Optimisation and clinical validation of a metagenomic third-generation sequencing approach for aetiological diagnosis in bronchoalveolar lavage fluid of patients with pneumonia

Shanshan Zhang,[a,b,l] Xi Li,[c,l] Xia Li,[d,e,l] Ying Fu,[f,l] Long Chen,[d,e] Wei Wang,[d,e] Qinqing Lin,[a] Hui Lou,[a] Yake Yao,[a] Wenyu Chen,[g] Cheng Zhong,[h] Jian Ye,[i] Yinan Yao,[a] Huimin Guo,[d,e,***] Yunsong Yu,[j,**] and Hua Zhou[a,k,*]

[a]Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[b]Department of Respiratory and Critical Care Medicine, Beilun People's Hospital, Ningbo, China
[c]Laboratory Medicine Centre, Department of Clinical Laboratory, Zhejiang Provincial People's Hospital, Affiliated People's Hospital, Hangzhou Medical College, Hangzhou, China
[d]Zhejiang Digena Diagnosis Technology Co., Ltd., Hangzhou, China
[e]Hangzhou Digena Medical Laboratory Co., Ltd., Hangzhou, China
[f]Department of Clinical Laboratory, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China
[g]Department of Respiratory Medicine, The First Hospital of Jiaxing, Affiliated Hospital of Jiaxing University, Jiaxing, China
[h]Department of Respiratory and Critical Care Medicine, Zhejiang Chinese Medicine and Western Medicine Integrated Hospital, Hangzhou, 310003, China
[i]Department of Respiratory and Critical Care Medicine, Zhejiang Hospital, Hangzhou, China
[j]Department of Infectious Disease, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, 310016, China
[k]Department of Respiratory Medicine, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200127, China

eBioMedicine
2025;116: 105752
Published Online xxx
https://doi.org/10.1016/j.ebiom.2025.105752

Summary

Background Metagenomic Third Generation Sequencing (mTGS), based on nanopore technology, has emerged as a promising tool for the rapid diagnosis of pneumonia pathogens. However, this technology currently lacks standardised technical protocols, quality control measures, and comprehensive performance evaluations for the simultaneous detection of bacteria, fungi, and viruses in clinical settings.

Methods We optimised the mTGS workflow by refining key parameters (cell wall lysis, fragment size selection, host DNA depletion, and sequencing depth) using reference samples and bronchoalveolar lavage fluid (BALF) from eight patients with pneumonia. These optimisations formed the basis for a standardised mTGS protocol. To assess the clinical diagnostic value of the optimised mTGS, a multicentre prospective cohort study involving 313 pneumonia-suspected patients was conducted. Each BALF sample was tested using conventional microbiological testing (CMTs), metagenomic next-generation sequencing (mNGS), pre-optimised mTGS, and optimised mTGS.

Findings The optimised mTGS protocol, based on the refined cell wall lysis, fragment size selection, no host DNA depletion, and 800 MB sequencing depth, achieved a tenfold increase in sensitivity compared with pre-optimised mTGS for detecting the species of *Bacillus subtilis*, *Mycobacterium tuberculosis*, *Mycobacterium avium*, *Cryptococcus neoformans*, and *Human papillomavirus* in reference samples. In the prospective cohort, 274 patients with a confirmed diagnosis of pneumonia were identified, yielding 376 distinct microbes. The mTGS identified more microbes than CMTs (314 vs. 115), with a 45.30% increase in sensitivity (84.70% vs. 39.40%, $P < 0.01$, Chi-square test/Fisher's exact test). Compared with pre-optimised mTGS, the sensitivity of optimised mTGS increased by 32.51% (84.70% vs. 52.19%, $P < 0.01$, Chi-square test/Fisher's exact test). mTGS showed comparable performance to mNGS (84.70% vs. 79.90%, $P = 0.14$,Chi-square test/Fisher's exact test), both significantly outperforming CMTs. mNGS was more sensitive for detecting *Non-tuberculous mycobacteria*, *Pneumocystis jirovecii*, and *Aspergillus* spp., while mTGS demonstrated higher sensitivity for *M. tuberculosis*, *Chlamydia psittaci*, and *Streptococcus pneumoniae*. The overall diagnostic agreement between mTGS and clinical diagnosis was 81.80%.

*Corresponding author.
**Corresponding author.
***Corresponding author. Zhejiang Digena Diagnosis Technology Co., Ltd., Hangzhou, China.
E-mail addresses: zhouhua1@zju.edu.cn (H. Zhou), yvys119@zju.edu.cn (Y. Yu), guohuimin@digena.com.cn (H. Guo).
[l]Authors contributed equally.

Interpretation We optimised and validated a standardised mTGS protocol that significantly improved the ability to detect pathogens in the BALF of patients with pneumonia. Optimised mTGS demonstrated comparable performance to mNGS, making it a promising tool for the aetiological diagnosis of pneumonia.

Funding The Research and Development Programme of Zhejiang Province (2023C03068, 2024C03187), the National Natural Science Foundation of China (82272338), the Key R&D Plan of the Ministry of Science and Technology (China) of China (2022YFC2504502).

Keywords: Metagenomic third-generation sequencing; Metagenomic next-generation sequencing; Bronchoalveolar lavage fluid; Pneumonia; Aetiological diagnosis

Research in context

Evidence before this study

The technical advance and popular use of shotgun metagenomic next-generation sequencing (mNGS) in pathogen identification have laid a good foundation for the trending application of metagenomic third-generation sequencing (mTGS) in respiratory infections. Currently, based on the rapid, real-time, and flexible sequencing characteristics of nanopore sequencing technology, there is a rising concern about the technical requirements for the standardised conduct of mTGS and the differences among mNGS, traditional assays in routine practice. At present, there is relatively little exploration on the improvement of mTGS methodology in clinical setting, and there is also a lack of research comparing mTGS, mNGS, and conventional microbiological tests (CMTs) in pathogen diagnosis using multicentre prospective pneumonia cohorts. However, the extent to which the mTGS methods contribute to clinical decision-making and their value in pathogen detection remain important questions.

Added value of this study

In this study, two progressive research studies were designed: the mTGS technical process was first established and optimised to determine the relevant quality control parameters. The analytical performance parameters (detection limit, sensitivity, PPV and NPV) were evaluated before and after the optimisation of mTGS. And then a total of 313 bronchoalveolar lavage fluid (BALF) samples were included in the clinical evaluation. Notably, the performance of mTGS was significantly improved over pre-mTGS for the detection of pneumonia pathogens, and the overall performance was equal to that of mNGS and significantly higher than CMTs. The mTGS technology provides an alternative and reliable means of detecting the pathogenicity of lung infections.

Implications of all the available evidence

Our data demonstrated that the optimised mTGS has an equivalent detection performance to mNGS and can be used as a broad-spectrum, rapid, accurate and simple method in pneumonia. Compared with mNGS, mTGS had lower sensitivity on some species like *Mycoplasma hominis* and *Aspergillus* spp., whereas had higher detection capability on the species of *Mycobacterium tuberculosis* and *Cryptococcus neoformans*. The results obtained from the analytical performance and clinical evaluation data will enhance overall understanding of the mTGS technology process, including cell wall lysis, fragment size selection, host DNA depletion conditions and sequencing data volume, and detection performance in clinical samples. mTGS achieves the same efficacy as mNGS in pathogen detection, and also has the natural advantages of easy operation, high specificity, and simple analysis, which has the potential to achieve on-demand and near-instant detection of pathogens in the clinical setting. These results suggest that mTGS has large clinical application prospects in the field of infections in the near future.

Introduction

Accurate aetiological diagnosis is essential for the effective treatment of pneumonia. Traditional microbial culture methods have a low detection rate (less than 30%) and time consuming (over 48 h), which makes them hard to meet the need for rapid and precise clinical pathogen diagnosis on bacteria, fungi or viruses.[1] Although serological tests and polymerase chain reaction (PCR) can detect specific microbes, they are limited to a narrow range of targets and cannot comprehensively identify pathogens.[2] Metagenomic next-generation sequencing (mNGS) has emerged as a powerful tool to overcome these limitations, enabling the rapid and comprehensive detection of a wide variety of microbes in clinical samples.[3,4] However, current mNGS technologies are hampered by short read lengths (50–100 bp), which can limit their ability to accurately distinguish between similar organisms. Additionally, mNGS is costly, complex, and not well-suited for rapid, point-of-care diagnostics.[5] Despite these challenges, mNGS has gained clinical traction, with established guidelines and standardised technical parameters for detecting

infectious pathogens.[6–12] Compared with mNGS, nanopore-based third-generation sequencing (mTGS) offers transformative advantages (including long reads, rapid sequencing, and real-time data analysis) in pneumonia pathogen detection, such as viruses, fungi, and *M. tuberculosis*. Nanopore sequencing analysis is able to be initiated before the sequencing run is completed, based on its single-molecule sequencing characteristic, reducing the report time from 48 to 72 h to 6 h.[13] The clinical application prospect of mTGS is promising in the pathogen diagnosis, providing a simpler, more cost-effective, and highly specific alternative to second-generation sequencing, with the potential for on-demand, near-instantaneous pathogen detection in clinical settings.[14,15] Despite its promise, mTGS still lacks a standardised technical framework and quality control processes, which limits its widespread clinical adoption. As a result, there is an urgent need for the development of standardised protocols to optimize mTGS for clinical pathogen identification.

Based on the current mTGS process, we did some improvement on the parameters and quality control procedures for mTGS in identifying microbes in bronchoalveolar lavage fluid (BALF) from patients with pneumonia. This will be achieved by using reference materials to fine-tune the method, followed by an initial verification of detection efficiency with a small cohort of clinical samples. Finally, we will confirm the standardised mTGS quality control process and parameters in a prospective cohort of patients with pneumonia.

Methods

Ethics approval

This study was approved by clinical research ethics committee of the First Affiliated Hospital, Zhejiang University School of Medicine (IIT20220389B and IIT20250192A). Informed consent was obtained from all participants.

Current process of mTGS

Building on previous studies,[16] we developed a nanopore metagenome sequencing process system (pre-mTGS) for analysing BALF samples (see Supplementary 1 for the detailed protocol). The key steps include: sample pretreatment, nucleic acid extraction and purification, DNA library preparation, sequencing, bioinformatics analysis, and determination of positive results.

Optimisation of mTGS using reference samples

Preparation of reference samples

A mixture of thirteen microbes, each at a concentration of 5×10^4 CFU/mL or copies/mL, was combined with THP1 cells (5×10^5 copies/mL) in PBS. The microbes included in the reference sample were: *Listeria monocytogenes*, *Pseudomonas aeruginosa*, *Bacillus subtilis*, *Escherichia coli*, *Salmonella enterica*, *Lactobacillus fermentans*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Saccharomyces cerevisiae*, *C. neoformans* (Purchased from ZymoBIOMICS,D6300), *M. tuberculosis*, *Mycobacterium avium* (Purchased from National Institutes for Food and Drug Control, 30030-201703), and *Human papillomavirus* (HPV, Purchased from BDSbiotech, BDS-BW-019). Using these reference samples, we evaluated the performance of various optimisation conditions. Each experiment was repeated three times.

Optimisation of cell wall lysis method

Previous studies have shown that the detection rate of mNGS is lower for thick-walled microbes, such as *Mycobacteria* and *Cryptococcus*.[17] To address this, we optimised the conditions for cell wall lysis and nucleic acid extraction as follows:1) Adjust Grinding Conditions: The grinding homogenizer was switched from a "horizontal vibration" model (Dinghaoyuan TL Smart, China) to a "three-dimensional high-speed vibration" model (Tiangen H24, China), which provides more effective disruption of thick cell walls.2) Increase Grinding Speed: The grinding speed was increased from 2000 rpm with 5 min of vibration to 5 m/s with 60 s of vibration (repeated for 3 cycles with an interval of 30 s). 3)Enhance Nucleic Acid Cleavage: Following the *Mycobacterium bovis* Nucleic Acid Test Kit manual (CapitalBio Technology, Chengdu, China), lysozyme (Lysozyme RT401) was added, and the cleavage temperature was raised from 56 °C to 100 °C. This adjustment improved the nucleic acid extraction efficiency from thick-walled microbes.

Optimisation of fragment size selection conditions

Previous studies have indicated that nanopore metagenome technology has a lower detection sensitivity for viruses (39.4%) compared with bacteria and fungi.[18] One reason for this is that viral genomes are smaller, typically consisting of only thousands to tens of thousands of base pairs. During extraction, grinding, and cleavage, viral DNA can be fragmented into smaller pieces. Fragments smaller than 1000 bp may be lost during library construction, resulting in poor detection of viral sequences.[18]

To optimise fragment size selection, we made the following adjustments:1) Adjust Repair Reagent: The repair enzyme with "terminal repair" ability (Vazyme N203) was replaced with a repair enzyme that possesses both "terminal repair" and "incision repair" abilities (Vazyme N210), which improves the ligation of shorter fragments to longer ones. 2) Increase Magnetic Bead Proportion: The proportion of magnetic beads (Vazyme N411) used during library purification was increased from 0.4 × to 0.7 × . This change helps retain fragments smaller than 1000 bp, improving the capture of smaller viral fragments.

Optimising sequencing data volume and host DNA depletion conditions

The proportion of host nucleic acids in a sample can significantly affect the sensitivity of microorganism identification, as host DNA occupies a large portion of the sequencing data. However, excessive host DNA depletion may also remove certain microbes, leading to missed detections.[19]

To determine the optimal sequencing data volume and host DNA depletion conditions, we designed three experimental groups. Group D001, the reference sample underwent no host DNA depletion. Group D002, the reference sample was divided into two equal parts (1:1 volume ratio), one part underwent host DNA depletion, while the other did not, and both parts were combined for nucleic acid extraction. Group D003, the reference sample was divided into two parts in a 1:4 volume ratio, with 1/5 of the sample undergoing no host DNA depletion and 4/5 undergoing host DNA depletion. Previous explorations revealed that complete host DNA depletion would significantly reduce the detection efficacy of Gram-negative bacterium, so further exploration of this condition was abandoned in this study. The library construction and sequencing protocols were consistent across all three groups.[13] Each group was sequenced with varying data volumes (400 MB, 600 MB, 800 MB, 1000 MB, and 1200 MB) to evaluate the impact of sequencing depth and host DNA depletion on detection efficiency.

Verification of detection efficiency using BALF from 8 patients with pneumonia

Compared with reference samples, patient BALF is more complex and diverse, with a greater impact on sequencing data volumes and host cell depletion. To further validate the optimal sequencing data volume and host DNA depletion conditions for mTGS in clinical samples, we used BALF from 8 patients with pneumonia with confirmed pathogen diagnoses. Three repeated experiments were conducted to assess the detection efficiency of mTGS in clinical samples. These experiments served as a pre-study to verify the performance of mTGS in detecting pathogens in clinical settings.

Validation of detection efficiency using a prospective cohort of patients with pneumonia

Patient enrolment

A total of 330 patients with clinically suspected pneumonia were enrolled from 7 hospitals of Zhejiang province, China from July 2023 to March 2025. The inclusion criteria were as follows: 1) Pneumonia was suspected based on positive radiographic findings (chest CT) and clinical symptoms, including onset or exacerbation of cough, sputum production, fever, shortness of breath, and elevated inflammatory markers (White Blood Cell [WBC] count, C-reactive protein [CRP], and procalcitonin [PCT]).[20,21] 2) Bronchoscopy was performed, with an adequate volume of BALF collected. 3) Informed consent was obtained. Seventeen patients were later excluded from the study (leaving 313 patients in total) due to incomplete medical records (10 patients) or unqualified BALF samples (7 patients). Sex stratification was not considered in the research design, and patients were enrolled continuously. Subsequently, demographic characteristics (including sex), laboratory findings, radiographic and pathological results, treatment course, and clinical outcomes were recorded for the 313 remaining patients.

Samples and tests

Bronchoscopy was performed on all patients, and the lesions were selected for alveolar lavage based on lung CT imaging. At least 2 tubes of BALF were collected from each patient, with the first (>15 mL) used for conventional microbiological testing (CMTs) and the second (>15 mL) evenly aliquoted into 3 portions for mNGS, pre-mTGS, and mTGS detection. Based on the specific condition of the patient, such as cytological testing, more BALF may be obtained.

CMTs

CMTs included bacterial, mycobacterial, and fungal cultures, as well as several, specialised stainings and antigen detection methods: Acid-fast bacilli stain for *Mycobacteria*, Modified acid-fast bacilli stain for *Nocardia*, Fluorescent antibody stain for *Pneumocystis jirovecii*, Calcofluor white stain for fungi, Galactomannan antigen test, PCR assays for various pathogens, including *M. tuberculosis* complex (GeneXpert, Cepheid, Sunnyvale, CA), *Mycoplasma pneumoniae*, *Chlamydia pneumoniae*, influenza A and B viruses, respiratory syncytial virus, human adenovirus, human parainfluenza virus, cytomegalovirus, and SARS-CoV-2 (Liferiver Biotechnology, Shanghai, China). CMTs were performed at the discretion of the clinician, and only a subset of these tests was conducted in each case.

mNGS

BALF samples (≥5 mL, collected in sterile sputum containers) were immediately transported to the molecular laboratory at the First Affiliated Hospital, Zhejiang University School of Medicine on dry ice. mNGS testing was performed within 8 h after sample collection. The detailed mNGS procedure and criteria for the judgement of positive results are described in the previous study reported by Zhou et al.[22]

mTGS

BALF samples (>5 mL, placed in sterile sputum containers) were immediately transported on dry ice to Digena Diagnostics Technology Co., Ltd. (Hangzhou, China) for processing. Both pre-mTGS and mTGS were performed on these samples.

Clinical adjudications

The final clinical diagnosis for each patient was considered the gold standard for evaluating the sensitivity and specificity of mTGS. This diagnosis was determined through a consensus process among five expert clinicians: three respiratory specialists, one clinical microbiologist, and one radiologist. The experts discussed and reviewed the patient's clinical history, imaging findings, results from CMTs, mNGS, mTGS, pathological results, treatment outcomes, and other relevant factors.

Oral commensal bacteria, which are commonly found in the human oropharynx and generally lack clinical significance, were carefully evaluated. Despite their frequent detection, these bacteria's clinical relevance was determined through a comprehensive discussion by the expert panel. Oral commensal bacteria could not be classified as clinically significant unless there was unanimous agreement among the experts.[23]

Bioinformatics processing and statistical analysis

Raw FAST5 files from GridIon were basecalled and debarcoded with using Basecalling v6.4.2 pipeline, and porechop v0.2.4 was used for separating the splice sequences, and NanoStat v1.6.0 for quality control. Filtlong v0.2.1 was used to remove low-quality reads generated by GridIon. Sequences retained were mapped to the reference database using minimap2 v2.26, and Genomic sequences of pathogens were validated with blast 2.14.0 against the NT database.

Sequencing results were analysed using R 4.3.2 software (University of Auckland, Auckland, New Zealand, Ross Ihaka). Measurement data conforming to a normal distribution are presented as mean ± standard deviation (SD), while data that did not meet normality assumptions are reported as the median (M) and interquartile range (P25, P75). The diagnostic performances of CMTs, mNGS, pre-mTGS, and mTGS were compared using Pearson's χ^2 test or Fisher's exact test, with clinical diagnosis serving as the reference standard. A two-tailed *P*-value of <0.05 was considered statistically significant. Visualisations, including Venn diagrams, heatmaps, and bar charts, were used to summarise and present the results for the 313 samples.

Role of funders

The funders did not play any role in the study design, data collection, management, analysis, interpretation, review, approval of the manuscript, or the decision to submit the manuscript for publication.

Results

Optimisation of mTGS experimental processes

Optimisation of cell wall lysis conditions

To evaluate the impact of optimised cell wall lysis conditions, we compared the detection performance of the reference samples before and after optimisation (denoted as C001 and C002). Under the initial C001 conditions, the detection of *M. tuberculosis*, *C. neoformans*, and *S. cerevisiae* was suboptimal. In contrast, under the optimised conditions (C002), these microbes were stably detected, with a significant increase in read counts ($P < 0.01$, Chi-square test/Fisher's exact test) (Fig. 1a). Importantly, no significant differences in detection were observed for the other 10 microbes. Based on these results, the C002 condition was selected as the optimal cell wall lysis process for mTGS and was consistently applied in subsequent experiments.

Optimisation of fragment size selection conditions

We next assessed the effects of fragment size selection optimisation by comparing the detection performance of reference samples under two conditions: R001 (pre-optimisation) and R002 (post-optimisation). Under R001, the sequencing fragments were predominantly in the 200–300 bp range. However, under R002, the fragment distribution shifted towards larger sizes, with the majority of reads centred around 1000 bp, and a significant increase in the number of longer fragments (>3 kb) (Fig. 1b). This shift in fragment size distribution was associated with improved detection sensitivity for larger DNA fragments. As a result, the R002 condition was identified as the optimal fragment size selection protocol for mTGS, and it was consistently used in subsequent analyses.

Optimisation of sequencing data volume and host DNA depletion conditions

To determine the optimal sequencing data volume for detecting the 13 pathogenic microbes in the reference sample, we analysed the detection performance under different data volumes. The results showed that the number of reads for each microbe increased linearly with the sequencing data volume. Notably, *E. faecalis*, *L. fermentans*, *M. avium*, and *P. aeruginosa* showed the most significant increases in read counts as data volume increased (Fig. 1c, Supplementary 2). Based on a balance of detection performance, analysis requirements, sequencing cost, and time, we determined that 800 MB of sequencing data provides the optimal data volume for mTGS and further clinical sample validation.

At the 800 MB sequencing data volume, we evaluated the impact of different host DNA depletion conditions on the detection of microbes. Under the D003 condition, the number of reads for *L.monocytogenes*, *S. aureus*, *Lactobacillus fermentum*, *B. subtilis*, *S. cerevisiae*, and *C. neoformans* increased significantly ($P < 0.05$, Chi-square test/Fisher's exact test) compared with the D001 and D002 conditions (Fig. 1d). Importantly, all 13 microbes were stably detected under all three host DNA depletion conditions.

We then summarised the results across different fragment size selection, host DNA depletion, and

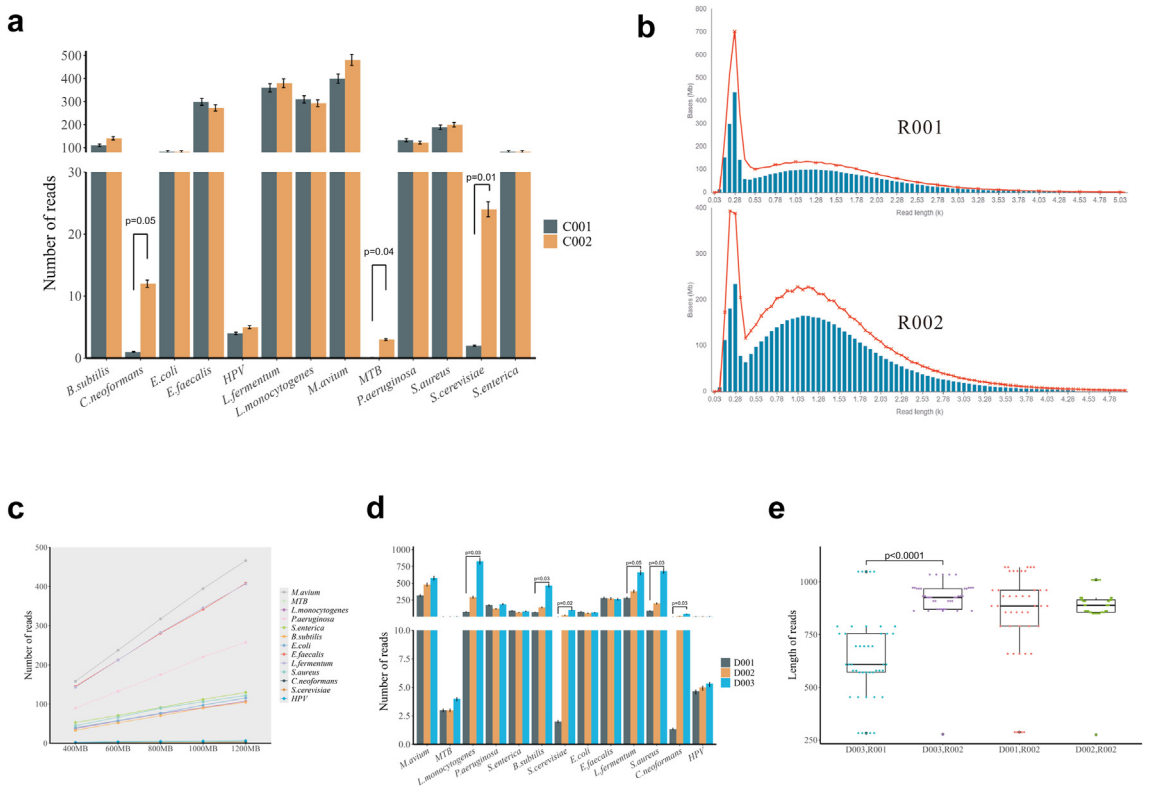

Fig. 1: **Detection of microbes in reference samples under different experimental conditions. a)** Detection efficiency of microbes before and after optimisation of cell wall lysis (C001 vs. C002). The reads for *M. tuberculosis*, *C. neoformans*, and *S. cerevisiae* were significantly increased under the C002 condition, compared with C001. **b)** Distribution of read lengths for sequenced fragments before (R001) and after optimisation of fragment size selection (R002). Under the R001 condition, the read length was concentrated between 200 and 300 bp, while under R002, the majority of reads were around 1000 bp, with a significant increase in long fragments (>3 kb). **c)** Detection of microbes under different sequencing data volumes without host DNA depletion. The number of reads detected for various microbes increased linearly with sequencing data volumes. *E. faecalis*, *L. fermentans*, *M. avium*, and *P. aeruginosa* showed the most significant increase in read count. **d)** Detection of microbes under different host DNA depletion conditions with 800 MB sequencing data. All 13 microbes were stably detected across three different conditions. The read counts for *L. monocytogenes*, *S. aureus*, *L. fermentum*, *B. subtilis*, *S. cerevisiae*, and *C. neoformans* significantly increased under the D003 condition. **e)** Average length of sequencing fragments for reference samples under different experimental conditions. There was a significant difference in R002 fragment size selection under D003 condition, but no significant difference in different host DNA depletion under R002 condition ($P > 0.05$, Chi-square test/Fisher's exact test).

sequencing data volume conditions. The analysis showed that under the R002 fragment size selection condition, the impact of host DNA depletion on the average read length was minimal. However, when using the D003 host DNA depletion condition, the sequencing fragments were most concentrated in size, yielding the best fragment analysis results (Fig. 1e). Compared with the R001 fragment size selection condition, the R002 condition under D003 host DNA depletion significantly increased the average read length (Fig. 1b and e). Based on these findings, we selected the following optimal conditions for mTGS: C002 for cell wall lysis, R002 for fragment size selection, and 800 MB of sequencing data. These conditions were then used for further clinical sample validation. Considering the complexity and potential risks associated with host DNA depletion, such as the possibility of removing some pathogenic microbes, the decision to use host DNA depletion as a pretreatment for mTGS should not be made before the small-scale clinical sample validation.

Clinical evaluation of mTGS

Results of clinical pre-experiment sequencing of 8 BALF samples

In the clinical pre-experiment with 8 BALF samples, we observed significant differences in the detection of microbes based on host DNA depletion treatment, while none differences were obtained on the cell wall lysis (C002) and fragment size selection (R002). The detection of low-abundance pathogens, such as *M. tuberculosis*, *C. psittaci*, and *P. jirovecii*, was notably improved when no host DNA depletion was applied (D001) (Fig. 2a).

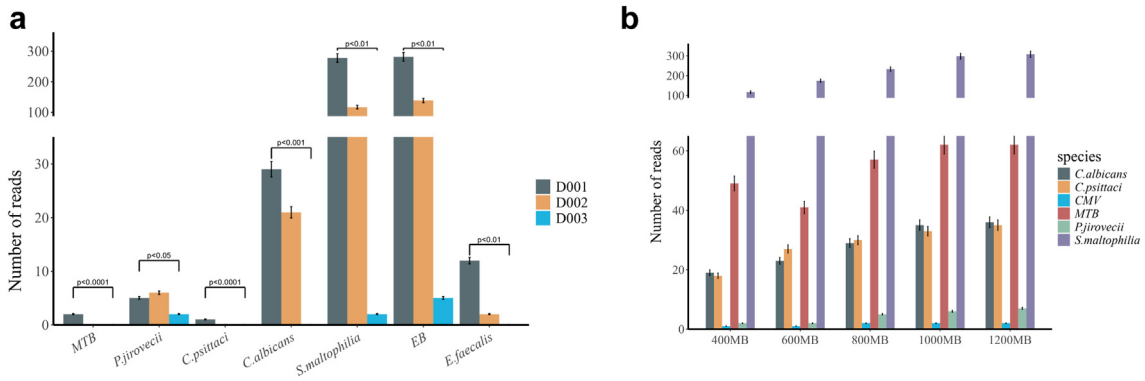

Fig. 2: **Detection of microbes in clinical pre-experimental samples (n = 8) under different experimental conditions. a)** Detection of microbes under different host DNA depletion conditions. The detection of *M. tuberculosis* ($P < 0.0001$, Chi-square test/Fisher's exact test), *C. psittaci* ($P < 0.0001$, Chi-square test/Fisher's exact test), *P. jirovecii* ($P < 0.05$, Chi-square test/Fisher's exact test), *C. albicans* ($P < 0.001$, Chi-square test/Fisher's exact test), *S. maltophilia* ($P < 0.01$, Chi-square test/Fisher's exact test), EB ($P < 0.01$, Chi-square test/Fisher's exact test), and *E. faecalis* ($P < 0.01$, Chi-square test/Fisher's exact test) was significantly improved without host DNA depletion processing. **b)** Detection of microbes under different sequencing data volumes. The number of reads detected for *P. jirovecii*, *C. albicans*, and *human cytomegalovirus* stabilised at 800 MB of sequencing data, and further increases in data volume did not significantly enhance detection.

Regarding sequencing data volumes, we found that pathogen detection remained stable at 800 MB, particularly for microbes like *P. jirovecii*, *C. albicans*, and CMV. Further increasing the data volume beyond 800 MB did not significantly increase the number of reads for these pathogens. Thus, 800 MB was determined to be the plateau for mTGS detection of clinical samples (Fig. 2b). Beyond this point, increasing the data volume further would only increase sequencing costs and time without adding substantial value, making it economically unfeasible. Based on these findings, we concluded that 800 MB sequencing data with no host DNA depletion (D001) is the optimal condition for clinical sample detection in mTGS.

mTGS detection process and sensitivity

Based on the results from the reference samples and the 8 clinical samples used in the pre-experiment, we confirmed that the optimal reaction system for mTGS detection in clinical samples involves C002 cell wall lysis, R002 fragment size selection, D001 (no host DNA depletion), and 800 MB sequencing data. These conditions were stable and effective for clinical sample analysis. In contrast, the basic process system without any optimisation (pre-mTGS) was used as the control for comparison. The experimental workflows for both methods are illustrated in Fig. 3.

The detection limits of both pre-mTGS and mTGS were assessed using the reference samples prepared as described in section 1.2.1. The results showed that mTGS significantly improved detection sensitivity compared with pre-mTGS. Specifically, the detection limit for *B. subtilis*, *M. tuberculosis*, *M. avium*, and *C. neoformans* was improved from 10^4 copies/mL to 10^3 copies/mL, while the detection limit for HPV improved from 10^5 copies/mL to 10^4 copies/mL—representing a sensitivity increase of more than 10-fold. For other pathogens, the detection limits were similar between the two groups (Table 1).

Prospective clinical evaluation study

Basic characteristics and clinical diagnosis of enrolled patients

Between July 2023 and March 2025, a total of 313 patients with suspected pulmonary infections were enrolled in this study. The cohort included 196 males (62.62%) and 117 females (37.38%), with a mean age of 58.3 ± 15.3 years. Of the 313 patients, 211 (70.7%) had one or more underlying diseases. The most common comorbidities included respiratory diseases in 67 patients (30.32%), followed by diabetes in 42 patients (19.00%). After treatment, 257 patients (82.11%) showed clinical or imaging improvement within 14 days, while 56 patients (17.89%) showed no improvement (Table 2).

Among the 313 enrolled patients, 274 were diagnosed with pulmonary infection (33 of whom were unable to identify the pathogen) and 39 were finally diagnosed as non-infectious diseases, including immune checkpoint inhibitor associated pneumonia (n = 6), lung cancer (n = 5), cryptogenic organising pneumonia (n = 5), pulmonary sequestration (n = 1), lymphoma (n = 2), vasculitis (n = 1), lung metastases (n = 1), and undetermined diagnoses (n = 18).

Results of microorganism detection by CMTs, mNGS, pre-mTGS, and mTGS

A comprehensive analysis was performed on the test results from 274 patients with confirmed pulmonary infections, identifying a total of 376 pathogens. The detection performance of CMTs, mNGS, pre-mTGS,

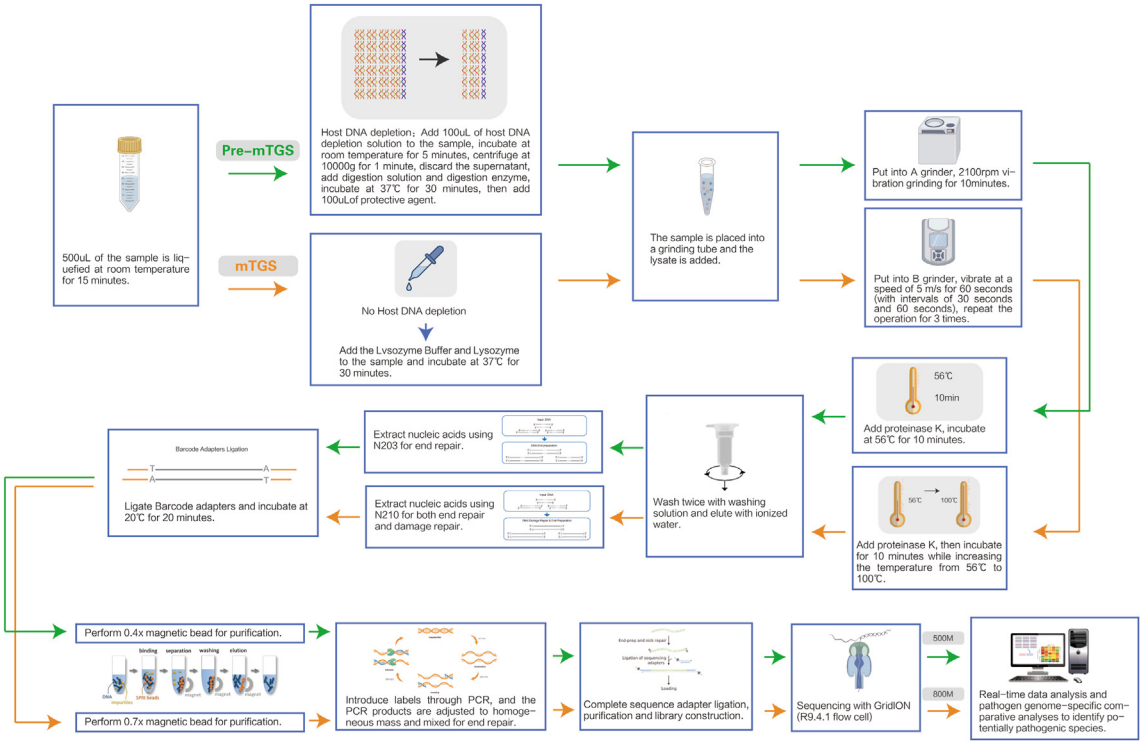

Fig. 3: **Schematic diagram of the experimental process for detecting clinical samples using pre-mTGS and mTGS.**

and mTGS was compared using these confirmed pathogens as a reference. The results showed variability in the detection of species and their abundance across different methods. Compared with standard CMTs, mTGS identified a broader range of pathogens, including more bacteria, fungi, viruses, and other types of pathogens. Specifically, mTGS detected 179 bacterial species, 93 fungal species, 22 viral species, and additional pathogens like *M. pneumoniae* and *C. psittaci* (compared with 67 bacteria, 37 fungi, 10 viruses, and only 1 additional pathogen with CMTs). When compared with pre-mTGS, mTGS detected more pathogens across all categories: bacteria (136 vs. 77), fungi (93 vs. 43), viruses (22 vs. 19), mycobacteria (43 vs. 17), and other pathogens (20 vs. 4). The overall detection performance of mTGS was comparable to that of mNGS, with mTGS identifying 136 bacterial species, 93 fungal species, 22 viral species, and 43 mycobacterial species, compared with 144 bacteria, 101 fungi, 25 viruses, and 42 mycobacteria identified by mNGS (Fig. 4a).

In terms of microorganism detection in BALF Samples, mTGS and mNGS showed similar detection performance, each with distinct advantages for certain pathogens. The most commonly detected bacteria in clinical samples were *M. tuberculosis*, *Non-tuberculous mycobacteria*, *P. aeruginosa*, and *H. influenzae* (Fig. 4b). For fungi, *P. jirovecii*, *Aspergillus* spp., *C. neoformans*, and *C. albicans* were the most frequently detected (Fig. 4b), while *CMV* and *EBV* were the most common virus identified (Fig. 4b). mNGS showed better sensitivity for detecting pathogens like NTM, *P. jiroveci* and *Aspergillus* spp. (Fig. 4b), while mTGS performed better in detecting *M. tuberculosis*, *C. psittaci*, and *S. pneumoniae* (Fig. 4b). Overall, the detection performance of mTGS

	Microbes	pre-mTGS (copies/mL)	mTGS (copies/mL)
G+ bacilli	*Listeria monocytogenes*	1×10^3	1×10^3
G+ bacilli	***Bacillus subtilis***	$\mathbf{1 \times 10^4}$	$\mathbf{1 \times 10^3}$
G+ bacilli	***Mycobacterium tuberculosis***	$\mathbf{1 \times 10^4}$	$\mathbf{1 \times 10^3}$
G+ bacilli	***Mycobacterium avium***	$\mathbf{1 \times 10^4}$	$\mathbf{1 \times 10^3}$
G+ bacilli	*Lactobacillus fermentum*	1×10^3	1×10^3
G+ cocci	*Enterococcus faecalis*	1×10^2	1×10^2
G+ cocci	*Staphylococcus aureus*	1×10^2	1×10^2
G- bacilli	*Escherichia coli*	1×10^2	1×10^2
G- bacilli	*Salmonella enterica*	1×10^3	1×10^3
G- bacilli	*Pseudomonas aeruginosa*	1×10^2	1×10^2
Fungi	*Saccharomyces erevisiae*	1×10^3	1×10^3
Fungi	***Cryptococcus eoformans***	$\mathbf{1 \times 10^4}$	$\mathbf{1 \times 10^3}$
Virus	**Human papillomavirus type 16**	$\mathbf{1 \times 10^5}$	$\mathbf{1 \times 10^4}$

The words in bold indicate the difference in detection limits between pre-mTGS and mTGS.

Table 1: **Detection limits of pre-mTGS and mTGS for various types of microbes in reference materials.**

Characteristics		
Race		
Han nationality	313	100%
Sex (%)		
Male	196	62.62%
Female	117	37.38%
Age (years)	58.3 ± 15.3	
Underlying diseases (%)		
Malignant haematologic disease	41	18.60%
Malignant solid tumour	22	9.95%
Diabetes	42	19.00%
Solid organ transplantation	8	3.62%
Hypertension	36	16.29%
Connective tissue disease	17	7.69%
Chronic kidney disease	9	4.07%
Previous inactive pulmonary tuberculosis	7	3.17%
Bronchiectasis	37	16.74%
Chronic obstructive pulmonary diseases	16	7.24%
Interstitial lung disease	7	3.17%
14-day clinical prognosis (%)		
Improved	257	82.11%
Not improved	56	17.89%

Table 2: **Basic characteristics of 313 enrolled patients.**

was significantly improved compared with pre-mTGS, demonstrating its enhanced sensitivity for a broader range of pathogens (Fig. 4b).

Clinical diagnostic performances of CMTs, mNGS, pre-mTGS, and mTGS

The clinical diagnostic performances of CMTs, mNGS, pre-mTGS, and mTGS were evaluated using 376 pathogen strains identified in 274 patients with confirmed infections. The results showed that mTGS demonstrated an overall clinical concordance of 81.80% with the final clinical diagnosis, with complete concordance in 70.80% of cases and partial concordance in 11.00%. Compared with pre-mTGS, the overall clinical concordance of mTGS was significantly improved by 35.80% (81.80% vs. 46.00%, $P < 0.01$, Chi-square test/Fisher's exact test). The concordance between mTGS and clinical diagnosis was comparable to that of mNGS (81.80% vs. 78.10%, $P = 0.14$, Chi-square test/Fisher's exact test), both significantly outperforming CMTs (81.80% vs. 38.00%, $P < 0.01$, Chi-square test/Fisher's exact test; 78.10% vs. 38.00%, $P < 0.01$, Chi-square test/Fisher's exact test) (Fig. 4c).

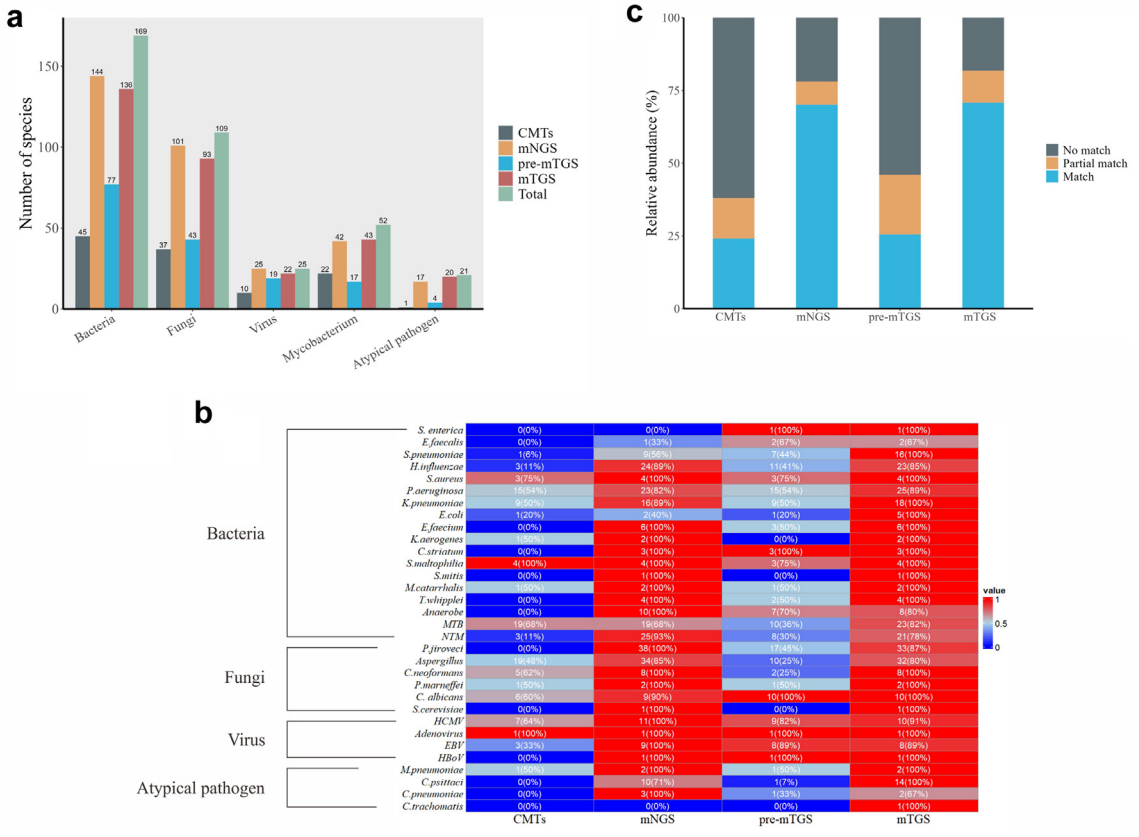

Fig. 4: **Detection efficiency of microbes in BALF using CMTs, mNGS, pre-mTGS, and mTGS (n = 313). a)** Overall detection of different microbial species using CMTs, mNGS, pre-mTGS, and mTGS. **b)** Distribution of microbes detected by the four methods. **c)** Consistency of the results from CMTs, mNGS, pre-mTGS, and mTGS with clinical diagnosis.

The sensitivity and negative predictive value (NPV) of mTGS were significantly higher than those of CMTs (84.70% vs. 39.40%, $P < 0.01$, Chi-square test/Fisher's exact test; 40.00% vs. 17.80%, $P < 0.01$, Chi-square test/Fisher's exact test). However, mTGS showed significantly lower specificity compared with CMTs (71.80% vs. 92.30%, $P < 0.01$, Chi-square test/Fisher's exact test), with comparable positive predictive values (PPV) (95.50% vs. 97.30%, $P = 0.32$, Chi-square test/Fisher's exact test). When compared with mNGS, the sensitivity and PPV of mTGS were similar (84.70% vs. 79.90%, $P = 0.14$, Chi-square test/Fisher's exact test; 95.50% vs. 96.90%, $P = 0.48$, Chi-square test/Fisher's exact test), but mNGS demonstrated superior specificity (82.10% vs. 71.80%, $P < 0.05$, Chi-square test/Fisher's exact test). However, NPV performance of mTGS is comparable to that of mNGS (40.00% vs. 36.80%, $P = 0.68$, Chi-square test/Fisher's exact test).

When compared with pre-mTGS, mTGS showed significant improvements in sensitivity and NPV (84.70% vs. 52.19%, $P < 0.01$, Chi-square test/Fisher's exact test; 40.00% vs. 19.63%, $P < 0.01$, Chi-square test/Fisher's exact test), while PPV remained unchanged. pre-mTGS had higher specificity than mTGS (82.10% vs. 71.80%, $P < 0.01$, Chi-square test/Fisher's exact test) (Table 3).

Discussion

Pneumonia still remained the leading cause of infectious disease-related deaths worldwide.[24] Delayed aetiological diagnosis can result in inappropriate use of broad-spectrum antibiotics, contributing to poorer clinical outcomes, extended hospital stays, and increased healthcare costs.[16] Therefore, rapid and accurate diagnosis is crucial for the targeted use of antibiotics and for improving treatment outcomes, and research on technologies of rapid pathogen diagnose has been a global hot topic in recent years.

	ID	NID	Sensitivity	Specificity	Positive predictive value	Negative predictive value
CMTs			39.40%	92.30%	97.30%	17.80%
+	108	3				
–	166	36				
mNGS			79.90%	82.10%	96.90%	36.80%
+	219	7				
–	55	32				
Pre-mTGS			52.19%	82.10%	95.33%	19.63%
+	143	7				
–	131	32				
mTGS			84.70%	71.80%	95.50%	40.00%
+	232	11				
–	42	28				

Table 3: Diagnostic performance of CMTs, mNGS, pre-mTGS, and mTGS in identifying pathogens.

In this study, we optimised the mTGS experimental process for BALF focusing on key technical innovations in cell wall lysis, fragment size selection, and host DNA depletion. Additionally, we refined quality control parameters, such as sequencing data volume, to establish a more standardised mTGS detection system. As a result, the detection limits for several pathogens, including *B. subtilis*, *M. tuberculosis*, *M. avium*, *C. neoformans*, and HPV, improved by up to tenfold compared with the baseline process.

Microbial diversity and structural differences significantly affect nucleic acid release efficiency. For instance, the thick cell walls of gram-positive bacteria and fungi can hinder the complete release of nucleic acids using conventional lysis methods, potentially leading to false negatives and reduced detection sensitivity.[17] Our study demonstrated that efficient cell wall lysis improved the detection efficiency for nine bacterial and fungal species.[25] Moreover, we found that the use of physical bead grinding for cell wall lysis negatively impacted the detection of RNA viruses.[26] To address this, we introduced an innovative dual lysis approach combining "physical vibration + enzymatic hydrolysis." This method not only enhanced the detection of thick-walled microbes such as *M. tuberculosis* and *C. neoformans*, but also maintained the integrity of pathogen detection across other species.

It is important to note that nanopore metagenomic sequencing has limitations in virus detection, primarily due to the small size and fragility of viral genomes during extraction.[18] To overcome this challenge, we optimised fragment size selection technology, allowing shorter sequences to be linked into longer fragments that are more effectively captured by nanopore sequencing. The average read length of the sequencing fragments increased to approximately 1000 bp, with a significant rise in the number of long fragments (>3 kb). This innovation has the potential to significantly enhance the accuracy and specificity of pathogen identification, particularly for low-abundance pathogens,[27] as further confirmed in our clinical validation.

Human-derived nucleic acids in samples are a major factor influencing the sensitivity of metagenomic pathogen detection. While host DNA depletion can enhance detection sensitivity,[26] excessive depletion may lead to the loss of specific or low-abundance pathogens.[25] To optimize this process, we explored different host DNA depletion strategies under varying sequencing data conditions. The pre-experimental results using standard samples showed that the detection efficiency of microbes was highest under the D003 condition. However, in clinical samples, the impact of host DNA depletion strategies on pathogen detection was more complex. Notably, the D001 condition significantly improved detection of pathogens such as *M. tuberculosis*, *C. psittaci*, and *P. jirovecii*. This discrepancy can be attributed to the simpler composition of the reference

samples, which primarily contained high-abundance human and microbial cells, leading to minimal interference. Additionally, pathogens like *P. jirovecii* and RNA viruses, which are prone to removal during host DNA depletion, were not present in the reference samples.[28] In contrast, clinical samples are more diverse and complex, with greater variability in human nucleic acid content and a higher level of potential interference. As such, the effectiveness of host DNA depletion varies depending on the abundance of the pathogen and the level of host DNA contamination.[25] Given these complexities and the risks associated with over-depletion, we opted not to perform host DNA depletion in the mTGS pre-processing system to preserve the accuracy and comprehensiveness of the results. Moderate host depletion strategy can improve the detection sensitivity of some microorganisms, but there is also a risk that some microorganisms will be missed due to host depletion, how to balance the retention of host depleted microorganisms will also be the focus of our next exploration direction.[26]

For data quality control in metagenomic sequencing, a standard of 20 million reads per sample is typically recommended for next-generation sequencing (NGS),[29] but no established guideline for nanopore metagenomic sequencing was set up. To address this research gap, we assessed the performance of microbial detection across different sequencing data volumes using reference and clinical pre-experimental samples. As the data volume increased, the number of reads detected for various microbes in the reference samples also increased. For clinical pre-experimental samples like *P. jirovecii*, *C. albicans*, and CMV, the number of detected reads stabilised at 800 MB of sequencing data, with no further improvement in detection when the data volume was increased. The sensitivity of results with 800 MB of data was comparable to the sensitivity of second-generation sequencing with 20 million reads. Based on these findings, we concluded that 800 MB of sequencing data is sufficient to meet the performance requirements for clinical microbial metagenomic identification. Moreover, the cost of sequencing a single sample with nanopore mTGS at this data volume has reduced to approximately 800 RMB, with a sequencing time of under 2 h. While increasing the data volume further would yield more sequence information, it would also significantly raise both the cost and time required, with only marginal improvements in pathogen detection. Therefore, increasing the data volume beyond 800 MB offers limited additional value and is not economically justified.

In summary, we have successfully established a robust mTGS process system by optimising key conditions, such as cell wall lysis, fragment size selection, sequencing data volume, and host DNA depletion. This optimised system has been validated for its performance in detecting microbes in BALF from a prospective cohort of patients with lung infections. Compared with CMTs, mTGS demonstrated a 45.30% increase in detection sensitivity. When compared with pre-optimised mTGS, the sensitivity improved by 32.51%. mTGS achieved sensitivity comparable to mNGS while maintaining a high level of consistency with clinical diagnoses. However, both mTGS and mNGS have specific advantages in detecting certain pathogens: mNGS had higher detection rates for NTM, *P. jirovecii* and *Aspergillus* species, whereas mTGS performed better for detecting *M. tuberculosis*, *C. psittaci,* and *S. pneumoniae*. This discrepancy is partly due to the host DNA depletion step in mNGS, which can remove certain intracellular bacteria, such as *Mycobacterium* species.[25]

Nanopore sequencing technology used in mTGS excels in identifying low-abundance pathogens, especially in cases where pathogen content is minimal. This is because nanopore sequencing does not require amplification and benefits from long read lengths. However, mNGS, with its reliance on shorter read lengths, may miss some low-abundance pathogens, leading to false negatives.[5] Overall, mTGS offers simpler experimental procedures, shorter sequencing times, and comparable costs to mNGS (Supplementary 3). Additionally, nanopore sequencing has distinct advantages for rapid pathogen detection and antimicrobial resistance analysis due to its ability to generate long reads and provide real-time sequencing and analysis.[14,15,30] Thus, mTGS is recommended as a promising tool for the rapid diagnosis of respiratory infections, particularly in the analysis of BALF.

However, this study has limitations as well. First, we did not perform a comparison of the timeliness and cost of testing for various methodologies due to the complexity of sample transport and cost measurement. Second, we did not explore RNA virus detection, which would be a key focus of future research. Furthermore, analysis of drug resistance genes was not included in the analysis, despite its significant role in guiding clinical antibiotic therapy. The inherently low abundance of pathogen-associated drug resistance limits the sensitivity of metagenomic detection for antimicrobial resistance. When targeting drug resistance gene analysis specifically, metagenomic sequencing requires 10 to 100 times greater sequencing depth compared with routine pathogen identification purposes profiling while increasing costs significantly.[31] To better assess the clinical utility of mTGS in diagnosing, treating, and predicting outcomes for pneumonia, more work like larger, multicentre prospective studies would be needed in the near future. Future research can also be extended to more heterogeneous populations, such as immunocompromised patients, children, or more comprehensive clinical samples, such as blood samples, cerebrospinal fluid, tissues, etc., to comprehensively

evaluate the practicality and robustness of mTGS in different clinical scenarios.

In conclusion, mTGS may serve as a new and a promising diagnostic tool for the rapid identification of pneumonia pathogens. Through ongoing optimisation, we aim for mTGS to provide more accurate, faster diagnostic information in clinical practice, ultimately improving patient treatment and prognosis.

In conclusion, we successfully established a standardised mTGS experimental process by optimising key parameters, significantly improving its performance compared with pre-optimised mTGS in the aetiological diagnosis of pneumonia. In a cohort of patients with pulmonary infections, mTGS demonstrated detection performance comparable to mNGS in BALF samples, with both methods significantly outperforming CMTs. The optimised mTGS system is a reliable and effective tool for the aetiological diagnosis of pulmonary infections.

Contributors

HZ, YSY and HMG conceived the study. SSZ, XL, QL, HL, YF, WYC, CZ and JY collected the clinical samples. XL, YF, LC, HMG, WW and YKY conducted the pre-mTGS and mTGS experiments. YKY, HZ, and YNY performed the mNGS experiments. XL, HL and YF contributed to bacterial cultures. SSZ, YKY, XL and HZ contributed to data analysis and interpretation. HZ, YF, XL and XL have verified the underlying data. SSZ, XL, XL, YF and HZ drafted the manuscript. HZ, HMG, YF, and YSY revised the final manuscript. HZ, YSY, and HMG supervised the study. All authors read and approved the final manuscript. All authors had access to the study data and had final responsibility for the decision to submit the manuscript for publication. HZ made the final decision to submit the manuscript.

Data sharing statement

The mTGS datasets have been submitted to NCBI Sequence Read Archive (SRA) under the bioproject accession number SUB15228714.

Declaration of interests

XL, LC, WW and HG are employees of Zhejiang Digena Diagnosis Technology CO., Ltd and Hangzhou Digena Medical Laboratory Co., Ltd. The other authors declare no conflicts of interest.

Acknowledgements

This study was supported by the the Research and Development Programme of Zhejiang Province (2023C03068, 2024C03187), the National Natural Science Foundation of China (82272338), the Key R&D Plan of the Ministry of Science and Technology (China) of China (2022YFC2504502).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2025.105752.

References

1 Rajapaksha P, Elbourne A, Gangadoo S, et al. A review of methods for the detection of pathogenic microorganisms. *Analyst*. 2019;144(2):396–411.
2 Holter JC, Müller F, Bjørang O, et al. Etiology of community-acquired pneumonia and diagnostic yields of microbiological methods: a 3-year prospective study in Norway. *BMC Infect Dis*. 2015;15:64.
3 Gu W, Miller S, Chiu CY. Clinical metagenomic next-generation sequencing for pathogen detection. *Annu Rev Pathol*. 2019;14:319–338.
4 Chiu CY. Viral pathogen discovery. *Curr Opin Microbiol*. 2013;16(4):468–478.
5 Lin Q, Yao Y, Li X, et al. The application of nanopore targeted sequencing for pathogen diagnosis in bronchoalveolar lavage fluid of patients with pneumonia: a prospective multicenter study. *Infect Dis (Lond)*. 2024;56(2):128–137.
6 Jing C, Chen H, Liang Y, et al. Clinical evaluation of an improved metagenomic next-generation sequencing test for the diagnosis of bloodstream infections. *Clin Chem*. 2021;67(8):1133–1143.
7 Chen H, Yin Y, Gao H, et al. Clinical utility of in-house metagenomic next-generation sequencing for the diagnosis of lower respiratory tract infections and analysis of the host immune response. *Clin Infect Dis*. 2020;71(Suppl 4):S416–S426.
8 Qi C, Hountras P, Pickens CO, et al. Detection of respiratory pathogens in clinical samples using metagenomic shotgun sequencing. *J Med Microbiol*. 2019;68(7):996–1002.
9 Liu H, Zhang Y, Chen G, et al. Diagnostic significance of metagenomic next-generation sequencing for community-acquired pneumonia in southern China. *Front Med (Lausanne)*. 2022;9:807174.
10 Wang J, Han Y, Feng J. Metagenomic next-generation sequencing for mixed pulmonary infection diagnosis. *BMC Pulm Med*. 2019;19(1):252.
11 Lin T, Tu X, Zhao J, et al. Microbiological diagnostic performance of metagenomic next-generation sequencing compared with conventional culture for patients with community-acquired pneumonia. *Front Cell Infect Microbiol*. 2023;13:1136588.
12 Zinter MS, Dvorak CC, Mayday MY, et al. Pulmonary metagenomic sequencing suggests missed infections in immunocompromised children. *Clin Infect Dis*. 2019;68(11):1847–1855.
13 Charalampous T, Kay GL, Richardson H, et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat Biotechnol*. 2019;37(7):783–792.
14 Sun X, Song L, Yang W, et al. Nanopore sequencing and its clinical applications. *Methods Mol Biol*. 2020;2204:13–32.
15 Satam H, Joshi K, Mangrolia U, et al. Next-generation sequencing technology: current trends and advancements. *Biology (Basel)*. 2023;12(7).
16 Webb BJ, Sorensen J, Jephson A, et al. Broad-spectrum antibiotic use and poor outcomes in community-onset pneumonia: a cohort study. *Eur Respir J*. 2019;54(1).
17 Miao Q, Ma Y, Wang Q, et al. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin Infect Dis*. 2018;67(suppl_2):S231–S240.
18 Guo Y, Li Z, Li L, et al. A dual-process of targeted and unbiased Nanopore sequencing enables accurate and rapid diagnosis of lower respiratory infections. *EBioMedicine*. 2023;98:104858.
19 Zhou H, Ouyang C, Han X, et al. Metagenomic sequencing with spiked-in internal control to monitor cellularity and diagnosis of pneumonia. *J Infect*. 2022;84(1):e13–e17.
20 Cao B, Huang Y, She DY, et al. Diagnosis and treatment of community-acquired pneumonia in adults: 2016 clinical practice guidelines by the Chinese Thoracic Society, Chinese Medical Association. *Clin Respir J*. 2018;12(4):1320–1360.
21 Torres A, Niederman MS, Chastre J, et al. International ERS/ESICM/ESCMID/ALAT guidelines for the management of hospital-acquired pneumonia and ventilator-associated pneumonia: Guidelines for the management of hospital-acquired pneumonia (HAP)/ventilator-associated pneumonia (VAP) of the European Respiratory Society (ERS), European Society of Intensive Care Medicine (ESICM), European Society of Clinical Microbiology and Infectious Diseases (ESCMID) and Asociación Latinoamericana del Tórax (ALAT). *Eur Respir J*. 2017;50(3).
22 Zhou H, Larkin PMK, Zhao D, et al. Clinical impact of metagenomic next-generation sequencing of bronchoalveolar lavage in the diagnosis and management of pneumonia: a multicenter prospective observational study. *J Mol Diagn*. 2021;23(10):1259–1268.
23 Graf EH, Simmon KE, Tardif KD, et al. Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: a systematic comparison to a commercial PCR panel. *J Clin Microbiol*. 2016;54(4):1000–1007.
24 Hill UG, Floto RA, Haworth CS. Non-tuberculous mycobacteria in cystic fibrosis. *J R Soc Med*. 2012;105 Suppl 2(Suppl 2):S14–S18.
25 Schlaberg R, Chiu CY, Miller S, et al. Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch Pathol Lab Med*. 2017;141(6):776–786.

26 Liu D, Zhou H, Xu T, et al. Multicenter assessment of shotgun metagenomics for pathogen detection. *EBioMedicine*. 2021;74:103649.
27 Matsuo Y, Komiya S, Yasumizu Y, et al. Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. *BMC Microbiol*. 2021;21(1):35.
28 Han D, Diao Z, Lai H, et al. Multilaboratory assessment of metagenomic next-generation sequencing for unbiased microbe detection. *J Adv Res*. 2022;38:213–222.
29 Consensus of clinical pathways of metagenomic next-generation sequencing test in diagnosis of lower respiratory tract infections in China. *Zhonghua Jie He He Hu Xi Za Zhi*. 2023;46(4):322–335.
30 Warburton PE, Sebra RP. Long-read DNA sequencing: recent advances and remaining challenges. *Annu Rev Genomics Hum Genet*. 2023;24:109–132.
31 Grumaz S, Stevens P, Grumaz C, et al. Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med*. 2016;8(1):73.