

## Identification Of The Microbiome Present In Sputum Samples From Patients With Clinical Tuberculosis By NGS (Next Generation Sequencing)

Ada Otilia Del Carpio-Sanz\*<sup>1#</sup>, Carmen Rosa Revilla-Mogrovejo<sup>1</sup>, Luis Fernando Luna-Paredes<sup>1</sup>, Karina Ttito-Velasquez<sup>1</sup>,<sup>a</sup> Rodrigo Alex Cáceres-Zegarra<sup>1</sup>, Tatiana Shessira Chávez-Arias<sup>1</sup>, Luis Carlos Lizárraga-Vargas<sup>1</sup>, Víctor Germán Torres-Díaz<sup>1,2</sup>, Hugo José Rojas-Flores<sup>1,3</sup>, Amalfi Rodolfo Gallegos-Morante<sup>3</sup>, Patricia Del Portillo-Obando<sup>4</sup>.

<sup>1</sup>-National University of San Agustín, Arequipa, Peru.

<sup>2</sup>-Hospital EsSalud III Yanahuara, Arequipa, Peru.

<sup>3</sup>-Regional Hospital III Honorio Delgado Espinoza, Arequipa, Peru.

<sup>4</sup>-Corporación Corpo Gen, Bogotá, Colombia.

<sup>#a</sup> Department of Human Morphology, School of Medicine, Universidad Nacional de San Agustín, Arequipa, Peru.

---

### Abstract

Tuberculosis is a reemerging disease that continues to be a public health problem worldwide with a high morbimortality rate. Infection with *M. tuberculosis* causes dysbiosis of the microbiota, condition that not only influences the latency of *M. tuberculosis* and the manifestation of the disease but also would be determinants of a greater progression of the disease and the failure of the recovery process. Another aspect to consider are the diagnostic methods which are varied and with limitations, however, thanks to the remarkable technological progress such as the New Generation Sequencing (NGS), there is currently access to the study of the microbiome that accompanies the human being in health conditions and associated with diseases such as tuberculosis. In the present study, the identification of the microbiome of samples from patients diagnosed with tuberculosis was performed, where 28 sputum samples were collected with clinical diagnosis of tuberculosis corresponding to: 20 without treatment, 6 with treatment, 2 Multidrug - Resistant and 8 samples from patients without tuberculosis (control). For the identification of *M. tuberculosis*, TB1, TB2 and TB3 primers were used. The molecular identification of the microbiome was carried out by Illumina MiSeq sequencing, using specific primers that target the V4 region of the 16S rRNA gene. According to the abundance of each study group, the following Phyla were presented: patients without tuberculosis (control) Firmicutes, Proteobacteria, Fusobacteria and Actinobacteria; patients with diagnosis of tuberculosis without treatment Proteobacteria, Bacteroidetes, Firmicutes and Actinobacteria; patients with diagnosis of tuberculosis who received treatment Firmicutes, Proteobacteria, Bacteroidetes and Proteobacteria; and patients with diagnosis of tuberculosis Multidrug-resistant phyla Proteobacteria, Actinobacteria,

Bacteroidetes, Firmicutes and Actinobacteria. The study identified the diversity and abundance at phylum and genus level by Next Generation Sequencing in individuals diagnosed with tuberculosis in untreated, treated and multidrug-resistant patients, finding a marked variation in the microbiome present in sputum samples associated with tuberculosis.

---

## Introduction

The term microbiome emerged in the last decade with great relevance in the field of medicine, although the concept itself is as old as humanity. The microbiome is understood as the diversity and totality of genes belonging to microorganisms that inhabit an individual, forming an ecosystem with the host, without which life would not be possible [1].

In the human intestine, hundreds of bacterial strains and more than 9 million genes have been found [2], which communicate closely with the human immune system and exert a significant influence on homeostasis, establishing a symbiotic relationship with the host [3–5]. Likewise, the lower respiratory tract and the lungs were considered a sterile environment, however, it is now known that these organs can also be affected by the intestinal microbiome, through the gut-lung axis, which involves a dialogue between the gut and the lungs [6–8]. This involves a cross-dialogue through the passage of bacteria, bacterial endotoxins, hormones and cytokines into the bloodstream from the gut to the lung and back again [3,9]. According to Nancy Gupta et al. (2018), several investigations have conclusively demonstrated that a healthy lower respiratory tract has an abundant microbiota similar to the predominant phyla detected in the healthy gut, evidencing Firmicutes, Bacteroidetes and Proteobacteria [6]. A functional and stable microbiota is also described, including up to 314 different species belonging to Bacteroidetes (*Prevotella* and *Bacteroides*), Firmicutes (*Veillonella*, *Streptococcus* and *Staphylococcus*), and Proteobacteria (*Pseudomonas*, *Haemophilus*, *Moraxella*, *Neisseria* and *Acinetobacter*) [10].

Recent research in metagenomics revealed that the gut microbiome is not just a passive bystander, but integrates environmental factors with genetic and immune signals that actively impact multiple host functions, including circadian rhythmicity, nutritional responses, metabolism, and immunity [11,12]. The maintenance of this interaction, regulates the immune response by inducing protective responses to pathogens and the utilization of regulatory pathways involved in sustained tolerance to innocuous antigens; influencing susceptibility to many diseases and immune-mediated disorders [13,14]. Among many other factors, the use of antibiotics, anti-ulcer and other drugs is known to alter the intestinal flora, greatly affecting the human microbiome [15,16]. Consequently, a dysbiotic microbiome cannot perform the functions mentioned above, being associated with a large number of health problems of metabolic and immunological etiology, as well as susceptibility to the development of infectious diseases [17,18].

Tuberculosis is a communicable disease caused by the bacillus *Mycobacterium tuberculosis*, it is considered one of the 10 leading causes of death worldwide and the leading cause of death from a single infectious agent (above HIV/AIDS). It is a global health problem that has a high morbidity and mortality rate [19]. It sickens more than 10 million people each year and kills 10-20% of them, about a quarter of the population [20,21]. Approximately, one quarter of the world's population is

latently infected with *M. tuberculosis* [22,23]. Globally, new TB cases have been reduced by 15%, however, multidrug-resistant patients have increased, with Peru being among a list of 30 countries with a high burden of multidrug-resistant patients. According to the Peruvian Epidemiological Bulletin, it is mentioned that the burden of TB in 2018 was 31,668 cases, presenting a total of 27,575 new cases [24].

The presence of *M. tuberculosis* causes dysbiosis of the airway microbiota in patients infected with this bacterium, reporting the presence of opportunistic bacteria in sputum samples of patients with tuberculosis prior to treatment; this suggests that the pulmonary microbial environment in patients infected with tuberculosis may become more susceptible to colonization by foreign and opportunistic microorganisms [25]. This may contribute to the pathophysiological processes associated with the disease, i.e., susceptibility, progression, resistance to treatment regimens and chronicity of pulmonary disease, as suggested by the study conducted by Maede Nakhaee, where the imbalance of the pulmonary microbiota and the presence of some bacterial species in TB patients not only influence the latency of *M. tuberculosis* and the manifestation of the disease, but may also be determinants of a greater progression of the disease and the failure of the recovery process [8].

Distinct findings have also been reported in characterizing the microbial diversity associated with *M. tuberculosis* in which the microbiome is not as varied [6]. However, a common finding in these studies is that, despite little variation, the pulmonary microbiota in TB patients differs from that of healthy individuals [26]. It should be noted that few studies report results on the microbiome associated with TB infection [6,27].

With the advent of the new techniques of massive non-culture-dependent gene sequencing, developed from the Human Microbiome Project, it has become possible to determine the types of microorganisms that are present in the different organs of the human body, both diseased and healthy, as well as the role they play in the interaction with host [1,8,28]. In addition, they make it possible to read in parallel innumerable DNA sequences or fragments (whole metagenomes) [29], techniques that are very useful in the clinic, since they allow to massively obtain DNA sequence data in a short time, allowing to expand knowledge in order to establish the denomination of normal and healthy microbiome, as well as some factors that could disrupt the human microbiome [1]. Such analyses have become accessible due to the remarkable progress in next-generation sequencing (NGS) in revealing the importance of the microbiome for human health, having been used to characterize the microbiome of the intestine, vagina, respiratory tract and oral cavity among others [27].

It becomes necessary, to identify the microbiome present in sputum samples of patients with clinical diagnosis of pulmonary tuberculosis by NGS, considering the possibility that changes in the pulmonary microbiome may also influence the inflammatory process that determines the formation of granulomas in *M. tuberculosis* infection [9]. The study also considers the evolution of the disease, its susceptibility-resistance to treatment and prognosis of TB. Thus, highlighting its impact on the health status of patients; promoting that this is examined and quantified to improve

the diagnosis, prognosis, therapy and prevention of infectious diseases including tuberculosis [30,31].

Based on this background, the research aims to identify the microbiome present in sputum samples in patients with clinical pictures of tuberculosis using NEW GENERATION SEQUENCING (NGS).

### Material and Methods

The study was carried out on sputum samples from patients with clinical diagnosis of tuberculosis from the Hospital III Regional Honorio Delgado Espinoza and EsSalud Yanahuara Arequipa – Peru. The confirmation of the presence of *M. tuberculosis* was provided by the respective laboratories using conventional methods, staining and culture. There was a total of 28 samples (20 untreated, 6 treated and 2 TB - multidrug resistant), and 8 control samples. The samples were transferred in cold chain to the Human Morphology Research Laboratory of the Faculty of Medicine of the Universidad Nacional de San Agustín de Arequipa - Peru.

### DNA extraction and identification of *M. tuberculosis* for testing of each sample

Bacterial DNA extraction was performed using the Silica method with the Macherey - Nagel DNA Extraction kit. The DNA was extracted and purified through a silica column according to the manufacturer's recommendations. First, the DNA adheres to the silica of the column, followed by washes with buffers containing 70% ethanol; finally, the DNA eluted in BE buffer was stored at -20 °C for later use.

For bacterial DNA testing, the 16S rRNA gene was amplified by polymerase chain reaction (PCR). Then, the identification of *M. tuberculosis* was carried out using the primers: TB1, TB2 and TB3 (Table 2) of the IS6110 region, for which a reaction was prepared with 15 ul of Mix TAG Polymerase, 1ul of Forward and 1ul of Reverse, 2 ul of DNA from each sample and ultrapure water to a final volume of 30 ul. The amplification conditions for TB1, TB2 and TB3 are shown in Table 2. For visualization of extracted DNA and the different amplicons, electrophoresis was performed on a 2% agarose gel stained with Syber Safe, using a 100-base pair (bp) reference marker (BIOLINE) (Major Science MP-310).

**Table 1.** Characteristics of the primers and their conditions for PCR.

First	Primer Sequence	Size	Cycles
TB1	5' CAA GGC TTC AAT TCC TCC GGT GAT GCC 3'	285 bp	94°C - 1 min
	5'-TGGTCCGGTTTTTCATCATACTCGGGCTGG-3'		70°C - 1 min 35 72°C - 0.4 min
TB2	5' TGG TCC GGT TCA TAC TCG GGC TGG 3' TCG	123 pb	94°C - 1 min
	TCG GGC TGG 3'		66°C - 1 min 35 72°C - 0.4 min
TB3	5'-CCTGCGAGCGTAGGCGCGTCCG-3'	541 bp	94°C - 1 min

	CTC GTC CAG CGC CGCCGC TTC GG - 3'		65°C - 1 min 35
	5'-GCGGGGCAAGGTACGTCAGAAC-3'		72°C - 0.4 min
	5'-ACG GTG GGT ACT AGGT TGT GGG TGG TTTC-3'		94°C - 1 min
<b>16SrRNA</b>	5'-TCT GCG ATT ACT AGC AGC GAC TCC GAC TTCA-3'	543 bp	60°C - 1 min 35 72°C - 0.4 min

Table 1 describes the primers used for the identification of *M. tuberculosis*, their sequence, size in base pairs (bp) and their respective conditions necessary for PCR amplification.

### Molecular sequencing of the microbiome

DNA was amplified using specific primers targeting the V4: 515F -806R (5'GTGCCAGCMGCCGCGGGTAA3') (5' GGACTACHVVGGGGTWTCTAAT 3') region of the bacterial 16s rRNA gene. The primers also carried the Illumina MiSeq sequencing adapter (16S Amplicon PCR). All PCR reactions were carried out with Phusion® High-Fidelity PCR Master Mix (New England Biolabs). For quantification of PCR products, the same volume of 1X loading buffer (contained SYB Green) was mixed with PCR products and electrophoresis was performed on a 2% agarose gel for detection. Samples with a bright main band between 400 - 450 bp were chosen. And 16S rRNA amplicons were purified with Qiagen Gel Extraction Kit (Qiagen, Germany). Library preparation was performed with Ion Xpress™ Plus fragment library kit (Thermo Fisher Scientific).

Finally, the library was sequenced on an Illumina Mi Seq platform (Illumina Mi Seq sequencing system, USA) at Novogene Inc. USA. The choice of this method over other next-generation sequencing (NGS) systems was chosen because of its ability to generate millions of reads at once.

### Bioinformatics analysis

The raw data were merged and filtered to obtain clean data for veracity. The effective data were used to make OTU (Operational Taxonomic Units) groupings and species annotation for the respective sequence of each OTU. The microbial composition and diversity of the samples was analyzed using QIIME v1.9.1 (Quantitative Insights into Microbial Ecology). Clustering (OTUs) was performed using the UCLUST and USEARCH algorithms. The alpha diversity of each of the samples was calculated from their richness (number of OTUs observed and estimated using Chao's index1), dominance (Simpson's index) and diversity (Shannon's index). Linked pair reads were merged using FLASH (V1.2.7), Quality filtering on the raw sequences was performed under specific filtering conditions to obtain the high-quality clean sequences, according to Qiime (V1.7.0). The sequences were compared with the reference database (Gold database) using the UCHIME algorithm (UCHIME algorithm), to detect chimera sequences.

### OTU and species annotation

Sequence analysis was performed using Uparse software, using all effective sequences. For each representative sequence, the Mothur software was contrasted with the SSUrRNA database and

the SILVA database for the identification of species in each taxonomic rank (Threshold: 0.8 ~ 1), (kingdom, phylum, class, order, family, genus and species). To obtain the phylogenetic relationship of all representative OTU sequences, the MUSCLE program (Version 3.8.31) was used. All the data obtained above were used for the subsequent analysis of alpha and beta diversity.

Bacterial species diversity analysis. All these indices in the samples were calculated with QIIME (Version 1.7.0) and displayed with R software (Version 2.15.3).

## Results

The sputum samples collected by the respective hospitals showed a positive diagnosis of tuberculosis and confirmation of the presence of *M. tuberculosis* was provided by the respective laboratories using conventional methods, staining and culture (Table 2).

### Determination of the presence of *Mycobacterium tuberculosis* by PCR

For the diagnosis of tuberculosis by PCR of sputum samples, three primers coded as TB1, TB2 and TB3 were used, with amplification bands of 285 bp, 123 bp and 541 bp respectively, the presence of these bands corresponds to the repetitive element of the insertion site IS6110 which will indicate whether the sample is positive or negative for *M. Tuberculosis*. The result of each of the primers with each working group (Group B, C and D) is expressed in percentages in Table 3, where it can be seen that all the samples were positive for *M. Tuberculosis*, but among the primers used TB2 with 123pb was the one that showed the highest percentage of amplification at the time of identification, since most of the samples processed were positive for this primer compared to primers TB1 and TB3, both for new patients, with treatment and Multidrug-Resistant. Taking into account these results, a good identifier of *M. Tuberculosis* could be considered.

**Table 2.** Characteristics of the patients, classified into groups

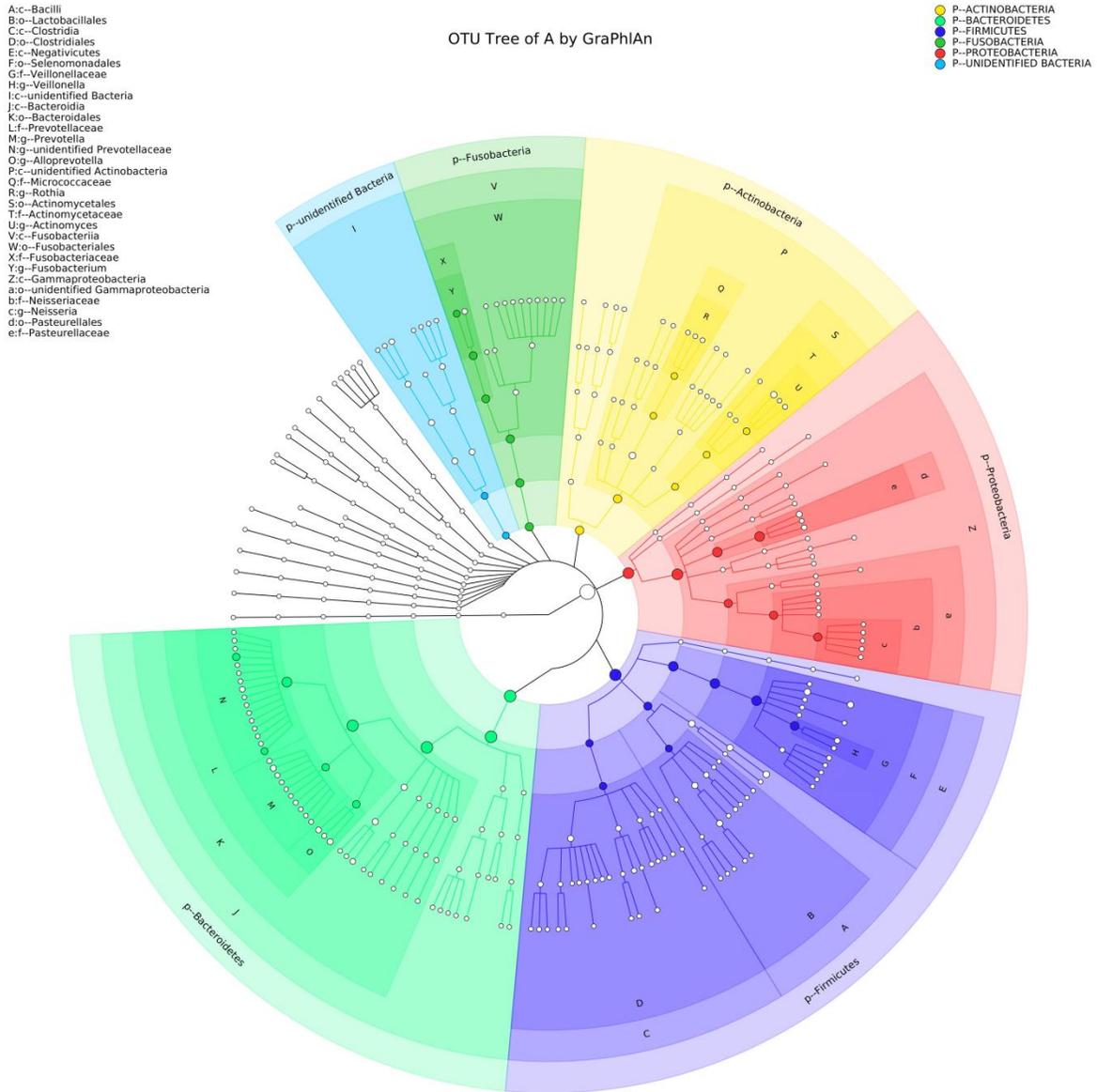
	Patients without treatment "Group B n= 20	Patients with "Group C" treatment n= 6	Multidrug-resistant patients "Group D". n= 2	Patients without tuberculosis "Group A n= 8
Age	45(22-85)	49(23-69)	53	(25-60)
Male	60%	17%	100%	50%
Female	40%	83%	0%	50%
Positive bacilloscopy (%)	100%	100%	100%	
Sputum Culture Positive (%)	85%	33%	100%	

In this study, the samples of patients diagnosed with tuberculosis ranged in age from 22 to 85 years. Group B is made up of 60% males and 40% females, with 100% positive cases for bacillus copy and 85% for culture. As for group C, 17% of samples corresponded to males and 83% to females, 100% of cases were positive for bacillus copy and 33% were positive for culture. For group D, 100% of cases corresponded to males and 100% of cases were positive for smear microscopy and culture. In group A (control), the ages ranged between 25 - 60 years, where 50% were women and 50% men.

**Table 3.** Identification of sputum samples of groups A, B, C and D by PCR using primers TB1, TB2 and TB3

<b>Primers</b>	<b>Patients without "B" treatment n=20</b>	<b>Patients with "C" treatment n=6</b>	<b>Multidrug-Resistant Patients "D". n=2</b>
TB1	85%	33%	100%
TB2	90%	83%	100%
TB3	95%	17%	100%

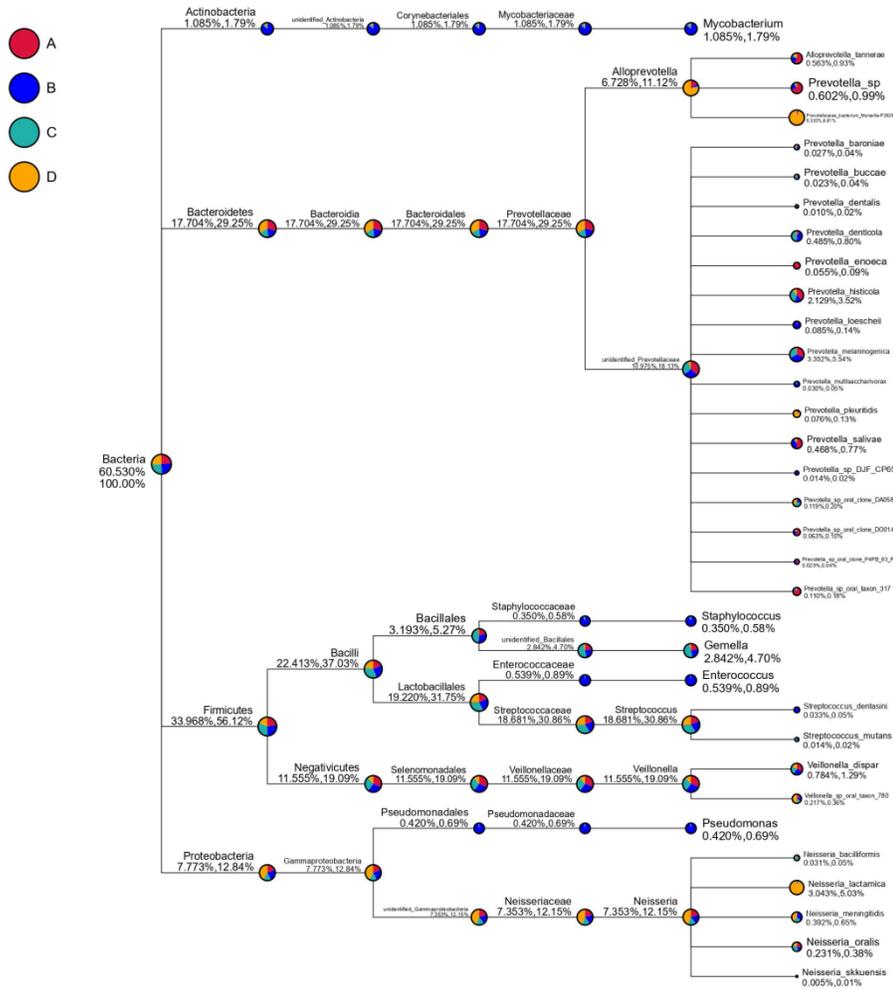
Using primers TB1, TB2 and TB3, all samples were positive for M. Tuberculosis. In the group of untreated patients "B", the primer with the highest identification percentage was TB3 (95%). In the group of patients with treatment "C", the primer with the highest identification percentage was TB2 (83%).



**Figure 1.** OTU tree of Group A.

To explore the dominant bacterial species, the results were annotated using the GraPhIAn diagram. The reading of the graph in the circle goes from inside to outside, where the names of different classification levels are observed, and the abundance of the species is proportional to the size of the circle, the different colors represent different Phyla.

The OTU tree (Figure 1) expresses the phyla present in sputum samples from patients without tuberculosis (control group). The most abundant phyla were Firmicutes and Bacteroidetes followed by Proteobacteria, Actinobacteria, Fusobacteria and a very small group of unidentified bacteria.

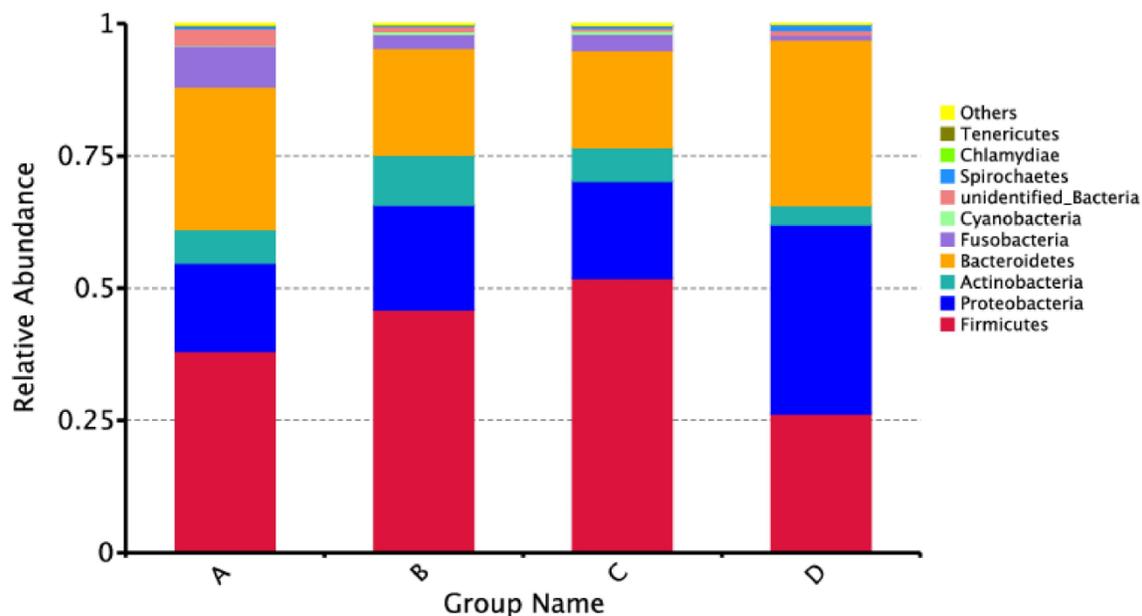


**Figure 2.** Taxonomy tree of sputum samples.

The taxonomic division of the sputum samples belonging to the groups (A) patients without tuberculosis, B) patients without treatment, C) with treatment, and D) multidrug-resistant) is observed.

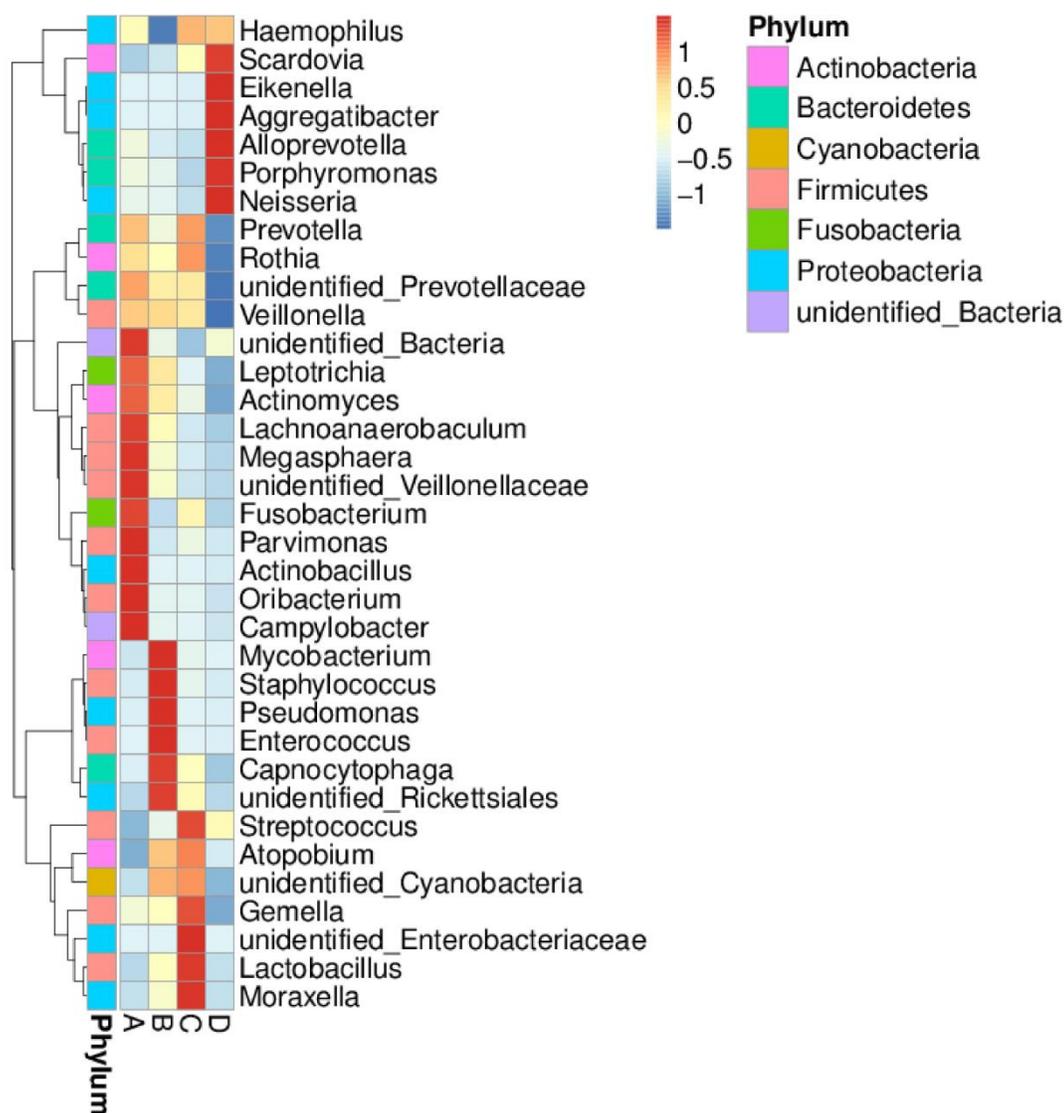
The taxonomy tree (Figure 2) shows the 10 main genera in high relative abundance, with each color representing a study group. The size of the circles represents the relative abundance of species. The first number below the taxonomic name represents the percentage in the whole taxon, while the second number represents the percentage in the selected taxon and its percentage of abundance with respect to the Microbiome.

The study groups, which are distributed by A, B, C and D, show that the most abundant genus is Prevotella with 18.13% and the least abundant is Staphylococcus with 0.58%. In group A, the dominant genera were Prevotella and Veillonella; while in groups B and C the dominant genera were Prevotella, Veillonella and Streptococcus. And in group D, the dominant genera were Neisseria and Alloprevotella (Fig. 2).



**Figure 3.** Differences in relation to the relative abundance of groups A, B, C and D.

Group A (control) shows the highest relative abundance of bacteria belonging to the phylum Firmicutes followed by Bacteroidetes. Groups B (patients without treatment) and C (patients with treatment) show the highest relative abundance of bacteria belonging to the phylum Firmicutes followed by Bacteroidetes and Proteobacteria, in comparison with group D (multidrug-resistant patients), in which the highest relative abundance is found in the phylum Proteobacteria, followed by Bacteroidetes and Firmicutes.



**Figure 4.** Cluster heat map showing abundance.

The similarity and difference of the microbial diversity of the sputum samples, classified in 4 groups: A, B, C and D; resulting in 35 genera and 7 predominant phyla.

In group A, the most abundant phylum is Firmicutes with the genera (Oribacterium, Parvimonas, Megasphaera, Lachnoanaerobaculum and unidentified Veionellaceae), also found in greater concentration are the phyla Proteobacteria (Acutebacillusc Actinobacillus) and Fusobacteria (Fusobacteriumand Leptotrichia). It is important to mention that there is an outstanding abundance of unidentified phyla (including the genus Campylobacter and a genus unknown to the database). However, the bacteria found in lower abundance belong to the phyla Actinobacteria (Atopobium) and Firmicutes (Streptococcus).

In group B, the Proteobacteria phyla are found in greater abundance with the genera (*Pseudomona*, *Rickettsiales*, unidentified), Bacteroidetes (*Capnocytophaga*), Firmicutes (*Staphylococcus*, *Enterococcus*) and Actinobacteria (*Mycobacterium*). On the other hand, the genus with the lowest abundance of bacteria is mainly *Haemophilus*.

In group C, the highest bacterial abundance was found in the phyla Proteobacteria with the genera (*Moraxella* and *Enterobacteriaceae* not identified) and Firmicutes (*Streptococcus*, *Gemella* and *Lactobacillus*). The Bacteroidetes phyla (*Alloprevotella* and *Porphyromonas*), Proteobacteria (*Neisseria*), and a group of unidentified bacteria were found in lower abundance.

In group D, a marked abundance of the phyla Proteobacteria is noted with the genera (*Aggregatibacter*, *Eikenella* and *Neisseria*), Bacteroidetes (*Alloprevotella* and *Porphyromonas*) and Actinobacteria (*Scardovia*). Within this group, the lowest concentration corresponds to the phyla Bacteroidetes (*Prevotella* and *Prevotellaceae* not identified), Firmicutes (*Veillonella*) and Actinobacteria (*Rothia*) (Figure 4).

## DISCUSSION

The study of microbiome related to infectious diseases such as tuberculosis is just being addressed in the environment. The 16S rRNA gene is used for the identification of bacterial DNA and presents hypervariable regions V1-V3 and V4-V5, although the V1-V3 region has a wider taxonomic coverage, it is known that the V4-V5 region provides a better identification of bacterial genera as reported by Drengeneset al.(2021)[32]. In the research, the V4 region was used to obtain the profile of the microbial community present in sputum samples, using high-throughput sequencing (NGS).

For the identification of the presence of *M. tuberculosis*, three primers of different repetitive elements of the IS6110 insertion site described in Table 2 were used, of which, primer TB2 (123bp) amplified for a greater number of samples compared to TB1 and TB3, results that agree with what was reported in the study done by Oguskuet al. (2004), where it was found that the IS6110 sequences in the genome of all strains isolated from tuberculosis patients in the state of Amazonas (Brazil), amplified in the 123 bp primer fragment [33]. The similarity between the two studies would be related to the demographic and cultural conditions shared by both countries in terms of population.

Figure 1 (Graphlam diagram) present, in the group of individuals without tuberculosis, the phylum level and in order of abundance: Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria and Fusobacteria. This is in agreement with the findings of O'Dwyer et al. (2016) and Dickson et al. (2016), and Finlay et al. (2019), where the phyla Bacteroidetes, Firmicutes, Proteobacteria and Actinobacteria are reported with greater predominance in samples from patients without tuberculosis [34-38].

Likewise, in the findings described by Yu et al. (2016), it is mentioned that the Proteobacteria phylum is more abundant in the healthy lung [39].

Figure 2 reports at the genus level the presence of: Prevotella, Veillonella, Streptococcus, Neiseria, Alloprevotella and Gemella, results concordant with that reported by Dickson et al. (2017) and Somariva et al. (2020) on sputum samples in healthy patients [40,41]. Although there is concordance in diversity at the phylum and genus level between these results and those described in other studies, it can be noted that there is a slight variation in their abundance, which may be mainly due to the endemic and demographic characteristics of the populations studied.

Figure 3 (bar graph) present an increase in the abundance of the phylum Firmicutes and decrease of Bacteroidetes in untreated and treated tuberculosis patients, in the latter group the variation was more noticeable. The predominance of Firmicutes in both groups agrees with that reported by Krishna et al. (2016) on Indian patients with untreated tuberculosis [27] and by Valdez Palomares et al. (2021) in Mexico City [42] on patients with untreated and treated tuberculosis. On the contrary, in this same study, an increase in the abundance of Bacteroidetes is also described in the mentioned groups [43]. This discrepancy is also present in the report by Hu Y et al. (2020), who mentions a higher abundance of Proteobacteria in relation to Firmicutes and Bacteroidetes in patients with pulmonary tuberculosis [44].

Knowing that members of the phylum Firmicutes produce short-chain fatty acids (SCFA) such as butyrate, which exerts effects on immunomodulation and anti-inflammatory response [45,46], it could be that during the dysbiosis produced by M. tuberculosis infection, there has been an increase in commensal bacteria that would generate the ideal environment for the development of Mycobacterium tuberculosis.

On the other hand, the discrepancy between the abundance of phyla found in our study in patients with treatment and that reported by other authors may be due to the antituberculosis treatment schedule which produces a profound alteration of the intestinal microbiome and therefore of the pulmonary microbiome according to the study conducted by Namasivayam et al. (2020) [47].

Figure 3 shows the alteration in the abundance of the pulmonary microbiome of multidrug-resistant patients, where the increase in the abundance and predominance of the Proteobacteria and Bacteroidetes phyla can be noted, as well as the marked decrease in Firmicutes phylum compared to the groups of patients without tuberculosis, with untreated tuberculosis. These results are consistent with those reported in the study by Luo et al. (2017), who reported a significant enrichment of Proteobacteria and decrease of Bacteroidetes and Firmicutes in multidrug-resistant patients [48]. Likewise, Wang et al. (2020), reported an altered microbiota and a loss of richness during treatment of multidrug-resistant patients that persists for 3 to 8 years after recovery and discontinuation of treatment [49,50].

The Cluster Heat map (Figure 4) displays the diversity and abundance at the genus level in the 4 groups. A greater diversity is observed in the group of patients diagnosed with tuberculosis without treatment, represented in order of abundance mainly by Streptococcus, Veillonella,

*Neisseria*, *Mycobacterium*, *Gemella*, *Enterococcus*, *Pseudomonas*, *Staphylococcus*, *Alloprevotella*. This diversity is maintained in patients with a diagnosis of tuberculosis with treatment but with a decrease in the abundance of these genera, due to antituberculosis treatment, with the exception of *Streptococcus* and *Gemella* which increased their abundance slightly [51-54]. On the contrary, in multidrug-resistant patients, this diversity has been lost, and the abundance of opportunistic bacteria such as *Alloprevotella*, *Neisseria*, *Aggregatibacter*, *Eikenella*, *Haemophilus* and *Porphyromonas* increased, the latter is found in a very high concentration and according to Zhou Y et al. (2015), this genus may be an important cofactor in the formation of lung lesions, because its proportion was significantly higher within the lung lesions in their group of patients with tuberculosis [55]. However, because the samples from group D are very scarce, further studies are suggested.

Finally, it is important to mention that within each group evaluated there were slight variations in the abundance of species, taking into consideration that the microbiome of each human being differs in terms of the types of microorganisms, the population of each species and the relationship between them; and that therefore, human health and disease are related not only to the microorganisms present, but also to the expression of the genes that they harbor and the host genes. It is believed that these variations in the abundance of the pulmonary microbiome of the groups evaluated may be related to individual characteristics related to factors such as: (1) type of delivery; (2) age; (3) type of breastfeeding, nutrition and/or malnutrition during infancy; (4) eating habits in adulthood; (5) use of medications such as anti-ulcer antibiotics, corticosteroids; (6) lifestyle and work environment, among others [56-60].

## Conclusion

The study identified the diversity and abundance at phylum and genus level using Next Generation Sequencing in individuals with tuberculosis, with diagnosis of tuberculosis without treatment, with treatment and Multidrug-Resistant patients, finding a marked variation in the Microbiome present in sputum samples associated with Tuberculosis.

## Acknowledgement

The research was completed thanks to the Universidad Nacional de San Agustín de Arequipa and the Vice-rectorate of Research together with UNSA Investiga for funding the development of this project, under contract IBAIB-09-2018-UNSA.

## Support Information

**S Fig 1. OTU tree of Group A.** The circle goes from inside to outside representing abundance at the Phylum and Order level (S Fig.1).

**S Fig 2. Taxonomy tree of sputum samples.** Represents the taxonomic division at the level of Phylum, Order, Family, Genus and Species in the evaluated groups (S Fig.2).

## References

1. Rivero JLG. Update in the knowledge of the microbiome in asthma. *Rev asthma* [Internet]. 2017 Dec 29 [cited 2021 Apr 2];2(2). Available from: <https://www.separcontenidos.es/revista3/index.php/revista/article/view/121/150>
2. Moreno del Castillo M, Valladares-García J, Halabe-Cherem J. Human microbiome. *Rev la Fac Med la UNAM* [Internet]. 2018 Nov [cited 2021 Apr 25];61(6):7-19. Available from: <http://www.scielo.org.mx/pdf/facmed/v61n6/2448-4865-facmed-61-06-7.pdf>
3. Hu Y, Feng Y, Wu J, Liu F, Zhang Z, Hao Y, et al. The Gut Microbiome Signatures Discriminate Healthy From Pulmonary Tuberculosis Patients [Internet]. Vol. 9, *Frontiers in Cellular and Infection Microbiology*. 2019. p. 90. Available from: <https://www.frontiersin.org/article/10.3389/fcimb.2019.00090>
4. Requena T, Velasco M. The human microbiome in sickness and in health. *Rev Clin Esp* [Internet]. 2021 Apr 1 [cited 2021 Apr 25];221(4):233-40. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0014256519301948?via%3Dihub>
5. Das B, Nair GB. Homeostasis and dysbiosis of the gut microbiome in health and disease. *J Biosci* [Internet]. 2019 Oct 1 [cited 2021 Apr 25];44(5). Available from: <https://pubmed.ncbi.nlm.nih.gov/31719226/>
6. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. *Cell Res* [Internet]. 2020;30(6):492-506. Available from: <https://doi.org/10.1038/s41422-020-0332-7>
7. Thaiss C, Zmora N, Levy M, Elinav E. The microbiome and innate immunity. *Nature* [Internet]. 2016 Jul 6 [cited 2021 Apr 25];535(7610):65-74. Available from: <https://pubmed.ncbi.nlm.nih.gov/27383981/>
8. Belkaid Y, Harrison OJ. Homeostatic Immunity and the Microbiota. *Immunity* [Internet]. 2017 Apr 18 [cited 2021 Apr 25];46(4):562-76. Available from: [/pmc/articles/PMC5604871/](https://pubmed.ncbi.nlm.nih.gov/27383981/).
9. Rooks MG, Garrett WS. Gut microbiota, metabolites and host immunity. *Nat Rev Immunol* [Internet]. 2016 Jun 1 [cited 2021 Apr 25];16(6):341-52. Available from: <https://pubmed.ncbi.nlm.nih.gov/27231050/>
10. Hufnagl K, Pali-Schöll I, Roth-Walter F, Jensen-Jarolim E. Dysbiosis of the gut and lung microbiome has a role in asthma. *Semin Immunopathol* [Internet]. 2020/02/18. 2020 Feb;42(1):75-93. Available from: <https://pubmed.ncbi.nlm.nih.gov/32072252>
11. Hao WZ, Li XJ, Zhang PW, Chen JX. A review of antibiotics, depression, and the gut microbiome. *Psychiatry Res* [Internet]. 2020 Feb 1 [cited 2021 Apr 25];284. Available from: <https://pubmed.ncbi.nlm.nih.gov/31791704/>
12. Langdon A, Crook N, Dantas G. The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Med* [Internet]. 2016 Apr 13;8(1):39. Available from: <https://pubmed.ncbi.nlm.nih.gov/27074706>
13. Aarnoutse R, Ziemons J, Penders J, Rensen SS, De Vos-Geelen J, Smidt ML. The clinical link between human intestinal microbiota and systemic cancer therapy. *Int J Mol Sci* [Internet]. 2019 Sep 1 [cited 2021 Apr 25];20(17). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6747354/>

14. Gupta N, Kumar R, Agrawal B. New Players in Immunity to Tuberculosis: The Host Microbiome, Lung Epithelium, and Innate Immune Cells [Internet]. Vol. 9, *Frontiers in Immunology* . 2018. p. 709. Available from: <https://www.frontiersin.org/article/10.3389/fimmu.2018.00709>
15. Naidoo C, Nyawo G, Wu B, Walzl G, Warren R, Segal L, et al. The microbiome and tuberculosis: state of the art, potential applications, and defining the clinical research agenda. *Lancet Respir Med* [Internet]. 2019 Mar 22;7. Available from: [https://doi.org/10.1016/S2213-2600\(18\)30501-0](https://doi.org/10.1016/S2213-2600(18)30501-0)
16. Nakhaee M, Rezaee A, Basiri R, Soleimanpour S, Ghazvini K. Relation between lower respiratory tract microbiota and type of immune response against tuberculosis. *Microb Pathog* [Internet]. 2018;120:161-5. Available from: <https://www.sciencedirect.com/science/article/pii/S0882401017311932>
17. Balcells ME, Yokobori N, Hong B, Corbett J, Cervantes J. The lung microbiome, vitamin D, and the tuberculous granuloma: A balance triangle. *Microb Pathog* [Internet]. 2019;131:158-63. Available from: <https://www.sciencedirect.com/science/article/pii/S0882401019302256>
18. Alarcón T, D'Auria G, Delgado S, Del Campo R, Ferrer M. Microbiological diagnosis of human papillomavirus infection 59 MICROBIOTA. In: Cercenado E, Cantón R, editors. *Procedures in Clinical Microbiology* [Internet]. 2nd ed. Madrid: Spanish Society of Infectious Diseases and Clinical Microbiology (SEIMC); 2016 [cited 2021 Apr 7]. Available from: <https://www.seimc.org/contenidos/documentoscientificos/procedimientosmicrobiologia/seimc-procedimientomicrobiologia59.pdf>
19. Matteelli A, Sulis G, Capone S, D'Ambrosio L, Migliori GB, Getahun H. Tuberculosis elimination and the challenge of latent tuberculosis. *Presse Med* [Internet]. 2017;46(2, Part 2):e13-21. Available from: <https://www.sciencedirect.com/science/article/pii/S0755498217300556>
20. Namasivayam S, Sher A, Glickman MS, Wipperman MF. The Microbiome and Tuberculosis: Early Evidence for Cross Talk. Garsin DA, editor. *MBio* [Internet]. 2018 Nov 7;9(5):e01420-18. Available from: <http://mbio.asm.org/content/9/5/e01420-18.abstract>
21. Daley CL. The Global Fight Against Tuberculosis. *Thorac Surg Clin* [Internet]. 2019;29(1):19-25. Available from: <https://www.sciencedirect.com/science/article/pii/S1547412718301373>
22. World Health Organization. Global tuberculosis report 2018 [Internet]. Geneva PP - Geneva: World Health Organization; 2018. Available from: <https://apps.who.int/iris/handle/10665/274453>
23. World Health Organization. Global Tuberculosis Report 2020 [Internet]. Geneva: World Health Organization; 2020 [cited 2021 Apr 7]. Available from: <https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf>
24. Ministry of Health - National Center for Epidemiology, Disease Prevention and Control. *Epidemiological Bulletin of Peru* [Internet]. Lima; 2019 Mar [cited 2021 Apr 7]. Available from: <https://www.dge.gob.pe/portal/docs/vigilancia/boletines/2019/10.pdf>
25. Hong BY, Maulén NP, Adami AJ, Granados H, Balcells ME, Cervantes J. Microbiome changes during tuberculosis and antituberculous therapy. *Clin Microbiol Rev* [Internet]. 2016 Oct 1

- [cited 2021 Apr 26];29(4):915-26. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5010754/>
26. Hong B, Paulson JN, Stine OC, Weinstock GM, Cervantes JL. Meta-analysis of the lung microbiota in pulmonary tuberculosis. *Tuberculosis* [Internet]. 2018;109:102-8. Available from: <https://www.sciencedirect.com/science/article/pii/S1472979217304481>
  27. Krishna P, Jain A, Bisen PS. Microbiome diversity in the sputum of patients with pulmonary tuberculosis. *Eur J Clin Microbiol Infect Dis* [Internet]. 2016;35(7):1205-10. Available from: <https://doi.org/10.1007/s10096-016-2654-4>
  28. Boers SA, Jansen R, Hays JP. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur J Clin Microbiol Infect Dis* [Internet]. 2019 Jun 1 [cited 2021 Apr 25];38(6):1059-70. Available from: <https://pubmed.ncbi.nlm.nih.gov/30834996/>
  29. Hernández M, Quijada NM, Rodríguez-Lázaro D, Eiros JM. Bioinformatics of next generation sequencing in clinical microbiology diagnosis. *Rev Argent Microbiol*. 2020 Apr 1;52(2):150-61.
  30. O'Toole RF, Gautam SS. The host microbiome and impact of tuberculosis chemotherapy. *Tuberculosis* [Internet]. 2018;113:26-9. Available from: <https://www.sciencedirect.com/science/article/pii/S1472979218303123>
  31. Hu Y, Yang Q, Liu B, Dong J, Sun L, Zhu Y, et al. Gut microbiota associated with pulmonary tuberculosis and dysbiosis caused by anti-tuberculosis drugs. *J Infect* [Internet]. 2019 Apr 1 [cited 2021 Apr 25];78(4):317-22. Available from: <http://www.journalofinfection.com/article/S0163445318302512/fulltext>
  32. Drengenes C, Eagan TML, Haaland I, Wiker HG, Nielsen R. Exploring protocol bias in airway microbiome studies: one versus two PCR steps and 16S rRNA gene region V3 V4 versus V4. *BMC Genomics* [Internet]. 2021 Dec 1 [cited 2021 Jun 9];22(1):1-15. Available from: <https://doi.org/10.1186/s12864-020-07252-z>
  33. Ogusku MM, Salem JI. Análise de diferentes primers utilizados na PCR visando ao diagnóstico da tuberculose no Estado do Amazonas. *J Bras Pneumol* [Internet]. 2004 Aug [cited 2021 Jun 10];30(4):343-9. Available from: <https://www.scielo.br/j/jbpneu/a/RLzwWqBTNDHCdYGbQknn8hm/?lang=pt>
  34. O'Dwyer DN, Dickson RP, Moore BB. The Lung Microbiome, Immunity, and the Pathogenesis of Chronic Lung Disease. *J Immunol* [Internet]. 2016 Jun 15 [cited 2021 Apr 26];196(12):4839-47. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4894335/>
  35. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The Microbiome and the Respiratory Tract. *Annu Rev Physiol* [Internet]. 2016 Feb 10 [cited 2021 Apr 26];78:481-504. Available from: <https://pubmed.ncbi.nlm.nih.gov/26527186/>
  36. Finlay BB, Finlay JM. *The Whole-Body Microbiome: How to Harness Microbes-Inside and Out-for Lifelong Health* Hardcover [Internet]. 1st ed. New York: The Experiment; 2019 [cited 2021 Jun 10]. 238-260 p. Available from: [https://books.google.com.pe/books?hl=es&lr=&id=HHVMDwAAQBAJ&oi=fnd&pg=PP1&ots=pCCssCwhNV&sig=tZlIvl6ifFGHhmrh\\_KeeZxWR9wk&redir\\_esc=y#v=onepage&q&f=false](https://books.google.com.pe/books?hl=es&lr=&id=HHVMDwAAQBAJ&oi=fnd&pg=PP1&ots=pCCssCwhNV&sig=tZlIvl6ifFGHhmrh_KeeZxWR9wk&redir_esc=y#v=onepage&q&f=false)

37. Cameron SJS, Lewis KE, Huws SA, Hegarty MJ, Lewis PD, Pachebat JA, et al. A pilot study using metagenomic sequencing of the sputum microbiome suggests potential bacterial biomarkers for lung cancer. *PLoS One* [Internet]. 2017 May 25;12(5):e0177062. Available from: <https://doi.org/10.1371/journal.pone.0177062>
38. Jandhyala SM. Role of the normal gut microbiota. *World Journal of Gastroenterology* [Internet]. 2015;21(29):8787. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528021/>
39. Yu G, Gail MH, Consonni D, Carugno M, Humphrys M, Pesatori AC, et al. Characterizing human lung tissue microbiota and its relationship to epidemiological and clinical features. *Genome Biol* [Internet]. 2016 Jul 28 [cited 2021 Apr 26];17(1):163. Available from: <http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1021-1>
40. Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Falkowski NR, Huffnagle GB, et al. Bacterial Topography of the Healthy Human Lower Respiratory Tract. *MBio* [Internet]. 2017 Feb 14;8(1):e02287-16. Available from: <https://pubmed.ncbi.nlm.nih.gov/28196961>
41. Sommariva M, Le Noci V, Bianchi F, Camelliti S, Balsari A, Tagliabue E, et al. The lung microbiota: role in maintaining pulmonary immune homeostasis and its implications in cancer development and therapy. *Cell Mol Life Sci* [Internet]. 2020 Jul 1 [cited 2021 Apr 26];77(14):2739-49. Available from: <https://doi.org/10.1007/s00018-020-03452-8>
42. Valdez-Palomares F, Muñoz Torrico M, Palacios-González B, Soberón X, Silva-Herzog E. Altered Microbial Composition of Drug-Sensitive and Drug-Resistant TB Patients Compared with Healthy Volunteers. *Microorganisms* [Internet]. 2021 Aug 18 [cited 2022 Jan 15];9(8):1762. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8398572/>
43. Vázquez-Pérez JA, Carrillo CO, Iñiguez-García MA, Romero-Espinoza I, Márquez-García JE, Falcón LI, et al. Alveolar microbiota profile in patients with human pulmonary tuberculosis and interstitial pneumonia. *Microb Pathog* [Internet]. 2020 Feb 1 [cited 2021 Apr 26];139. Available from: <https://pubmed.ncbi.nlm.nih.gov/31715320/>
44. Hu Y, Kang Y, Liu X, Cheng M, Dong J, Sun L, et al. Distinct lung microbial community states in patients with pulmonary tuberculosis. *Sci China Life Sci* [Internet]. 2020 Oct 1 [cited 2021 Apr 26];63(10):1522-33. Available from: <https://pubmed.ncbi.nlm.nih.gov/32303963/>
45. Mori G, Morrison M, Blumenthal A. Microbiome-immune interactions in tuberculosis. *PLoS pathogens* [Internet]. 2021 Apr 1;17(4):e1009377. Available from: <https://pubmed.ncbi.nlm.nih.gov/33857251/>
46. Lu Y, Fan C, Li P, Lu Y, Chang X, Qi K. Short chain fatty acids prevent high-fat-diet-induced obesity in mice by regulating g protein-coupled receptors and gut Microbiota. *Sci Rep* [Internet]. 2016 Nov 28 [cited 2021 Apr 26];6(1):1-13. Available from: [www.nature.com/scientificreports/](http://www.nature.com/scientificreports/)
47. Namasivayam S, Diarra B, Diabate S, Sarro Y dit S, Kone A, Kone B, et al. Patients infected with *Mycobacterium africanum* versus *Mycobacterium tuberculosis* possess distinct intestinal microbiota. *PLoS Negl Trop Dis* [Internet]. 2020 May 13;14(5):e0008230. Available from: <https://doi.org/10.1371/journal.pntd.0008230>
48. Luo M, Liu Y, Wu P, Luo D-X, Sun Q, Zheng H, et al. Alternation of Gut Microbiota in Patients with Pulmonary Tuberculosis. *Front Physiol* [Internet]. 2017 Nov 17 [cited 2021 Apr

- 26];8(NOV):822. Available from: <http://journal.frontiersin.org/article/10.3389/fphys.2017.00822/full>
49. Wang J, Xiong K, Zhao S, Zhang C, Zhang J, Xu L, et al. Long-Term Effects of Multi-Drug-Resistant Tuberculosis Treatment on Gut Microbiota and Its Health Consequences. *Front Microbiol* [Internet]. 2020 Jan 30 [cited 2021 Apr 26];11:53. Available from: <https://www.frontiersin.org/article/10.3389/fmicb.2020.00053/full>
50. Wipperman MF, Fitzgerald DW, Juste MAJ, Taur Y, Namasivayam S, Sher A, et al. Antibiotic treatment for Tuberculosis induces a profound dysbiosis of the microbiome that persists long after therapy is completed. *Sci Rep* [Internet]. 2017 Dec 1 [cited 2021 Apr 26];7(1):1-11. Available from: [www.nature.com/scientificreports/](http://www.nature.com/scientificreports/)
51. World Health Organization. Guidelines for treatment of drug-susceptible tuberculosis and patient care [Internet]. Geneva; 2017 [cited 2021 Apr 26]. Available from: <http://apps.who.int/iris/bitstream/handle/10665/255052/9789241550000-eng.pdf?sequence=1>
52. Ministry of Health. General Directorate of People's Health. National Health Strategy for the Prevention and Control of Tuberculosis. Norma técnica de salud para la atención integral de las personas afectadas por tuberculosis [Internet]. Lima; 2013 Dec [cited 2021 Apr 26]. Available from: <http://www.tuberculosis.minsa.gob.pe/portaldpctb/recursos/20180308083418.pdf>
53. Original A. *RevPeruMedExp Salud Publica*. 2016 [cited 2020 Dec 8]; Available from: <http://www.scielo.org.pe/pdf/rins/v33n4/a02v33n4.pdf>
54. Farga V, Peña C, Lagos M, Carlos Rodríguez J. TUBERCULOSIS SECTION Recommendations in the treatment of drug-susceptible tuberculosis (According to ATS/CDC/IDSA Guidelines. 2016) Guidelines for the treatment of drug-susceptible tuberculosis. According to official ATS, CDC and IDSA Guidelines 2016 [Internet]. Available from: <https://scielo.conicyt.cl/pdf/rcher/v33n1/art08.pdf>
55. Zhou Y, Lin F, Cui Z, Zhang X, Hu C, Shen T, et al. Correlation between either *Cupriavidus* or *Porphyromonas* and primary pulmonary tuberculosis found by analysing the microbiota in patients' bronchoalveolar lavage fluid. *PLoS One* [Internet]. 2015 May 22 [cited 2021 Apr 26];10(5). Available from: <https://pubmed.ncbi.nlm.nih.gov/30536353/>
56. Campanella V, Syed J, Santacroce L, Saini R, Ballini A, Inchingolo F. Oral probiotics influence oral and respiratory tract infections in pediatric population: A randomized double-blinded placebo-controlled pilot study. *Eur Rev Med Pharmacol Sci* [Internet]. 2018;22(22):8. Available from: <https://pubmed.ncbi.nlm.nih.gov/30536353/>
57. Al Bataineh MT, Hamoudi RA, Dash NR, Ramakrishnan RK, Almasalmeh MA, Sharif HA, et al. Altered respiratory microbiota composition and functionality associated with asthma early in life. *BMC Infect Dis* [Internet]. 2020 Sep 22 [cited 2021 Apr 26];20(1):697. Available from: <https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-020-05427-3>
58. de Steenhuijsen Piters WAA, Huijskens EGW, Wyllie AL, Biesbroek G, van den Bergh MR, Veenhoven RH, et al. Dysbiosis of upper respiratory tract microbiota in elderly pneumonia patients. *The ISME Journal*. 2015 Jul 7;10(1):97-108.

59. Lee J, Kim S, Lee M, Kim B, Song W, Park H, et al. Different upper airway microbiome and their functional genes associated with asthma in young adults and elderly individuals. *Allergy*. 2018 Nov 22;74(4):709-19.
60. Meléndez Rentería NP, Aguilar CN, Nevárez Moorillón GV, Rodríguez Herrera R. Prebiotic compounds: from molecules to humans. *Journal of the Venezuelan Society of Microbiology* [Internet]. 2011 Jun 1;31(1):6-12. Available from: [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S1315-25562011000100003](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1315-25562011000100003)