, alpha diversity analysis revealed that species richness as well the evenness between the NTM and
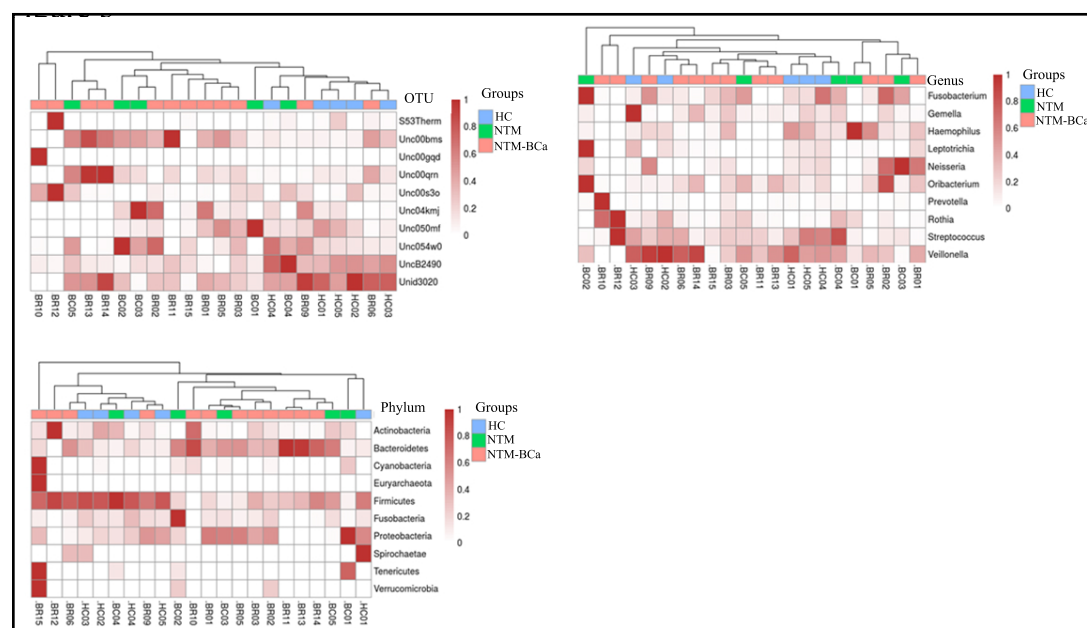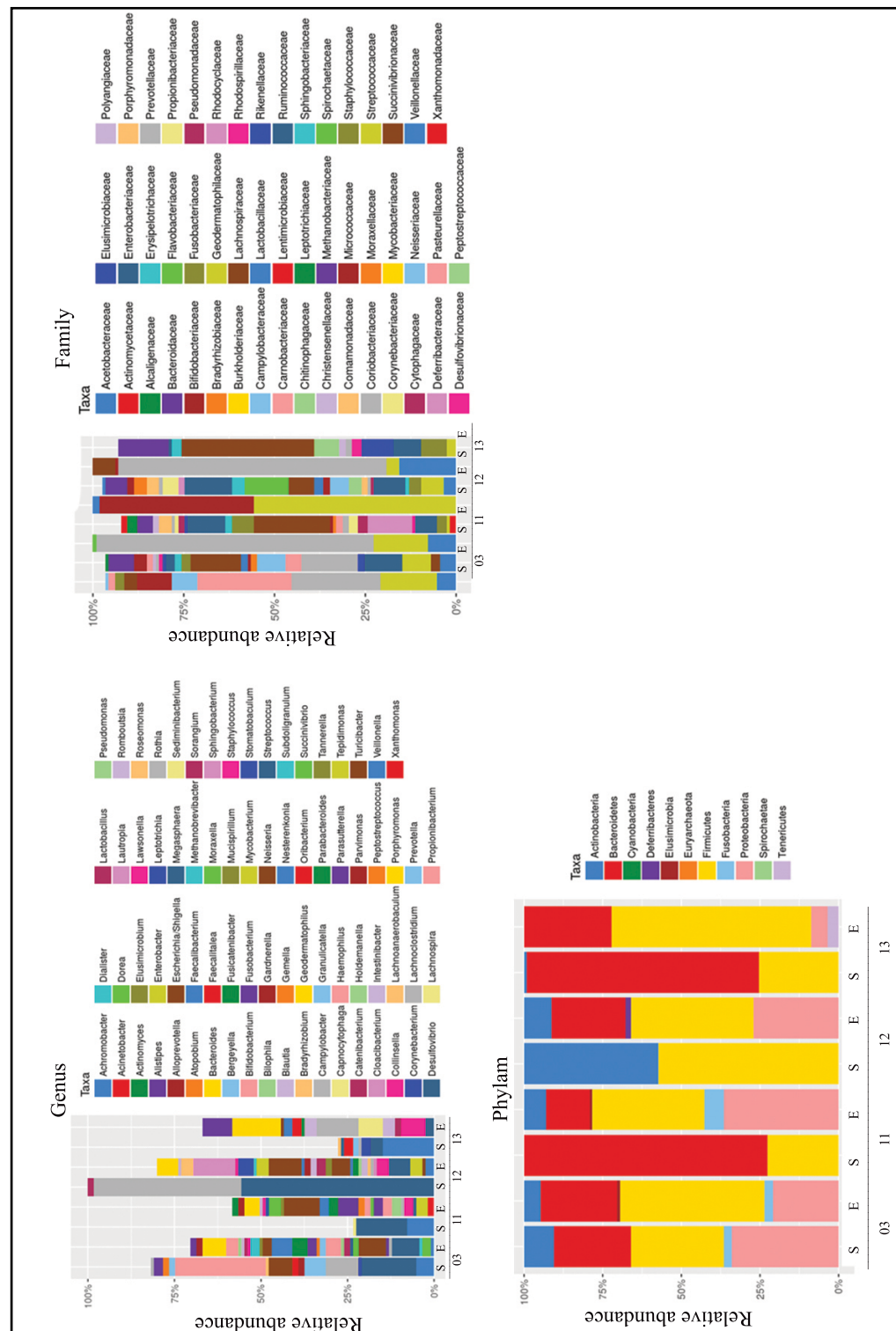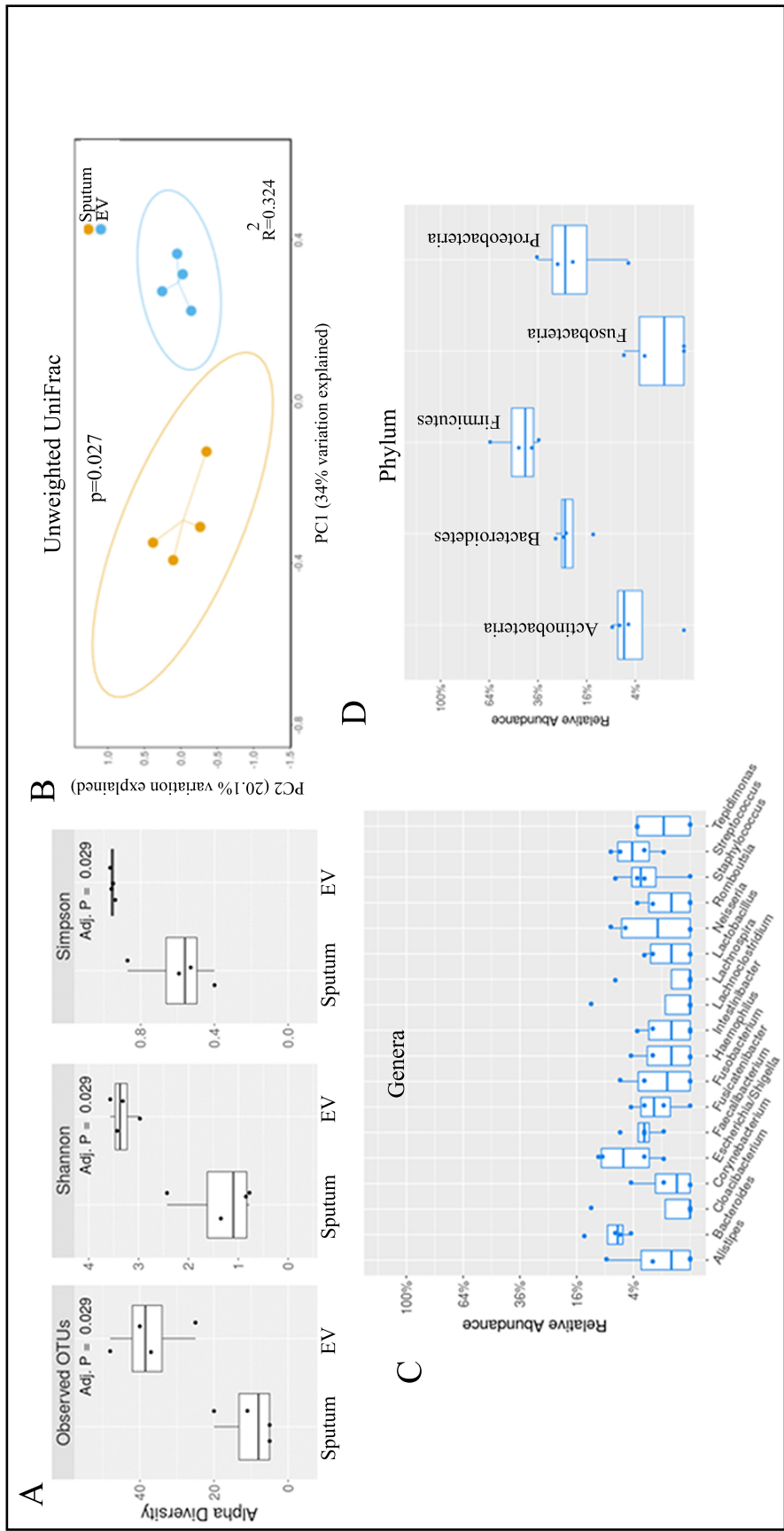

Fig. 5. Heatmaps and clustering of individual sputum microbiota at OTU, genus and phylum level. The heatmaps were generated based on the abundance of the top 10 Taxa. A minimum average abundance of 1% was considered to determine relative abundance at various taxonomic levels. HC: Healthy controls; NTM: Women with NTM lung disease only; NTM-BCa: Women with both NTM lung disease and breast cancer.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

272

Philley et al.: Microbiome of Sputa

HC group was significantly different (p=0.03-0.05; Fig. 4D). Similarly, species richness was significantly higher (p=0.02, Fig. 3E) in the NTM-BCa group compared to the HC but not for the sample evenness (p=0.19, Fig. 3E). However, we did not observe any considerable difference in alpha diversity analysis, among the NTM and NTM-BCa groups (p=0.53-0.88, Fig. 4F). The richness of various bacterial species among different groups has been illustrated in the Heatmaps at OTU, genus and phylum level (Fig. 5). Regarding the relative abundance of the OTUs, *Unc02v3h* and *UncB2490* were significantly lower in the NTM-BCa group compared to the HC subjects (Fig. 3A, p=0.007-0.017). No significant difference was observed in the OTU distribution between HC *vs.* NTM (Fig. 3A, p=0.06-0.08) or NTM *vs.* NTM-BCa groups (Fig. 3A, p=0.09-0.1). At the generic level, the dominating genera in the HC subjects were *Streptococcus* (28%), *Veilonella* (21%), *Haemophilus* (20%), *Rothia* (5%), *Nisseria* (4%), *Leptotrichia* (3%), *Prevotella* (2%) and *Fusobacterium* (2%) (Fig. 2B). In the NTM group, *Haemophilus* (20%), *Streptococcus* (16%), *Neisseria* (10%), *Veilonella* (9%), *Rothia* (7%), *Fusobacterium* (3%), *Leptotrichia* (3%) and *Prevotella* (1%) were the predominant taxa (Fig. 3B). In the NTM-BCa group, the most abundant genera were *Streptococcus* (13%), *Veilonella* (12%), *Haemophilus* (10%), *Rothia* (8%), *Neisseria* (7%), *Prevotella* (6%) and *Fusobacterium* (3%) (Fig. 3B). When we compared the generic distribution among the various groups, a significantly lower frequency of *Leptotrichia* (0.4% vs. 2.8%, p=0.006, Mann-Whitney test) and *Streptococcus* (13% vs. 28%, p=0.01, Mann-Whitney test) were evident in the BCa-NTM group compared to the HC subjects (Fig. 3B). While comparing the genera distribution between the HC and the NTM group, we observed a significant decline in *Veilonella* in the NTM group (9% vs. 21%, p=0.03, Mann-Whitney test) (Fig. 3B). No considerable difference in genera distribution was observed between the NTM and NTM-BCa groups (p=0.16-0.95) (Fig. 3B). Other than the top 8 genera described above across all the samples, the rest were present in less than 1% abundance. At the phyla level, the majority of the taxa in the HC group belonged to *Firmicutes* (52%) *Proteobacteria* (24%) and *Bacteriods* (13%) (Fig. 3C). Compared to the HC, the top phyla in the NTM group were *Proteobacteria* (28%), *Firmicutes* (27%) and *Bacteriods* (13%) (Fig. 3C). In the NTM-BCa group the top phyla were *Bacteriods* (36%), *Firmicutes* (33%) and *Proteobacteria* (19%) (Fig. 3C). While comparing the phyla distribution, *Firmicutes* and *Fusobacteria* appear to be dominant (p=0.04, Mann-Whitney test) in the HC subjects compared to the NTM-BCa group (Fig. 3C). No considerable difference in the distribution of phyla was observed between the HC *vs.* NTM or the NTM *vs.* NTM-BCa groups (=0.29-0.80) (Fig. 3C).
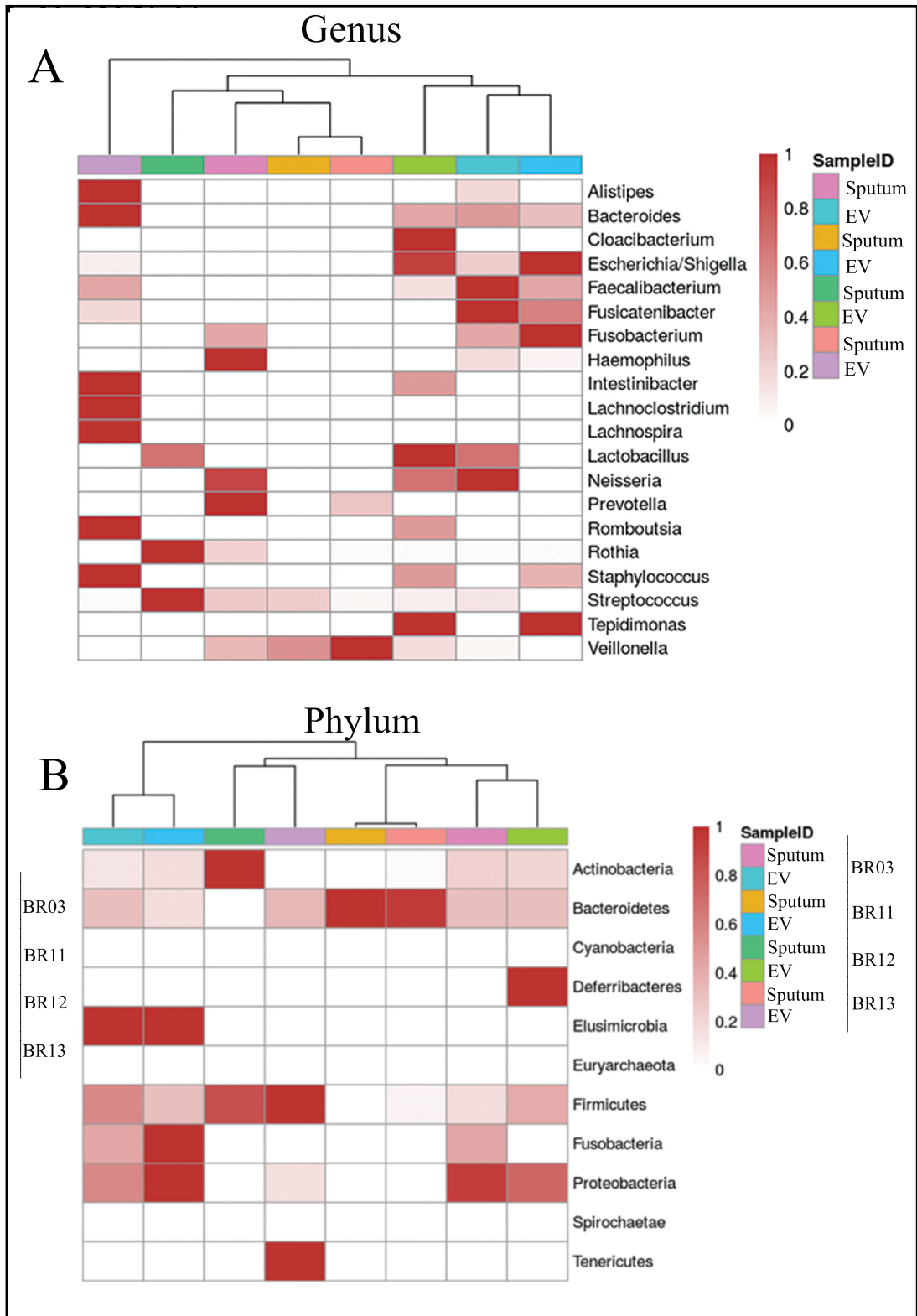
Microbiome niche in matched sputum and circulating EVs of NTM-BCa patients

Recent studies detected genomic DNA in the EVs [25, 38]. We analyzed circulating EVs from 4 available NTM-BCa subjects for microbial DNA detection and determined the relative abundance of various taxa in matched sputum and the EV samples. The pairwise analysis of the microbiota of the matched sputa and circulating EVs revealed the presence of a total of 74 genera from 50 families and 11 phyla in these samples considering average 0.05% abundance (Fig. 6). Both alpha and beta diversity analyses demonstrated significant differences in taxa abundance and richness between EVs and the matched sputa (Fig. 7A-B). Considering a minimum 1% average abundance, 18 genera were identified in the EVs, which include *Bacteroides* (7%), *Escherichia/Shegella* (6%), *Streptococcus* (4%), *Allistipes* (3%), *Cloacbacterium* (3%), *Faecalibacterium* (3%), *Lachnoclostridium* (3%)*, Staphylococcus* (3%), *Nisseria* (3%), *Fusicatenibacter* (2%), *Fusobacterium* (2%), *Lachnospira* (2%) and *Tepidimonus* (2%) (Fig. 7C). Other genera, including *Corynebacterium, Haemophilus, Intestinibacter, Lactobacillus* and *Romboutsia* were detected with a range of 1-1.5% abundance. The top phyla in the EVs were *Firmicutes* (46%), *Bacteriodetes* (22%), *Proteibacteria* (22%), A*ctinobacteria* (5%) and *Fusobateria* (2%) (Fig. 7D). The Heatmaps at genus (Fig. 8A) and phylum level (Fig. 8B) further illustrate the distinct abundance of the microbiota in the matched sputum and EV samples.

Philley et al.: Microbiome of Sputa

Fig. 6. Microbiota of matched EVs and sputum. The relative abundance of bacterial taxa identified in paired EV and sputum samples has been represented as stacked bar plots. Each bar represents an individual and each colored box, a bacterial taxon, at phylum, family and class. The height of a colored box represents the relative abundance at various taxonomic levels as indicated. S: Sputum; E: Extracellular vesicles. The respective patient's number from NTM-BCa group has been shown under each paired sputum and EV samples. A minimum average abundance of 0.05% was considered for calculating the entire taxa abundance.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

274

Philley et al.: Microbiome of Sputa

Fig. 7. Alpha and beta diversity ordination plots. (A) Measurement of the intra-sample diversity by alpha diversity method. The alpha diversity metrics use the counts (richness) and distribution (evenness, Shannon Index) of the OTUs within a sample as the basic values for these calculations. (B) Measurement and comparison of beta diversity in various groups using weighted UniFrac (phylogenitic) distance metrics. Data were plotted using PCoA ordination method utilizing the OTUs as the basic units. The percentage of variability of each principal coordinates (PC) has been explained on the axis of the plot. Each colored circle represents a sample from a group, as indicated. The probability distribution has been represented by the ellipses. (C-D) Exclusive abundance of various genera and phyla in the circulating exosomes. A minimum average abundance of 1% was considered for the categorization of the top taxa.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

275

Philley et al.: Microbiome of Sputa

Fig. 8. Heatmaps and clustering of paired EV and sputum microbiota samples at generic (A) and phylum (B) level. A minimum average abundance of 1% was considered to determine relative abundance at various taxonomic levels. S: Sputum; E: Extracellular vesicles; BR: Women with NTM lung disease and breast cancer.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

276

Philley et al.: Microbiome of Sputa

Discussion

Treatment of NTM lung disease is a daunting task because of the recurrent episodes of acquired drug resistance [3]. Moreover, increasing incidences of NTM among individuals with a history or current diagnosis of malignant disease such as BCa, making it even more challenging therapeutically. Clearly, a better understanding of the molecular mechanism of the NTM susceptibility/progression and characterization of the associated risk factors is warranted. The microbiome research has received considerable attention in this decade because of the increasing evidences implicating their critical role in inflammatory, autoimmune and malignant diseases [5, 39]. Dysbiosis or imbalance in microbiota community could influence tumorigenesis and therapeutic response by modulating the genotoxic response and alteration of the microenvironment and metabolism [39]. Thus, host-microbiome niche and their interplay appear to play a pivotal role in preventing or promoting various inflammatory diseases and neoplasm. Characterizing and understanding of the molecular biological and pathophysiological connections between the overall microbiome and MAC species in the NTM infected individuals could be helpful in risk assessment and better therapeutic management.

Our study establishes that other than *Mycobacterium*, DNA from many other unique pathogenic microbes was present in the sputa of the women with NTM infection. For example, *Fusobacteria*, found in higher abundance in the NTM-BCa and NTM subjects are a pathogenic oropharyngeal flora in periodontal and gingival diseases and associated with colorectal cancer development [5]. In a recent study, *Fusobacterium* was also found in higher abundance in the BCa compared to the benign breast tissues [8]. On the other hand, studies on the oral cavity microbiota in human papilloma virus positive or negative oral squamous cell carcinomas and Fanconi Anemia, identified *Streptococcus, Nesseria, Veilonella, and Haemophilus* as the dominating genera [40-41]. In another study of sputum microbiota in patients with pulmonary tuberculosis from the Indian subcontinent, *Streptococcus, Nesseria, Veilonella* has been identified as the most abundant microbes [42]. In another microbiota analysis of sputum from recurrent and treatment failure tuberculosis, abundance of *Streptococcus* was noted [9]. The sputa as well the circulating EVs from our patient cohort were also enriched with these specific genera, further suggesting their functional importance in the pathogenesis of NTM lung disease.

Recent studies uncovered a functional link between estrogen and microbiota [11-12]. Estrogens are steroid hormone, which activate their cognate receptors ERα and ERβ upon binding and associated with BCa tumorigenesis [12]. Estrogen like compounds when consumed, can be metabolized to its active form by specific microbiota and on the other hand, estrogen like compounds may promote growth and proliferation of certain microbes [12]. The microbial genes, whose products are capable of metabolizing estrogen have been identified and known as the estrobolome [43]. A certain group of microbes, including *Bacteroids, Faecalibacterium, Alistipes, Fusobacterium, Prevotella, Staphylococcus, Streptococcus* harbor estrobolome with β-glucaronidase and/or β-galactosidase activity, which could increase the intestinal reabsorption of estrogen [11-12]. This increase in estrogen metabolite is strongly associated with microbial diversity compared to the parent estrogen [12]. On the other hand, increased parent estrogen concentration compared to estrogen metabolite is associated with an increased risk of BCa [12]. Of note, all the estrobolome containing microbes described above were identified in the sputa or EVs of various groups of patients who had undergone estrogen therapy. Further functional studies are warranted to establish the functional role of the estrobolome containing microbes in NTM pathogenesis detected in these women.

Emerging studies also established a crucial role of EVs in host pathogen interaction during microbial infection and malignant transformation [13-19]. We recently identified expression of some BCa and obesity associated molecules exclusively in the circulating EVs of these NTM-BCa subjects [2]. Moreover, circulating EVs from these patients not only induced inflammatory cytokine production by normal human T cells, but also epithelial to mesenchymal transition in normal mammary epithelial cells, when co-cultured [2]. Other

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

277

Philley et al.: Microbiome of Sputa

than the cellular proteins, we also detected human mitochondrial DNA and human papilloma virus-16 associated E7 protein in the circulating EVs of prostate and head and neck cancer patients respectively [25, 44]. Recent studies also detected mutant *KRAS* and *TP53* DNA in the circulating EVs of pancreatic cancer patients and genomic DNA in prostate cancer patients derived EVs respectively [38, 45]. We have detected DNA of numerous unique microbes in the circulating EVs from some NTM-BCa patients, including the ones with the estrobolome activity. For example, *Alistipes* detected in the EVs has estrobolome activity [11] and recently been found in higher abundance in nipple aspirate fluid (NAF) of BCa patients [5]. The additional microbiota detected in the EVs compared to the sputa could be due to the fact that the EVs are secreted by various different types of cells. Importantly, irrespective of the cellular resources, these EVs harbored bacterial DNA from diverse microbial community, including the ones with the estrobolome activity, implicating their possible role in NTM pathogenesis.

To our knowledge, this is the first study of microbiome analysis in sputa and the circulating EVs from women with NTM-BCa and NTM disease, which confirms the existence of a diverse microbial community with estrobolome activity representing the oral cavity and breast cancer tissue microbiome to some extent. Our study also suggests that NTM might not be a "MAC" disease alone. Many other players from the diverse microbiome community as identified through this pilot study could be associated with NTM pathogenesis and therapeutic resistance. Molecular interaction between the "MAC" and "other microbiome/estrobolome" derived factors, and transport of this information through the EVs might potentially initiate and drive NTM pathogenesis. In concert with the MAC, monitoring and surveillance of the overall microbiome niche in a patient specific manner are warranted and could aid in better therapeutic design and risk assessment of the NTM infected individuals.

Acknowledgements

Disclosure Statement

The authors have no competing interests.

References

1   Macovei L, McCafferty J, Chen T, Teles F, Hasturk H, Paster BJ, Campos-Neto A: The hidden 'mycobacteriome' of the human healthy oral cavity and upper respiratory tract. J Oral Microbiol 2015;7:26094.

2   Philley JV, Kannan A, Griffith DE, Devine MS, Benwill JL, Wallace RJ Jr, Brown-Elliott BA, Thakkar F, Taskar V, Fox JG, Alqaid A, Bains H, Gupta S, Dasgupta S: Exosome secretome and mediated signaling in breast cancer patients with nontuberculous mycobacterial disease. Oncotarget 2017;8:18070-18081.

3   Griffith DE, Aksamit TR: Understanding nontuberculous mycobacterial lung disease: it's been a long time coming. F1000Res 2016;5:2797.

4   Hosgood HD 3rd, Sapkota AR, Rothman N, Rohan T, Hu W, Xu J, Vermeulen R, He X, White JR, Wu G, Wei F, Mongodin EF, Lan Q: The potential role of lung microbiota in lung cancer attributed to household coal burning exposures. Environ Mol Mutagen 2014;55:643-651.

5   Chan AA, Bashir M, Rivas MN, Duvall K, Sieling PA, Pieber TR, Vaishampayan PA, Love SM, Lee DJ: Characterization of the microbiome of nipple aspirate fluid of breast cancer survivors. Sci Rep 2016;6:28061.

6   Goedert JJ, Jones G, Hua X, Xu X, Yu G, Flores R, Falk RT, Gail MH, Shi J, Ravel J, Feigelson HS: 2015. Investigation of the association between the fecal microbiota and breast cancer in postmenopausal women: a population-based case-control pilot study. J Natl Cancer Inst 2015;107:pii:djv147.

7   Coburn B, Wang PW, Diaz Caballero J, Clark ST, Brahma V, Donaldson S, Zhang Y, Surendra A, Gong Y, Elizabeth Tullis D, Yau YC, Waters VJ, Hwang DM, Guttman DS: Lung microbiota across age and disease stage in cystic fibrosis. Sci Rep 2015:5:10241.

8   Hieken TJ, Chen J, Hoskin TL, Walther-Antonio M, Johnson S, Ramaker S, Xiao J, Radisky DC, Knutson KL, Kalari KR, Yao JZ, Baddour LM, Chia N, Degnim AC: The Microbiome of Aseptically Collected Human Breast Tissue in Benign and Malignant Disease. Sci Rep 2016:6:30751.

9   Wu J, Liu W, He L, Huang F, Chen J, Cui P, Shen Y, Zhao J, Wang W, Zhang Y, Zhu M, Zhang W, Zhang Y: Sputum microbiota associated with new, recurrent and treatment failure tuberculosis. PLoS One 2013;8:e83445.

10   Abreu MT, Peek RM Jr: Gastrointestinal malignancy and the microbiome. Gastroenterology 2014;146:1534-1546.

11   Kwa M, Plottel CS, Blaser MJ, Adams S: The Intestinal Microbiome and Estrogen Receptor-Positive Female Breast Cancer. J Natl Cancer Inst 2016; DOI:10.1093/jnci/djw029.

12   Chen KL, Madak-Erdogan Z: Estrogen and Microbiota Crosstalk: Should We Pay Attention? Trends Endocrinol Metab 2016;27:752-755.

13   Schorey JS, Bhatnagar S: Exosome function: from tumor immunology to pathogen biology. Traffic 2008;9:871-881.

14   Schorey JS, Cheng Y, Singh PP, Smith VL: Exosomes and other extracellular vesicles in host-pathogen interactions. EMBO Rep 2015;16:24-43.

15   Kahlert C, Kalluri R: Exosomes in tumor microenvironment influence cancer progression and metastasis. J Mol Med (Berl) 2013;91:431-437.

16   Greening DW, Gopal SK, Xu R, Simpson RJ, Chen W: Exosomes and their roles in immune regulation and cancer. Semin Cell Dev Biol 2015;40:72-81.

17   Driscoll O: Expanding on exosomes and ectosomes in cancer. New Eng J Med 2015;372:2359-2362.

18   Melo SA, Sugimoto H, O'Connell JT, Kato N, Villanueva A, Vidal A, Qiu L, Vitkin E, Perelman LT, Melo CA, Lucci A, Ivan C, Calin GA, Kalluri R: Cancer Exosomes Perform Cell-Independent MicroRNA Biogenesis and Promote Tumorigenesis. Cancer Cell 2014;26:707-721.

19   Zhao H, Yang L, Baddour J, Achreja A, Bernard V, Moss T, Marini JC, Tudawe T, Seviour EG, San Lucas FA, Alvarez H, Gupta S, Maiti SN, Cooper L, Peehl D, Ram PT, Maitra A, Nagrath D: Tumor microenvironment derived exosomes pleiotropically modulate cancer cell metabolism. Elife 2016;5:e10250.

20   Wallace RJ Jr, Brown BA, Griffith DE, Girard WM, Murphy DT, Onyi GO, Steingrube VA, Mazurek GH: Initial clarithromycin monotherapy for Mycobacterium avium-intracellulare complex lung disease. Am J Respir Cri Care Med 1994;149:1335–1341.

21   Wallace RJ Jr, Brown BA, Griffith DE, Girard WM, Murphy DT: Clarithromycin regimens for pulmonary Mycobacterium avium complex. The first 50 patients. Am J Respir Crit Care Med 1996;153:1766–1772.

22   Wallace RJ Jr, Brown-Elliott BA, McNulty S, Philley JV, Killingley J, Wilson RW, York DS, Shepherd S, Griffith DE: Macrolide/Azalide therapy for nodular/bronchiectatic mycobacterium avium complex lung disease. Chest 2014;146:276-282.

23   Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, Gordin F, Holland SM, Horsburgh R, Huitt G, Iademarco MF, Iseman M, Olivier K, Ruoss S, von Reyn CF, Wallace RJ Jr, Winthrop K: ATS Mycobacterial Diseases Subcommittee; American Thoracic Society; Infectious Disease Society of America. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med 2007;175:367-416.

24   Bange FC, Böttger EC: Improved decontamination method for recovering mycobacteria from patients with cystic fibrosis.Eur J Clin Microbiol Infect Dis 2002;21:546-548.

25   Philley JV, Kannan A, Qin W, Sauter ER, Ikebe M, Hertweck KL, Troyer DA, Semmes OJ, Dasgupta S: Complex-I alteration and enhanced mitochondrial fusion are associated with prostate cancer progression. J Cell Physiol 2016;231:1364-1374.

26   Human Microbiome Project. A framework for human microbiome research. Nature 2012;486: 215-221.

27   Human Microbiome Project. Structure, function and diversity of the healthy human microbiome. Nature 2012;486:207-214.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2019;52:263-279

DOI: 10.33594/000000020
Published online: 28 February 2019

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

279

Philley et al.: Microbiome of Sputa

28 Goodrich JK, Di Rienzi SC, Poole AC, Koren O, Walters WA, Caporaso JG, Knight R, Ley RE: Conducting a microbiome study. Cell 2014;158:250-262.

29 Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R: Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 2012;6:1621-1624.

30 Luna PN, Hasegawa K, Ajami NJ, Espinola JA, Henke DM, Petrosino JF, Piedra PA, Sullivan AF, Camargo CA Jr, Shaw CA, Mansbach JM: The association between anterior nares and nasopharyngeal microbiota in infants hospitalized for bronchiolitis. Microbiome 2018;6:2.

31 Edgar RC: Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;26:2460-2461.

32 Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, et al.: QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7:335-336.

33 Urbaniak C, Gloor GB, Brackstone M, Scott L, Tangney M, Reid G: The Microbiota of Breast Tissue and Its Association with Breast Cancer. Appl Environ Microbiol 2016;82:5039-5048.

34 Lahti L, Shetty S: Tools for microbiome analysis in R. 2017;Version 0.99.71.

35 Edgar RC: "UPARSE: Highly accurate OTU sequences from microbial amplicon reads". Nat Methods 2013:10:996-998.

36 Quast CE, Pruesse PY, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO: The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013;41:D590-596.

37 Lozupone C, Knight R: UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 2005;71:8228-8235.

38 Yang S, Che SP, Kurywchak P, Tavormina JL, Gansmo LB, Correa de Sampaio P, Tachezy M, Bockhorn M, Gebauer F, Haltom AR, Melo SA, LeBleu VS, Kalluri R: Detection of Mutant KRAS and TP53 DNA in Circulating Exosomes from Healthy Individuals and Patients with Pancreatic Cancer. Cancer Biol The 2017;18:158-165.

39 Guerrero-Preston R, Godoy-Vitorino F, Jedlicka A, Rodríguez-Hilario A, González H, Bondy J, Lawson F, Folawiyo O, Michailidi C[1], Dziedzic A, Thangavel R, Hadar T, Noordhuis MG, Westra W, Koch W, Sidransky D:. 16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment. Oncotarget 2016;7:51320-51334.

40 Al-Hebshi NN, Nasher AT, Idris AM, Chen T: Robust species taxonomy assignment algorithm for 16S rRNA NGS reads: application to oral carcinoma samples. J Oral Microbiol 2015;7:28934.

41 Furquim CP, Soares GM, Ribeiro LL, Azcarate-Peril MA, Butz N, Roach J, Moss K, Bonfim C, Torres-Pereira CC, Teles FR: The Salivary Microbiome and Oral Cancer Risk: A Pilot Study in Fanconi Anemia. J Dent Res 2017;96:292-299.

42 Krishna P, Jain A, Bisen PS: Microbiome diversity in the sputum of patients with pulmonary tuberculosis. Eur J Clin Microbiol Infect Dis 2016;35:1205-1210.

43 Plottel CS, Blaser MJ: Microbiome and malignancy. Cell Host Microbe 2011;10:324.

44 Kannan A, Hertweck KL, Philley JV, Wells RB, Dasgupta S: Genetic Mutation and Exosome Signature of Human Papilloma Virus Associated Oropharyngeal Cancer. Sci Rep 2017;7:46102.

45 Lázaro-Ibáñez E, Sanz-Garcia A, Visakorpi T, Escobedo-Lucea C, Siljander P, Ayuso-Sacido A, Yliperttula M: Different gDNA content in the subpopulations of prostate cancer extracellular vesicles: apoptotic bodies, microvesicles, and exosomes. Prostate 2014;74:1379-1390.