EBioMedicine

Optimized DNA extraction and purification method for characterization of bacterial and fungal communities in lung tissue samples

--Manuscript Draft--

Manuscript Number:	EBIOM-D-19-03672
Article Type:	Article (Original Research)
Keywords:	Microbiome; 16S rRNA gene; environmental contamination
Corresponding Author:	Andres Moya, Professor Fundacio per al Foment de la Investigacio Sanitaria i Biomedica SPAIN
First Author:	Vicente Pérez-Brocal, PhD
Order of Authors:	Vicente Pérez-Brocal, PhD
	Fabien Magne
	Susana Ruiz-Ruiz
	Carolina A. Ponce
	Rebeca Bustamante
	Viviana San Martin
	Mireya Gutierrez
	Gianna Gatti
	Sergio L. Vargas
	Andres Moya, Professor
Manuscript Region of Origin:	SPAIN
Abstract:	Background : The human lungs harbors a very poor microbial community, making it necessary to develop methods to enhance the recovery of nucleic acids isolated from bacteria and fungi, leading to a more efficient analysis of the lung tissue microbiota. Here we describe five different extraction protocols including pre-treatment, bead-beating and/or Phenol:Chloroform:Isoamyl alcohol steps, applied to non-selected lung tissue samples from autopsied individuals. Results : The resulting total DNA yield and quality, bacterial and fungal DNA amount and the microbial community structure were analyzed by DNA quantitation, qPCR and Illumina amplicon sequencing of bacterial 16S rRNA and fungal ITS genes, respectively. Bioinformatic modeling revealed that a large part of microbiome data from lung tissue samples is composed of microbial contaminants, although our blank controls clustered separately from biological samples. After removal of contaminant sequences, the effects of DNA extraction protocols on the microbial community were assessed. The major differences among samples could be attributed to inter-individual variations rather than DNA extraction protocols. However, inclusion of the bead-beater and Phenol:Chloroform:Isoamyl alcohol steps resulted in changes in the relative abundance of some bacterial and fungal taxa. Furthermore, inclusion of a pre-treatment step increased microbial DNA concentration but not diversity. Conclusions : The addition of a pre-treatment, bead-beater and Phenol:Chloroform:Isoamyl alcohol steps affects the characterization of both fungal and bacteria communities inhabiting the lung. In particular, the pre-treatment step might be determinant, as it increases the microbial DNA concentration and it may contribute to eliminate microbial DNA fragments resulting from dead microorganisms in lung tissue samples, which would make the microbial profile closer to the actual one living in the lung. Therefore, our data represent a comprehensive understanding of the potential impact of technical variation on the lung microbiome, which will help to limit preventable bias and to avoid controversy in the future on the live microorganisms inhabiting the lung.

Optimized DNA extraction and purification method for characterization of bacterial and fungal communities in lung tissue samples

Vicente Pérez-Brocal[1,2*], Fabien Magne[3,*,§], Susana Ruiz-Ruiz[1,2*], Carolina A. Ponce[3], Rebeca Bustamante[3], Viviana San Martin[4], Mireya Gutierrez[4], Gianna Gatti[4], Sergio L. Vargas[3,§] and Andrés Moya[1,2,5 §]

[1]Department of Genomics and Health, Foundation for the Promotion of Health and Biomedical Research of Valencia Region (FISABIO-Public Health), València, Spain; perez_vicbro@gva.es, ruiz_susrui@gva.es, Andres.Moya@uv.es.

[2]CIBER in Epidemiology and Public Health (CIBEResp), Madrid, Spain; perez_vicbro@gva.es, ruiz_susrui@gva.es, Andres.Moya@uv.es.

[3]Microbiology and Mycology Program, Biomedical Sciences Institute (ICBM), University of Chile School of Medicine, Independencia 1027, Santiago 8380453, Chile; fabienmagne@med.uchile.cl, cponce@med.uchile.cl, rebebus@gmail.com, svargas@med.uchile.cl

[4]Médico Legal Institute of Chile, Av. La Paz 1012, Independencia, Santiago 8380454, Santiago, Chile; ggatti@sml.gob.cl, vsanmartin@sml.gob.cl, mgutierrez@sml.gob.cl, ggatti@sml.cl

[5]Institute for Integrative Systems Biology (I2SysBio), University of Valencia and Spanish National Research Council (CSIC), València, Spain. Andres.Moya@uv.es

*These co-authors contributed equally to the work.

[§]Corresponding authors: Andres.Moya@uv.es, fabienmagne@med.uchile.cl, svargas@med.uchile.cl

Abstract

Background: The human lungs harbors a very poor microbial community, making it necessary to develop methods to enhance the recovery of nucleic acids isolated from bacteria and fungi, leading to a more efficient analysis of the lung tissue microbiota. Here we describe five different extraction protocols including pre-treatment, bead-beating and/or Phenol:Chloroform:Isoamyl alcohol steps, applied to non-selected lung tissue samples from autopsied individuals.

Results: The resulting total DNA yield and quality, bacterial and fungal DNA amount and the microbial community structure were analyzed by DNA quantitation, qPCR and Illumina amplicon sequencing of bacterial 16S rRNA and fungal ITS genes, respectively. Bioinformatic modeling revealed that a large part of microbiome data from lung tissue samples is composed of microbial contaminants, although our blank controls clustered separately from biological samples. After removal of contaminant sequences, the effects of DNA extraction protocols on the microbial community were assessed. The major differences among samples could be attributed to inter-individual variations rather than DNA extraction protocols. However, inclusion of the bead-beater and Phenol:Chloroform:Isoamyl alcohol steps resulted in changes in the relative abundance of some bacterial and fungal taxa. Furthermore, inclusion of a pre-treatment step increased microbial DNA concentration but not diversity.

Conclusions: The addition of a pre-treatment, bead-beater and Phenol:Chloroform:Isoamyl alcohol steps affects the characterization of both fungal and bacteria communities inhabiting the lung. In particular, the pre-treatment step might be determinant, as it increases the microbial DNA concentration and it may contribute to eliminate microbial

DNA fragments resulting from dead microorganisms in lung tissue samples, which would make the microbial profile closer to the actual one living in the lung. Therefore, our data represent a comprehensive understanding of the potential impact of technical variation on the lung microbiome, which will help to limit preventable bias and to avoid controversy in the future on the live microorganisms inhabiting the lung.

Keywords: Microbiome, 16S rRNA gene, environmental contamination

Background

Despite the fact that lungs have traditionally been considered to be sterile, recent studies have revealed that the respiratory airways harbor a complex microbial community. While it is suspected that the airway microbiome is involved in colonization resistance of respiratory pathogens, as well as in the development and integrity of the immune system, its role is not yet well understood [1]. Therefore, the airway microbiome and particularly the lung-associated microbiome may provide new perspectives in understanding the pathogenesis of airway infections and chronic lung disease.

However, analysis of the microbiome of lung tissue samples remains a technical challenge for researchers. Like all microbiome research based on culture-independent techniques, data quality is based on the effectiveness of extracting microbial DNA and removing polymerase chain reaction (PCR) inhibitors. As previously shown, different DNA extraction methods can generate different microbial profiles in the same fecal samples, mainly due to cell wall disruption [2-4]. Additionally, as the lung microbiome contains a low biomass, the microbial DNA extracted is not always sufficient for detection by PCR. Moreover, when human genomic DNA (gDNA) is present in higher concentrations than microbial DNA [5], it can inhibit the polymerase enzyme used in PCR. Thus, it is crucial to develop an efficient method of extracting DNA from microorganisms inhabiting the lung.

Collectively, the presence of both microbial DNA and human gDNA can affect the efficiency and sensitivity of subsequent PCR reactions, which may distort the apparent composition of microbial communities obtained using PCR-based techniques such as high-throughput sequencing [6,7].

It is known that samples with low microbial biomass are more sensitive to DNA contaminants introduced during sample processing from consumables, laboratory reagents, commercial DNA extraction kits or the laboratory environment [8-11]. Although many microbiome studies analyze low microbial biomass samples, they do not routinely analyze negative controls from DNA extractions and/or sequencing; some studies have even reported statistically noteworthy taxa that overlap with those observed in negative controls [12,8,13]. Thus, microbial contaminants in low microbial biomass samples, such as lung tissue samples, can greatly alter the relative abundance of the microbial communities under analysis; however, this is often less of a concern for molecular microbiologist.

Many investigations on the airway microbiome use DNA extraction methods that have not been previously validated. Thus, here we present validation of a new protocol for DNA extraction for use when analyzing the microbiome of lung tissue samples. The present study compares five protocols for assessing bacterial and fungal DNA recovery, subsequent microbial PCR detection, and the resulting bacterial and fungal community structures in order to define the optimal extraction method for use in lung microbiome studies.

Results

Quality and quantity of genomic DNA from five extraction protocols

Lung tissue specimens collected from seven autopsied individuals were processed for DNA extraction following the five methods described in the methods section (**Table 1**). The concentration and quality of the extracted DNA was assessed using a NanoDrop™ spectrophotometer (**Figure 1A**). Both protocol 2, including the bead-beating, and protocol 3, including the Phenol:Chloroform:Isoamyl alcohol step, had slightly increased DNA yields compared to the protocol 1 (the QIAamp DNA Mini kit, Qiagen); however these

finding were not significant. Combining these two additional steps, as in protocol 4, significantly increased DNA concentrations (1,104.8 ± 522.7 ng/µl) compared to the protocols 1 (525.4 ± 460.2 ng/µl, p=0.003) and 2 (557.4 ± 245.2 ng/µl, p=0.036). Protocol 5 (686.2 ± 172.3 ng/µl) also showed a relative, but non-significant, increase in DNA compared to protocol 1.

In terms of DNA quality, protocol 1 had the lowest 260/280-absorbance ratio (mean ± SD) as compared to all the others protocol (**Figure 1B**). Compared to protocol 1 (1.39 ± 0.25), the 260/280-absorbance ratio was higher in protocols 3 (1.59 ± 0.16, p= 0.090), 4 (1.73 ± 0.10, p=0.012) and 5 (1.63 ± 0.12, p=0.036), which all include a Phenol:Chloroform:Isoamyl alcohol step.

Evaluation of human, bacterial, and fungal DNA in the five extraction protocols

We evaluated the content of human, bacterial, and fungal genomes in our extracted DNA through qPCR amplification of human β-actin, 16S rRNA and 18S rRNA genes respectively. We observed that the different DNA extraction methods did not affect the *C*t obtained through qPCR amplification of the β-actin gene, suggesting that the level of human DNA was equivalent for all the DNA extraction protocols (**Figure 2A**). However, the *C*t of the 16S rRNA gene (mean ± SD) was significantly decreased in samples processed with protocol 5 (22.9 ± 2.2) compared to all others (26.3 ± 2.2 for the protocol 1; 25.3 ± 2.5 for the protocol 2; 26.3 ± 1.9 for the protocol 3; and 26.1 ± 2.4 for the protocol 4) (**Figure 2B**).

We detected low quantities of fungi in lung tissue samples, as the *C*t values of 18S RNA gene ranged from 37.3 to 40 (**Figure 3B**). Although there were no significant

differences, the *C*t of the 18S RNA gene tended to be decreased in samples processed with protocol 5 (38.9 ± 1.0) compared to others (39.6 ± 0.7 for protocol 1; 39.2 ± 0.9 for protocol 2; 40.0 ± 0.0 for protocol 3; and 39.7 ± 0.5 for protocol 4).

Impact of bacterial and fungal contamination of DNA

It is known that laboratory contamination can occur during the DNA extraction process, and that such contamination can compromise the results of NGS sequencing, particularly in low microbial mass samples, such as lung tissue samples.

Thus, we analyzed sequencing data of negative controls. Sequencing of the 16S rRNA and ITS genes was carried out on four negative controls using the Illumina MiSeq platform, which yielded a total of 22,756 bacterial reads (mean ± SD, 5,689 ± 3,268 sequences per sample) and 73,037 fungal reads (mean ± SD, 18,259 ± 5,228 sequences per sample). Overall, we detected 55 bacterial families belonging to 8 phyla in negative controls, represented by Proteobacteria (75.3%), Firmicutes (12.9%), and Actinobacteria (7.6%), followed by Bacteroidetes, Fusobacteria, Deinococcus-Thermus, Cyanobacteria and WPS-2 in lower abundance (<3% each) (**Figure 3A**). In addition, we detected archaeal taxa of the phylum Euryarchaeota (0.19%). In negative controls, 8 out of every 10 reads were assigned as Pseudomonaceae (38.5% of the reads), Rhizobiaceae (24.2%), Streptococcaceae (8.6%) or Nocardiaceae (4.3%) (**Figure 3B** and **Additional file 1: Table S1**). For fungi (**Figure 3C & D**), we identified 13 fungal families belonging to 3 phyla, namely: Ascomycota (86.2%), Basidiomycota (9.2%) and Chytridiomycota (0.23%). Nearly 9 out of every 10 reads corresponded to Aspergillaceae (53.9%), Nectriaceae (17.1%), Malasseziaceae (9.0%) or Cladosporiaceae (6.2%) (**Additional file 1: Table S1**).

To evaluate the contamination level of our samples, we identified contaminant reads in our lung tissue samples, reads mapping 100% identity with those of negative controls. For bacteria, the total number of reads removed after clustering sample reads with control reads at 100% of identity was 337,935 (representing 61.7% of the reads before clustering), reaching 9,655 ± 13,707 (mean ± SD) reads per sample. For fungi, those figures were lower, with 168,683 (representing 8.7% of the reads) and 4,820 ± 7,295 reads per sample. It is noteworthy that the negative controls contained only 1,580 reads (6.9%) showing 100% identity with at least one sample, whereas in fungi this figure was even lower, with only 258 reads (0.4%). Those reads were removed from the samples. Despite the relatively high number of reads obtained for the negative controls, most of them were actually unique and restricted to negative samples, as a consequence the impact of the removal of reads due to potential contamination on the samples was relatively low. To ensure that the microorganisms detected in our lung tissue samples were not completely a result of contamination, we compared the microbial community of our samples with that of negative controls by performing a CCA based on Bray-Curtis distances for bacterial and fungal communities (**Additional file 2 Figure S1 A and B**). While some lung tissue samples were close to the negative control, we observed that the majority of lung tissue samples and negative controls clustered separately for both bacteria and fungi, which was confirmed by the Adonis test ($p = 0.003$ for bacteria and $p = 0.017$ for fungi).

Lung microbial community diversity

Sequencing of the V3–V4 regions of the 16S rRNA gene was carried out on the 35 samples processed from the 7 individuals by the 5 distinct DNA extraction methods, which yielded a total of 1,005,778 raw reads with a mean read count of 28,737 reads per sample, and a

range of 4,717 to 137,755 reads. The number of reads per sample did not differ between DNA extraction protocols (data not shown). Similarly, the different extraction protocols did not affect the sequencing of the ITS regions in these samples, resulting in a total of 2,620,221 raw reads with a mean read count of 74,863 reads per sample, ranged from 18,150 to 256,441 reads (data not shown).

After sequence processing, including filtering, merging, chimera, and host reads removal, there were 547,468 reads (mean 15,642, ranging from 2,632 to 88,859 reads) for bacteria, and 1,945,654 reads (mean 55,590, ranging from 10,685 to 201,390 reads) for fungi.

After removal of sequences that clustered with negative control reads at 100% identity (see above), the composition and abundance of the remaining 209,533 bacterial and 1,776,971 fungal "contaminant-free" reads is summarized in **Figure 4**.

For bacteria (**Figure 4A & B**), 54 families were identified, represented mainly by the phyla Firmicutes (68.7%), Proteobacteria (17.5%), Bacteroidetes (7.7%), Fusobacteria (3.5%) and Actinobacteria (1.70%), as well as Cyanobacteria, Dependentiae and Patescibacteria (<1% of the reads each). The three most abundant families found, Bacillaceae (28.7%), Veillonelaceae (11.8%), and Streptococcaceae (10.6%), accounted for half of the total sample reads. For a complete record of bacterial families present see **Additional file 3–Table S2**.

In the case of fungi (**Figure 4C & D**), we identified 38 fungal families, divided mainly between two phyla, Ascomycota (32.6%) and Basidiomycota (11.7%); however, most reads could not be assigned to any phylum (55.7%). Unlike bacteria, no dominant families were identified for fungi, with the most abundant ones Hypocreales_fam_Incertae_sedis (6.6%), Aspergillaceae and Pneumocystidaceae (3.7%

each), Didymellaceae (3.6%) and Malasseziaceae (3.5%). These families accounted for only slightly more than 2 of every 10 reads, likely due to the lack of taxonomical identification. For a complete record of fungal families identified see **Additional file 3 – Table S2**.

Impact of DNA extraction protocol on the bacterial community

We investigated the effect of the various extraction protocols on bacterial community detection in lung tissue samples. In the **Figure 5A**, CCA analysis shows that DNA extraction protocols did not drive significantly differences in bacterial community profiles. Rather, samples clustered together by individual more than by extraction protocol **Figure 5B**. Furthermore, there were no significant differences in the Shannon index between the different DNA extraction protocols **(Additional file 4 –Figure S2)**.

Nevertheless we observed that variations in the DNA extraction protocols did drive changes in the relative abundance of some taxa (**Figure 6** and **Additional file–Table S3**). The DESEQ2 software was used to analyze family-level fold changes in taxa between the distinct protocols. Overall, the addition of a bead-beating step (protocols 2 and 4) improved the detection of family-level taxa (40 and 37 families detected with protocols 2 and 4 respectively) compared to protocol 1 (34 families detected). Based on $\log_2$ transformed relative abundance, twenty-two taxa were increased in protocol 2 compared to protocol 1. We observed increases ($\log_2$ fold-change >1) in families belonging to the phylum Firmicutes, such as Bacillaceae ($\log_2$ fold-change +26.2, $p<0.001$), Clostridiaceae_1 ($\log_2$ fold-change +3.5) and Streptococcaceae ($\log_2$ fold-change +1.0). Also, in the phylum Bacteroidetes families such as Weeksellaceae ($\log_2$ fold-change +21.3, $p<0.001$) and Porphyromonadaceae ($\log_2$ fold-change +1.2) were increased in protocol 2; in the phylum

Actinobacteria the family Actinomycetaceae (log_2 fold-change +7.3, p=0.010) was increased. Additionally, we also observed increases in the phylum Proteobacteria, including the families Halomonadaceae (log_2 fold-change +21.8, p<0.001), Rhizobiaceae (log_2 fold-change +4.8), Unknown_Family (log_2 fold-change +2.3), Sphingomonadaceae (log_2 fold-change +2.0), Xanthobacteraceae (log_2 fold-change +1.5) and Burkholderiaceae (log_2 fold-change +1.0).

The Phenol:Chloroform:Isoamyl alcohol step also affected the detection of some family taxa (34 families detected), although to lesser extent than bead beating. Comparisons between protocols 3 and 1 showed that the the Phenol:Chloroform:Isoamyl alcohol step improved the detection of 18 families of the phylum Fusobacteria, including Fusobacteriaceae (log_2 fold-change +23.7, p<0001); the phylum Bacteroidetes, including the family Porphyromonadaceae (log_2 fold-change +2.3); the phylum Actinobacteria, including Atopobiaceae (log_2 fold-change +1.4); the phylum Firmicutes, including Clostridiaceae_1 (log_2 fold-change +2.0) and Carnobacteriaceae (log_2 fold-change +1.8); and the phylum Proteobacteria, including Beijerinckiaceae (log_2 fold-change +4.4), Rhizobiaceae (log2 fold-change +4.4), Moraxellaceae (log_2 fold-change +2.2), Reyranellaceae (log_2 fold-change +1.5) and Pasteurellaceae (log_2 fold-change +1.5).

The combination of the bead-beater and Phenol:Chloroform:Isoamyl alcohol steps (Protocol 4) appeared to have a synergistic effect on the abundance of bacterial taxa (37 families detected). Comparisons of the relative abundance observed in protocols 4 and 1 revealed increased log_2 fold changes in 20 bacterial families; these were the primarily the same families (18/20) that were increased in protocols 2 and/or 3. Increase family-level taxa (log_2 fold-change >1) belonged to the phylum Actinobacteria, including Atopobiaceae (log_2 fold-change +5.9, p=0.048) and Micrococcaceae (log_2 fold-change +2.7); the phylum

Bacteroidetes, including Weeksellaceae (log2 fold-change +2.8), Porphyromonadaceae (log_2 fold-change +2.1) and Prevotellaceae (log_2 fold-change +1.2); the phylum Firmicutes, including Clostridiaceae_1 (log_2 fold-change +3.6), Carnobacteriaceae (log_2 fold-change +3.3), Leuconostocaceae (log_2 fold-change +2.0), Lactobacillaceae (log_2 fold-change +1.9), Streptococcaceae (log_2 fold-change +1.2), Staphylococcaceae (log_2 fold-change +1.2) and Veillonellaceae (log_2 fold-change +1.2); the phylum Fusobacteria, such as Leptotrichiaceae (log_2 fold-change +1.3); and the phylum Proteobacteria, such as Rhizobiaceae (log_2 fold-change +3.4), Sphingomonadaceae (log_2 fold-change +2.9) and Neisseriaceae (log_2 fold-change +1.1).

Fewer family taxa were identified in protocol 5 (31 families detected) as compared to the other protocols; however the pre-treatment step did increase the detection of a few family-level taxa. Comparisons between protocol 4 and 5 showed an average log_2 fold-change increase (>1) in 13 families, belonging to the phylum Firmicutes, including Lactobacillaceae (log_2 fold-change +3.2), Ruminococcaceae (log_2 fold-change +3.0), Family_XI (log_2 fold-change +1.0) and Staphylococcaceae (log_2 fold-change +1.0); the phylum Bacteroidetes, including Weeksellaceae (log_2 fold-change +1.1); the phylum Proteobacteria, including Unknown_Family (log_2 fold-change +6.1); and the phylum Fusobacteria, including Fusobacteriaceae (log_2 fold-change +6.1).

As compared to protocol 1, protocol 5 increased detection for 15 families. Among those, we observed an increased family abundance (log_2 fold-change >1) in the phylum Firmicutes, including Lactobacillaceae (log_2 fold-change +3.5), Carnobacteriaceae (log_2 fold-change +2,8), Staphylococcaceae (log_2 fold-change +2.3) and Family_XI (log_2 fold-change +1,8); the phylum Actinobacteria, including Micrococcaceae (log_2 fold-change +1.6); the phylum Bacteroidetes, including Weeksellaceae (log_2 fold-change +4.3),

Porphyromonadaceae (log_2 fold-change +2.5) and Prevotellaceae (log_2 fold-change +1.5); the phylum Fusobacteria, including Fusobacteriaceae (log_2 fold-change +6,2); and the phylum Proteobacteria, including Unknown_Family (log_2 fold-change +1.1).

In contrast, we observed that 13 family taxa were decreased in protocol 5 compared to protocol 1 (log_2 fold-change < 0), primarily belonging to the phylum Proteobacteria (9/13), as well as the phyla Firmicutes (1/13), Bacteroidetes (1/13), Dependentiae (1/13), and Fusobacteria (1/13). The majority of these families (11/13) were also decreased in protocols 2, 3 and/or 4.

Impact of DNA extraction protocol on the fungal community

Similarly, we investigated whether the DNA extraction protocols affected the composition of fungal community. CCA analysis showed that samples tend to cluster according by extraction protocol, however this grouping is not statistically significant, suggesting that the DNA extraction method affects the fungal community (**Figure 5C**). In addition, we observed that samples cluster according to the subjects (**Figure 5D)**.

Compositional comparison revealed that the number of fungal families identified ranged from 9 in protocol 5 to 24 in protocol 2, with figures for the other protocols of 13 (protocol 4), 15 (protocol 3), and 17 (protocol 1). We noted that the bead-beater step improved fungal retrieval at the family level (protocol 2), although this increase was not found for protocol 4, which includes the bead-beater and Phenol:Chloroform:Isoamyl alcohol steps.

To determine the effect of bead beating, we compared changes in the relative abundance of taxa between protocols 2 and 1. The following families showed a log_2fold change increase in protocol 2: Malasseziaceae (log_2 fold-change +6.2, p=0.033) from the

phylum Basidiomycota and Cladosporiaceae (log_2 fold-change +3.5), Dipodascaceae (log2 fold-change +2.6) and Aspergillaceae (log_2 fold-change +1.1) from the phylum Ascomycota (**Figure 7** and **Additional file–Table S3**).

In contrast, the Phenol:Chloroform:Isoamyl alcohol step appeared to induce fewer changes (protocol 3) compared to protocol 1, as similar number of family taxa were recovered (15 and 17 families detected with protocols 3 and 1, respectively). However, changes were observed in the phylum Basidiomycota, including the related family Malasseziaceae (log_2 fold-change +5.8, p=0.044) and in the phylum Ascomycota, including Hypocreales_fam_Incertae_sedis (log_2 fold-change +5.2) and Cladosporiaceae (log_2 fold-change +0.9).

We observed that the combination of the bead-beater and Phenol:Chloroform:Isoamyl alcohol steps, protocol 4, as compared to protocol 1 increases the families Malasseziaceae (log_2 fold-change +6.2, p=0.033) from the phylum Basidiomycota, which was also increased protocols 2 and 3. In addition, the family Saccharomycetales_fam_Incertae_sedis (log_2 fold-change +28.3, p<0.001) from the phylum Ascomycota was also increased.

Although protocol 5 recovered fewer family taxa than protocol 1 (9 versus 15 families detected, respectively), comparison between protocols 5 and 4 revealed that the pre-treatment step seems to improve retrieval of the families Aspergillaceae (log_2 fold-change +1.6), Pneumocystidaceae (log_2 fold-change +0.8), Didymellaceae (log_2 fold-change +0.6) and Saccharomycetales_fam_Incertae_sedis (log2 fold-change +0.1) from the phylum Ascomycota; and Malasseziaceae (log2 fold-change +0.4) from the phylum Basidiomycota.

All family taxa increased in protocol 5 compared to protocol 1 were also increased by the bead-beating, the Phenol:Chloroform:Isoamyl alcohol or the pre-treatment steps: Didymellaceae (log_2 fold-change +29.3, $p<0.001$), Saccharomycetales_fam_Incertae_sedis (log2 fold-change +28.9, $p<0.001$), Aspergillaceae (log2 fold-change +1.1) and Pneumocystidaceae (log2 fold-change +0.1) from the phylum Ascomycota; and Malasseziaceae (log2 fold-change +7.2, p=0,012) from the phylum Basidiomycota.

Discussion

Here we provide the first description of the bacterial and fungal community inhabiting the lungs in healthy human subjects. Furthermore, we evaluated the effect of five different DNA extraction protocols on the quantity and quality of nucleic acids, as well as on the bacterial and fungal community composition. We show that additional steps, including bead-beating, Phenol:Chloroform:Isoamyl alcohol, and pre-treatment affect the characterization of the microbial community in lung tissue samples; data that should considered in the future lung microbiome studies.

Over the last decade, the development of culture-independent techniques for microbiological analysis has uncovered the existence of a microbial community in the lung, which was previously considered to be sterile. The lower respiratory tract of healthy subjects was mainly colonized by anaerobic bacteria, such as the phylum Bacteroidetes inlduing the family Prevotellaceae and the phylum Firmicutes including the families Veillonellaceae and Streptococcaceae, as well as aerobic bacteria, such as the phylum Proteobacteria including the families Pseudomonaceae, Neisseriaceae, and Pasteurellaceae. [14,15]. However, these data are based on studies whose samples were contaminated by

336 microorganisms of the upper respiratory tract (i.e. sputum, bronchial aspirate, lung biopsy
337 and bronchoalveolar lavage), as during their collection they transited through the
338 oropharynx; thus, they do not represent the commensal communities inhabiting the lung
339 [16]. Except for sputum, the collection of these samples is ethically difficult to perform in
340 healthy subjects, as they require invasive procedures, making the characterization of the
341 healthy lung microbiome extremely difficult. We show that lung tissue samples from
342 healthy subjects were composed mainly of the phylum Firmicutes, largely represented by
343 the families Anaerobacillaceae, Streptococcaceae, Clostridiaceae and Veillonellaceae. We
344 also detected the phyla Bacteroidetes (mainly represented by the family Prevotellaceae) and
345 Proteobacteria (mainly represented by the families Neisseriaceae and Pasteurellaceae) in
346 lower concentrations. We observed great diversity in the microbiome inhabiting the lung,
347 which had been poorly characterized in the past. Although a large number of fungal
348 sequences in our samples were not able been identified, specifically due to a lack of
349 reference genomes [17], the fungal families we were able to identify came primarily from
350 the phyla Ascomycota and Basidiomycota, as well as Saccharomycetae, Saccharomycetales
351 family Incertae sedis, and Aspergillaceae. Interestingly, species of some identified fungal
352 families such as *Candida*, *Malassezia* and *Saccharomyces*, have been previously reported
353 to cause lung infection, suggesting that they may be also be present in the lungs as
354 commensals [18].

355 However, the study of the lung microbiota remains a technical challenge for
356 researchers due to the low microbial biomass inhabiting the human lung, estimated at
357 approximately 2.2×10^3 bacterial genomes per cm^2 [14,19]. Consequently, low microbial
358 biomass samples are more sensitive to inherent contamination during the DNA extraction
359 process, affecting sequencing data and interpretation of the microbiota. This is even more

of a concern for fungi, which were present in lower concentrations in our samples as compared to bacteria. This problem was well illustrated in a study by Lauder and collaborators, who characterized the microbiota of the placenta, which is also composed of a low microbial biomass [20]. The authors were unable to detect significant differences between the placental samples and contamination controls. In contrast, our study shows a significant separation between lung tissue samples and environmental contamination using CCA based on Bray-Curtis distances, (Adonis test, $p<0.05$).

However, we did identify potential contaminants in our samples with bacterial contamination (67.7% of the sequence reads) more common than fungal contamination (8.7% of sequence reads). Contaminants were identified through a computational approach, based on 100% identity sequences present in both negative controls and lung tissue samples. This strategy appears more successful in removing sequences, which is vital as the lung microbiota can be contaminated by the bacterial content of inhaled air. It is likely that some bacterial species of the lung microbiota are the same species found in contaminant indoor air. Nevertheless, it is probable that some bacterial species in the lung microbiota were removed through our strategy. As the length of sequencing reads was only about 460 bp, distinct microbial strains present in the lung microbiota and in the contaminants could have 100% identity sequences. Thus, we probably overestimated the level of contamination in our samples. Nevertheless, we considered it better to analyzing lung tissue samples that do not contain DNA contaminations, as the residing microbiota in the lung is still being discovered.

In microbiome studies, it is well known that DNA extraction protocols are directly related to the quality of sequencing and taxonomic identification of microorganisms. The need for high concentration and quality of extracted DNA is of special concern in low

biomass samples, such as lung tissue samples. In addition to the low microbial concentration, lung tissue samples contain a larger part of host eukaryote as opposed to microbial DNA, making it harder to amplification and sequencing microbial DNA. In this context choosing and adequate DNA extraction protocol ensures not only an efficient recovery of all microorganisms present in the samples, but also the quality of the extracted DNA.

In our study, the addition of a bead-beating step in protocol 2 did not significantly increase the recovery of extracted DNA compared with protocol 1. However, it is known that the mechanical action of the beads on the microbial wall improves microbial disruption. We note that the bead beating step mainly favors the detection of taxa from the phyla Firmicutes (*e.g.* the families Bacillaceae and Clostridiaceae_1) and Actinobacteria (*e.g.* Actinomycetaceae and Atopobiaceae) that are gram positive bacteria and known to have many layers of peptidoglycan in their thick cell wall, which is not easily destroyed. Moreover, although the Proteobacteria was reduced, we also retrieved increased abundances in some gram-negative bacteria from the phyla Bacteroidetes (*e.g.* the families Weeksellaceae and Porphyromonadaceae) and Proteobacteria (*e.g.* the families Halomonadaceae, Rhizobiaceae and Sphingomonadaceae). Previously, DNA extraction methods based only on enzymatic treatment without physical disruption showed reduced recovery of Gram-positive bacteria and relatively elevated levels of Gram negatives [21,22]. It is noteworthy that yeasts and other fungi often have a cell wall that is harder to lyse than bacterial cell walls and DNA extraction kits are generally not optimized for fungal DNA extractions. Specifically, we observed a higher abundance of the families Malasseziaceae and Aspergillaceae in samples processed with bead beating. It has been previously shown extraction protocols that employ bead beating increase DNA yields of

genera begonging to the Aspergillaceae family [23] and improve the recovery of the Malasseziaceae family [24]. In addition to the families Aspergillaceae and Malasseziaceae, our study revealed that the bead-beating step also increased the recovery of other fungal taxa, such as the families Cladosporiaceae and Dipodascaceae.

On the other hand, the additional of Phenol:Chloroform:Isoamyl alcohol step in the protocol 3 lead to modestly changes in the bacterial and fungal communities. However we observed that the Phenol:Chloroform:Isoamyl alcohol step ensured a better quality of the extracted DNA (ratio 260/280-absorbance) compared with protocol 1. While the phenol can aid in disrupting the cell wall by denaturing protein and lipids, it also permits the removal of PCR inhibitors by separating Nucleic Acids from other compounds.

When the beat-beating and phenol chloroform steps were included in the same protocol (protocol 4), their respective effects tend to be additive. Protocol 4 produced the highest concentration and the best quality of extracted DNA. Nevertheless, the bacterial and fungal DNA, estimated by qPCR, was not significantly elevated as compared to protocol 1. Moreover, the majority of bacterial and fungal taxa increased in protocol 4 were also increased in protocol 2 (beat-beating) and/or the protocol 3 (phenol chloroform).

Subsequently, we experimented with the addition of a pre-treatment step that consisted in disrupting the lung tissue in PBS 1X under agitation, followed by a filtering step on gauze and centrifugation to harvest microbes. Compared to protocol 1, we did not observe a loss of total DNA in lung tissue samples processed with the protocol 5. In contrast a higher concentration of bacterial and fungal DNA, determined by qPCR, was observed in protocol 5 compared with all other protocols. These results were expected, as the filtering and centrifugation steps in pre-treatment may both diminish the concentration of human DNA and concentrate microorganisms. Nevertheless, the estimation of human

DNA in the protocol 5 was not significantly different from other protocols. Although a better concentration of microbial DNA is observed in the protocol 5, fewer bacterial and fungal taxa were observed, suggesting that the addition of the pre-treatment step resulted in the loss of microorganisms. On the other hand, this may have resulted from the remove of microbial DNA from dead and/or broken-down microorganisms. No methodologies based on sequence analysis of polymerase chain reaction amplicons from microbial markers (such as 16SrRNA and ITS genes in our study) can differentiate living, dead, or ruptured bacteria, as all of these generate the same positive signals [25]. In our study, the centrifugation performed during the pre-treatment step contributed to pellet the live microorganisms and remove others particles in the supernatant such as free DNA, resulting in the enrichment of living microbes. Although we did not show that the pre-treatment step decreases the proportion of dead microorganisms, the protocol 5 including the pre-treatment step could be a very important issue. In effect, controversy exists regarding the presence of live commensal bacteria in the lung (as well as in the placenta) since the detected DNA sequences may be the result of the breakdown of microorganisms and not live/reproducing microbial community members.

Conclusions

Using lung samples from autopsied subjects, we describe for the first time the bacterial and fungal communities inhabiting the lung of healthy subjects. Collectively, our study shows that low microbial biomass samples, such as lung tissue samples, require particular care since high relative contaminate sequences could be retrieved. In addition, we found that the DNA extraction protocols affected the detection and abundance of bacterial and fungal

taxa, which may influence the interpretation of results, even though a small percentage of the total estimated microbial communities were affected. The bead-beating and Phenol:Chloroform:Isoamyl alcohol steps improved the detection of bacterial and fungal community through an efficient lysis of microorganisms and a higher quality of extracted DNA, respectively. Additionally, we found that the addition of a pre-treatment step improves the amplification of microbial DNA, but it may also eliminate microbial DNA fragments resulting from dead microorganisms in lung tissue samples. Therefore, the microbial community profiles from samples processed with protocol 5, which included a pre-treatment step, may be closer to those inhabiting the lung. As lung microbiome studies are needed to determine whether pathogenic relationships between the microbiome and lung disease exist, this protocol, including the pre-treatment step described in our study, may represent a good alternative to avoid controversy as to the live microorganisms inhabiting the lung.

Methods

Study design and sample collection

Lung tissue samples were obtained from individuals, four children aged 4.7 ± 2.8 months and three adults aged 42.4 ± 17.4 years, whose autopsy was legally required at the Servicio Medico Legal (Chilean coroner's office) in Santiago. Subjects were selected on the basis of unexpected death at home (in the case of children) and unexpected or violent death occurring in the street (in the case of adults). Additionally inclusion criteria were: not having been admitted to the hospital, absence of known immunocompromising conditions, and absence of obvious pulmonary disease on macroscopic examination. Autopsy diagnosis

was established on the basis of clinical history, results of post-mortem laboratory tests, and gross findings. Medical information (including age, date of death, autopsy findings, and autopsy diagnoses) was collected from the coroner's report. Lung biopsy specimens were anonymized after medical information from the autopsy was obtained.

For each autopsied subjects, the right upper lobe was removed using sterile equipment and stored at −80°C in a sterile plastic bag until processing for analysis. The right upper lobe samples were transported to a biosafety cabinet, where they were removed from the bag and placed on a large sterile plate. For each sample, the pleura were carefully removed in order to access untouched tissue using separate sterile equipment. Small samples were obtained from deep lung tissue, cut into small pieces, divided into five aliquots (0.4 g) and frozen at -80ºC for DNA extraction.

DNA extraction protocols

In this study, we compared five DNA extraction protocols (**Table 1**) based on modifications to the QIAamp DNA Mini kit (Qiagen) used to detect Pneumocystis in lung tissue samples [26]. Protocol 1 was based on the use of the QIAmp kit alone. The altered protocols were as follows: in protocol 2 we introduced a bead-beating step; in protocol 3 we added a Phenol:Chloroform:Isoamyl alcohol step; in protocol 4 we added both a bead-beating and Phenol:Chloroform:Isoamyl alcohol step; and in protocol 5 we added a pre-treatment step followed by the bead-beating and Phenol:Chloroform:Isoamyl alcohol steps. In all DNA extraction protocols, the DNA was eluted with 100-µL of elution buffer supplied with the kit.

Protocol 1: Small pieces obtained from the right upper lobe were homogenized in 200-µL of PBS using the Ultra Turrax® homogenizer (Biospec Products Inc.). DNA was extracted and purified from 200-µL of this solution using the QIAamp DNA Mini kit (Qiagen), according to the manufacturer's instructions. Briefly, the homogenate was mixed with 180-µL ATL buffer and 20 µL of protease and incubated for 30 min at 56ºC. Then 200-µL AL lysis buffer was added and the mixture was incubated for 10 min at 70ºC. The remaining steps consisted of washing binding DNA to the column, according to the manufacturer's recommendations.

Protocol 2: DNA extraction was performed following the procedure in protocol 1 with an additional bead-beating step to increase the disruption of microbial cells. The only changes to protocol 1 are the following: i) samples were homogenized using the Ultra Turrax® homogenizer (Biospec Products Inc.) ii) 400 µL of sterilized zirconia beads (0.1 and 0.5 mm) were added in the tubes that were then homogenized using a Mini-Beadbeater-8 (BioSpec Product) at medium speed for 2x3 min, iii) the tubes were centrifuged at 4 °C for 10 min (15,000 g) and the supernatant was transferred to a new 2-mL tube.

Protocol 3: This protocol consisted in extracting DNA as in the protocol 1, supplemented with an additional Phenol:Chloroform:Isoamyl alcohol step. As in the protocol 1, tissue disruption and lysis step were performed using the QIAamp DNA Mini kit (Qiagen). Next, we added a volume of Phenol:Chloroform:Isoamyl alcohol (25:24:1) followed by vortexing for 1 min and centrifugation at 4 °C for 10 min (10,000 xg). We removed ~600-µL of the upper aqueous solution avoiding picking up any of the phenol/chloroform/isoamyl alcohol phase and it was used to purify the extracted DNA. Then DNA was washed and

precipitated as in the protocol 1 using the column of the QIAamp DNA kit (Qiagen), following the manufacturer's recommendations.

Protocol 4: It consisted in the addition of both bead-beating and Phenol:Chloroform:Isoamyl alcohol steps to the protocol 1. The tissue disruption and lysis steps were performed as in the protocol 1, using the QIAamp DNA Mini kit (Qiagen). After the last lysis step, we added 400 µL of sterilized zirconia beads (0.1 and 0.5 mm) and a volume of Phenol:Chloroform:Isoamyl alcohol (25:24:1) followed by vortexing for 1 min. Next, the samples were homogenized using a Mini-Beadbeater-8 at middle speed for 2x3 min and centrifuged at 4 °C for 10 min (15,000 xg). The supernatant corresponding to the aqueous solution was transferred to a new 2-mL tube and used for the washing and purification of the DNA, as performed in the protocol 1.

Protocol 5: This protocol was performed according to the procedure of the protocol 4. However, small pieces of lung tissue were not homogenized with the Ultra Turrax® homogenizer (Biospec Products Inc.), instead they were homogenized by magnetic stirrer agitation in 20 mL of sterile PBS (pH 7.2) in ice pack–covered screw-capped flasks for 30 min. The homogenate was then filtered using sterile gauze and the stirring flasks were washed with sterile PBS to collect any remnants of the specimens. The filtrate was centrifuged at 4°C for 10 min (2,900 xg) and the pellet was reconstituted in 200 µL of sterile PBS (pH 7.2). DNA was extracted in the same way as the protocol 4, which included the addition of both bead-beating and Phenol:Chloroform:Isoamyl alcohol steps.

DNA extraction controls

Blank samples consisting of buffer supplied with the kit were processed together with the lung tissue samples. Four blank samples were processed with the Ultra Turrax® homogenizer and DNA was extracted according to the protocols 1-4; resulting in one blank sample per extraction protocol. One blank sample was processed with the pre-treatment step and processed according to the protocol 5.

DNA quantitation and quality assessment

DNA concentrations were initially measured using the Qubit double-stranded DNA (dsDNA) BR assay kit on a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA). As high DNA concentrations may exceed the limit of Qubit capacity for PCR reactions, dilutions were also prepared and measured. In addition, since the extractions can contain proteins and other organic molecules that are known to affect downstream procedures such as DNA amplifications in PCR, we also determined DNA purity by measuring the concentration of undiluted DNA as well as absobence ratios at 260/280 using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Pittsburgh, PA, USA).

Estimation of human, bacterial and fungal DNA levels in DNA extracted from lung tissue samples

Levels of human, bacterial and fungal DNA present in DNA extracted from lung tissue samples were assessed using qPCR method by amplifying the human β-actin, the bacterial 16S rRNA, and the 18S rDNA gene of fungi, respectively. qPCRs were carried out in 10 µL reactions containing 2 µL of diluted template (~10 ng/µl) or water (negative template control), 2X LightCycler® 480 SYBR Green I Master (Roche Diagnostics), 0.3 µM each of the forward and reverse primers: the forward 5'-TTGTTACAGGAAGTCCCTTGCC-3'

and the reverse 5'-ATGCTATCACCTCCCCTGTGTG-3' to amplify the human β-actin; the forward 515F 5'-GTGCCAGCMGCCGCGGTAA-3' and reverse 5'-CTTGTGCGGKCCCCCGYCAATTC-3' to amplify the V4 hyper variable region of the 16S rRNA gene of bacteria [27]; and the forward 5'-TTAGCATGGAATAATRRAATAGGA-3' and reverse 5'-TCTGGACCTGGTGAGTTTCC-3' to amplify the V4 (partial) and V5 variable regions of the 18S rDNA of fungi [28], and 2.4 µL of water. PCR amplification consisted of an initial heating step at 95°C for 5 min, followed by 40 cycles at 95°C for 20 s, and 59°C (for the β-actin gene), 58ºC and 57ºC for 30 s, the 16S rRNA and 18S RNA gene region, respectively, and 72ºC for 25 s. PCR reactions were performed on the Roche® LightCycler 480 instrument (Roche Diagnostic, Switzerland). All samples and controls were run in duplicate. The human, bacterial, and fungal DNA in lung tissue samples was estimated by real-time PCR. All samples were adjusted to 10 ng/µL using the Qubit 4 Fluorometer (Invitrogen), allowing for comparison of *C*t values.

16S rRNA gene and ITS amplification, library construction, and sequencing

For bacterial 16S rRNA gene amplification, primers (forward primer 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3′ and reverse primer 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3′) spanning the V3/V4 hypervariable regions were used. PCR conditions were as follows: 3 min of initial denaturation at 95 °C followed by 25 cycles of denaturation (30 s at 95 °C), annealing (30 s at 55 °C) and elongation (30 s at 72 °C), with a final extension at

72°C for 5 min. Internal controls of extraction and amplification were also analyzed together with the samples.

As for the fungi, an internal transcribed spacer (ITS) region was amplified. A pre-amplification with primers ITS1-F 5'-TAGAGGAAGTAAAAGTCGTAA-3' and ITS2-R_KYO2 5'-TTYRCTRCGTTCTTCATC-3' spanning the small subunit and the 5.8S region of the rRNA operon, respectively, was carried out. PCR conditions were as follows: initial denaturation at 95°C for 3 min, followed by 30 cycles of denaturation (30 s at 95°C), annealing (30 s at 56°C) and elongation 20 s at 72 °C, with a final extension at 72°C for 5 min.

A second PCR amplification with internal primers (ITS1-FInt 5'-*TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG*GGAAGTAAAAGTCGTAACAAGG-3', positions 1737-1758, and ITS2_RInt: 5'-*GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG*CTRYGTTCTTCATCGDT-3', positions 2026-2042) containing the adapters sequence (in italics), was performed on 10.5 µl of the primary PCR. Reaction conditions were: initial denaturation at 95°C for 2 min, followed by 28 cycles of denaturation (30 s at 95°C), annealing (30 s at 58°C) and elongation 30 s at 72 °C, with a final extension at 72°C for 5 min.

The resulting products were verified in a 1.4% agarose gel run in 1x TBE and purified amplicons were quantified using a Qubit® 3.0 Fluorometer (Thermo Fisher Scientific). Next, dual indices were attached to both ends of the PCR products using Nextera XT Index Kit (Illumina). Equimolar amounts of DNA per sample were pooled and sequenced using the MiSeq® reagent kit version 3 (Illumina), on a MiSeq desktop sequencer (2 x 300 bp paired-end reads) (Illumina).

Bioinformatic processing

Demultiplexed raw sequencing data, delivered as forward and reverse fastq files containing reads in matched order, and free of primer, adapter and linker sequences were the input files for the DADA2 pipeline [29], which was used to analyse the quality profiles, for filtering and trimming to remove Ns, expected errors and low quality tails. After learning the error rates with the DADA2 algorithm, a dereplication step was used to reduce computation time by collapsing redundant reads into unique ones, but counting them. Next, using the dereplicated data, true sequence variants were inferred. Paired reads were then merged by aligning denoised forward and reverse reads, provided they had a minimum overlapping set in 15 identical bases. Merged reads were used to construct the amplicon sequence variant table, and chimeric sequences were identified and removed. Before taxonomy assignment, there was a step of removal of human sequences linked to bowtie2-2.3.4.2 [30] against the reference human genome database GRCh38.p11, Dec 2013, using very sensitive parameters. The unaligned reads were used to assign taxonomy implemented by the naïve Bayesian classifier method, using the Silva reference database for bacteria and Unite database for fungi, extending the assignment to species level, when possible, with the silva_species_assignment_v132.fa.gz file for bacteria and sh_general_release_dynamic_01_12_2017.fasta files for fungi, respectively, and complemented with a blastn search. In the end, counts were obtained for operational taxonomic units (OTUs), and collapsed to the family level.

Contaminated sequences assessment

The bacterial and fungal read sequences obtained in lung tissue samples, clustering with the DADA2 pipeline at 100% identity with those present in the negative controls, were removed from the analysis of samples, for each group, respectively. The proportion of removed sequences was calculated for each lung tissue sample.

Bacterial and fungal community composition, abundance, and diversity analysis

The table containing the collapsed taxonomy up to the species level was converted into Biom format, using the QIIME pipeline version 1.9.0 [31] for composition and absolute and relative abundance analyses, as well as for ecological diversity.

For within sample diversity, or alpha diversity, 1,000 rarefactions of 2,600 random reads per sample in the case of bacteria; and 12,000 for fungi, with replacement, were carried out and the alpha diversity was calculated using the Shannon diversity index. As for diversity between samples, or beta diversity, variation was assessed using canonical correspondence analysis (CCA), implemented in R version 3.1.0 [32], on a Bray-Curtis dissimilarity matrices generated with the QIIME pipeline.

Statistical analyses

Diversity within groups of samples was analysed using a pairwise comparison according to the method and individual, using the qiime script compare_alpha_diveristy.py, with the default non-parametric t-tests (i.e Monte Carlo permutations). Boxplots were also generated by this script [33]. Fold change in relative abundance of family level taxa between distinct protocols was determined using the R DESEQ2 statistical software [34] within the phyloseq package [35]. Additional analyses, such as Wilcoxon-Mann-Whitney non-parametrical tests for groups of samples were conducted using R version 3.5.1 [35].

666

Additional files

Additional file 1: Table S1. Total count and percentage of bacterial and fungal families identified in negative controls during the DNA extraction process.

Additional file 2: Figure S1. Fungal and bacterial microbiomes of lung tissue samples cluster separately from those of negative controls. CCA plot of bacterial (**A**) and fungal (**B**) microbiomes according to sample type.

Additional file 3: Table S2. Total counts and percentages of bacterial and fungal families identified in lung tissue samples.

Additional file 4: Figure S2. Boxplots showing the Shannon diversity index for each extraction protocol. M1 to M5: Protocols 1 to 5.

Additional file 5: Table S3. Percentage of bacterial and fungal families identified per sample in lung tissue samples, grouped by individual and by extraction method.
Green color intensity is proportional to the relative abundance of a family in each sample.

List of abbreviations

CCA: Canonical Correspondence Analysis; *C*t: Cycle threshold; gDNA: genomic DNA; ITS: internal transcribed spacer; OTU: Operational Taxonomic Unit; PCR: polymerase chain reaction; SD: standard deviation.

Declarations

Ethics approval and consent to participate

The Ethics Committee of the Facultad de Medicina at the Universidad de Chile approved this study.

Consent for publication

Not applicable.

Availability of data and materials

The sequence data produced by 16S rRNA and fungal 18S rRNA sequencing are deposited in EBI Short Read Archive repository (https://www.ebi.ac.uk/ena) under the study accession number PRJEB31011 with accession numbers for bacteria from ERS3088332 to ERS3088370 and for fungi from ERS3088371 to ERS3088409.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants to AM from the Spanish Ministry of Economy and Competitiveness (projects SAF2012-31187, SAF2013-49788-EXP, SAF2015-65878-R), Carlos III Institute of Health (projects ELAC2014/HID-0254, PIE14/00045 and AC15/00022), Generalitat Valenciana (project PrometeoII/2014/065 and Prometeo/2018/A/133), Asociación Española contra el Cancer (project AECC 2017-1485)

and co-financed by the European Regional Development Fund (ERDF). This project has received funding from ERANet LAC (ELAC2014/HID-0254) lead by SLV. This project was also funded by grant from the National Fund for Science and Technology (Fondecyt, Chile) Grant number 1140412 to SLV and Grant number 1191311 to FM.

Author's contributions

VPB, FM and, SRR contributed equally to this work. VSM and MG recruited autopsied subjects and collected the lung samples. FM, CP, SRR, and RB conducted the experimental work. VPB and FM performed the bioinformatic and statistical analyses. VPB and FM wrote the first draft of the manuscript. All authors read and approved the manuscript.

Acknowledgments

We would like to thank the staff of the Genomics and Health Area, in particular to Nuria Jimenez for conducting the sequencing of the DNA, and Javier Pons and Alejandro Artacho for their advice in computational analyses. We thank Anne Lagomarcino for editing the language in this manuscript.

References

1. Marsland BJ, Yadava K, Nicod LP. The airway microbiome and disease. Chest 2013;144:632–7.
2. Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM, Louis P, Thomson JM; UK IBD Genetics Consortium, Satsangi J, Flint HJ, Parkhill J, Lees CW, Hold GL. The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. PLoS One 2014;9:e88982.
3. Wagner Mackenzie B, Waite DW, Taylor MW. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. Front Microbiol. 2015;6:130.
4. Lim MY, Song EJ, Kim SH, Lee J, Nam YD. Comparison of DNA extraction methods for human gut microbial community profiling. Syst Appl Microbiol. 2018;41:151–7.
5. Trung NT, Hien TT, Huyen TT, Quyen DT, Van Son T, Hoan PQ, Phuong NT, Lien TT, Binh MT, Van Tong H, Meyer CG, Velavan TP, Song le H. Enrichment of bacterial DNA for the diagnosis of blood stream infections. BMC Infect Dis. 2016;16:235.
6. Knapp M, Hofreiter M. Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives. Genes (Basel) 2010;1:227–43.
7. Thomsen PF, Willerslev E. Environmental DNA-an Emerging Tool in Conservation for Monitoring Past and Present Biodiversity. Biol Conserv. 2015;183:4–18.

8. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014;12;12:87.
9. Gefrides LA, Powell MC, Donley MA, Kahn R. UV irradiation and autoclave treatment for elimination of contaminating DNA from laboratory consumables. Forensic Sci Int Genet. 2010;4:89–94.
10. Goodrich JK, Di Rienzi SC, Poole AC, Koren O, Walters WA, Caporaso JG, Knight R, Ley RE. Conducting a microbiome study. Cell 2014;158:250–62.
11. Mohammadi T, Reesink HW, Vandenbroucke-Grauls CM, Savelkoul PH. Removal of contaminating DNA from commercial nucleic acid extraction kit reagents. J Microbiol Methods 2005;61:285–8.
12. Bhatt AS, Freeman SS, Herrera AF, Pedamallu CS, Gevers D, Duke F, Jung J, Michaud M, Walker BJ, Young S, Earl AM, Kostic AD, Ojesina AI, Hasserjian R, Ballen KK, Chen YB, Hobbs G, Antin JH, Soiffer RJ, Baden LR, Garrett WS, Hornick JL, Marty FM, Meyerson M. Sequence-based discovery of Bradyrhizobium enterica in cord colitis syndrome. N Engl J Med. 2013;369:517–28.
13. Xuan C, Shamonki JM, Chung A, Dinome ML, Chung M, Sieling PA, Lee DJ. Microbial dysbiosis is associated with human breast cancer. PLoS One 2014;9:e83744.
14. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, Davies J, Ervine A, Poulter L, Pachter L, Moffatt MF, Cookson WO. Disordered microbial communities in asthmatic airways. PLoS One 2010;5:e8578.

15. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med. 2011;184:957–63.
16. Goddard AF, Staudinger BJ, Dowd SE, Joshi-Datar A, Wolcott RD, Aitken ML, Fligner CL, Singh PK. Direct sampling of cystic fibrosis lungs indicates that DNA-based analyses of upper-airway specimens can misrepresent lung microbiota. Proc Natl Acad Sci USA 2012;109:13769–74.
17. Tipton L, Ghedin E, Morris A. The lung mycobiome in the next-generation sequencing era. Virulence 2017;8:334–41.
18. Denning DW, Chakrabarti A. Pulmonary and sinus fungal diseases in non-immunocompromised patients. Lancet Infect Dis. 2017;17:e357-e366.
19. Mathieu E, Escribano-Vazquez U, Descamps D, Cherbuy C, Langella P, Riffault S, Remot A, Thomas M. Paradigms of Lung Microbiota Functions in Health and Disease, Particularly, in Asthma. Front Physiol. 2018;9:1168.
20. Lauder AP, Roche AM, Sherrill-Mix S, Bailey A, Laughlin AL, Bittinger K, Leite R, Elovitz MA, Parry S, Bushman FD. Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. Microbiome 2016;4:29.
21. Zoetendal EG, Ben-Amor K, Akkermans AD, Abee T, de Vos WM. DNA isolation protocols affect the detection limit of PCR approaches of bacteria in samples from the human gastrointestinal tract. Syst Appl Microbiol. 2001;24:405–10.
22. Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS One 2012;7: e33865.

797 23. Fredricks DN, Smith C, Meier A. Comparison of six DNA extraction methods for
798 recovery of fungal DNA as assessed by quantitative PCR. J Clin Microbiol.
799 2005;43:5122–8.
800 24. Dupuy AK, David MS, Li L, Heider TN, Peterson JD, Montano EA, Dongari-
801 Bagtzoglou A, Diaz PI, Strausbaugh LD. Redefining the human oral mycobiome
802 with improved practices in amplicon-based taxonomy: discovery of Malassezia as a
803 prominent commensal. PLoS One 2014;9:e90899.
804 25. Kliman HJ. Comment on "the placenta harbors a unique microbiome". Sci Transl
805 Med. 2014;6:254le4.
806 26. Vargas SL, Ponce CA, Gallo M, Pérez F, Astorga JF, Bustamante R, Chabé M,
807 Durand-Joly I, Iturra P, Miller RF, Aliouat el M, Dei-Cas E. Near-universal
808 prevalence of Pneumocystis and associated Increase in mucus in the lungs of Infants
809 with sudden unexpected death. Clin Infect Dis. 2013;56:171–9.
810 27. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh
811 PJ, Fierer N, Knight R. Global patterns of 16S rRNA diversity at a depth of millions
812 of sequences per sample. Proc Natl Acad Sci USA 2011;108 Suppl 1:4516–22.
813 28. Borneman J, Hartin RJ. PCR primers that amplify fungal rRNA genes from
814 environmental samples. Appl Environ Microbiol. 2000;66: 4356–60.
815 29. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA., Holmes SP.
816 DADA2: high-resolution sample inference from Illumina amplicon data. Nat.
817 Methods 2016;13 581–3.
818 30. Langmead B, and Salzberg, SL. Fast gapped-read alignment with Bowtie 2. Nat
819 Methods 2012;9:357–9.

31. 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, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;2:335–6.
32. R Development Core Team, R: A language and environment for statistical computing. 2013, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
33. McArdle BH, Anderson MJ. Fitting multivariate models to community data: a comment based on distance-based redundancy analysis. Ecology 2001;82, 290–297.
34. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
35. McMurdie PJ, Holmes S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013;8:e61217.

835

836 **Table 1.** DNA extraction protocols used in this study.

837

Protocol	Pre-treatment step	Phenol:Chloroform:Isoamyl alcohol step	Bead-beating step	QIAamp DNA Mini kit (Qiagen)
1	-	-	-	+
2	-	-	+	+
3	-	+	-	+
4	-	+	+	+
5	+	+	+	+

838

839

840

Figure legends

Figure 1: Comparison of the yield and purity of extracted DNA between protocols. (**A**) DNA yield and (**B**) DNA quality (n = 7 samples/protocol). DNA yield is expressed as DNA concentration (ng/µl) normalized by quantity of lung tissue used for DNA extraction. Solid black lines indicate the median, and the lower and upper bounds of the box represent the 25 and 75% quartiles. Outliers, defined as falling outside the 10% and 90% quartiles, are indicated with black circles. Significant differences specified in the figure are based on ANOVA or the Kruskal-Wallis tests.

Figure 2: Evaluation of human, bacterial and fungal DNA by extraction protocol. Boxplot of Ct values for β-actin gene (**A**), 16S rRNA (**B**) and 18S rRNA gene (**C**) obtained with the five extraction protocols. Solid black lines indicate medians, and the lower and upper bounds of the box represent the 25 and 75% quartiles. Outliers, defined as falling outside the 10% and 90% quartiles, are indicated with black circles. Significant differences based on the Kruskal-Wallis test are specified in the figure.

Figure 3: Relative abundance of phyla and families identified in negative controls during the DNA extraction process. Sequencing of the 16S rRNA gene (**A**) and the ITS region (**B**) carried out on four negative controls using the Illumina MiSeq platform. A complete list of taxa is provided in the **Additional file 1: Table S1**.

Figure 4: Relative abundance of phyla and families identified in lung tissue samples. Sequencing of the 16S rRNA gene (**A**) and the ITS region (**B**) carried out on 7 lung tissue

samples using the Illumina MiSeq platform. A complete list of taxa is provided in **Additional file 3–Table S2**.

Figure 5: Canonical correspondence analysis (CCA) plots of bacterial and fungal microbiomes according to the DNA extraction protocol (A and C respectively) and according to individuals (B and D respectively).

Figure 6: Bacterial changes associated with bead beating (comparison of protocol 2 with 1); the Phenol:Chloroform:Isoamyl alcohol step (comparison of protocol 3 with 1); bead-beating and the Phenol:Chloroform:Isoamyl alcohol steps (comparison of protocol 4 with 1); pre-treatment steps (comparison of protocol 5 with 4); and pre-treatment, the bead-beating and the Phenol:Chloroform:Isoamyl alcohol steps (comparison of protocol 5 with 1). Average of $\log_2$ fold-cahnge in relative taxon abundance was calculated using the DESEQ2 software.

Figure 7: Fungal changes associated with bead beating (comparison of protocol 2 with 1); the Phenol:Chloroform:Isoamyl alcohol step (comparison of protocol 3 with 1); the bead-beating and the Phenol:Chloroform:Isoamyl alcohol steps (comparison of protocol 4 with 1), the pre-treatment steps (comparison of protocol 5 with 4); and the pre-treatment, the bead-beating and the Phenol:Chloroform:Isoamyl alcohol steps (comparison of protocol 5 with 1). Average of $\log_2$ fold-change in relative taxon abundance was calculated using the DESEQ2 software.

Figure 1

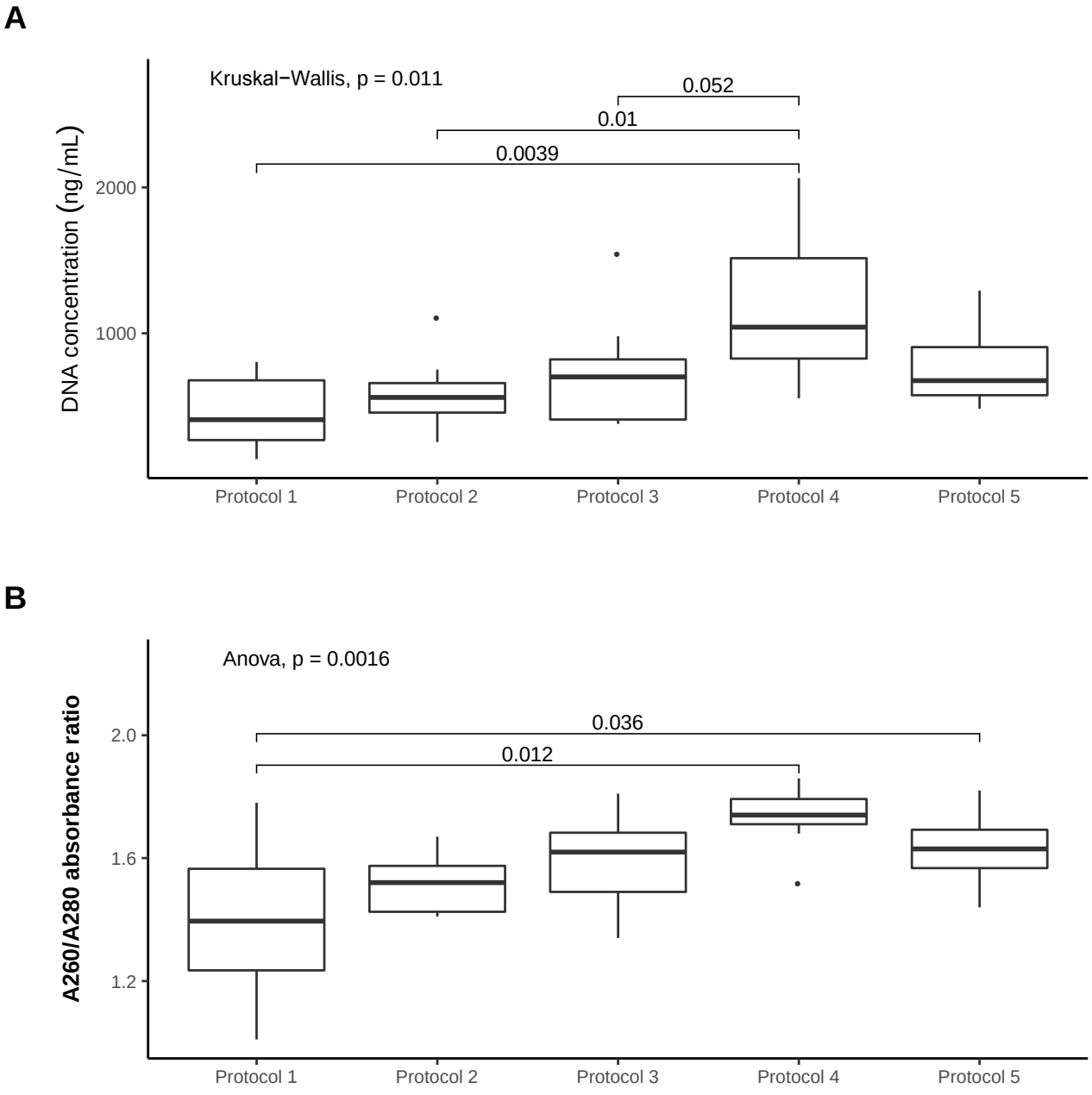


Figure 2

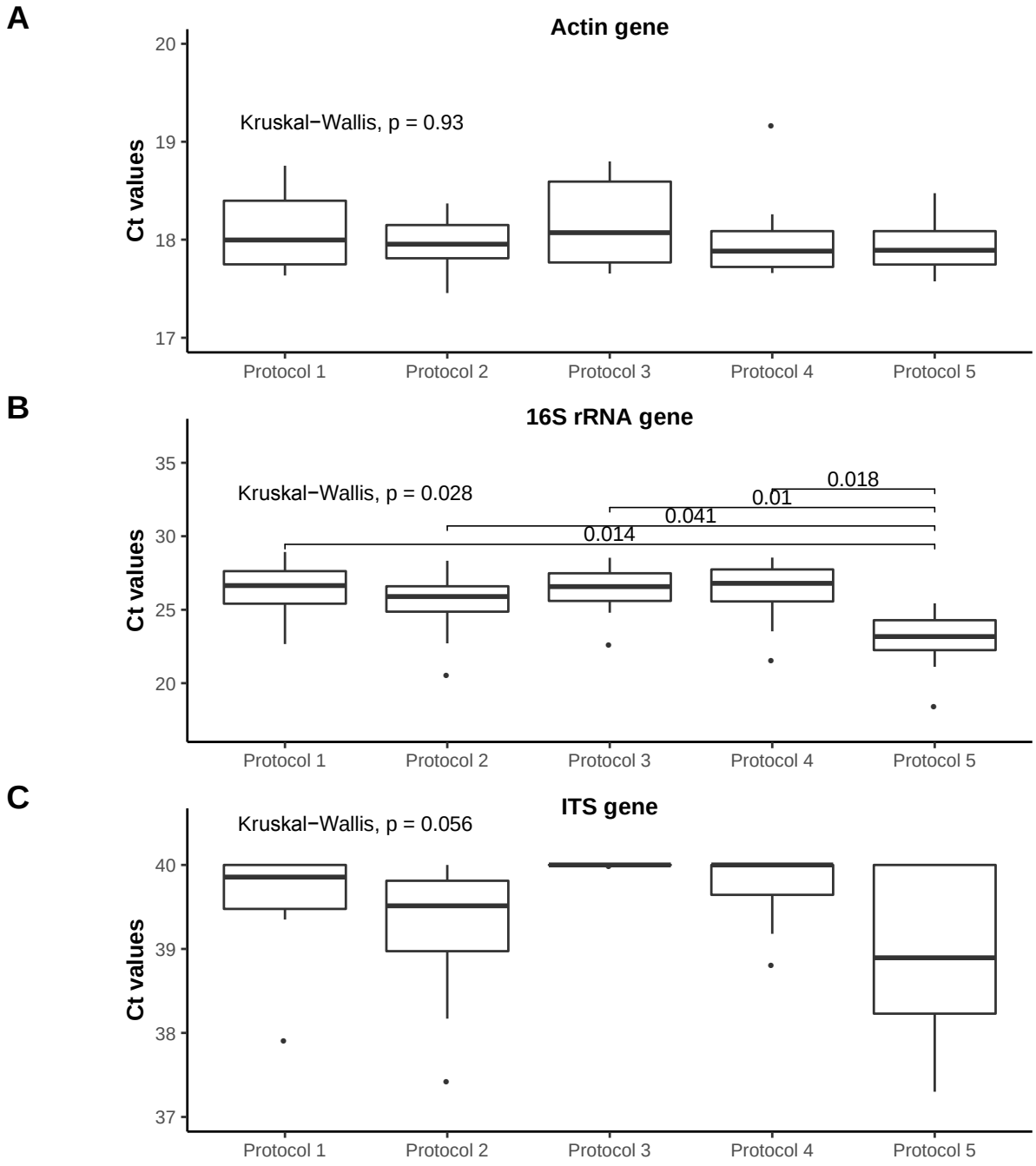


Figure 3

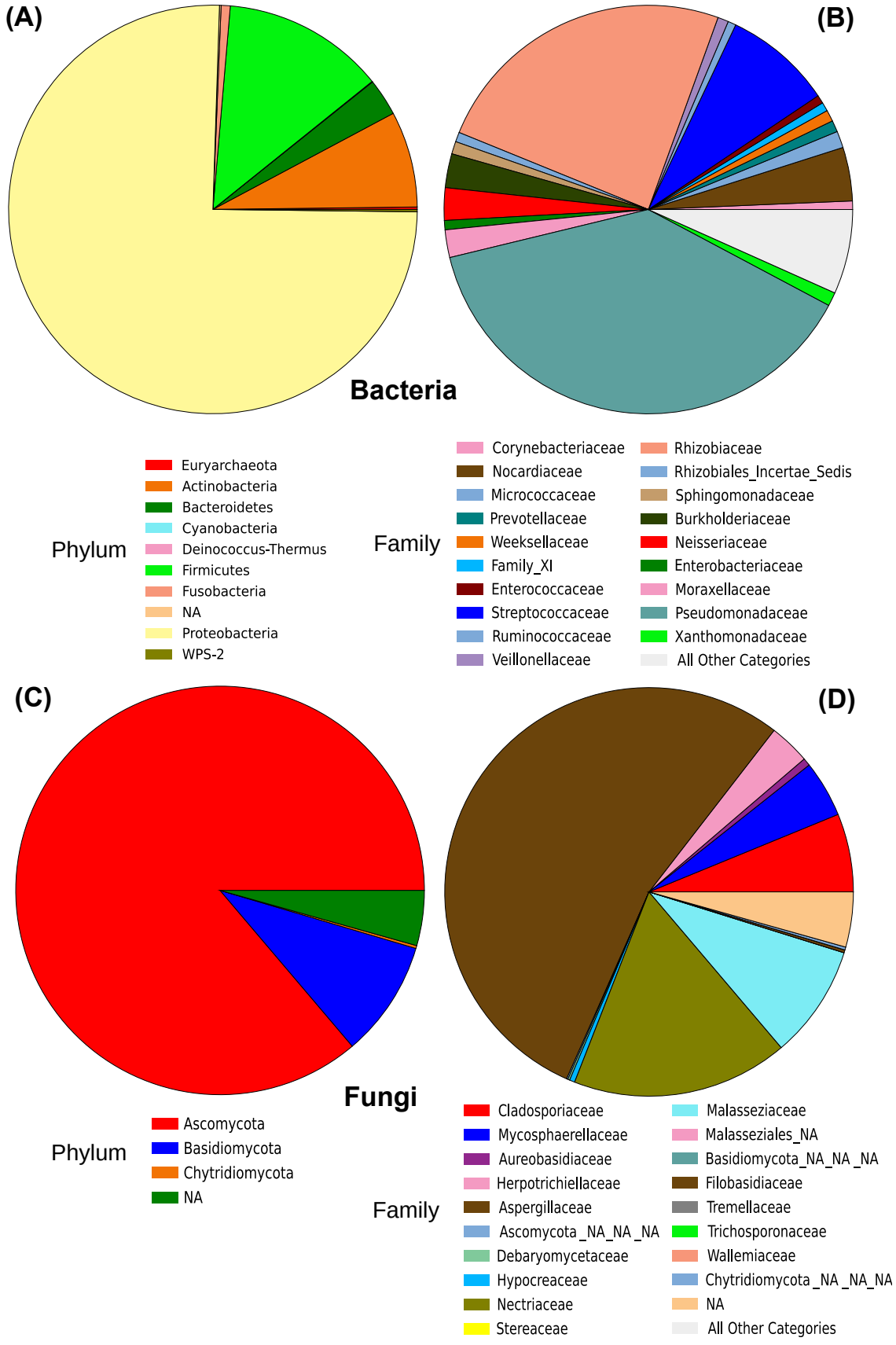


Figure 4

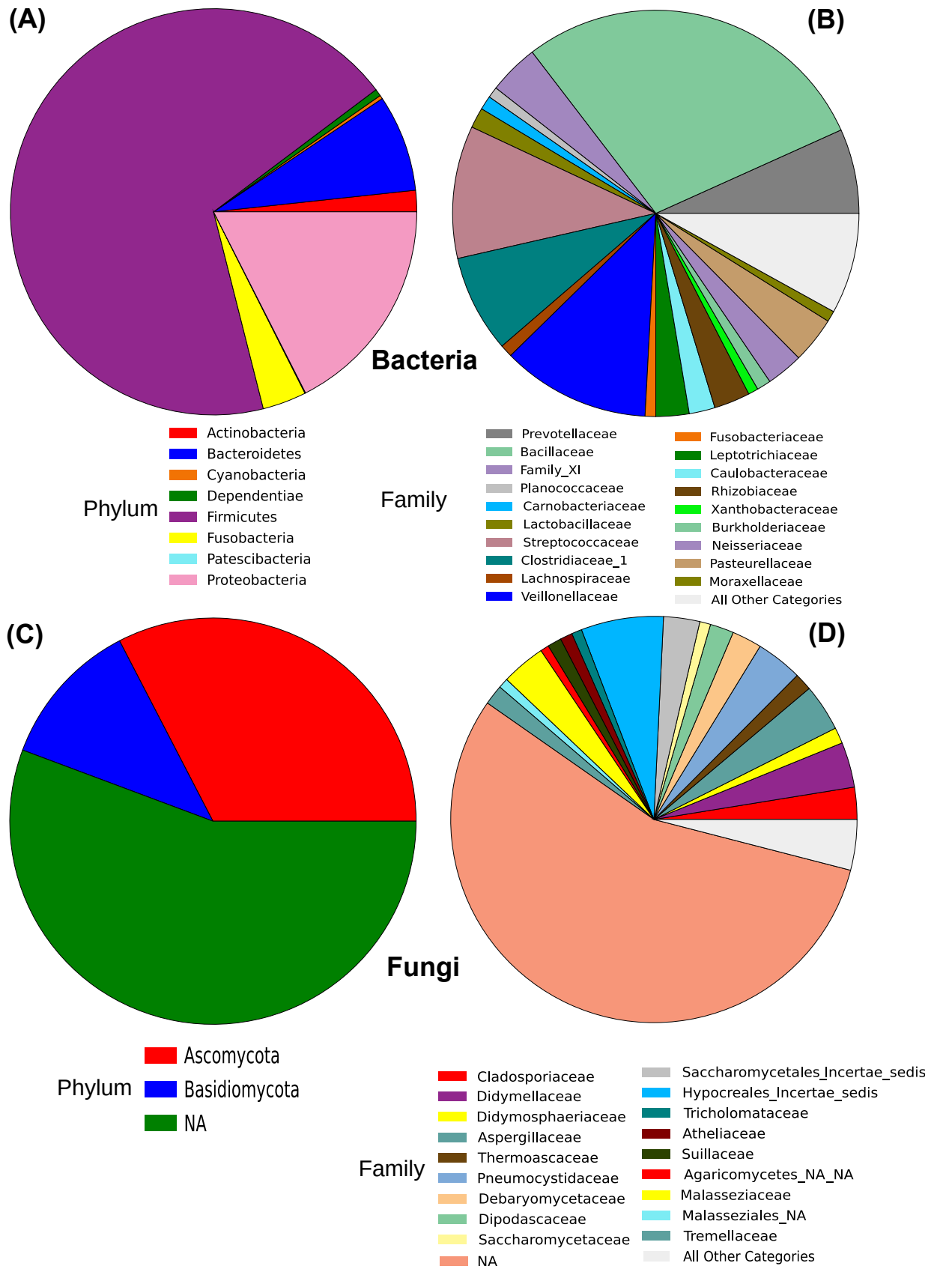


Figure 5

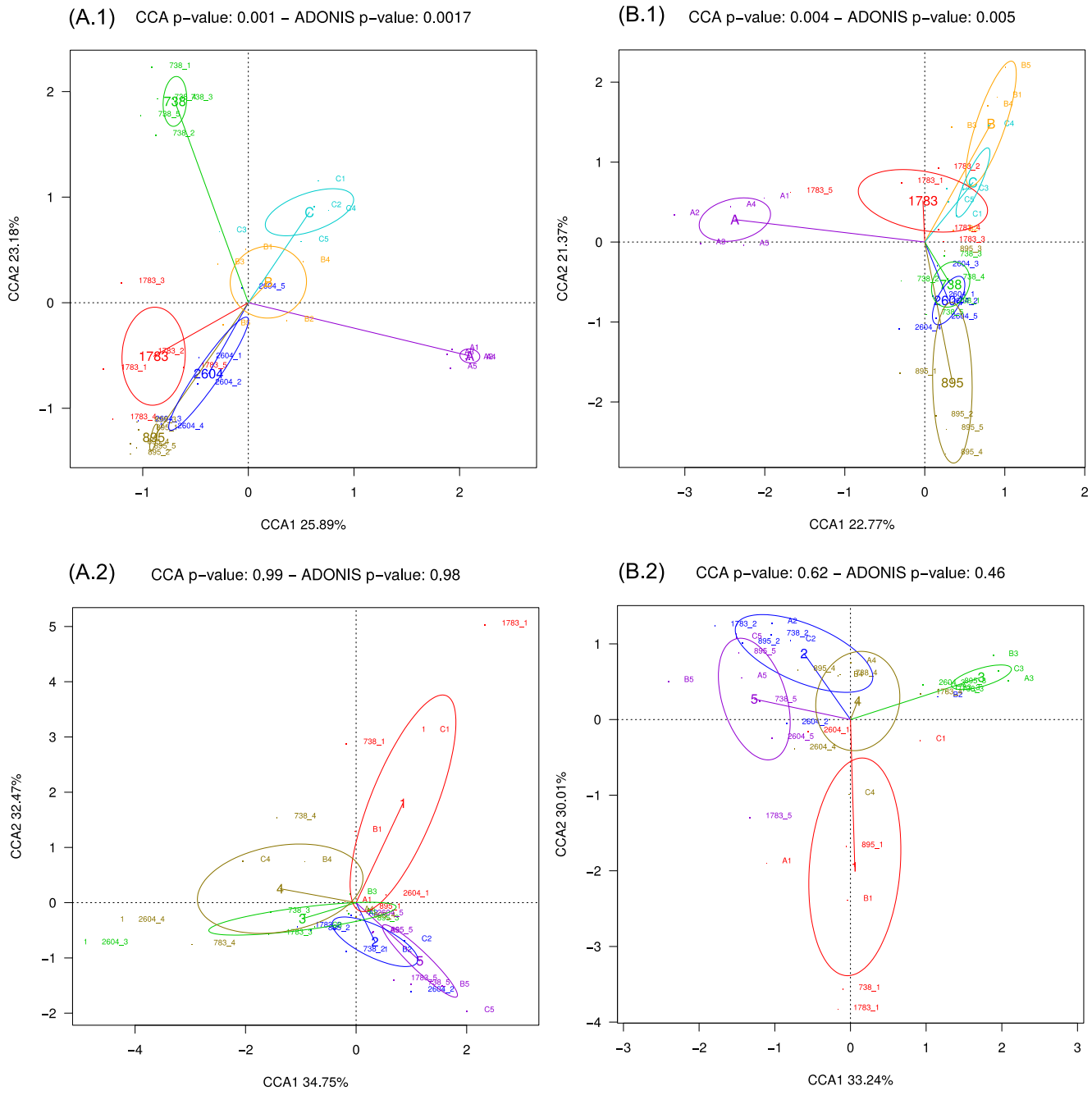


Figure 6

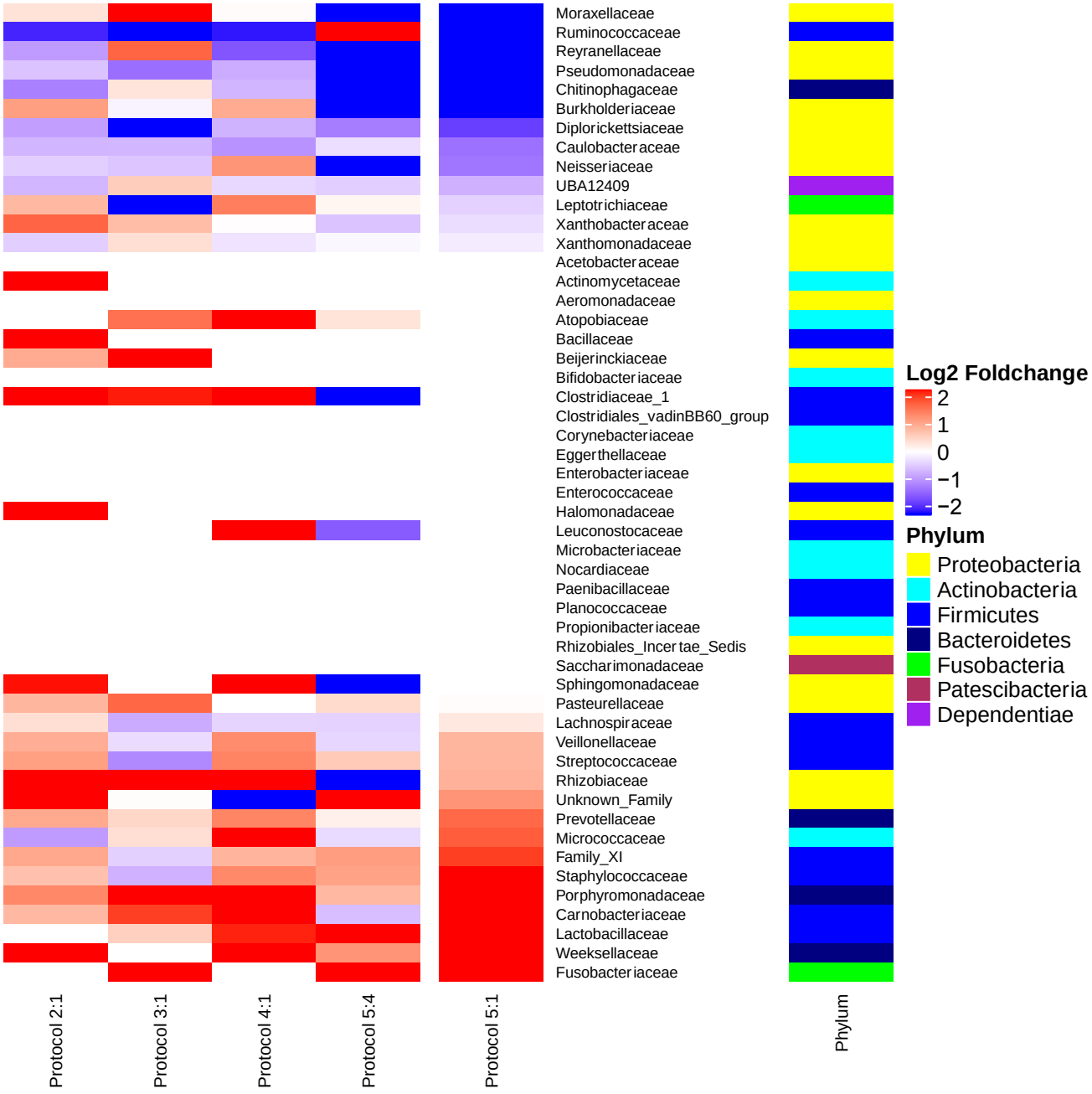
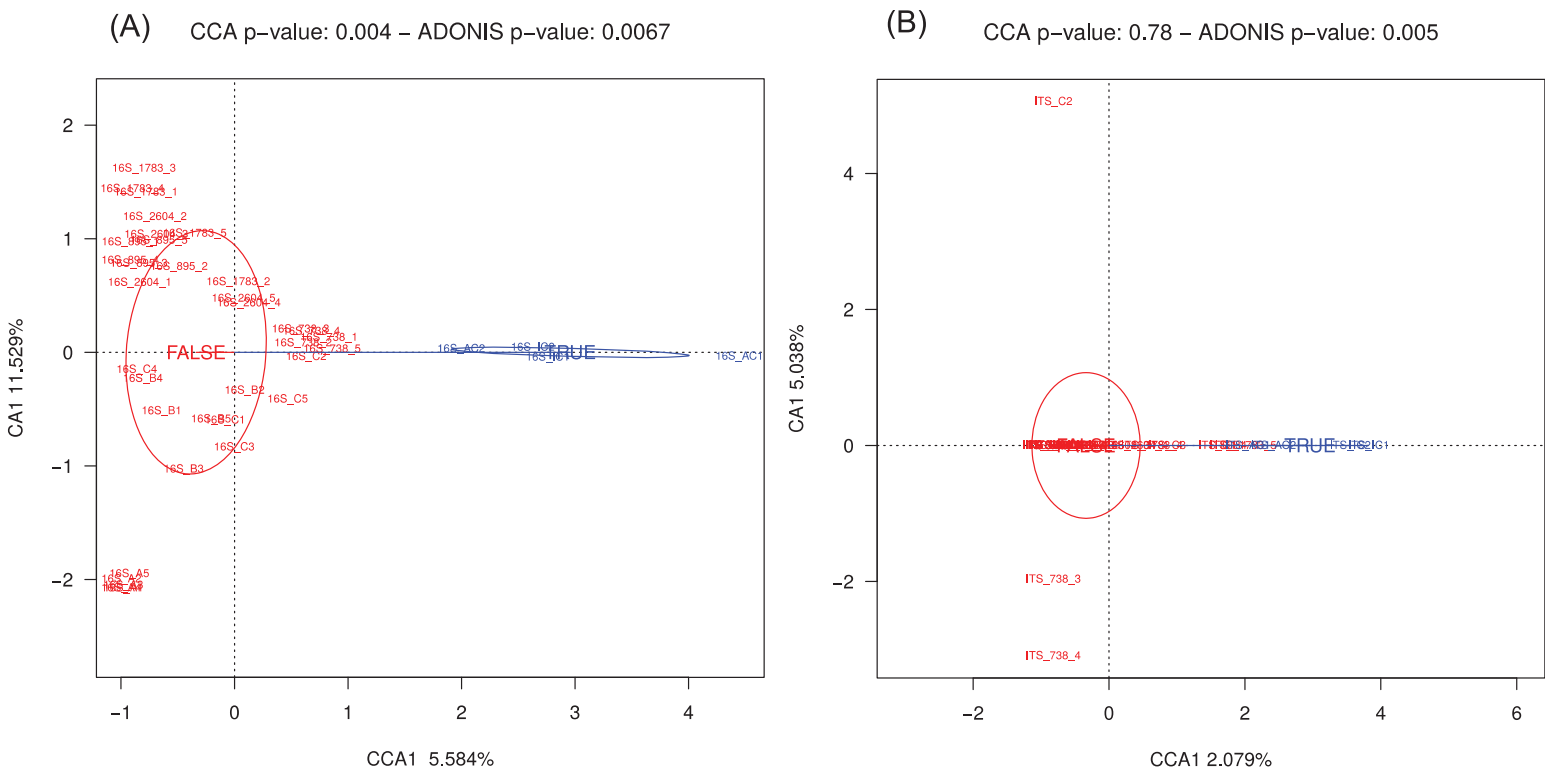
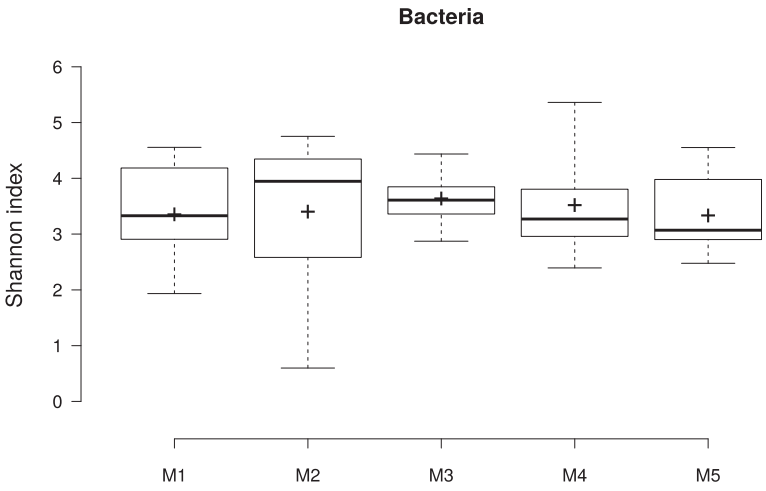


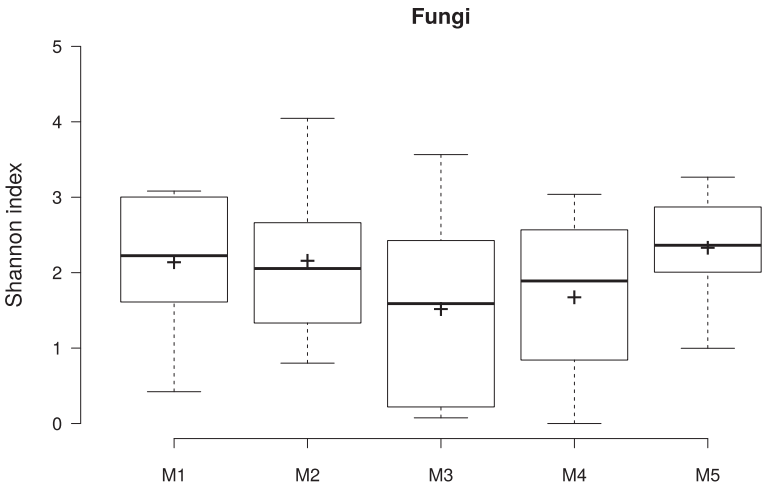
Figure 7

Additional file 2: Figure S1



Additional file 4: Figure S2.





Group1	Group2	G1 mean	G1 std	G2 mean	G2 std	t stat	p-value
M1	M2	3.356	0.694	3.404	1.289	-0.08	0.931
M1	M3	3.356	0.694	3.64	0.348	-0.896	0.402
M1	M4	3.356	0.694	3.522	0.753	-0.396	0.729
M1	M5	3.356	0.694	3.336	0.561	0.055	0.959
M2	M3	3.404	1.289	3.64	0.348	-0.434	0.745
M2	M4	3.404	1.289	3.522	0.753	-0.194	0.888
M2	M5	3.404	1.289	3.336	0.561	0.118	0.911
M3	M4	3.64	0.348	3.522	0.753	0.35	0.741
M3	M5	3.64	0.348	3.336	0.561	1.129	0.283
M4	M5	3.522	0.753	3.336	0.561	0.485	0.642

Group1	Group2	G1 mean	G1 std	G2 mean	G2 std	t stat	p-value
M1	M2	2.14	0.821	2.158	0.936	-0.037	0.964
M1	M3	2.14	0.821	1.519	1.199	1.047	0.336
M1	M4	2.14	0.821	1.675	1.011	0.875	0.392
M1	M5	2.14	0.821	2.331	0.669	-0.443	0.644
M2	M3	2.158	0.936	1.519	1.199	1.03	0.333
M2	M4	2.158	0.936	1.675	1.011	0.86	0.407
M2	M5	2.158	0.936	2.331	0.669	-0.369	0.716
M3	M4	1.519	1.199	1.675	1.199	-0.244	0.8
M3	M5	1.519	1.199	2.331	0.669	-1.45	0.179
M4	M5	1.675	1.011	2.331	0.669	-1.327	0.242

Additional file 1: Table S1

Bacteria

Family	counts	%
p__Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae	44	0,19
p__Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae	20	0,09
p__Actinobacteria;c__Actinobacteria;o__Corynebacteriales;f__Corynebacteriaceae	153	0,67
p__Actinobacteria;c__Actinobacteria;o__Corynebacteriales;f__Dietziaceae	3	0,01
p__Actinobacteria;c__Actinobacteria;o__Corynebacteriales;f__Nocardiaceae	966	4,25
p__Actinobacteria;c__Actinobacteria;o__Frankiales;f__Geodermatophilaceae	18	0,08
p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Microbacteriaceae	146	0,64
p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Micrococcaceae	300	1,32
p__Actinobacteria;c__Actinobacteria;o__Propionibacteriales;f__Propionibacteriaceae	122	0,54
p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae	2	0,01
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae	14	0,06
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Dysgonomonadaceae	14	0,06
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Paludibacteraceae	9	0,04
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae	14	0,06
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae	207	0,91
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Tannerellaceae	48	0,21
p__Bacteroidetes;c__Bacteroidia;o__Chitinophagales;f__Chitinophagaceae	120	0,53
p__Bacteroidetes;c__Bacteroidia;o__Cytophagales;f__Hymenobacteraceae	36	0,16
p__Bacteroidetes;c__Bacteroidia;o__Flavobacteriales;f__Weeksellaceae	206	0,91
p__Cyanobacteria;c__Melainabacteria;o__Obscuribacterales;f__NA	6	0,03
p__Deinococcus-Thermus;c__Deinococci;o__Deinococcales;f__Deinococcaceae	4	0,02
p__Firmicutes;c__Bacilli;o__Bacillales;f__Family_XI	158	0,69
p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae	27	0,12
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Carnobacteriaceae	53	0,23
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae	149	0,65
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae	1.949	8,56
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae_1	21	0,09
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Family_XI	118	0,52
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae	34	0,15
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae	34	0,15
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae	146	0,64
p__Firmicutes;c__Erysipelotrichia;o__Erysipelotrichales;f__Erysipelotrichaceae	45	0,20
p__Firmicutes;c__Negativicutes;o__Selenomonadales;f__Veillonellaceae	194	0,85
p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Fusobacteriaceae	113	0,50
p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Leptotrichiaceae	50	0,22
p__NA;c__NA;o__NA;f__NA	29	0,13
p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae	93	0,41
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae	7	0,03
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Devosiaceae	20	0,09
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae	54	0,24
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae	5.549	24,38
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiales_Incertae_Sedis	174	0,76
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Xanthobacteraceae	83	0,36
p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae	30	0,13
p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae	226	0,99
p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae	611	2,69
p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Neisseriaceae	585	2,57
p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae	167	0,73
p__Proteobacteria;c__Gammaproteobacteria;o__Gammaproteobacteria_Incertae_Sedis;f__Unknown_Family	15	0,07
p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae	29	0,13
p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae	498	2,19
p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae	8.749	38,45
p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Rhodanobacteraceae	6	0,03
p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae	246	1,08
p__WPS-2;c__NA;o__NA;f__NA	42	0,18
Total	22.756	100

Fungi

Family	counts	%
p__Ascomy	4.523	6,19
p__Ascomy	3.270	4,48
p__Ascomy	443	0,61
p__Ascomy	2.367	3,24
p__Ascomy	39.349	53,88
p__Ascomy	78	0,11
p__Ascomy	109	0,15
p__Ascomy	304	0,42
p__Ascomy	12.517	17,14
p__Basidior	6.562	8,98
p__Basidior	166	0,23
p__Chytridi	166	0,23
p__NA;c__N	3.183	4,36
Total	73.037	100

Additional file 3: Table S2.

Bacteria

Family	counts	%
p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae	1.363	0,65
p__Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae	63	0,03
p__Actinobacteria;c__Actinobacteria;o__Corynebacteriales;f__Corynebacteriaceae	0	0,00
p__Actinobacteria;c__Actinobacteria;o__Corynebacteriales;f__Nocardiaceae	240	0,11
p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Microbacteriaceae	0	0,00
p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Micrococcaceae	1.525	0,73
p__Actinobacteria;c__Actinobacteria;o__Propionibacteriales;f__Propionibacteriaceae	0	0,00
p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Atopobiaceae	218	0,10
p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Eggerthellaceae	149	0,07
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae	1.254	0,60
p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae	14.112	6,73
p__Bacteroidetes;c__Bacteroidia;o__Chitinophagales;f__Chitinophagaceae	413	0,20
p__Bacteroidetes;c__Bacteroidia;o__Flavobacteriales;f__Weeksellaceae	388	0,19
p__Cyanobacteria;c__Melainabacteria;o__Obscuribacterales;f__NA	601	0,29
p__Dependentiae;c__Babeliae;o__Babeliales;f__UBA12409	1.181	0,56
p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae	60.086	28,68
p__Firmicutes;c__Bacilli;o__Bacillales;f__Family_XI	8.391	4,00
p__Firmicutes;c__Bacilli;o__Bacillales;f__Paenibacillaceae	225	0,11
p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae	1.829	0,87
p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae	1.260	0,60
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Carnobacteriaceae	2.326	1,11
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae	113	0,05
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae	3.393	1,62
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae	129	0,06
p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae	22.115	10,55
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae_1	16.267	7,76
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiales_vadinBB60_group	140	0,07
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Family_XI	0	0,00
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae	2.095	1,00
p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae	760	0,36
p__Firmicutes;c__Negativicutes;o__Selenomonadales;f__Veillonellaceae	24.770	11,82
p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Fusobacteriaceae	1.768	0,84
p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Leptotrichiaceae	5.537	2,64
p__Patescibacteria;c__Saccharimonadia;o__Saccharimonadales;f__Saccharimonadaceae	112	0,05
p__Proteobacteria;c__Alphaproteobacteria;o__Acetobacterales;f__Acetobacteraceae	166	0,08
p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae	4.312	2,06
p__Proteobacteria;c__Alphaproteobacteria;o__Reyranellales;f__Reyranellaceae	138	0,07
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae	637	0,30
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae	6.020	2,87
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiales_Incertae_Sedis	0	0,00
p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Xanthobacteraceae	1.670	0,80
p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__NA	195	0,09
p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae	1.550	0,74
p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Aeromonadaceae	215	0,10
p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae	2.346	1,12
p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Neisseriaceae	6.130	2,93
p__Proteobacteria;c__Gammaproteobacteria;o__Diplorickettsiales;f__Diplorickettsiaceae	109	0,05
p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae	0	0,00
p__Proteobacteria;c__Gammaproteobacteria;o__Gammaproteobacteria_Incertae_Sedis;f__Unknown_Family	734	0,35
p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales;f__Halomonadaceae	1.041	0,50
p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae	7.726	3,69
p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae	1.812	0,86
p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae	1.280	0,61
p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae	629	0,30
Total	209.333	100

Fungi

Family	counts	%
p__Ascomy	45.438	2,56
p__Ascomy	3.395	0,19
p__Ascomy	1.963	0,11
p__Ascomy	63.873	3,59
p__Ascomy	22.573	1,27
p__Ascomy	5.898	0,33
p__Ascomy	66.311	3,73
p__Ascomy	24.760	1,39
p__Ascomy	1.220	0,07
p__Ascomy	158	0,01
p__Ascomy	2.714	0,15
p__Ascomy	65.623	3,69
p__Ascomy	42.260	2,38
p__Ascomy	33.431	1,88
p__Ascomy	5.843	0,33
p__Ascomy	15.314	0,86
p__Ascomy	51.193	2,88
p__Ascomy	116.573	6,56
p__Ascomy	6.988	0,39
p__Ascomy	3.870	0,22
p__Basidion	1.402	0,08
p__Basidion	2.753	0,15
p__Basidion	14.489	0,82
p__Basidion	17.353	0,98
p__Basidion	19.688	1,11
p__Basidion	5.310	0,30
p__Basidion	12.110	0,68
p__Basidion	3.445	0,19
p__Basidion	2.234	0,13
p__Basidion	9.790	0,55
p__Basidion	668	0,04
p__Basidion	62.585	3,52
p__Basidion	14.087	0,79
p__Basidion	9.066	0,51
p__Basidion	28.307	1,59
p__Basidion	4.477	0,25
p__Basidion	222	0,01
p__NA;c__N	989.587	55,69
Total	1.776.971	100

Additional file 5: Table S3

Bacteria	Grouped per subject																																
Family	1783					2604					738					895					A					B					C				
	1783_1	1783_2	1783_3	1783_4	1783_5	2604_1	2604_2	2604_3	2604_4	2604_5	738_1	738_2	738_3	738_4	738_5	895_1	895_2	895_3	895_4	895_5	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	C1	C2	C3	C4	C5
Actinomycetaceae	0	0,07	0	0	0	0	2,4	0	0	0	0	0	0	0	0	0	1,77	0	1,36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bifidobacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6,99	0	0	0	0	0	0	0
Corynebacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nocardiaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4,92	0	0	0	0	0	1,65	7,49	0	0
Microbacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Micrococcaceae	2,31	0	6	0	2,5	3,37	0,09	0,02	25,51	7,54	0	0	0	0	5,16	0	0,74	0	0	0,69	1,59	0,6	3,01	1,75	13,67	1,23	0	0	0	0,71	0	0	0	0	0
Propionibacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Atopobiaceae	0	0	0	0	0	0	0	0	1,26	0	0	0	0	0	0	0	0	0,13	0,42	1,36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Eggerthellaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21,38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Porphyromonadaceae	0	0	0	0	0	0	1,1	0,05	7,81	2,79	0	0	4,83	0	0	2,03	2,02	2,21	0	5,02	0	0	0	0	0	0	0	0	0	0	0	12,21	0,35	1,32
Prevotellaceae	0	2,46	0	2,95	1,2	5,99	4,34	1,65	8,67	10,61	0	0	0	0	6,17	27,02	27,25	24,34	22,11	23,8	0	0,25	0	0	0	0	0	4,77	0	0	0	0	0	2,17
Chitinophagaceae	0	0	7,93	2,36	0	0	0	0	0	0	2,19	7,85	3,07	1,07	0	0	0,74	0	0	0	0	0	0,21	0	0	5,37	0	0	0,32	0	0	0	2,03	0
Weeksellaceae	0	0,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,13	0	0	0	0	0	0	0	0	0,22	0	0	4,47	0	0	0,6
NA	0	0	4,16	3,2	0	2,24	0	0	0,43	0	0	0	4,29	0	2,58	0	0,18	0	0	0,33	0,29	0	0,63	0	0	7,79	2,59	4,11	0,9	0	0	0,87	0,98	2,73	0
UBA12409	2,25	0	9,39	2,53	0	1,62	0	0	0,09	8,1	9	0	6,21	17,14	7,89	0,13	0,14	0,27	0	0,31	0,9	0	0,95	0,09	0	2,78	1,47	4,77	0,96	1,01	4,27	7,61	9,03	2,55	0
Bacillaceae	0	30,8	0	0	0	0	53,04	69,54	0	0	0	0	0	0	0	0	0	0	13,12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Family_XI	0	3,48	2,93	0	11,45	27,93	5,05	0	7,68	0	0	0	0	0	4,02	3,19	3,05	0	7,05	10,09	1,3	0,98	1,19	0,75	0	1	11,31	21,86	2,48	33,84	0	0	0,57	0,99	10,38
Paenibacillaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9,31	0	0	0
Planococcaceae	0	0	0	3,12	0	0	0	5,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcaceae	0	0	0	0	1,46	0	0	0	0	0	0	0	0	0	0	0	0	0,71	0,72	1,11	8,32	4,69	4,38	4,09	4,16	1,73	3,63	0	0,32	0	0	0,74	0	2,67	0
Carnobacteriaceae	5,38	0	0	0	5,41	0	0,5	1,86	2,17	0	0	0	0	0	0	1,08	0	0	0	0	0	1,2	1,61	2,26	1,86	0	0	0	12,14	23,51	0	0	0	0	4,52
Enterococcaceae	7,69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lactobacillaceae	0	0	0,69	0,34	5,88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5,88	0	0	15	1,49	10,66	41,01	48,49
Leuconostocaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1,74	0	0	0,23	2,35
Streptococcaceae	0	0,02	0	6,66	9,21	13,34	6,66	0	3,12	26,26	0	0	5,37	0	0	4,3	3,9	4,91	2,28	3,06	80,93	90,9	82,59	89,68	79,43	14,44	29,36	5,33	33,65	6,18	15,52	33,43	5,13	19,26	17,61
Clostridiaceae_1	6,94	31,79	24,79	36,39	18,69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,35	1,75	0,16	0	0	0	2,22	0	0	0	0	0	0	0
Clostridiales_vadinBB60_group	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12,21	0	0	0	0
Family_XI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lachnospiraceae	0	0	0,69	0	0	6,61	0	0	0,78	0	0	0	9,97	0	0	6,33	8,89	2,79	3,21	6,9	0,43	0	0	0	0	0	0	0	0	0	0	0	0	0	0,4
Ruminococcaceae	10,62	0	0	0	0	0	1,29	0	1,34	0	13,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,31
Veillonellaceae	20,22	15,21	9,16	26,2	33,52	19,33	0,5	0,04	17,05	6,42	0	0	0	11,43	0	29,68	33,91	38,07	28,45	28,21	0,9	0,57	0,7	0,51	0,28	14,66	0	9,66	21,26	13,37	0	0	6,43	3,42	9,73
Fusobacteriaceae	0	0,33	5,77	0	2,39	0	0	4,53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1,62	0	0	0	0	0
Leptotrichiaceae	0	0	0	0	0,26	2,49	5,14	0,02	4,95	4,19	0	0	0	0	0	8,52	5,99	9,43	7,49	11,42	0,4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Saccharimonadaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4,63	0	0	0
Acetobacteraceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Caulobacteraceae	8,03	0	15,09	7,58	2,13	12,72	0	0	1,39	17,04	53,73	88,74	25,46	20,54	30,42	4,07	2,34	5,58	0	0	2,24	0,44	0,53	0,37	0	15,66	18,48	11,21	3,6	3,24	17,18	15,43	14,73	12,47	0,92
Reyranellaceae	0,34	0	0	0	0	0	0	0	0	0	0	0	5,52	0	0	0	0,28	0	0	0	0,33	0	0,28	0	0	0	0	0	0,26	0	0	0,08	1,71	0	0
Beijerinckiaceae	0	0	0	0	0	0	0	0	0	0	0	0,68	0	0	0	0	0	0,58	1,85	0	0	0	0	0	0	0	0	0,55	0	0	0	0	2,03	0	0
Rhizobiaceae	0	4,6	0	0	0	0	4,35	2,55	0	0	2,06	0	11,73	16,07	4,3	0	0	0	3,43	0	0	0	0	0	0	0,55	4,75	0	0,42	0	0	0,62	0	0	0
Rhizobiales_Incertae_Sedis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xanthobacteraceae	0	0	0	0,51	1,93	0	3,09	0	0	2,51	3,47	0	8,21	6,61	0	0,98	0,67	0	0	0	0,58	0	0,18	0	0	0,36	7,68	0	0,7	0	0	1,28	6,84	1,68	0
NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4,16	0	0,28	1,46	0	0	0	0	0	0	0	1,5	0	0	0,3	0	0	2,07	0	1,57	0
Sphingomonadaceae	0	1,53	0	0	0	4,36	0,86	0	0	0	0	0	0	0	0	0	0	0	1,5	0	0	0	0	0	0	0	0	0	0,22	0	0	0	0	0	0
Aeromonadaceae	0	0,46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Burkholderiaceae	0	0,83	0,92	1,77	0,83	0	1,98	0	0	0	9,38	0	0	0	0	0,69	0	0	2,5	0	0	0	0	0	0	2,09	3,11	5,11	0,92	0	0	3,1	2,77	1,57	0
Neisseriaceae	31,86	1,65	0,54	0	0	0	2,44	7,01	12,67	0	0	2,73	3,68	19,64	2,01	0	1,59	0	0	0,82	0	0	0	0	0	7,1	1,99	4	16,88	12,56	0	0	0	0	0
Diplorickettsiaceae	0	0	0	1,68	0	0	0	0	0	0	0	0	0	0	0	0,77	0	0	0	0	0,14	0	0	0	0	0,59	2,42	0	0,14	0,71	0	0	0	0	0
Enterobacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Unknown_Family	1,91	1,19	1,77	0	0	0	0	0	0	4,47	0	0	2,22	0	5,74	0	0	0	0	0	0,11	0	0	0	0	0	0	0	0	0	0	1,45	0	0	0
Halomonadaceae	0	1,2	0	0	0	0	1,35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pasteurellaceae	2,11	3,28	2	1,26	2,19	0	5,01	5,08	4,34	0	0	0	0,46	0	6,17	11,2	6,24	7,08	2,48	6,72	0	0	0,6	0,33	0,59	8,74	3,02	3,11	1,64	3,24	0	0	19,45	0	0,36
Moraxellaceae	0	0,61	8,16	0	0	0	0,81	2,15	0,74	0	0	0	1,15	0	0	0	0	0	0,76	0	0	0	0,21	0	0	2,19	0	0	0	0	6,89	0	0	1,22	0
Pseudomonadaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,74	0	1,34	0	1,19	0	0	5,46	10,19	7,33	1,86	0	23,37	6,74	0	6,26	0,85
Xanthomonadaceae	0,34	0	0	3,45	0,94	0	0	0	0	10,06	6,68	0	7,82	7,5	0	0	0	2,43	0	0	0,18	0	0	0	0	1,82	0	3,11	0,8	0	3,84	5,01	0	0	0

Fungi	Grouped per subject																																		
Family	1783					2604					738					895					A					B					C					
	1783_1	1783_2	1783_3	1783_4	1783_5	2604_1	2604_2	2604_3	2604_4	2604_5	738_1	738_2	738_3	738_4	738_5	895_1	895_2	895_3	895_4	895_5	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	C1	C2	C3	C4	C5
Cladosporiaceae	0	3,11	0	0	0	39,72	0	0	30,27	19,8	0	0	0	0	0	0	0	0	0	0	0	1,59	0	11,11	3,33	0	0	14,27	0	0	0	0	35,6	0	0
Mycosphaerellaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1,02	44,4	0	0	0	0	0	0	0	0	0	0	0	0
o__Capnodiales;f__NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27,78	0	0	0	0	0	0	0	0	0	0	0	0
Didymellaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	42,17	0	71,99	84,25	0	0	0	0	43,1	0	0	0	0	0	0	0	0	0	0
Didymosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	57,84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichomeriaceae	93,44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Aspergillaceae	0	0	0	0	0	0	35,05	0	18,81	0	0	59,88	0	0	0	41,34	0	0	0	15,75	0	0	27,82	0	18,78	0	0	0	0	0	0	0	0	0	25,78
Thermoascaceae	0	0	0	0	0	0	0	13,45	0	0	0	0	0	0	0	2,34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sclerotiniaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4,84	0	0	0	0	0	0	0	0	0	0	0	0	0
Rhytismataceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,63	0	0	0	0	0	0	0	0	0	0	0	0	0
p__Ascomycota;c__NA;o__NA;f__NA	0	0	0	0	0	0	0	0	0	0	1,79	0	0	0	0	0	0	0	0	0	0	2,84	0	0	0	0	0	0	0	0	0	0	0	21,79	0
Pneumocystidaceae	0	0	0	0	0	53,15	56,66	4,08	50,92	80,2	0	0	0	0	62,42	19,77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Debaryomycetaceae	0	0	0	0	80,93	0	0	0	0	0	0	0	0	0	0	0	0	0	6,05	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dipodascaceae	0	0	0	0	0	0	0	0	0	0	94,13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,02	0	0	0	0
Pichiaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18,88	0	0	0
Saccharomycetaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10,59	0	0	0	0,28	
Saccharomycetales_fam_Incertae_sedis	0	0	0	100	19,07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53,91	73,94	
Hypocreales_fam_Incertae_sedis	0	13,51	9,73	0	0	0	0	5,36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20,3	0	22,4	0	0	0	0
Nectriaceae	6,56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10,55	0	36,77	0	0	0	0	0	0	0	0	0	0	0
c__Sordariomycetes;o__NA;f__NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2,7	0	0	0	0
Crepidotaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5,56	0	0	0	0	0	0	0	0	0	0	0	0	0
Stephanosporaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10,91	0	0	0	0	0	0	0	0	0	0	0	0	0
Tricholomataceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2,69	0	45,84	0	49,26	0	0	0	0	0	0	0	0	0
Atheliaceae	0	11,4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Suillaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	81,91	0	0	0	0	0	0	0	0
Schizoporaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21,05	0	0	0	0	0	0	0	0	0	0	0	0	0
c__Agaricomycetes;o__NA;f__NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14,34	0	0	0	0	0	29,9	0	0	0	0	0	0	0
Meruliaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12,13	0	0	0	0	0	0	0
Steccherinaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8,85	0	0	0	0	0	0	0	0	0	0	0	0	0
Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	34,47	0	0	0	0	0	0	0
Stereaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6,28	0	0	0	0	0	0	0	0	0	0	0
Malasseziaceae	0	0	0	0	0	0	8,3	12,2	0	0	0	0	0	100	37,58	0,5	0	0	21,96	0	0	14,26	0	0	34,79	0	18,09	9,23	0	0	0	0	0	0	0
o__Malasseziales;f__NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45,52	0	0	0
Filobasidiaceae	0	0	0	0	0	0	0	0	0	0	0	0	2,65	0	0	36,06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tremellaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50,74	0	0	0	0	99,98	0	0	0	0
Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
Wallemiaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,88	0	0	0	0	0	0	0	0	0	0	0	0	0
k__Fungi;p__NA;c__NA;o__NA;f__NA	0	71,98	90,27	0	0	7,13	0	64,92	0	0	4,08	40,12	97,35	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	79,7	0	64,31	0	0	24,3	0

Bacteria	Grouped per method																																		
Family	M1							M2							M3							M4							M5							
	1783_1	2604_1	738_1	895_1	A1	B1	C1	1783_2	2604_2	738_2	895_2	A2	B2	C2	1783_3	2604_3	738_3	895_3	A3	B3	C3	1783_4	2604_4	738_4	895_4	A4	B4	C4	1783_5	2604_5	738_5	895_5	A5	B5	C5
Actinomycetaceae	0	0	0	0	0	0	0	0,07	2,4	0	1,77	0	0	0	0	0	0	0	0	0	0	0	0	0	1,36	0	0	0	0	0	0	0	0	0	0
Bifidobacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6,99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Corynebacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nocardiaceae	0	0	0	0	0	4,92	0	0	0	0	0	0	0	1,65	0	0	0	0	0	0	7,49	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Microbacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Micrococcaceae	2,31	3,37	0	0	1,59	1,23	0	0	0,09	0	0,74	0,6	0	0	6	0,02	0	0	3,01	0	0	0	25,51	0	0	1,75	0	0	2,5	7,54	5,16	0,69	13,67	0,71	0
Propionibacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Atopobiaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,13	0	0	0	0	1,26	0	0,42	0	0	0	0	0	0	1,36	0	0	0
Eggerthellaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21,38	0	0	0	0
Porphyromonadaceae	0	0	0	2,03	0	0	0	0	1,1	0	2,02	0	0	0	0	0,05	4,83	2,21	0	0	12,21	0	7,81	0	0	0	0	0,35	0	2,79	0	5,02	0	0	1,32
Prevotellaceae	0	5,99	0	27,02	0	0	0	2,46	4,34	0	27,25	0,25	0	0	0	1,65	0	24,34	0	4,77	0	2,95	8,67	0	22,11	0	0	0	1,2	10,61	6,17	23,8	0	0	2,17
Chitinophagaceae	0	0	2,19	0	0	5,37	0	0	0	7,85	0,74	0	0	0	7,93	0	3,07	0	0,21	0	0	2,36	0	1,07	0	0	0,32	2,03	0	0	0	0	0	0	0
Weeksellaceae	0	0	0	0	0	0	0	0,5	0	0	0	0	0	4,47	0	0	0	0	0	0	0	0	0	0	0	0	0,22	0	0	0	0	0,13	0	0	0,6
NA	0	2,24	0	0	0,29	7,79	0	0	0	0	0,18	0	2,59	0,87	4,16	0	4,29	0	0,63	4,11	0,98	3,2	0,43	0	0	0	0,9	2,73	0	0	2,58	0,33	0	0	0
UBA12409	2,25	1,62	9	0,13	0,9	2,78	4,27	0	0	0	0,14	0	1,47	7,61	9,39	0	6,21	0,27	0,95	4,77	9,03	2,53	0,09	17,14	0	0,09	0,96	2,55	0	8,1	7,89	0,31	0	1,01	0
Bacillaceae	0	0	0	0	0	0	0	30,8	53,04	0	0	0	0	0	0	69,54	0	0	0	0	0	0	0	0	13,12	0	0	0	0	0	0	0	0	0	0
Family_XI	0	27,93	0	3,19	1,3	1	0	3,48	5,05	0	3,05	0,98	11,31	0	2,93	0	0	0	1,19	21,86	0,57	0	7,68	0	7,05	0,75	2,48	0,99	11,45	0	4,02	10,09	0	33,84	10,38
Paenibacillaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	9,31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Planococcaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5,5	0	0	0	0	0	3,12	0	0	0	0	0	0	0	0	0	0	0	0	0
Staphylococcaceae	0	0	0	0	8,32	1,73	0	0	0	0	0	4,69	3,63	0,74	0	0	0	0,71	4,38	0	0	0	0	0	0,72	4,09	0,32	2,67	1,46	0	0	1,11	4,16	0	0
Carnobacteriaceae	5,38	0	0	1,08	0	0	0	0	0,5	0	0	1,2	0	0	0	1,86	0	0	1,61	0	0	0	2,17	0	0	2,26	12,14	0	5,41	0	0	0	1,86	23,51	4,52
Enterococcaceae	7,69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lactobacillaceae	0	0	0	0	0	0	15	0	0	0	0	0	0	1,49	0,69	0	0	0	0	5,88	10,66	0,34	0	0	0	0	0	41,01	5,88	0	0	0	0	0	48,49
Leuconostocaceae	0	0	0	0	0	0	1,74	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,23	0	0	0	0	0	0	0	2,35
Streptococcaceae	0	13,34	0	4,3	80,93	14,44	15,52	0,02	6,66	0	3,9	90,9	29,36	33,43	0	0	5,37	4,91	82,59	5,33	5,13	6,66	3,12	0	2,28	89,68	33,65	19,26	9,21	26,26	0	3,06	79,43	6,18	17,61
Clostridiaceae_1	6,94	0	0	0	0	0	0	31,79	0	0	0	0,35	0	0	24,79	0	0	0	1,75	2,22	0	36,39	0	0	0	0,16	0	0	18,69	0	0	0	0	0	0
Clostridiales_vadinBB60_group	0	0	0	0	0	0	12,21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Family_XI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lachnospiraceae	0	6,61	0	6,33	0,43	0	0	0	0	0	8,89	0	0	0	0,69	0	9,97	2,79	0	0	0	0	0,78	0	3,21	0	0	0	0	0	0	6,9	0	0	0,4
Ruminococcaceae	10,62	0	13,5	0	0	0	0	0	1,29	0	0	0	0	0	0	0	0	0	0	0	0	0	1,34	0	0	0	0	0	0	0	0	0	0	0	0,31
Veillonellaceae	20,22	19,33	0	29,68	0,9	14,66	0	15,21	0,5	0	33,91	0,57	0	0	9,16	0,04	0	38,07	0,7	9,66	6,43	26,2	17,05	11,43	28,45	0,51	21,26	3,42	33,52	6,42	0	28,21	0,28	13,37	9,73
Fusobacteriaceae	0	0	0	0	0	0	0	0,33	0	0	0	0	0	0	5,77	4,53	0	0	0	0	0	0	0	0	0	0	0	0	2,39	0	0	0	0	1,62	0
Leptotrichiaceae	0	2,49	0	8,52	0,4	0	0	0	5,14	0	5,99	0	0	0	0	0,02	0	9,43	0	0	0	0	4,95	0	7,49	0	0	0	0,26	4,19	0	11,42	0	0	0
Saccharimonadaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	4,63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Acetobacteraceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,52	0	0	0	0	0	0	0	0
Caulobacteraceae	8,03	12,72	53,73	4,07	2,24	15,66	17,18	0	0	88,74	2,34	0,44	18,48	15,43	15,09	0	25,46	5,58	0,53	11,21	14,73	7,58	1,39	20,54	0	0,37	3,6	12,47	2,13	17,04	30,42	0	0	3,24	0,92
Reyranellaceae	0,34	0	0	0	0,33	0	0	0	0	0	0,28	0	0	0,08	0	0	5,52	0	0,28	0	1,71	0	0	0	0	0	0,26	0	0	0	0	0	0	0	0
Beijerinckiaceae	0	0	0	0	0	0	0	0	0	0,68	0	0	0	0	0	0	0	0,58	0	0,55	2,03	0	0	0	1,85	0	0	0	0	0	0	0	0	0	0
Rhizobiaceae	0	0	2,06	0	0	0,55	0	4,6	4,35	0	0	0	4,75	0,62	0	2,55	11,73	0	0	0	0	0	0	16,07	3,43	0	0,42	0	0	0	4,3	0	0	0	0
Rhizobiales_Incertae_Sedis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xanthobacteraceae	0	0	3,47	0,98	0,58	0,36	0	0	3,09	0	0,67	0	7,68	1,28	0	0	8,21	0	0,18	0	6,84	0,51	0	6,61	0	0	0,7	1,68	1,93	2,51	0	0	0	0	0
NA	0	0	0	0	0	1,5	0	0	0	0	0,28	0	0	2,07	0	0	0	1,46	0	0	0	0	0	0	0	0	0,3	1,57	0	0	4,16	0	0	0	0
Sphingomonadaceae	0	4,36	0	0	0	0	0	1,53	0,86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1,5	0	0,22	0	0	0	0	0	0	0	0
Aeromonadaceae	0	0	0	0	0	0	0	0,46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Burkholderiaceae	0	0	9,38	0,69	0	2,09	0	0,83	1,98	0	0	0	3,11	3,1	0,92	0	0	0	0	5,11	2,77	1,77	0	0	2,5	0	0,92	1,57	0,83	0	0	0	0	0	0
Neisseriaceae	31,86	0	0	0	0	7,1	0	1,65	2,44	2,73	1,59	0	1,99	0	0,54	7,01	3,68	0	0	4	0	0	12,67	19,64	0	0	16,88	0	0	0	2,01	0,82	0	12,56	0
Diplorickettsiaceae	0	0	0	0,77	0,14	0,59	0	0	0	0	0	0	2,42	0	0	0	0	0	0	0	0	1,68	0	0	0	0	0,14	0	0	0	0	0	0	0,71	0
Enterobacteriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Unknown_Family	1,91	0	0	0	0,11	0	0	1,19	0	0	0	0	0	1,45	1,77	0	2,22	0	0	0	0	0	0	0	0	0	0	0	0	4,47	5,74	0	0	0	0
Halomonadaceae	0	0	0	0	0	0	0	1,2	1,35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pasteurellaceae	2,11	0	0	11,2	0	8,74	0	3,28	5,01	0	6,24	0	3,02	0	2	5,08	0,46	7,08	0,6	3,11	19,45	1,26	4,34	0	2,48	0,33	1,64	0	2,19	0	6,17	6,72	0,59	3,24	0,36
Moraxellaceae	0	0	0	0	0	2,19	6,89	0,61	0,81	0	0	0	0	0	8,16	2,15	1,15	0	0,21	0	0	0	0,74	0	0,76	0	0	1,22	0	0	0	0	0	0	0
Pseudomonadaceae	0	0	0	0	1,34	5,46	23,37	0	0	0	0	0	10,19	6,74	0	0	0	0	1,19	7,33	0	0	0	0	0,74	0	1,86	6,26	0	0	0	0	0	0	0,85
Xanthomonadaceae	0,34	0	6,68	0	0,18	1,82	3,84	0	0	0	0	0	0	5,01	0	0	7,82	2,43	0	3,11	0	3,45	0	7,5	0	0	0,8	0	0,94	10,06	0	0	0	0	0

Fungi

Family	Grouped per method																														
	M1							M2							M3							M4							M5						
	1783_1	2604_1	738_1	895_1	A1	B1	C1	1783_2	2604_2	738_2	895_2	A2	B2	C2	1783_3	2604_3	738_3	895_3	A3	B3	C3	1783_4	2604_4	738_4	895_4	A4	B4	C4	1783_5	2604_5	738_5	895_5	A5	B5	C5
Cladosporiaceae	0	39,72	0	0	0	0	0	3,11	0	0	0	1,59	0	0	0	0	0	0	0	14,27	35,6	0	30,27	0	0	11,11	0	0	0	19,8	0	0	3,33	0	0
Mycosphaerellaceae	0	0	0	0	0	0	0	0	0	0	0	1,02	0	0	0	0	0	0	44,4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
o__Capnodiales;f__NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27,78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Didymellaceae	0	0	0	0	0	0	0	0	0	0	42,17	0	0	0	0	0	0	0	0	0	0	0	0	0	71,99	0	0	0	0	0	0	84,25	43,1	0	0
Didymosphaeriaceae	0	0	0	0	0	0	0	0	0	0	57,84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichomeriaceae	93,44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Aspergillaceae	0	0	0	41,34	0	0	0	0	35,05	59,88	0	0	0	0	0	0	0	0	27,82	0	0	0	18,81	0	0	0	0	0	0	0	0	15,75	18,78	0	25,78
Thermoascaceae	0	0	0	2,34	0	0	0	0	0	0	0	0	0	0	0	13,45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sclerotiniaceae	0	0	0	0	0	0	0	0	0	0	0	4,84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rhytismataceae	0	0	0	0	0	0	0	0	0	0	0	0,63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p__Ascomycota;c__NA;o__NA;f__NA	0	0	1,79	0	0	0	0	0	0	0	0	2,84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21,79	0	0	0	0	0	0	0
Pneumocystidaceae	0	53,15	0	19,77	0	0	0	0	56,66	0	0	0	0	0	0	4,08	0	0	0	0	0	0	50,92	0	0	0	0	0	0	80,2	62,42	0	0	0	0
Debaryomycetaceae	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6,05	0	0	0	80,93	0	0	0	0	0	0
Dipodascaceae	0	0	94,13	0	0	0	0	0	0	0	0	0	0	0,02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pichiaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18,88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Saccharomycetaceae	0	0	0	0	0	0	10,59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,28
Saccharomycetales_fam_Incertae_sedis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	53,91	19,07	0	0	0	0	0	73,94
Hypocreales_fam_Incertae_sedis	0	0	0	0	0	0	22,4	13,51	0	0	0	0	0	0	9,73	5,36	0	0	0	0	0	0	0	0	0	0	20,3	0	0	0	0	0	0	0	0
Nectriaceae	6,56	0	0	0	0	0	0	0	0	0	0	10,55	0	0	0	0	0	0	0	0	0	0	0	0	0	36,77	0	0	0	0	0	0	0	0	0
c__Sordariomycetes;o__NA;f__NA	0	0	0	0	0	0	2,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Crepidotaceae	0	0	0	0	0	0	0	0	0	0	0	5,56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Stephanosporaceae	0	0	0	0	0	0	0	0	0	0	0	10,91	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tricholomataceae	0	0	0	0	0	49,26	0	0	0	0	0	2,69	0	0	0	0	0	0	0	0	0	0	0	0	0	45,84	0	0	0	0	0	0	0	0	0
Atheliaceae	0	0	0	0	0	0	0	11,4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Suillaceae	0	0	0	0	0	0	0	0	0	0	0	0	81,91	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Schizoporaceae	0	0	0	0	0	0	0	0	0	0	0	21,05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
c__Agaricomycetes;o__NA;f__NA	0	0	0	0	0	0	0	0	0	0	0	14,34	0	0	0	0	0	0	0	29,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Meruliaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12,13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Steccherinaceae	0	0	0	0	0	0	0	0	0	0	0	8,85	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	34,47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Stereaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6,28	0	0	0	0	0	0	0	0	0
Malasseziaceae	0	0	0	0,5	0	0	0	0	8,3	0	0	14,26	18,09	0	0	12,2	0	0	0	9,23	0	0	0	100	21,96	0	0	0	0	0	37,58	0	34,79	0	0
o__Malasseziales;f__NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45,52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Filobasidiaceae	0	0	0	36,06	0	0	0	0	0	0	0	0	0	0	0	0	2,65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tremellaceae	0	0	0	0	0	50,74	0	0	0	0	0	0	0	99,98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
Wallemiaceae	0	0	0	0	0	0	0	0	0	0	0	0,88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
k__Fungi;p__NA;c__NA;o__NA;f__NA	0	7,13	4,08	0	0	0	64,31	71,98	0	40,12	0	0	0	0	90,27	64,92	97,35	100	0	0	0	0	0	0	0	0	79,7	24,3	0	0	0	0	0	0	0