scientific reports

Check for updates

OPEN

Clinical metagenomics analysis of bacterial and fungal microbiota from sputum of patients suspected with tuberculosis infection based on nanopore sequencing

Pawarisa Terbtothakun[1,2], Suthida Visedthorn[2], Pavit Klomkliew[2], Prangwalai Chanchaem[2], Vorthon Sawaswong[3], Pavaret Sivapornnukul[2], Samitanan Sunantawanit[2], Ariya Khamwut[2], Suwatchareeporn Rotcheewaphan[2,4], Pornchai Kaewsapsak[2] & Sunchai Payungporn[2]✉

Tuberculosis (TB) remains a significant global health challenge, demanding rapid and comprehensive diagnostics for effective treatment. Secondary infections further complicate TB infection, worsening outcomes. Conventional diagnostics are hindered by prolonged turnaround times, high costs, and inability to detect co-infections. This study utilizes full-length 16S rDNA and internal transcribed spacer (ITS) amplicon sequencing based on Oxford Nanopore Technologies (ONT) to analyze clinical metagenomics of sputum microbiota from patients suspected with TB Infection. Our findings highlight the potential of ONT for profiling microbial communities associated with TB infection. The MTB group exhibited a significant abundance of *Mycobacterium tuberculosis* (*M. tuberculosis*) and *Stenotrophomonas maltophilia*. In contrast, *Prevotella melaninogenica*, *Veillonella parvula*, *Corynebacterium striatum* and *Pseudomonas aeruginosa* were more abundant in the negative samples. Fungal analysis revealed *Candida orthopsilosis* was enriched in MTB samples, while *Aureobasidium leucospermi* and *Wallemia muriae* predominated in negative samples. Correlation network analysis revealed *M. tuberculosis* exhibits positive and negative correlations with other microbial species, suggesting cooperative and competitive interactions that may influence microbial community dynamics and disease progression in TB patients. This study demonstrates the promise of ONT-based clinical metagenomics for rapid, comprehensive detection of bacterial and fungal co-infections, addressing limitations of conventional diagnostics and improving outcomes.

Keywords *Mycobacterium tuberculosis*, Metagenomics, 16S rDNA, Internal transcribed spacer (ITS), Oxford nanopore technologies (ONT)

Tuberculosis (TB) remains a significant global health concern despite being a preventable and treatable illness[1]. *Mycobacterium tuberculosis* (*M. tuberculosis*) causes TB, which spreads through airborne bacteria expelled when patients with active TB cough or exhale. Over 10 million people are diagnosed annually, with numbers rising since 2021[1]. An estimated 25% of the global population has been exposed to TB[2]. The risk of active TB is highest in the first two years post-infection, diminishes over time, and certain infections may cure spontaneously[3]. About 90% of those who develop TB are adults, with a higher incidence in men[1]. In 2023, 10.8 million new TB cases were reported, representing a slight increase from 10.7 million in 2022. The number of cases was higher than the 10.4 million reported in 2021 and 10.1 million in 2020[1].

Accurate and timely diagnosis is crucial for managing infections caused by *M. tuberculosis*. TB infection poses significant diagnostic challenges due to slow culture growth, varied clinical presentations, and limitations in existing diagnostic tests[4]. Co-infections with other bacteria and fungi further complicate the clinical course of

[1]Medical Biochemistry Program, Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand. [2]Faculty of Medicine, Department of Biochemistry, Center of Excellence in Systems Microbiology, Chulalongkorn University, 1873 Rama IV Road, Pathumwan, Bangkok 10330, Thailand. [3]Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand. [4]Department of Microbiology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand. ✉email:

TB, exacerbating disease severity and impeding treatment outcomes[5]. Impaired immunity can lead to secondary bacterial infections, which worsen respiratory function and increase the risk of systemic complications. Bacterial pathogens such as *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Escherichia coli*, and *Haemophilus influenzae* significantly affect TB management, increasing disease severity and impairing recovery[6–8]. Fungal infections, particularly *Candida* spp. and *Aspergillus* spp., are also increasingly observed in TB patients with weakened immune systems[6]. Rapid diagnosis of these infections is essential for timely treatment and improved outcomes.

Conventional diagnostic methods are associated with various limitations, including prolonged processing times, high costs, and the inability to identify co-infections, which can compromise their effectiveness in accurately diagnosing infections. Delays in diagnostic results can adversely impact treatment, potentially resulting in unfavorable patient outcomes[9]. Consequently, there is an urgent need for strategies to improve the detection and characterization of these pathogens. Nanopore technology has played a crucial role in the field of clinical metagenomics by enabling rapid, real-time sequencing of microbial genomes directly from patient specimens[10]. This technology offers a distinct advantage by providing high-throughput, long-read sequencing that can simultaneously identify mycobacterial infections, detect co-infections, and analyze microbial communities[11]. Furthermore, nanopore sequencing is cost-effective and adaptable to various clinical settings, making it a promising tool for enhancing diagnostic accuracy and improving patient outcomes[12].

This study aimed to utilize ONT for metagenomic analysis and investigate the bacterial and fungal microbiota in sputum samples from patients suspected of TB infection.

Results

Patient characteristics

Following application of the exclusion criteria described in the Methods section, the final cohort consisted of 97 samples. Sputum samples obtained from individuals with suspected TB infection were categorized into two groups, including 56 MTB and 41 negative samples. Demographic and categorical data for these samples, including age, gender, and specimen group, are presented in Table 1 and Supplementary Table S1.

Comparative analysis of bacterial diversity in sputum samples from patients with and without tuberculosis infection

Bacterial microbiota was identified in two sample groups using 16S rDNA sequencing. Sequencing was performed on the high-throughput MinION™ platform (ONT). The sequencing yielded 1,467,408 raw reads (mean ± SD: 15,446 ± 9,401 reads/sample), of which 13,422 ± 8,460 reads/sample were classified. The sequencing depth significantly differed between MTB (22,297 ± 10,005 reads/sample) and negative (9,963 ± 2,575 reads/sample) groups (unpaired *t*-test, $P < 0.0001$). Rarefaction analysis was conducted to determine if the sequencing depth was sufficient for accurate classification of all samples into bacterial taxa and operational taxonomic units (OTUs). Bacterial diversity in sputum samples was assessed to compare MTB and negative groups. Alpha and beta diversity metrics were analyzed to evaluate differences in bacterial community structure. Alpha diversity was measured using the Chao1 and Shannon indices. The Chao1 index (species richness measure) differed significantly between MTB and negative groups (Mann-Whitney test, $P = 0.0281$; Fig. 1A). In contrast, the Shannon index that accounts for species richness and evenness showed no significant difference between the two groups (Mann-Whitney test, $P = 0.2947$; Fig. 1B). Beta diversity was assessed using Principal Coordinate Analysis (PCoA) based on Bray-Curtis dissimilarity (Fig. 1C). The analysis revealed distinct clustering of bacterial communities between two groups. PERMANOVA confirmed a significant difference in bacterial community composition ($F = 3.6638$, $R^2 = 0.0371$, $P = 0.001$). These results demonstrate compositional differences in bacterial communities between MTB and negative groups.

Relative abundance of bacterial microbiota based on 16S rDNA amplicon sequencing

The relative abundances of bacterial microbiota are presented in Fig. 1D and E. The major bacterial phyla identified in sputum samples were consistent between two groups (Fig. 1D). The four predominant phyla included Firmicutes, Proteobacteria, Actinobacteria, and Bacteroidetes. Among these, Firmicutes and Proteobacteria were the most abundant, contributing the majority of the total bacterial composition. Minor bacterial phyla, such as Deferribacteres, Fusobacteria, Spirochaetes, Tenericutes, and Verrucomicrobia, were detected in some samples at low relative abundances. The relative abundances of bacterial species are shown in Fig. 1E, highlighting the 20 most abundant species. *M. tuberculosis* was detected only in the MTB group, while *Stenotrophomonas maltophilia* had a higher relative abundance in the MTB group than in the negative

Characteristics		MTB group ($n = 56$)	Negative group ($n = 41$)	*P*-value
Age (Mean ± SD)		61.20 ± 16.68	63.71 ± 19.24	0.4943
Gender	Male	35	30	0.3758
	Female	21	11	

Table 1. The characteristics of patients suspected of tuberculosis infection. Age: reported as Mean ± SD, with statistical significance tested using an unpaired *t*-test. Gender: statistical significance tested using a Chi-square test without Yates' correction.

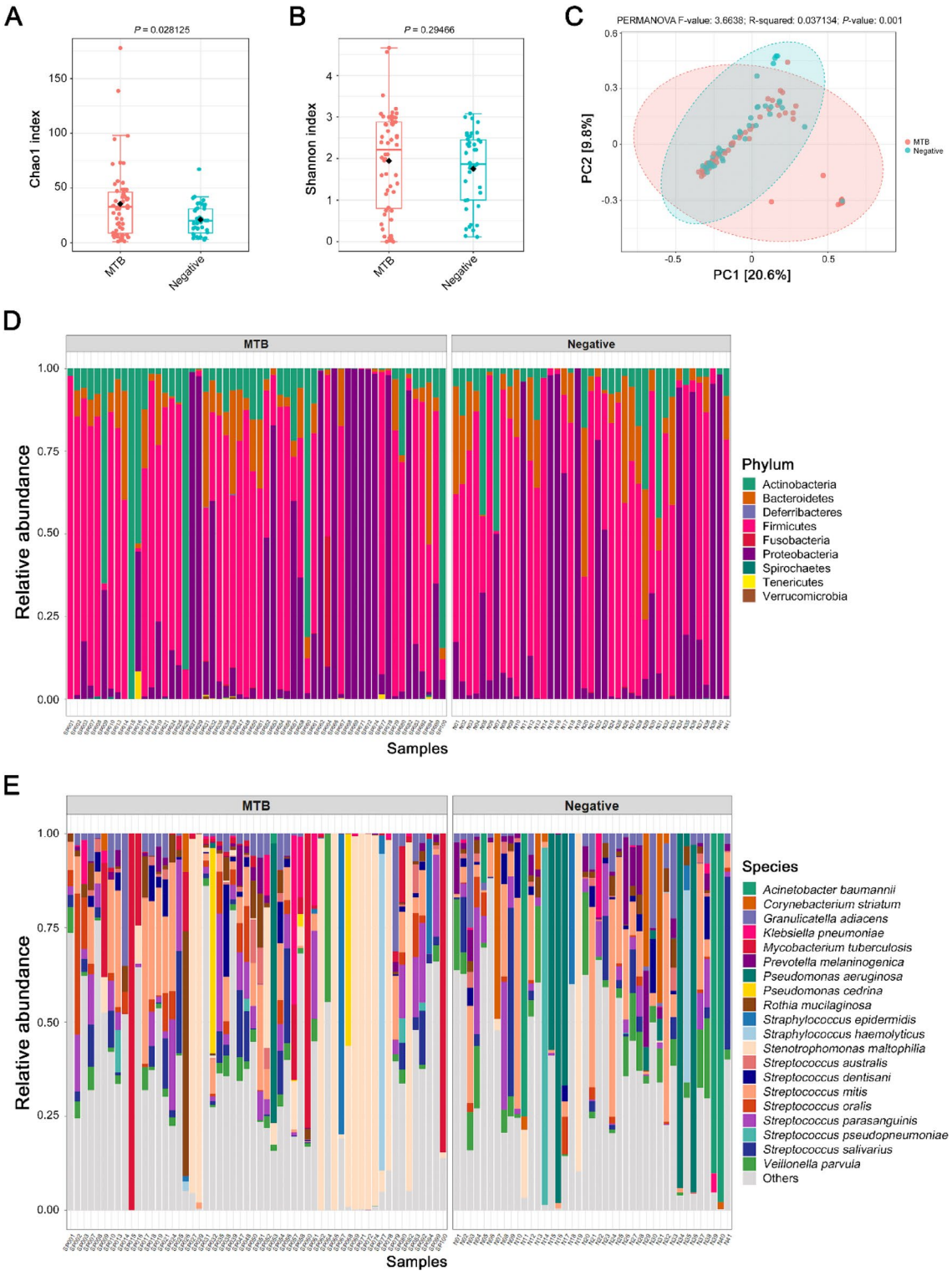

Fig. 1. Bacterial diversity in MTB and negative samples. (A) Chao1 and (B) Shannon indices are shown as scatterplots. β-diversity analysis (C). Relative abundance of bacteria at (D) phylum level and (E) top 20 species level. The taxa below the top 20 were classified as 'Others'.

group. Conversely, *Prevotella melaninogenica*, *Veillonella parvula*, *Corynebacterium striatum*, and *Pseudomonas aeruginosa* were more abundant in the negative group.

Differential abundance analysis of significant bacterial species in sputum samples

Understanding the differences in bacterial composition among MTB and negative sputum samples is crucial for identifying key microbial signatures associated with tuberculosis infection. The Venn diagram (Fig. 2A) illustrates the overall distribution of bacterial taxa between the two groups, showing that 350 taxa were unique to the MTB group and 33 were exclusive to the negative group, while 148 taxa were shared between the two groups. Detailed information regarding these taxes is provided in Supplementary Table S3. Differential abundance analysis further revealed significant differences in bacterial taxa across these groups (Fig. 2B and C). Species abundance was compared using linear discriminant analysis effect size (LEfSe) with thresholds of a nominal *P*-value cutoff < 0.05 and LDA score ± 4 (Fig. 2B and Supplementary Table S2). Relative abundance plots (Fig. 2C) confirmed these findings, showing that *Stenotrophomonas maltophilia* and *M. tuberculosis* were significantly more abundant in MTB samples (Mann-Whitney test, $P < 0.05$). Conversely, *Prevotella melaninogenica*, *Veillonella parvula*, *Corynebacterium striatum* and *Pseudomonas aeruginosa* exhibited higher abundance in negative samples (Mann-Whitney test, $P < 0.05$).

Comparative analysis of fungal diversity in sputum samples from patients with and without tuberculosis infection

Fungal microbiota was successfully identified in 55 of 56 culture-positive and 40 of 41 culture-negative samples using ITS sequencing. This sequencing generated 1,208,256 raw reads (mean ± SD: 12,718 ± 12,436 reads/sample), with a mean of 9,709 ± 8,729 classified reads/sample. The sequencing depth significantly differed between MTB (10,405.45 ± 12,684.69 reads/sample) and negative (15,898.90 ± 11,491.45 reads/sample) groups (unpaired *t*-test, $P = 0.0328$). The alpha diversity of fungal species was assessed using the Chao1 and Shannon indices (Fig. 3A and B). The Chao1 index showed no significant difference between both groups (Mann-Whitney test, $P = 0.8457$; Fig. 3A), indicating similar species richness across both groups. Similarly, the Shannon index revealed no significant difference between the groups (Mann-Whitney test, $P = 0.4397$; Fig. 3B). Beta diversity was evaluated using PCoA based on Bray-Curtis dissimilarity (Fig. 3C). The PCoA analysis demonstrated distinct clustering of fungal communities between the MTB and negative groups. PERMANOVA revealed significant differences in fungal community composition between groups ($F = 3.1601$, $R^2 = 0.0329$, $P = 0.007$), suggesting that tuberculosis infection status affects fungal microbiota structure.

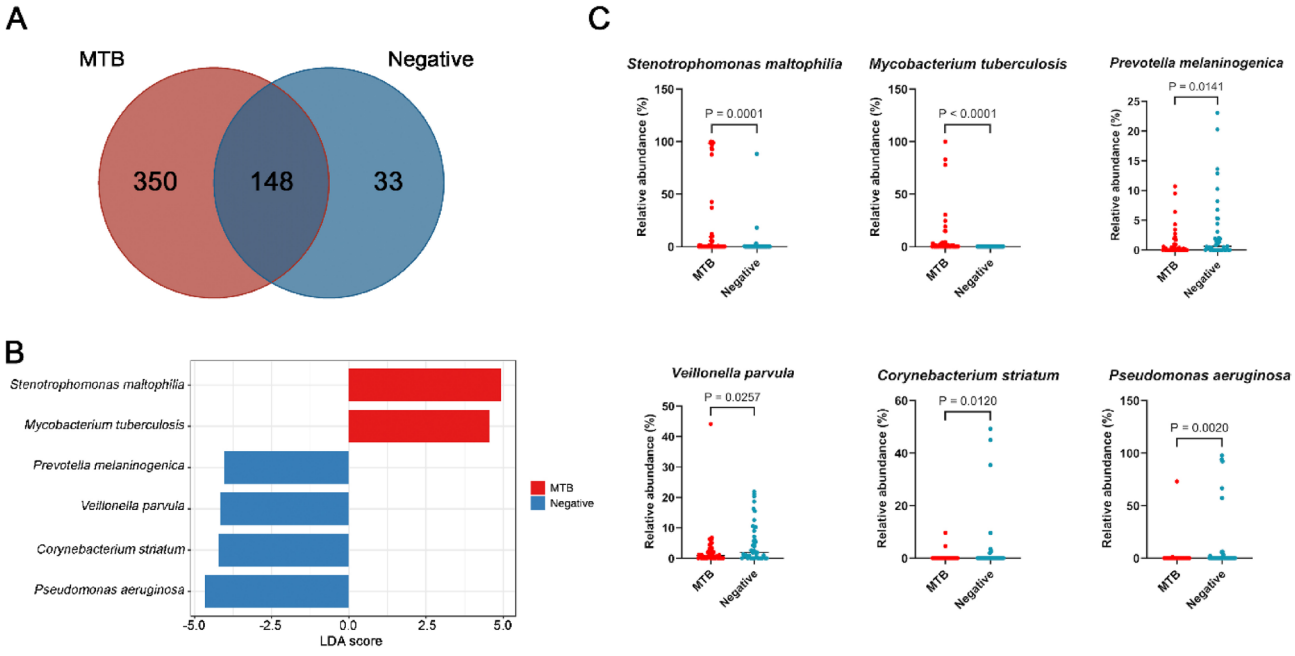

Fig. 2. Differential abundance analysis was performed using Linear Discriminant Analysis Effect Size (LEfSe) on MTB and negative sputum samples. (A) Venn diagram showing the number of bacterial species identified through the metagenomic approach (16S rDNA sequencing) in MTB samples (red) and negative samples (blue). (B) Bar plots display differentially abundant bacterial taxa across various taxonomic ranks. The LDA score indicates the effect size and ranking of each taxon with significant differential abundance (LDA score ± 4, $P < 0.05$). (C) Relative abundance of the most prevalent bacterial species. Scatter dot plots illustrate significant differences in the top abundant bacteria between MTB group (red) and negative group (blue), determined by a Mann-Whitney test ($P < 0.05$).

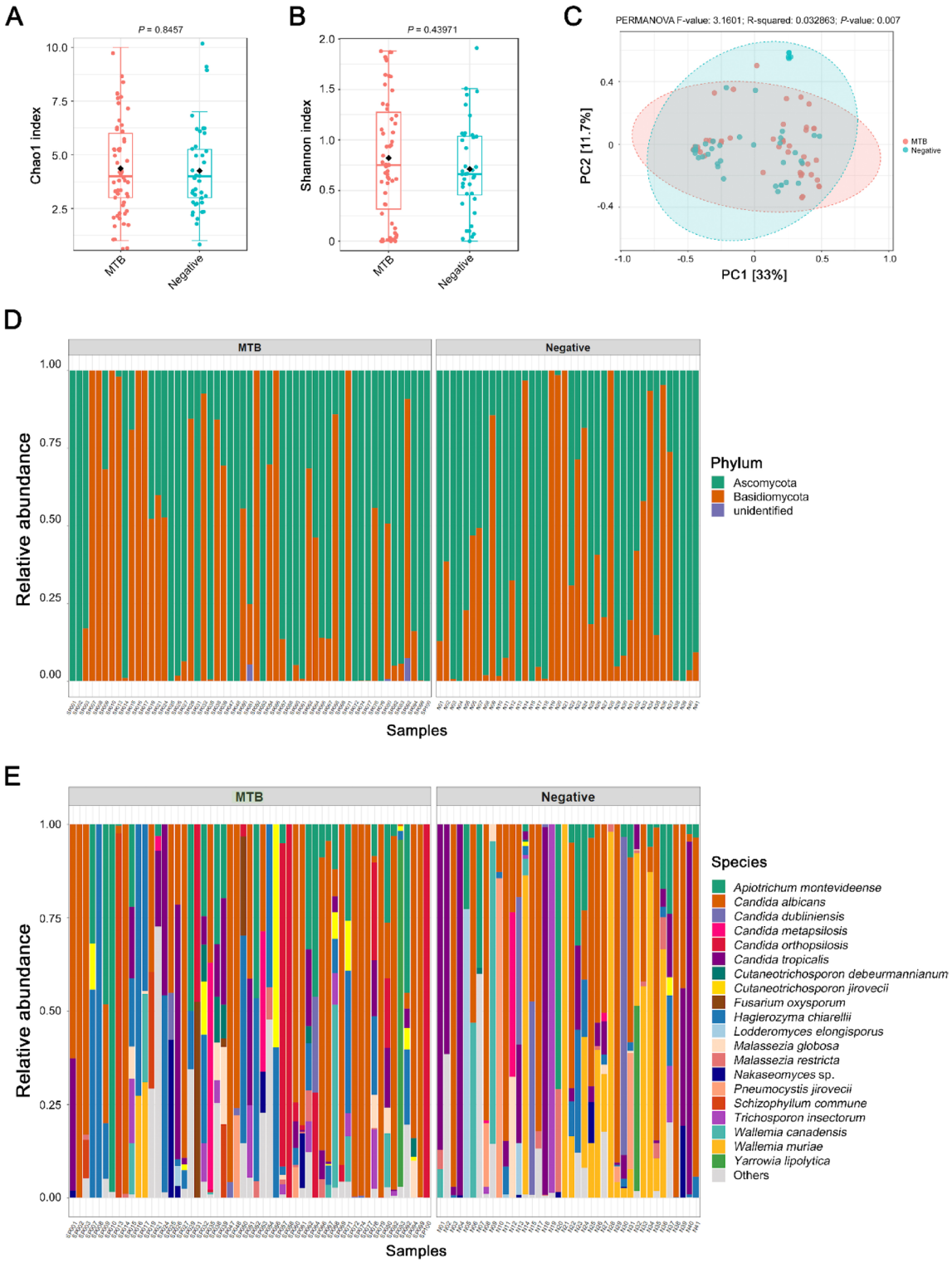

Fig. 3. Fungal diversity in MTB and negative samples. (A) Chao1 and (B) Shannon indices are shown as scatterplots. (C) β-diversity analysis. (D) Relative abundance of fungi at (D) phylum level and (E) top 20 species level. The taxa below the top 20 were classified as 'Others'.

Relative abundance of fungal microbiota based on ITS amplicon sequencing

The relative abundances of fungal microbiota are shown in Fig. 3D and E. The major fungal phyla (Fig. 3D) identified in sputum samples were consistent between two groups, with Ascomycota and Basidiomycota being the predominant phyla. The relative abundances of fungal species are presented in Fig. 3E, highlighting the top 20 most abundant species. The most prevalent fungal species identified across both groups included *Candida*

albicans, *Candida tropicalis*, and *Apiotrichum montevideense*. Notably, *Candida orthopsilosis* was detected only in the MTB group, while *Cutaneotrichosporon jirovecii* showed a higher abundance in the MTB group. Conversely, *Wallemia muriae* was more abundant in the negative group.

Differential abundance analysis of significant fungal species in sputum samples

The Venn diagram (Fig. 4A) illustrates the overall distribution of fungal taxa among the MTB and negative groups, showing that 67 taxa were unique to the MTB group and 19 were exclusive to the negative group, while 31 taxa were shared among all groups. Detailed information regarding these taxa is provided in Supplementary Table S4. The differential abundance analysis revealed significant variations in fungal species between both groups. Figure 4B highlights fungal taxa with notable differences in relative abundance, as determined by LEfSe. *Candida orthopsilosis* was exclusively found in the MTB group, whereas *Aureobasidium leucospermi* and *Wallemia muriae* were enriched in negative group. Additional details are provided in Supplementary Table S2. Relative abundance plots (Fig. 4C) supported these findings. *Candida orthopsilosis* was exclusively detected in the MTB group, whereas *Wallemia muriae* was significantly more abundant in the negative group, and *Aureobasidium leucospermi* was only found in the negative group. These differences were statistically significant (Mann-Whitney test, $P < 0.05$).

Co-occurrence network analysis of bacterial and fungal microbiota in sputum samples

The co-occurrence network analysis of bacterial and fungal microbiota in sputum samples revealed a complex microbial ecosystem (Fig. 5). The network was constructed using the SparCC algorithm at the species level, with a correlation threshold of 0.3 and a significance threshold of $P < 0.05$. In the network, nodes represent individual taxa, with the red portion indicating MTB samples and the blue portion representing negative samples. Edges denote significant correlations, classified as positive (green) or negative (orange). The correlation network analysis demonstrated distinct interaction patterns between bacterial and fungal taxa in the MTB and negative groups. *M. tuberculosis* exhibited positive correlations with *Stenotrophomonas maltophilia*, *Prevotella copri*, and *Klebsiella pneumoniae*, suggesting potential synergistic interactions that may contribute to microbial co-occurrence in the MTB group. In contrast, *M. tuberculosis* showed negative correlations with *Actinomyces odontolyticus* and *Prevotella melaninogenica*, indicating potential competitive interactions. Notably, *Stenotrophomonas maltophilia*, which was highly abundant in the MTB group, exhibited multiple correlations with diverse microbial species. Positive associations were observed with bacterial taxa, including *M. tuberculosis*, *Prevotella copri*, and *Klebsiella pneumoniae*, as well as fungal taxa, such as *Haglerozyma chiarellii* and *Apiotrichum montevideense*. These findings suggest that specific fungal species may co-exist with MTB-associated bacteria, potentially shaping the

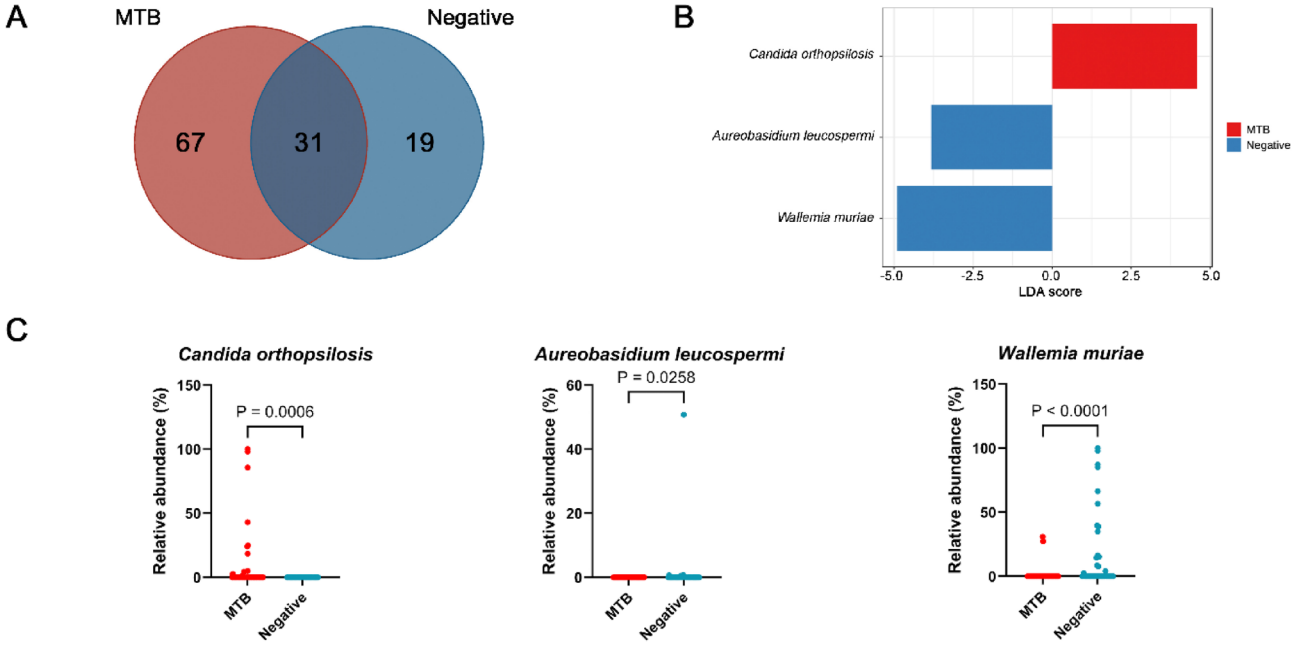

Fig. 4. Differential abundance analysis was performed using Linear Discriminant Analysis Effect Size (LEfSe) on MTB and negative sputum samples. (A) Venn diagram showing the number of fungal species identified through the metagenomic approach (ITS sequencing) in MTB samples (red) and negative samples (blue). (B) Bar plots display differentially abundant fungal taxa across various taxonomic ranks. The LDA score indicates the effect size and ranking of each taxon with significant differential abundance (LDA score ± 3, $P < 0.05$). (C) Relative abundance of the most prevalent fungal species. Scatter dot plots illustrate significant differences in the top abundant fungi between MTB group (red) and negative group (blue), determined by the Mann-Whitney test ($P < 0.05$).

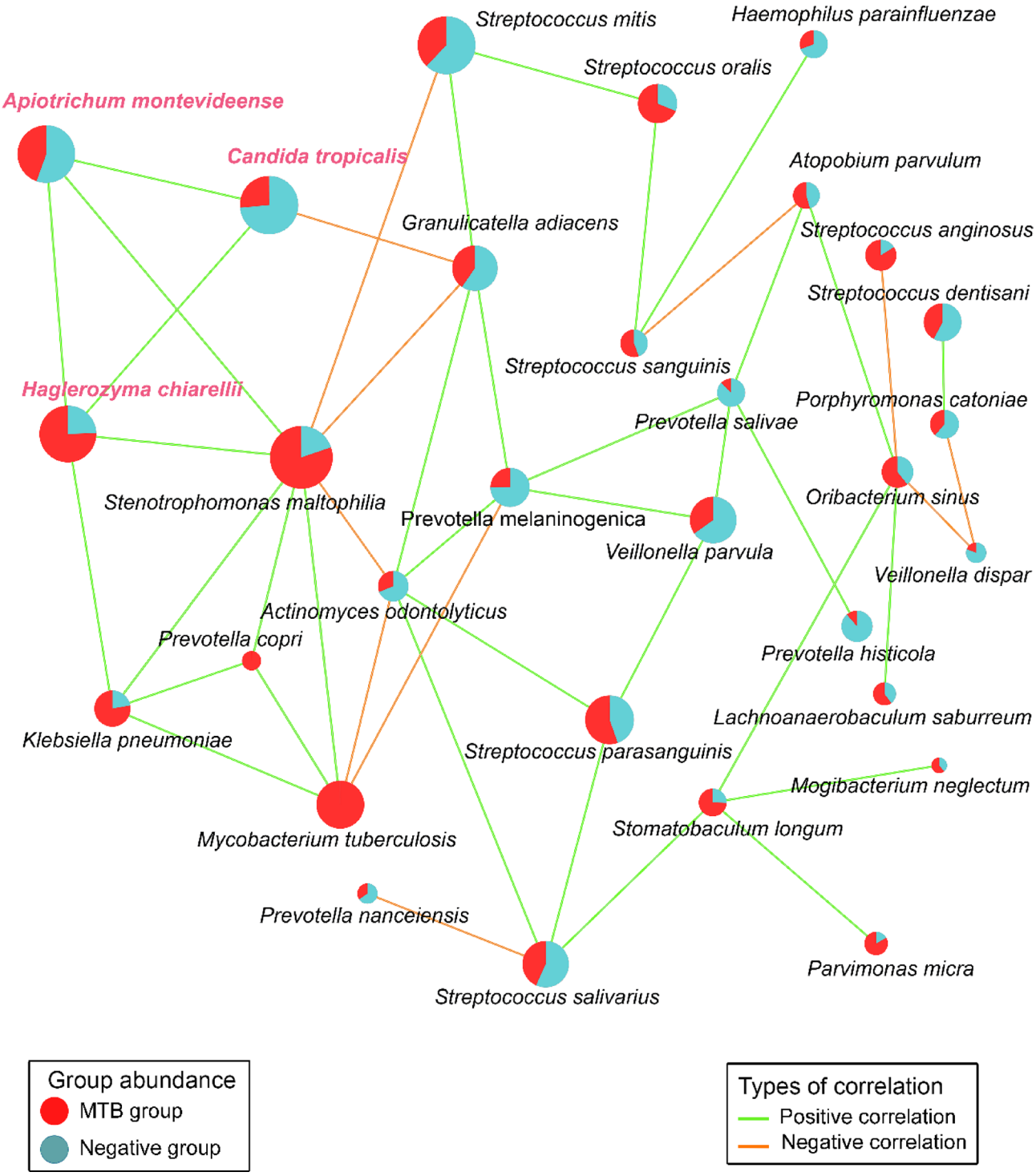

Fig. 5. Correlation network of bacterial and fungal taxa in MTB and negative sputum samples. The network visualizes interactions between bacterial and fungal taxa identified in MTB (red nodes) and negative (blue nodes) sputum samples. Each node represents an individual taxon, with node size reflecting its abundance. Edges indicate significant correlations, with green lines representing positive correlations and orange lines indicating negative correlations. Fungal taxa are labeled in pink, while bacterial taxa are labeled in black.

respiratory microbiome in tuberculosis infection. In contrast, *Stenotrophomonas maltophilia* exhibited negative correlations with *Streptococcus mitis* and *Granulicatella adiacens*, both of which were more abundant in the negative group.

Discussion

Amplicon sequencing with the ONT platform represents a culture-free approach offering numerous advantages. Key benefits include the generation of long-read data, the ability to accurately classify organisms at the species level, and a rapid processing workflow[11]. Furthermore, this method is cost-effective, attributed to the reusability of flow cells and its fast turnaround time. This study demonstrates the application of ONT-based clinical metagenomics for detecting bacterial and fungal microbiota through full-length 16S rDNA and ITS sequencing. Of the 56 MTB samples, 22 (39.29%) did not yield detectable *M. tuberculosis* DNA using ONT sequencing. While 16S rDNA sequencing is a well-established and robust method for microbiome analysis, one major limitation is its inability to differentiate between closely related species, particularly within the *Mycobacterium* genus, where genetic similarity is high[13]. Factors such as lower bacterial load or sample degradation could theoretically influence the results. The retrospective design of the study resulted in incomplete clinical data, including antibiotic usage, smoking history, HIV status, and comorbidities, which may have confounded microbiome analysis.

Alpha diversity analysis of bacterial communities revealed significant differences between MTB and negative groups, suggesting complex interactions within the existing bacterial community. This enrichment in MTB group could result from multiple factors. *M. tuberculosis* may alter the environment, thereby promoting the growth of certain bacterial taxa. Moreover, the host immune response to tuberculosis infection might indirectly influence the bacterial community structure[14,15]. Although an increase in richness was observed, the absence of a significant difference in the Shannon index suggests that while the number of bacterial species is impacted, the relative abundance distribution (evenness) remains similar between the groups. This indicates that tuberculosis infection predominantly affects species richness without disrupting the overall community balance. Beta diversity analysis of bacterial communities highlights the impact of *M. tuberculosis* infection on community composition. The distinct clustering of MTB and negative samples demonstrates compositional changes in bacterial communities caused by the infection. These changes may involve variations in the relative abundance of key bacterial taxa, the presence or absence of specific indicator species, or both. The analysis of bacterial microbiota identified four dominant phyla: Firmicutes, Proteobacteria, Actinobacteria, and Bacteroidetes, consistent with previous reports on the respiratory microbiome[16,17]. At the species level, *Streptococcus mitis*, *Streptococcus salivarius*, and *Veillonella parvula* were among the most abundant in both groups, reflecting their roles as core members of the respiratory microbiota and normal oral flora. However, these bacteria can act as opportunistic pathogens under certain conditions[18–20]. *M. tuberculosis* was only detected in the MTB group, validating the culture results. *Stenotrophomonas maltophilia* is a multidrug-resistant organism and one of the most common strains associated with intensive care unit-acquired and nosocomial infections[21]. In the present study, it exhibited a significantly higher relative abundance in the MTB group. A previous report also highlighted that the culture-based isolation rate of *Stenotrophomonas maltophilia* in TB patients indicates it is a major co-occurring species with *M. tuberculosis*[22]. Differential abundance analysis revealed distinct bacterial taxa associated with each group. *Stenotrophomonas maltophilia* and *Mycobacterium tuberculosis* were significantly enriched in MTB samples, whereas *Prevotella melaninogenica*, *Veillonella parvula*, *Corynebacterium striatum* and *Pseudomonas aeruginosa* were more abundant in negative samples. This *M. tuberculosis*-associated dysbiosis, characterized by depletion of commensals and enrichment of opportunistic pathogens, could have implications for respiratory tract susceptibility to secondary infections.

Fungal communities responded differently to tuberculosis infection. Similar fungal richness and evenness indices suggest that tuberculosis infection does not affect fungal community structure. However, significant differences in fungal beta diversity indicate compositional shifts, with distinct fungal taxa differentiating MTB from negative samples. Analysis of fungal communities revealed that Ascomycota and Basidiomycota were the predominant phyla in both groups, consistent with prior studies of the respiratory mycobiome[23]. Analysis revealed that *Candida orthopsilosis* was only detected in the MTB group. *Candida orthopsilosis*, a human fungal pathogen within the *Candida parapsilosis* species complex[24], may have a specific association with tuberculosis infection, as evidenced by its enrichment in MTB samples. Previous studies have also reported co-infections involving *Candida* species in tuberculosis patients[25,26]. In contrast, *Wallemia muriae* and *Aureobasidium leucospermi* were significantly enriched in the negative group, highlighting distinct fungal community structures.

The co-occurrence network analysis highlights intricate interactions within the respiratory microbiota influenced by tuberculosis infection. *M. tuberculosis* showed positive correlations with *Stenotrophomonas maltophilia*, *Prevotella copri and Klebsiella pneumoniae*, suggesting that synergistic interactions may be driven by shared metabolic pathways, immune evasion, or resource availability[27,28]. This co-occurrence may promote infection persistence or worsen disease progression. Conversely, *M. tuberculosis* displayed negative correlations with *Actinomyces odontolyticus* and *Prevotella melaninogenica*, indicating competitive interactions. While *Actinomyces odontolyticus* is a rare respiratory pathogen[29], co-infections with *M. tuberculosis* have been scarcely reported[30]. Previous study reported an increase in *Prevotella melaninogenica* abundance following anti-TB treatment[31]. Similarly, Hu et al. observed elevated *Prevotella* levels after successful TB treatment[32]. The reduced abundance of *Prevotella melaninogenica* in untreated TB patients may impair lung function and exacerbate inflammation[33]. In negative samples, fungal taxa like *Candida tropicalis* and *Apiotrichum montevidense* were more abundant. *Apiotrichum montevidense* correlated positively with *Haglerozyma chiarellii*, *Candida tropicalis* and *Stenotrophomonas maltophilia*. Interestingly, *Haglerozyma chiarellii* was high abundant in MTB samples. These findings suggest that specific fungal species may either co-exist with or compete against MTB-associated bacteria, potentially influencing the composition of the respiratory microbiome in tuberculosis infection. The fungal enrichment in culture-negative samples may reflect opportunistic growth in patients presenting with symptoms that often resemble mycobacterial infections, as fungal infections are known to mimic tuberculosis and other clinical conditions[34].

Our study demonstrates the effectiveness of ONT-based amplicon sequencing in analyzing bacterial and fungal microbiota at the species level in tuberculosis infection. MTB group exhibited higher bacterial diversity and distinct community structures, while fungal diversity remained stable, showing some compositional changes. These results reveal complex microbial interactions influenced by *M. tuberculosis*, highlighting their potential impact on infection and disease progression.

Methods

Clinical experiment statement

This study was approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University (COA no. 1094/2024; IRB no. 0451/67; Bangkok, Thailand). All methods were carried out in accordance with relevant guidelines and regulations. Informed consent was obtained from all participants included in the study.

Study cohort

Leftover sputum samples from patients with suspected TB infection at King Chulalongkorn Memorial Hospital (Bangkok, Thailand) between 2019 and 2024 were used in this study. A total of 161 samples were divided into two cohorts, including MTB and negative groups based on sputum culture and the GenoType® Mycobacterium CM/AS/NTM-DR line probe assay (Hain Lifescience GmbH, Germany) to identify the mycobacterial species. The inclusion criteria included participants aged ≥18 years, with sputum samples from patients suspected of tuberculosis infection based on clinical presentations or findings from additional investigations (e.g., chest X-rays) as determined by physicians. Only sputum samples with bacterial culture results were included. Exclusion criteria included patients aged < 18 years, patients infected with non-tuberculous mycobacteria (NTM), sputum samples without culture results, and samples with a volume of less than 200 µL. The final number of samples was 97, consisting of 41 negative samples and 56 MTB samples.

Mycobacterial culture

Clinical specimens were processed using the sodium hydroxide-N-acetyl-L-cysteine-sodium citrate (NaOH/NALC-Na citrate) method[35]. Briefly, specimens were decontaminated and digested by mixing them with an equal volume of NaOH/NALC-Na citrate solution (2% NaOH, 0.5% NALC, and 1.45% Na-citrate) for 15 min at room temperature. Phosphate buffer (pH 6.8) was then added to adjust the final volume to 40 mL, followed by mixing through inversion and centrifugation at 3,000 ×g for 15 min at 12 °C. After centrifugation, the supernatant was carefully discarded, and the remaining sediment was resuspended in 2.5 mL of fresh phosphate buffer. A 500 µL aliquot of the sediment was inoculated into mycobacterial growth indicator tubes and incubated in the BD BACTEC MGIT 960 System (Becton Dickinson, USA) for up to six weeks or until a positive signal was detected. The GenoType® Mycobacterium CM/AS/NTM-DR line probe assay (Hain Lifescience GmbH, Germany) was performed according to the manufacturer's instructions to identify mycobacterial species.

DNA extraction

Genomic DNA was extracted from sputum samples using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, USA) and a FastPrep-24™ 5G homogenizer (MP Biomedicals, USA). Each 200 µL sputum sample was first treated with 0.1% dithiothreitol (final concentration of 0.01% or 0.65 mM) and vortexed for 20 s to ensure thorough mixing. The specimen was then incubated on a shaker at 200 rpm and 37°C for 20 min, following the protocol to achieve a final elution volume of 50 µL per sample The concentration of genomic dsDNA was measured using NanoPhotometer® C40 (Implen, Germany). The extracted DNA was stored at − 20 °C for future analysis.

Bacterial amplification

The full-length bacterial 16S rDNA (V1-V9 regions) was amplified using primers modified from a previous study[36], incorporating 3' specific target sequences (italized) and 5' adaptors. The primers used were 16S_27 F: 5'-TTTCTGTTGGTGCTGATATTGC*AGRGTTYGATYMTGGCTCAG*–3' and 16S_1492R: 5'-ACTTGCCT GTCGCTCTATCTTC*CGGGY* TACCTTGTTACGACTT–3'. PCR amplification was performed using the Ultra HiFidelity PCR Kit (Tiangen®, China). The 10 µL reaction mixture consisted of 5 µL of 2×Ultra HiFi mix, 2 µL of PCR Enhancer, 0.25 µM of each forward and reverse primer, 1.5 µL of DEPC-treated water, and 10 ng of nucleic acid template. The amplification conditions were as follows: initial denaturation at 98 °C for 30 s, followed by 30 cycles of 98 °C for 10 s, 60 °C for 10 s, and 72 °C for 45 s, with a final extension at 72 °C for 5 min. Amplicon barcoding was performed using a five-cycle PCR with barcode primers from the PCR Barcoding Expansion 1–96 kit (EXP-PBC096; ONT, UK). The barcoding PCR conditions included an initial denaturation at 98 °C for 30 s, followed by five cycles of 98 °C for 10 s, 60 °C for 10 s, 72 °C for 50 s, and a final extension at 72 °C for 5 min. The PCR products were separated by 1% agarose gel electrophoresis using RedSafe™ Nucleic Acid Staining Solution, then purified with a QIAquick Gel Extraction Kit (QIAGEN, Germany) according to the manufacturer's guidelines.

Fungal amplification

The full-length fungal ITS region (ITS1 and ITS2) was amplified using primers adapted from a previous study[37,38], incorporating 3' specific target sequences (italized) and 5' adaptors as follows: ITS_1 F: 5'-TTTCTG TTGGTGCTGATATTGCTCCGTAGGTGAACCTGCGG–3' and ITS_4R 5'-ACTTGCCTGTCGCTCTATCTT CTCCTCCGCTTATTGATATGC–3'. PCR amplification was carried out using the KOD One™ PCR Master Mix (Toyobo, Japan). The total 10 µL PCR reaction contained 5 µL of KOD One™ PCR Master Mix, 0.3 µL of each 10 µM forward and reverse primer, 3.4 µL of DEPC-treated water, and 10 ng of genomic DNA. The PCR cycling conditions consisted of an initial denaturation at 98 °C for 2 min, followed by 35 cycles of 98 °C for 10 s, 60 °C for

10 s, and 68 °C for 10 s, with a final extension at 68 °C for 5 min. For amplicon barcoding, a five-cycle PCR was performed using the PCR Barcoding Expansion 1–96 kit (ONT, UK). The barcoding PCR conditions included an initial denaturation at 98 °C for 2 min, followed by five cycles of 98 °C for 10 s, 60 °C for 10 s, and 68 °C for 10 s, with a final extension at 68 °C for 5 min. After amplification, the PCR products were analyzed by 1% agarose gel electrophoresis and subsequently purified using a QIAquick Gel Extraction Kit (QIAGEN, Germany).

Library preparation for nanopore sequencing

The DNA library was then purified using the QIAquick PCR Purification Kit (Qiagen, Germany) and measured for quantity by Qubit® dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific, USA) and measured on a Qubit® 4.0 Fluorometer (Thermo Fisher Scientific, USA). The library was pooled and size-selected using 0.5 × (for 16S) and 0.8 × (for ITS) Agencourt AMPure XP beads (Beckman Coulter, USA). A total of 1 µg of the barcoded library was ligated to nanopore adaptors using the SQK-LSK114 Ligation Sequencing Kit (ONT, UK) and subsequently sequenced on a flow cell R10.4.1 with the MinION Mk1C sequencer (ONT, UK).

Bioinformatic analysis

The FAST5 files from Nanopore sequencing were basecalled using Guppy basecaller v6.5.7 (ONT, UK) in super-accuracy mode with a minimum Q-score of 15. The quality of the obtained FASTQ sequences was evaluated with MinIONQC v1.4.1[39]. Demultiplexing and adaptor trimming of the FASTQ reads were conducted with Porechop v0.2.4 (available at https://github.com/rrwick/Porechop). Each sample's filtered reads were subsequently used for reads clustering, polishing, and taxonomy identification via the NanoCLUST pipeline[40]. Taxonomic categorization was performed using the Blastn algorithm, with reference sequences obtained from the RDP version 2.1.3 (16S rDNA gene sequences) and SILVA version 138.1 (ITS region sequences) databases. Total Sum Scaling (TSS) normalization was applied to account for sequencing depth variations across samples, ensuring that the relative abundance of taxa was comparable for downstream analyses. For filtering, all taxa were retained without exclusion of low-abundance taxa to preserve potentially biologically relevant rare taxa. The bacterial and fungal abundance outputs were further examined using QIIME2 (https://qiime2.org/) v2021.2[41], which generated a collapsed taxonomy table. In addition, MicrobiomeAnalyst v2.0 was used to assess the microbiota's alpha and beta diversity[42]. Differential abundance analysis was performed using the LEfSe method, with an LDA score threshold of ± 4 and a *P*-value of 0.05 for bacteria and ± 3 with a *P*-value of 0.05 for fungi[43]. The co-occurrence network analysis of bacterial and fungal microbiota within sputum samples was carried out using the Sparse Correlations for Compositional Data (SparCC)[44] method (correlation threshold 0.3, *P*-value threshold 0.05) with the plugin in MicrobiomeAnalyst[42].

Data availability

The sequencing datasets from the current study are publicly available in the NCBI Sequence Read Archive (SRA), BioProject ID: PRJNA1213270 for 16S bacterial datasets while BioProject ID: PRJNA1214521 for ITS fungal datasets. The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 19 February 2025; Accepted: 9 May 2025
Published online: 22 May 2025

References

1. World Health Organization. Global Tuberculosis Report 2024. (2024).
2. Houben, R. M. & Dodd, P. J. The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. *PLoS Med.* **13**, e1002152. https://doi.org/10.1371/journal.pmed.1002152 (2016).
3. Menzies, N. A. et al. Progression from latent infection to active disease in dynamic tuberculosis transmission models: a systematic review of the validity of modelling assumptions. *Lancet Infect. Dis.* **18**, e228–e238. https://doi.org/10.1016/s1473-3099(18)30134-8 (2018).
4. Gopalaswamy, R., Shanmugam, S., Mondal, R. & Subbian, S. Of tuberculosis and non-tuberculous mycobacterial infections – a comparative analysis of epidemiology, diagnosis and treatment. *J. Biomed. Sci.* **27** https://doi.org/10.1186/s12929-020-00667-6 (2020).
5. Shu, W. et al. A case series of co-infection in *Mycobacterium tuberculosis* and other pathogens: insights from nanopore sequencing. *Egypt. J. Bronchol.* **18** https://doi.org/10.1186/s43168-024-00270-5 (2024).
6. Aldriwesh, M. G. et al. Coinfection of pulmonary tuberculosis with other lower respiratory tract infections: A retrospective cross-sectional study. *Ann. Thorac. Med.* **17**, 229–236. https://doi.org/10.4103/atm.atm_200_22 (2022).
7. Attia, E. F. et al. Tuberculosis and other bacterial co-infection in Cambodia: a single center retrospective cross-sectional study. *BMC Pulm Med.* **19** https://doi.org/10.1186/s12890-019-0828-4 (2019).
8. Ngekeng, S. et al. High prevalence of bacterial pathogens in sputum of tuberculosis suspected patients in Buea. *Microbiol. Res. J. Int.* **11**, 1–8. https://doi.org/10.9734/BMRJ/2016/22426 (2015).
9. Dong, B. et al. Improved conventional and new approaches in the diagnosis of tuberculosis. *Front. Microbiol.* **13**, 924410. https://doi.org/10.3389/fmicb.2022.924410 (2022).
10. Charalampous, T. et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat. Biotechnol.* **37**, 783–792. https://doi.org/10.1038/s41587-019-0156-5 (2019).
11. Ciuffreda, L., Rodríguez-Pérez, H. & Flores, C. Nanopore sequencing and its application to the study of microbial communities. *Comput. Struct. Biotechnol. J.* **19**, 1497–1511. https://doi.org/10.1016/j.csbj.2021.02.020 (2021).
12. Dippenaar, A. et al. Nanopore sequencing for mycobacterium tuberculosis: a critical review of the literature, new developments, and future opportunities. *J. Clin. Microbiol.* **60**, e0064621. https://doi.org/10.1128/jcm.00646-21 (2022).
13. Beye, M., Fahsi, N., Raoult, D. & Fournier, P. E. Careful use of 16S rRNA gene sequence similarity values for the identification of *Mycobacterium* species. *New. Microbes New. Infect.* **22**, 24–29. https://doi.org/10.1016/j.nmni.2017.12.009 (2018).
14. Naidoo, C. C. et al. The Microbiome and tuberculosis: state of the art, potential applications, and defining the clinical research agenda. *Lancet Respir Med.* **7**, 892–906. https://doi.org/10.1016/S2213-2600(18)30501-0 (2019).

15. Hong, B. Y. et al. Microbiome changes during tuberculosis and antituberculous therapy. *Clin. Microbiol. Rev.* **29**, 915–926. https://doi.org/10.1128/cmr.00096-15 (2016).
16. Sahin-Yilmaz, A. & Naclerio, R. M. Anatomy and physiology of the upper airway. *Proc. Am. Thorac. Soc.* **8**, 31–39. https://doi.org/10.1513/pats.201007-050RN (2011).
17. Huffnagle, G. B., Dickson, R. P. & Lukacs, N. W. The respiratory tract Microbiome and lung inflammation: a two-way street. *Mucosal Immunol.* **10**, 299–306. https://doi.org/10.1038/mi.2016.108 (2017).
18. Jin, D. J. & Zhou, Y. N. *In Methods in Enzymology*Vol. 273300–319 (Academic, 1996).
19. Zhou, L. N., Ge, X. L., Dong, T. T., Gao, H. Y. & Sun, B. H. Antibacterial steroidal alkaloids from Holarrhena antidysenteriaca. *CJNM* **15**, 540–545. https://doi.org/10.1016/S1875-5364(17)30080-8 (2017).
20. Delorme, C., Abraham, A. L., Renault, P. & Guédon, E. Genomics of *Streptococcus salivarius*, a major human commensal. *Infect. Genet. Evol.* **33**, 381–392. https://doi.org/10.1016/j.meegid.2014.10.001 (2015).
21. Mojica, M. F. et al. Clinical challenges treating *Stenotrophomonas maltophilia* infections: an update. *JAC-Antimicrobial Resist.* **4** https://doi.org/10.1093/jacamr/dlac040 (2022).
22. Li, Y. et al. Effect of *Stenotrophomonas maltophilia* on tuberculosis. *Microbiol. Spectr.* **11**, e0094423. https://doi.org/10.1128/spectrum.00944-23 (2023).
23. Nash, A. K. et al. The gut mycobiome of the human Microbiome project healthy cohort. *Microbiome* **5** https://doi.org/10.1186/s40168-017-0373-4 (2017).
24. Tavanti, A., Davidson, A. D., Gow, N. A., Maiden, M. C. & Odds, F. C. *Candida orthopsilosis* and *Candida metapsilosis* spp. Nov. To replace *Candida parapsilosis* groups II and III. *J. Clin. Microbiol.* **43**, 284–292. https://doi.org/10.1128/jcm.43.1.284-292.2005 (2005).
25. Amala, S. E., Hanson, A. & Wokem, G. N. *Candida* co-infection with *Mycobacterium tuberculosis* in tuberculosis patients and antifungal susceptibility of the isolates. *J. Tuberc Res.* **8**, 53–65 (2020).
26. Chen, X. H. et al. Tuberculosis infection might increase the risk of invasive candidiasis in an inmunocompetent patient. *Rev. Inst. Med. Trop. Sao Paulo.* **57**, 273–275. https://doi.org/10.1590/s0036-46652015000300016 (2015).
27. Warner, D. F. *Mycobacterium tuberculosis* metabolism. *Cold Spring Harb Perspect. Med.* **5** https://doi.org/10.1101/cshperspect.a021121 (2014).
28. Llibre, A. et al. Tuberculosis alters immune-metabolic pathways resulting in perturbed IL-1 responses. *Front. Immunol.* **13**, 897193. https://doi.org/10.3389/fimmu.2022.897193 (2022).
29. Wang, L. et al. Pulmonary lesions associated with sputum culture-positive actinomycetes: report of one case. *Ann. Transl Med.* **7** (23), 793. https://doi.org/10.21037/atm.2019.12.38 (2019).
30. Balis, E. et al. Presentation of pulmonary tuberculosis and actinomyces co-infection as a lung mass: a literature review and unique case report. *Monaldi Arch. Chest Dis.* **7;89** (3). https://doi.org/10.4081/monaldi.2019.1180 (2019).
31. Xiao, G. et al. Insights into the unique lung microbiota profile of pulmonary tuberculosis patients using metagenomic next-generation sequencing. *Microbiol. Spectr.* **10** (1), e0190121. https://doi.org/10.1128/spectrum.01901-21 (2022).
32. Hu, Y. et al. Metagenomic analysis of the lung Microbiome in pulmonary tuberculosis - a pilot study. *Emerg. Microbes Infect.* **9** (1), 1444–1452 (2020).
33. O'Neill, K. et al. Reduced bacterial colony count of anaerobic bacteria is associated with a worsening in lung clearance index and inflammation in cystic fibrosis. *PLoS One.* **20** (5), e0126980. https://doi.org/10.1371/journal.pone.0126980 (2015).
34. Ekeng, B. E. et al. Pulmonary and extrapulmonary manifestations of fungal infections misdiagnosed as tuberculosis: the need for prompt diagnosis and management. *J. Fungi (Basel).* **8** https://doi.org/10.3390/jof8050460 (2022).
35. Sawatpanich, A. et al. Diagnostic performance of the anyplex MTB/NTM real-time PCR in detection of *Mycobacterium tuberculosis* complex and nontuberculous mycobacteria from pulmonary and extrapulmonary specimens. *Heliyon* **8**, e11935. https://doi.org/10.1016/j.heliyon.2022.e11935 (2022).
36. Sawaswong, V. et al. Full-length 16S rDNA sequencing based on Oxford nanopore technologies revealed the association between gut-pharyngeal microbiota and tuberculosis in cynomolgus macaques. *Sci. Rep.* **14**, 3404. https://doi.org/10.1038/s41598-024-53880-w (2024).
37. White, T. et al. Vol. 31 315–322 (1990).
38. Gardes, M. & Bruns, T. D. ITS primers with enhanced specificity for basidiomycetes–application to the identification of mycorrhizae and rusts. *Mol. Ecol.* **2**, 113–118. https://doi.org/10.1111/j.1365-294x.1993.tb00005.x (1993).
39. Lanfear, R., Schalamun, M., Kainer, D., Wang, W. & Schwessinger, B. MinIONQC: fast and simple quality control for minion sequencing data. *Bioinform* **35**, 523–525. https://doi.org/10.1093/bioinformatics/bty654 (2019).
40. Rodríguez-Pérez, H., Ciuffreda, L. & Flores, C. NanoCLUST: a species-level analysis of 16S rRNA nanopore sequencing data. *Bioinform* **37**, 1600–1601. https://doi.org/10.1093/bioinformatics/btaa900 (2021).
41. Bolyen, E. et al. Reproducible, interactive, scalable and extensible Microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857. https://doi.org/10.1038/s41587-019-0209-9 (2019).
42. Chong, J., Liu, P., Zhou, G. & Xia, J. Using Microbiomeanalyst for comprehensive statistical, functional, and meta-analysis of Microbiome data. *Nat. Protoc.* **15**, 799–821. https://doi.org/10.1038/s41596-019-0264-1 (2020).
43. Segata, N. et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* **12** (R60). https://doi.org/10.1186/gb-2011-12-6-r60 (2011).
44. Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* **8**, e1002687. https://doi.org/10.1371/journal.pcbi.1002687 (2012).

Acknowledgements

The authors would like to thank all members of the Center of Excellence in Systems Microbiology (CESM), Faculty of Medicine, Chulalongkorn University, for their technical support and assistance with the experiments. The authors also extend their gratitude to all the staff of the Department of Microbiology, Faculty of Medicine, Chulalongkorn University, for their support in providing samples. We extend our gratitude to the Chulalongkorn University Center for Artificial Intelligence in Medicine (CU-AIM) for providing high-performance computing (HPC) clusters that facilitated the bioinformatics analysis.

Author contributions

P.T. conceived the study design, conducted the experiments, and drafted the manuscript. S.V. and P.Kl. performed the bioinformatics analysis. P.C. and S.S. contributed to the experimental works. V.S. analyzed and visualized the data. P.S. and A.K. reviewed and revised the manuscript. S.R. provided resources and supervised the study. P.Ka. supervised the study and reviewed the manuscript. S.P. led the study design and conceptualization, supervised the study, reviewed and revised the manuscript, and secured funding. All authors reviewed and approved the final manuscript.

Funding
This research project is supported by the Second Century Fund (C2F), Chulalongkorn University; Thailand Science research and Innovation Fund, Chulalongkorn University (Grant Number: HEA_FF_68_100_3000_016); the Ratchadapisek Sompotch Fund, Faculty of Medicine, Chulalongkorn University (Grant Number: GA68/047); The 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund); the National Science, Research and Innovation Fund (NSRF) via the Program Management Unit for Human Resources and Institutional Development, Research and Innovation (Grant Number: B05-F640122).

Declarations

Competing interests
The authors declare no competing interests.

Ethical approval
This study related to humans was approved by the Institutional Review Board (IRB) of the Faculty of Medicine, Chulalongkorn University with IRB No. 0451/67.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-01905-3.

Correspondence and requests for materials should be addressed to S.P.

Reprints and permissions information is available at www.nature.com/reprints.

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