frontiers | Research Topics

# Targeted metagenomics in pathogen detection

**Edited by**
Qing Wei, Edwin Kamau and Jiemin Zhou

**Published in**
Frontiers in Cellular and Infection Microbiology

ISSN 1664-8714
ISBN 978-2-8325-6398-4
DOI 10.3389/978-2-8325-6398-4

## About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

# Targeted metagenomics in pathogen detection

**Topic editors**

Qing Wei — Genskey Co. Ltd, China

Edwin Kamau — Tripler Army Medical Center, United States

Jiemin Zhou — Vision Medicals Co, Ltd, China

**Citation**

Wei, Q., Kamau, E., Zhou, J., eds. (2025). *Targeted metagenomics in pathogen detection*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6398-4

# Table of contents

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Editorial
PUBLISHED 16 May 2025
DOI 10.3389/fcimb.2025.1612802



OPEN ACCESS

EDITED AND REVIEWED BY
Nahed Ismail,
University of Illinois Chicago, United States

*CORRESPONDENCE
Qing Wei
vubwqing@hotmail.com

RECEIVED 16 April 2025
ACCEPTED 05 May 2025
PUBLISHED 16 May 2025

CITATION
Zhou J, Kamau E and Wei Q (2025) Editorial: Targeted metagenomics in pathogen detection.
*Front. Cell. Infect. Microbiol.* 15:1612802.
doi: 10.3389/fcimb.2025.1612802



# Editorial: Targeted metagenomics in pathogen detection

Jiemin Zhou[1], Edwin Kamau[2] and Qing Wei[3,4]*

[1]Medical Marketing Center, Vision Medicals Co, Ltd, Guangzhou, China, [2]Department of Pathology and Area Laboratory Services, Tripler Army Medical Center, Honolulu, HI, United States, [3]Shanghai Cinopath Medical Laboratory Co., Kindstar Globalgene Technology Inc., Shanghai, China, [4]Department of Research and Development, Kindstar Global Precision Medicine Institute, Wuhan, China

KEYWORDS

**metagenomics, next generation sequencing, clinical microbiology, infectious dieseases, targeted metagenomics**

Editorial on the Research Topic
Targeted metagenomics in pathogen detection

## Introduction

Infectious diseases are the leading cause of morbidity and mortality worldwide, accounting for approximately 25.5% of global deaths (Diseases and Injuries, 2020). The low sensitivity of conventional diagnostic methods and long turnaround times pose significant challenges for timely and accurate diagnosis, which is critical for improving patient prognosis. Unbiased metagenomics, a high-throughput and non-targeted technology used to analyze all genomic information in a sample, has been widely used to diagnose various infections, such as bloodstream infections, abdominal cavity infections, and central nervous system infections. Metagenomics has become a promising detection method for infectious diseases. While targeted metagenomics, a modified technique, focuses on sequencing specific genes or microbial communities, providing more focused data on selected regions or species, and it allows for the selective enrichment and sequencing of specific microbial species or communities within complex samples, such as those found in clinical settings. This technique is particularly useful when traditional culture-based methods fail to detect the causative pathogen or when multiple pathogens are present in the same sample. Therefore, targeted metagenomics has the potential to revolutionize the diagnosis and treatment of infectious diseases in clinical settings, providing a more personalized approach to healthcare.

## Goals of targeted metagenomics

The primary objective of targeted metagenomics is to identify and characterize microbial communities in clinical samples—such as blood, urine, and sputum—to aid in the diagnosis and treatment of infectious diseases. Additionally, targeted metagenomics has applications in other fields, including agriculture and forensic medicine. This technique

seeks to lower the cost of conventional metagenomics, which has been proposed as a method to detect all potential pathogens in clinical samples. In addition, targeted metagenomics seeks to improve the performance of current approaches, such as 16S/18S rRNA-based amplicon sequencing and custom-designed primer pools. Establishing guidelines to standardize the workflow of targeted metagenomics across different sequencing platforms is also a key objective. Ultimately, targeted metagenomics aims to offer an affordable, accurate, and fast approach for precise pathogen detection.

## Clinical applications of targeted metagenomics

Immunocompromised patients are more susceptible to infections by rare (Zhan et al., 2021), regional (Ramirez et al., 2020), and emerging (El Zein et al., 2020; Fishman, 2023) pathogens, posing significant challenges for the clinical application of targeted metagenomics. Liu et al. retrospectively enrolled 546 immunocompetent and immunocompromised patients with suspected community-acquired pneumonia to evaluate the performance of metagenomics and targeted metagenomics. The total coincidence rate of targeted metagenomics was much higher than that of metagenomics, with final comprehensive clinical diagnoses as the reference standard. However, there were few negative cases with non-infectious diseases, resulting in slight bias in calculating specificity. Sun et al. assessed the performance of targeted metagenomics in diagnosing pulmonary infections in HIV-infected patients, finding an 86.7% concordance rate for the detection of main pathogens, while it was a small-sample, single-center research study, which might limit the accuracy of the study. Yang et al. investigated diagnostic value of targeted metagenomics in cancer patients with pneumonia and found its sensitivity can reach up to 84.6%. They also pointed out that the main limitation was sample size was small. Using targeted metagenomics, Zhang et al. provided a case report on *Mycobacterium bovis* infection in an infant, and patients with lung abscesses caused by *Parvimonas micra* were successfully diagnosed by Zhang et al.

Beyond pathogen detection, Zou et al. performed hybridization capture-based targeted metagenomics on patients undergoing allogeneic hematopoietic stem cell transplantation at different intervals to monitor medication efficacy, providing significant reference for treatment strategies. They also proposed that targeted metagenomics can be used to rule out infections. However, RNA viruses were considered. Shi et al. used 16S rRNA amplicon sequencing to investigate the succession of microbial communities in intensive care units treated with bacteriophage, finding that the relative abundance of target pathogens decreased while overall species diversity remained stable. Future research should focus on long-term observation of pathogen dynamics and mutations in bacterial phage receptor sites following phage treatment.

## Technical improvement of targeted metagenomics

Targeted metagenomics with high sensitivity has reduced the economic burden on patients, and its extensive application can be expected (Sun et al., 2025). Given the high sensitivity of PCR and the high throughput of mNGS, targeted metagenomics can detect pathogens with predesigned primers in the panel (Huang et al., 2023; Li et al., 2021). However, adding more primers targeting a broader range of pathogens to the panels can produce more primer dimer species, reducing the mapping rate (Xie et al., 2022) and increasing the likelihood of missing certain pathogens. Considering the epidemiology of pathogens characterized by geographical specificity (Ramirez et al., 2020), rarity (Zhan et al., 2021), and novelty (El Zein et al., 2020; Fishman, 2023), Liu et al. proposed the designing and developing regional targeted metagenomics (Xia et al., 2023; Xie et al., 2022) should be performed, and an era of widespread application of regional targeted metagenomics in diagnosing and monitoring infections with high sensitivity and low economic burden on patients can be expected.

## Challenges and limitations of targeted metagenomics

A published study explored the feasibility of capture hybridization-based targeted metagenomics and multiplex PCR-based targeted metagenomics in distinguishing lower respiratory tract infections in clinical practice. Although these methods can decrease costs with high detection ability, they have disadvantages, including long research and development cycles, limited targets, and the need to accumulate enough samples for sequencing (Yin et al., 2024). Zhao et al. provided a comprehensive review of the application of metagenomics in diagnosing infectious diseases, summarizing the advantages and disadvantages of targeted metagenomics, while health economic evaluations of metagenomics should be conducted.

## Future directions of targeted metagenomics

Given the current application scenarios, future studies can be carried out as follows:

1. Development and improvement of compatible primer pools to enhance amplification efficiency of blood samples, addressing severe nucleic acid fragmentation due to broad-spectrum antibiotics.
2. Large-scale cohort studies to evaluate host responses to different infections, determining host biomarkers for assisted diagnosis in targeted metagenomics. Combining immune repertoire analysis to characterize immunological exhaustion signatures and establishing an infection-immunity interaction model.

3. Application of syndromic panels for targeted metagenomics in diagnosing infections in blood culture, respiratory specimens, stool, and cerebrospinal fluid.
4. Technical innovation in primer design, primer dimer cleanup, turnaround time optimization, fast sample preparation, and sequencing protocols to promote quick application of targeted metagenomics in clinical settings.

## Concluding remarks

We hope that published research will inspire continued exploration and innovation, ultimately advancing our ability to detect, track, and control infectious diseases globally. We extend our gratitude to all authors who contributed their innovative work to this Research Topic. We also thank the reviewers for their rigorous and constructive feedback, which significantly enhanced the quality of the published research. Additionally, we thank the editorial team at Frontiers for their unwavering support and guidance throughout the process.

## Author contributions

JZ: Writing – review & editing, Writing – original draft. EK: Writing – original draft, Writing – review & editing. QW: Writing – review & editing, Writing – original draft.

## Conflict of interest

Author JZ was employed by the company Vision Medicals Co, Ltd. Author QW was employed by the company Shanghai Cinopath Medical Laboratory Co., Kindstar Globalgene Technology Inc.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note



## Author disclaimer

The views expressed in this study are those of the authors and do not necessarily reflect the official policy or position of the Defense Health Agency, Department of Defense, nor the U.S. Government. This work was prepared as part of official duties. Title 17, U.S.C., Section 105 provides that copyright protection under this title is not available for any work of the U.S. Government. Title 17, U.S.C., Section 101 defines a U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person's official duties.

## References

Diseases, G. B. D., and Injuries, C. (2020). Global burden of 369 diseases and injuries in 204 countries and territories 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396, 1204–1222. doi: 10.1016/S0140-6736(20)30925-9

El Zein, S., Hindy, J.-R., and Kanj, S. S. (2020). Invasive saprochaete infections: an emerging threat to immunocompromised patients. *Pathog. (Basel Switzerland)* 9. doi: 10.3390/pathogens9110922

Fishman, J. A. (2023). Next-generation sequencing for identifying unknown pathogens in sentinel immunocompromised hosts. *Emerging Infect. Dis.* 29, 431–432. doi: 10.3201/eid2902.221829

Huang, C., Huang, Y., Wang, Z., Lin, Y., Li, Y., Chen, Y., et al. (2023). Multiplex PCR-based next generation sequencing as a novel, targeted and accurate molecular approach for periprosthetic joint infection diagnosis. *Front. In Microbiol.* 14, 1181348. doi: 10.3389/fmicb.2023.1181348

Li, B., Xu, L., Guo, Q., Chen, J., Zhang, Y., Huang, W., et al. (2021). GenSeizer: a multiplex PCR-based targeted gene sequencing platform for rapid and accurate identification of major mycobacterium species. *J. Clin. Microbiol.* 59. doi: 10.1128/JCM.00584-20

Ramirez, J. A., Musher, D. M., Evans, S. E., Dela Cruz, C., Crothers, K. A., Hage, C. A., et al. (2020). Treatment of community-acquired pneumonia in immunocompromised adults: A consensus statement regarding initial strategies. *Chest* 158, 1896–1911. doi: 10.1016/j.chest.2020.05.598

Sun, N., Zhang, J., Guo, W., Cao, J., Chen, Y., Gao, D., et al. (2025). Comparative analysis of metagenomic next-generation sequencing for pathogenic identification in clinical body fluid samples. *BMC Microbiol.* 25, 165. doi: 10.1186/s12866-025-03887-8

Xia, H., Zhang, Z., Luo, C., Wei, K., Li, X., Mu, X., et al. (2023). MultiPrime: A reliable and efficient tool for targeted next-generation sequencing. *iMeta* 2, e143. doi: 10.1002/imt2.v2.4

Xie, N. G., Wang, M. X., Song, P., Mao, S., Wang, Y., Yang, Y., et al. (2022). Designing highly multiplex PCR primer sets with Simulated Annealing Design using Dimer Likelihood Estimation (SADDLE). *Nat. Commun.* 13, 1881. doi: 10.1038/s41467-022-29500-4

Yin, Y., Zhu, P., Guo, Y., Li, Y., Chen, H., Liu, J., et al. (2024). Enhancing lower respiratory tract infection diagnosis: implementation and clinical assessment of multiplex PCR-based and hybrid capture-based targeted next-generation sequencing. *EBioMedicine* 107, 105307. doi: 10.1016/j.ebiom.2024.105307

Zhan, Y., Xu, T., He, F., Guan, W.-J., Li, Z., Li, S., et al. (2021). Clinical evaluation of a metagenomics-based assay for pneumonia management. *Front. In Microbiol.* 12, 751073. doi: 10.3389/fmicb.2021.751073

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Case Report
PUBLISHED 12 February 2024
DOI 10.3389/fcimb.2024.1341236



OPEN ACCESS

EDITED BY
Edwin Kamau,
Tripler Army Medical Center, United States

REVIEWED BY
Shashank Ganatra,
Oregon Health &Science University, United States
Jacinta Bustamante,
UniversitéParis Cité, France
Taj Ali Khan,
Kohat University of Science and Technology, Pakistan
Evan Ewers,
Tripler Army Medical Center, United States

*CORRESPONDENCE
Sisi Wang
wssmeitan@163.com

RECEIVED 20 November 2023
ACCEPTED 29 January 2024
PUBLISHED 12 February 2024

CITATION
Zhang H, Liao Y, Zhu Z, Liu H, Li D and Wang S (2024) Assistance of next-generation sequencing for diagnosis of disseminated Bacillus Calmette-Guerin disease with X-SCID in an infant: a case report and literature review.
*Front. Cell. Infect. Microbiol.* 14:1341236.
doi: 10.3389/fcimb.2024.1341236



# Assistance of next-generation sequencing for diagnosis of disseminated Bacillus Calmette-Guerin disease with X-SCID in an infant: a case report and literature review

Haiyang Zhang[1,2], Yi Liao[2,3], Zhensheng Zhu[4], Hanmin Liu[1,2], Deyuan Li[1,2] and Sisi Wang[1,2]*

[1]Department of Pediatrics, West China Second University Hospital, Sichuan University, Chengdu, China, [2]Key Laboratory of Birth Defects and Related Diseases of Women and Children, Sichuan University, Ministry of Education, Chengdu, China, [3]Department of Radiology, West China Second University Hospital, Sichuan University, Chengdu, China, [4]Depertment of Bioinformation, Hugobiotech Co., Ltd., Beijing, China

*Bacille Calmette-Guérin* (BCG) is a live strain of *Mycobacterium bovis* (*M.bovis*) for use as an attenuated vaccine to prevent *tuberculosis* (TB) infection, while it could also lead to an infection in immunodeficient patients. *M.bovis* could infect patients with immunodeficiency via BCG vaccination. Disseminated BCG disease (BCGosis) is extremely rare and has a high mortality rate. This article presents a case of a 3-month-old patient with disseminated BCG infection who was initially diagnosed with hemophagocytic syndrome (HPS) and eventually found to have X-linked severe combined immunodeficiency (X-SCID). *M.bovis* and its drug resistance genes were identified by metagenomics next-generation sequencing (mNGS) combined with targeted next-generation sequencing (tNGS) in blood and cerebrospinal fluid. Whole exome sequencing (WES) revealed a pathogenic variant in the common γ-chain gene (*IL2RG*), confirming X-SCID. Finally, antituberculosis therapy and umbilical cord blood transplantation were given to the patient. He was successfully cured of BCGosis, and his immune function was restored. The mNGS combined with the tNGS provided effective methods for diagnosing rare BCG infections in children. Their combined application significantly improved the sensitivity and specificity of the detection of *M.bovis*.

KEYWORDS

**Bacillus Calmette-Guérin, next-generation sequencing, whole exome sequencing, immunodeficiency, vaccine**

## Introduction

*Bacille Calmette-Guérin* (BCG) is a live strain of *Mycobacterium bovis* (*M.bovis*) for use as an attenuated vaccine to prevent *tuberculosis* (TB) infection. It remains the only vaccine against TB in general use in China. Localized adverse reactions, including hypersensitivity, abscess formation, and regional lymphadenitis following BCG vaccination, are common and self-limiting. Disseminated BCG disease (BCGosis) is extremely rare, while it can lead to high mortality in infants with the immunodeficient disease (Amanati et al., 2017; Lange et al., 2022). The early symptoms of this infection are very insidious and not specific, so a timely diagnosis could contribute to early treatment and be life-saving. High-throughput sequencing technology may have potential advantages in the diagnostic field of BCGosis. Metagenomics next-generation sequencing (mNGS) can detect a wide variety of organisms, but cannot achieve comprehensive detection of drug resistance genes. Targeted next-generation sequencing (tNGS) has a higher specificity than mNGS, and the species and origin of the bacteria can often be specifically identified. Through specific capture techniques, different drug resistance genes can be captured by tNGS (Beviere et al., 2023). Here, we report a rare case of severe X-linked severe combined immunodeficiency (X-SCID) with disseminated BCG infection that was initially presented with hemophagocytic syndrome (HPS) and eventually received umbilical cord blood transplant (UCBT). We identified the presence of infection quickly and found its resistance gene through mNGS combined with tNGS. Through this case, we proposed that mNGS combined with tNGS could effectively and rapidly identify BCGosis in potentially immunodeficient patients.

## Case presentation

A 3-month-old male infant had sudden, unprovoked bouts of high fever in October 2022. Anti-infective treatment with ceftriaxone for 3 days was ineffective. Miliary red papules appeared all over the body gradually He was transferred to the pediatric intensive care unit (PICU) of West China Second University on October 14, 2022. The patient did not have diarrhea. He had no history of recurrent infection since birth. His parents denied a family history of *Mycobacterium tuberculosis* (MTB) infection, immunodeficiency, or consanguineous marriage. There's no history of sudden death in infancy in his immediate family. He was born at term and received an intradermal injection of 0.1 ml of BCG vaccine on the left upper arm on day 1 after birth. The growth and development of the child were within the normal range.

The physical examination at admission showed temperature of 38.6°C, respiratory rate of 56 times/min, pulse rate of 190 times/min, blood pressure of 73/59 mmHg, height of 60 cm, and weight of 6.8 kg. Scattered red papules were observed on the pale skin of his chest and abdomen. The BCG scar had crusted over without redness or suppuration. No thrush was found in the buccal mucosa. No enlarged superficial lymph nodes were palpable. The three-concave sign was positive, and the breath sounds of both lungs were rough without moist rales. An abdominal examination revealed hepatosplenomegaly. The liver was enlarged, with its lower edge 6 cm below the right costal margin. A palpable splenic edge was felt 8 cm below the left costal margin. Laboratory studies at the time of admission revealed hemoglobin (Hb) of 87 g/L, white blood cell count (WBC) of $4.7\times10^9$/L, lymphocyte of $1.44\times10^9$/L, neutrophils of $3.21\times10^9$/L, and platelet count of $58\times10^9$/L. CRP was 148.8 mg/L. PCT was 2.05 ng/ml. Coagulation tests suggested a hypercoagulable state (D-dimer was 27.95 mg/L, fibrin degradation products were 76.6 ug/L). The polymerase chain reaction (PCR) of Epstein-Barr virus (EBV) and Cytomegalovirus (CMV) was negative. The purified protein derivative (PPD) skin test (72 h) and T-cell spot test for tuberculosis infection (T-SPOT) were negative. A chest X-ray revealed slight infiltrate in both lungs. His initial diagnosis at admission was sepsis, severe pneumonia, and coagulopathy.

The patient received nasal high-flow ventilation immediately after admission. Anticoagulant therapy and empirical antibiotic treatment with intravenous imipenem (120 mg/kg/day, q6h) were administered to him. On the 2nd day of admission, the inflammatory markers were further increased (PCT > 150 ng/ml). The level of ferritin in the serum was 3022.70 ng/ml (normal range: 10-291 ng/ml). Dexamethasone (0.5 mg/kg/day, qd) was given as an empirical treatment for hemophagocytic syndrome (HPS). At the same time, the patient developed irritability and repeated moaning, indicating mild disturbance of consciousness. Subsequently, the patient began to show positive signs of meningeal irritation including projectile vomiting, increased muscle tone, and positive neck stiffness. We performed a cerebrospinal fluid (CSF) test because of his severe infection and worsening state of consciousness. The cytological and biochemical results of the CSF were normal. In addition, we applied samples of peripheral blood and CSF (2 ml each) for mNGS. On day 3, the results of natural killer (NK) cell activity were shown as 24.25%, and the soluble CD25 (sCD25) was 14784 pg/ml. The criteria of the HPS diagnostic criteria were fulfilled (Henter et al., 2007). HPS was confirmed. Etoposide (150 mg/m$^2$/day, biw) was implemented with reference to the hemophagocytic lymphohistiocytosis-1994 protocol (Henter et al., 1997). On day 4, Xpert Mycobacterium tuberculosis and rifampicin resistance detection (Xpert MTB/RIF) from the sputum revealed positive for MTB, but negative for RIF. MTB-PCR was weakly positive in sputum. For this patient, we collected the deep sputum samples by nasotracheal suction. MTB complex (23178 reads in blood and 15710 reads in CSF) was detected via mNGS (Hugobiotech, Beijing, China, Table 1). A four-tuberculostatic drug regimen included isoniazid (H, 15 mg/kg/day, peros), rifampin (R, 15 mg/kg/day, peros), pyrazinamide(Z, 33 mg/kg/day, peros) and ethambutol (E, 19 mg/kg/day, peros). On day 7, *M.bovis* with resistance to pyrazinamide was identified by tNGS of the blood (Hugobiotech, Beijing, China). The drug resistance gene was *pncA* (Table 1). Levofloxacin(L, 15 mg/kg/day, peros) was added for anti-tuberculosis treatment, and pyrazinamide was discontinued. Immunologic studies showed T cells and NK cells were very low, while B cells were present. The immunoglobulin profile was within the lower range (Table 2). The chest computed tomography (CT) revealed the absence of a thymus and showed a hazy opacity (Figure 1). To sum up, the infant developed severe BCGosis and

TABLE 1 The data of mNGS and tNGS.

| Method | Sample | Pathogen | Sequence number (reads) | Confidence | Relative abundance (%) | Drug resistance gene |
|---|---|---|---|---|---|---|
| mNGS | CSF | MTB | 15710 | high | 10.81 | no |
| mNGS | blood | MTB | 23178 | high | 99.55 | no |
| tNGS | blood | *M. Bovis* | 4459 | high | no | *pncA* |

TABLE 2 Immunological data before/after hematopoietic stem cell transplantation (humoral immunity and cellular immunity).

| | Pre-transplant | Post-transplant | | |
|---|---|---|---|---|
| | | 3 months | 4 months | 6 months |
| Humoral immunity: lymphocyte subsets ($\times10^9$/L) | | | | |
| CD3+ cells | 0.03 | 0.1 | 1.56 | 2.73 |
| CD4+ cells | 0.01 | 0.08 | 0.49 | 0.84 |
| CD8+ cells | 0.01 | 0.03 | 1.06 | 1.88 |
| CD4/CD8 | 1 | 2.67 | 0.46 | 0.45 |
| CD19+ cells | 0.97 | 0.01 | 0.01 | 0.02 |
| CD56+ cells (NK cells) | 0.01 | 0.23 | 1.05 | 1.03 |
| Cellular immunity: immunoglobulins (g/L) | | | | |
| IgG | 1.7 | 20.6 | 16.9 | 18.2 |
| IgA | 0.07 | 0.15 | 0.07 | 0.17 |
| IgM | 0.16 | 1.23 | 4.16 | 1.31 |

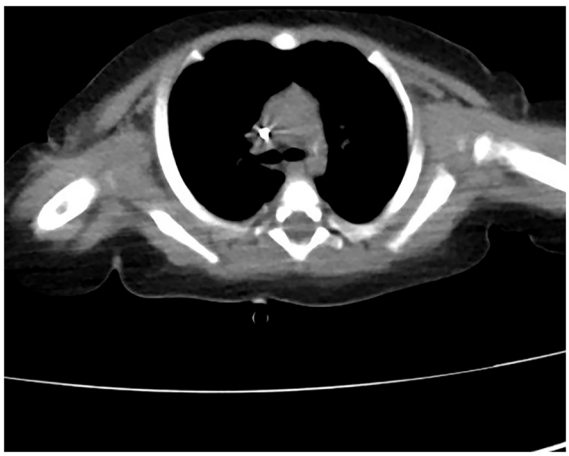


FIGURE 1
The chest CT revealed the absence of a thymus.

progressed rapidly, even developing complicated HPS. People with normal immune systems are usually not susceptible to *M.bovis*. The clinical findings, the imaging absence of thymus, and decreased cellular and humoral immune function suggested that the child may have primary immunodeficiency (PID). The clinical standard for diagnosis of PID was to conduct high-throughput sequencing of the whole genome exon. Therefore, 2 mL of peripheral blood from each of the family members was identified for a gene mutation through whole exome sequencing (WES). The results of the WES revealed a pathogenic variant (Exon 3: c.391C > T; p.Gln131Ter, acquired from the mother), in the common γ-chain gene (*IL2RG*), confirming X-SCID. This variant was a nonsense hemizygous variant, which caused the Glutamine at 131th changed the stop codon. The variant was not collected in healthy population database. Previous study reported this variant in association with X-linked SCID, it was predicted to cause loss of normal protein function either through protein truncation or nonsense-mediated mRNA decay. So, finally it was classified as pathogenic. In this case, the T cell counts were less than $0.05\times10^9$/L, and the pathogenic gene was identified. According to the 2022 diagnostic criteria from the Primary Immune Deficiency Treatment Consortium's (PIDTC's) (Dvorak et al., 2023), we identified the case as a classical SCID.

1 month after the diagnosis of X-SCID, the patient received UCBT from unrelated donors (matched for 8/10 HLA alleles) after a conditioning chemotherapy regimen that included busulfan and cyclophosphamide. Regular infusions of immunoglobulin and etanercept were required in the post-transplant course. The HREL-based anti-tuberculosis regimen was used for 9 months. 6 months after UCBT, the patient was free from recurrent infections.

During follow-up, there were no abnormalities in blood routine or ferritin. Re-examination of immunologic studies showed T cell counts (CD3, CD4 and CD8) and NK cell counts (CD56) returned to normal, while B cell counts (CD19) remained low (Table 2). After UCBT, we quantitatively assessed the post-transplant mosaicism status of this patient at regular intervals over 8 months. The final assessment results showed that donor cells accounted for 100% and T cells accounted for 99.82% of the peripheral blood after transplantation, both of which showed complete chimerism (Table 3). Mosaicism status was defined according to the rate of donor chimerism (DC) after transplantation. DC≥95% indicated a complete chimeric state (Locatelli et al., 2013). We have summarized the timeline of patient treatment and disease progression in Figure 2.

## mNGS and tNGS methods

The patient presented with sepsis at the beginning of admission. Pathogen cultures were performed on blood, sputum and CSF samples from the patient. All samples were cultured for 14 days, but the results were negative. In order to find the agent of infection as soon as possible, we tested the patient for mNGS. Blood and CSF samples were transported to Hugobiotech Co., Ltd. (Beijing) for nucleic acid extraction and mNGS. Blood samples (2-4 ml from the patient) were collected in the Cell-Free DNA BCT STRECK and then stored or shipped between 6-35°C for mNGS detection immediately. CSF samples (2-3 mL) were collected, then sterile sealed, stored at -20°C, and transported on dry ice for mNGS detection immediately. We extracted and purified DNA from 200 ul of plasma according to the instructions (YGZZ015, Hugobiotech, Beijing, China) for the QlAamp DNA Micro Kit (50) #56304. Qubit 3.0 fluorimeter (Q33216, Invitrogen, USA) and agar-gel electrophoresis (UVC1-1100, Major Science, USA) were performed to verify DNA concentration and quality. QIAseqTM Ultralow Input Library Kit was used to construct the DNA libraries for mNGS. We pooled qualified libraries with different barcode labeling and sequenced them using the Illumina Nextseq 550 sequencing platform (Illumina, San Diego, USA) and SE75bp sequencing strategy. By removing adapters, low-quality, low-complexity, short reads, and adapter-related data from mNGS sequencing data, high-quality data were obtained. We removed human reads by mapping them to the human reference genome using SNAP software and aligned the remaining reads to the Burrows-Wheeler Alignment (BWA) database. The National Center of Biotechnology Information (NCBI) provided the genomes for the database. The microbial composition of the sample was finally determined. The reads mapped to the MTB genome with coverage of 14.8487% in the blood (Figure 3A) and 29.1761% in the CSF (Figure 3B). The distribution of species-specific reads aligns with the MTB in the blood and CSF, respectively. The average length of each read was 75 bp.

TABLE 3 The post-transplant rate of DC.

| Detection time (Post-transplant) | The rate of DC (%) | |
|---|---|---|
| | Peripheral blood | T cell |
| 2 weeks | 99.36 | 93.12 |
| 4 weeks | 70.1 | 81.6 |
| 2 months | 85.95 | 89.16 |
| 3 months | 95.86 | 95.67 |
| 4 months | 98.42 | 99.82 |
| 5 months | 100 | 99.25 |
| 6 months | 99.85 | 99.63 |
| 8 months | 100 | 99.82 |

After MTB was found in mNGS in both blood and CSF, tNGS were further performed for patients in order to screen out the species, subgroups and drug resistance genes of MTB. The DNA library for tNGS was constructed according to the operation manual of the General Kit for Identification and Drug Resistance Gene Detection of MTB Complex Group (YGZZ016, Hugobiotech, Beijing, China). Library quality control was performed with a Qubit 3.0 fluorimeter (Invitrogen, Q33216) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA). Qualified DNA libraries with different barcodes were pooled and sequenced using the Illumina Nextseq 550 sequencing platform and SE75bp sequencing strategy. After the sequencing data was disassembled, the splices, low-quality simple repeats, and N-sequence data were removed to obtain high-quality sequencing data, and the human genome data was filtered through the BWA. The remaining sequencing data was compared with the dedicated microbial database blast to complete the species identification analysis of the target pathogen. The database was originally from NCBI. After manual collection and processing, it contained 49 kinds of microorganisms, including 10 types of MTB complex group and 39 kinds of nontuberculous mycobacteria (NTM). Blastn was used to compare the fragments with resistance markers to the reference sequences of drug resistance genes to find polymorphic loci (SNPS) and obtain relevant information on the corresponding drug resistance genes. Through the above tNGS detection process, we accurately identified the *M. Bovis*. The drug resistance gene was pncA.

We used sterile deionized water as a negative template control (NTC) and synthesized fragments with known quantities as a positive template control (PC). As quality control steps, NTC and PC were included in each wet lab procedure and bioinformatics analysis. In the case of *Cryptococcus* and MTB, positive mNGS results were considered when at least 1 unique read was mapped to species level and absent in NTC or when the ratio of reads per million (RPM) between sample and NTC ($RPM_{sample}/RPM_{NTC}$) > 5 as $RPM_{NTC}\neq0$. The above process ensured that mNGS combined with tNGS had significant sensitivity and specificity in detecting MTB.

## WES methods

Exons cover most of the genetic information related to protein coding, which plays a vital role in the normal function and health of

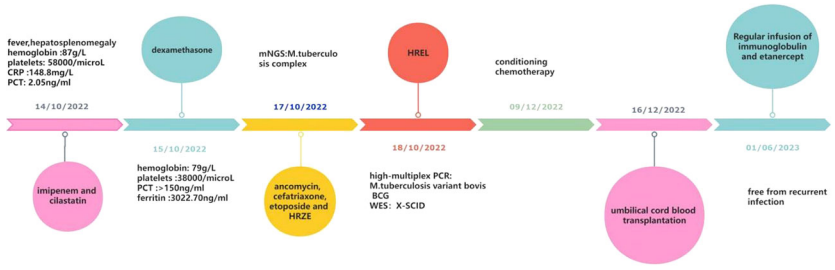



FIGURE 2
The treatment timeline of the patient.

life. WES generally has high sensitivity and can detect many types of mutations, including single nucleotide variations (SNVs) and small insertions or deletions (Indels). In this case, the genomic DNA obtained from all available family members was used for WES and Sanger Sequencing (Biosune). Then WES was performed based on the GenCap® Whole Exon Gene Capture Probe V6.0. Multiple software programs were used to read the alignment to the human reference genome GRCh37/hg19 and annotate variants. The average coverage depth for the WES is 138.92 X, and the base coverage (≥20 X) in the target region is 97.01%. The identified mutations were assessed by referencing the Human Gene Mutation Database (HGMD), the Single Nucleotide Polymorphism Database (dbSNP), the Online Mendelian Inheritance in Man (OMIM), and ClinVar databases. Subsequently, all identified mutations were filtered based on clinical attributes, inherent patterns, and data from the Genome Aggregation Database (gnomAD) and the Exome Aggregation Consortium (ExAC) databases.

PCR samples were visualized on an agarose gel in the ABI PRISM 3730 genetic analyzer (Thermo Fisher Scientific, USA), which performed purification and sequencing using the terminated cycle sequencing method. The mutation sites were identified by comparing the DNA sequence with the NCBI database. A pair of primers were designed to amplify exon 3 of the *IL2RG* gene (NM_000206.3). The forward primer sequence was 5'-TACCTCCTCCCTTCCCATCA-3'. The reverse primer sequence was 5'-AAAGTCCAGAAGTGCAGCCA-3', and the amplified fragment length was 298 bp. Purified PCR products with a size of 298 bp were sequenced by Sanger Sequencing. DNASTAR (Madison) software was used to analyze the sequencing results.

Finally, a flash analysis platform was used to analyze the pathogenicity of variation, and the possible variation loci were determined. The pathogenicity of variation loci was also analyzed according to ACMG (American College of Medical Genetics and Genomics) genetic variation classification criteria and guidelines (Richards et al., 2015). Considering the proband's clinical features, SCID associated genes was searched and obtained from OMIM database (https://www.omim.org/). All variants located in exon and classic splicing regions was obtained for further analysis. Variants with a minor allele frequency of <0.01 in the gnomAD/ExAC database were obtained for following pathogenic classification. The IL-12/23/ISG15-IFN-γ and other related genes associating with X-linked SCID were carefully considered. Through the above WES and Sanger Sequencing process, we accurately identified the mutation in the *IL2RG* gene (Exon 3: c.391C>T; p.Gln131Ter) in the patient (Figure 4). The genetic detection of the father was wild-type, and the genetic detection of the mother was with respect to the mutation in the *IL2RG* gene (Exon 3: c.391C>T; p.Gln131Ter. Figure 4).

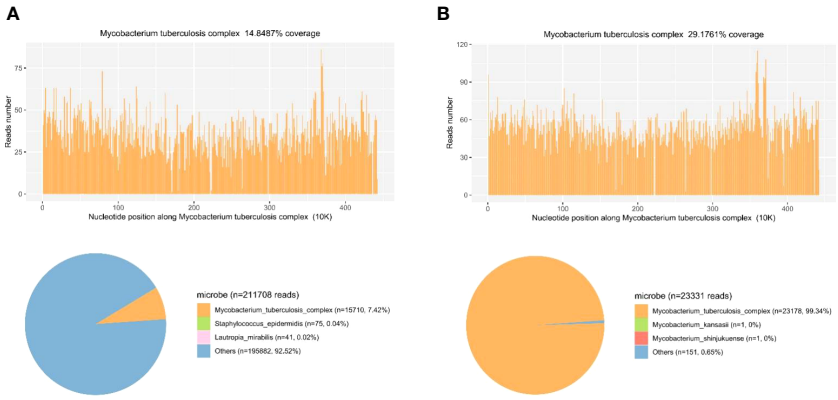



FIGURE 3
Diagnosis of MTB infection using mNGS. **(A)** The reads mapped to the MTB genome with coverage of 14.8487% in the blood. The distribution of species-specific reads aligns with the MTB in the blood. **(B)** The reads mapped to the MTB genome with coverage of 29.1761% in the CSF. The distribution of species-specific reads aligns with the MTB in the CSF.

10.3389/fcimb.2024.1341236

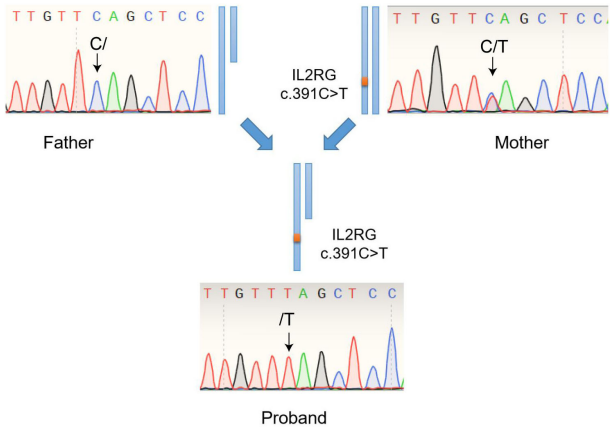



FIGURE 4
Genetic detection of the variant in the *IL2RG* gene (Exon 3: c.391C>T; p.Gln131Ter) among the patient, his father (wild-type) and his mother (carrier).

## Systematic review

We searched PubMed, Web of Science databases, Embase, and Medline from 1993 to 2023. The keywords were "X-SCID" and "BCG disease". We extracted the following details from each article: first author, vaccination, clinical manifestation, detection method, complication, therapy and outcome. We summarized a total of 7 articles from our literature search in Table 4 (Culic et al., 2004; Huang et al., 2006; Jaing et al., 2006; Bacalhau et al., 2011; Shi et al., 2020; Maron et al., 2021; Liu et al., 2022).

# Discussion

The BCG is a live attenuated vaccine strain of *M.bovis* that is commonly used to prevent MTB infection. Once BCG is inoculated into the body, it can cause an extremely mild asymptomatic infection, thereby inducing the body to produce memory T lymphocytes and ultimately achieving the purpose of preventing tuberculous meningitis and disseminated MTB infection in infants and young children. The BCG vaccination is considered safe for competently immune infants. However, in patients with immunosuppression or BCG treatment for bladder cancer, postvaccination or posttreatment adverse reactions may occur, which may manifest as local adverse reactions, lymphadenitis, disseminated BCG disease, immune reconstructive inflammatory syndrome, and osteomyelitis (Toskas et al., 2022). *M.bovis* is naturally resistant to pyrazinamide (de la Morena et al., 2017). From a multicentre study, the incidence of BCGosis and BCGtis in SCID was 34% and 19%, respectively (Marciano et al., 2014). The overall mortality of BCGosis ranged from 50% to 71% (Amanati et al., 2017; Ong et al., 2020). While a potentially life-threatening BCGosis with an overall fatality rate of 50% and higher could arise in PID, which was 0.06 to 1.56 per million vaccinated children (Amanati et al., 2017; Lange et al., 2022). Early detection and appropriate treatment are lifesaving for them. However, it often took 1.5 to 2 months from the early onset to the diagnosis of MTB infection, probably due to a lack of exposure and poor specificity of symptoms in children (Ong et al., 2020; Thomas et al., 2022).

In this case, The infant was predisposed to *M.bovis* due to SCID. The known gene defects that cause SCID include *IL2RG, ADA, IL7R, RAG1, RAG2, JAK3, DCLRE1C, PTPRC, BCL11B, RMRP, CD3E, CD247, NHEJ1, CORO1A, LIG4, PRKDC, LAT, RAC2, AK2* and *TTC7A* (Tangye et al., 2020). Considering the proband's clinical features, SCID associated genes was obtained from OMIM database (https://www.omim.org/). The IL-12/23/ISG15-IFN-γ and other related genes associating with SCID were carefully considered. All variants located in exon and classic splicing regions was obtained for further analysis. Variants with a minor allele frequency of <0.01 in the gnomAD/ExAC database were obtained for following pathogenic classification according to the ACMG guideline. Finally, the proband was diagnosed with X-SCID by WES. PID is a group of heterogeneous diseases dominated by T/B cell defects accompanied by other cell defects to varying degrees. X-SCID is due to defects in the *IL2RG*. IL2 was discovered as a T cell growth factor that can potently boost the cytolytic activity of NK cells. This receptor subunit is shared by several different cytokine receptor complexes. Pathogenic variants in this gene often cause absent T cells and NK cells with nonfunctional B cells via the blockade of multiple cytokines (Spolski et al., 2018). X-SCID usually manifests as extreme susceptibility to infections, which may lead to death in the first few months of life unless immunologic reconstitution is achieved (Lange et al., 2022). For patients with X-SCID, the tubercle bacilli continue to proliferate, even leading to BCGosis after the BCG vaccine due to an ineffective cell-mediated immune response. In China, BCG is regularly recommended for newborns in the first month of their lives without screening for PID. In the majority of immunocompromised populations, BCG might be a major cause of infection and an obstacle to future immune reconstitution (Zhang et al., 2011; Norouzi et al., 2012). Further screening should be performed for children with repeated infections, even if the screening results for latent TB infection are negative. Children with X-SCID had the highest survival rate after receiving transplantation before 3.5 months (Hardin et al., 2022). Key clinical information on the incidence, prevalence, treatment status and long-term prognosis of X-SCID in China is still incomplete. Due to the large population base and high testing costs, neonatal screening for X-SCID is not widespread in China. X-SCID network registration and data collection for clinical epidemiology are further needed in China. For infants with a family history, immunological screening should be performed before BCG vaccination. The family histories included a history of immunodeficiency in the immediate family or an unexplained early death. If the parents are recently married, the child should also be included in the scope of PID screening. The ultimate cure for X-SCID is suggested to be hematopoietic stem cell transplantation. After transplantation, regular monitoring of humoral immunity, cellular immunity and mosaicism status is required to comprehensively evaluate the transplantation effect and immune reconstitution (Locatelli et al., 2013). Among them, the CD4/CD8 T cell ratio is an indispensable monitoring indicator. The reversed CD4/CD8 T cell ratio usually indicates immunosuppressive status or reduced rejection after

TABLE 4 BCG disease in X-SCID patients in recent 30 years.

| Author | Age | BCG vaccination | Clinical manifestation | Screening for TB | Chest imaging | Diagnostic methods | Drug resistance | Complicated with HPS | Antituberculosis drugs | Immune reconstitution | Outcome |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Liu S (Richards et al., 2015) | 5m | vaccinated at birth | fever, cough and respiratory insufficiency | not mentioned | absence of athymus, opacities | PCR | no | no | H、R、E、Lfx | HSCT | alive |
| Maron G (Liu et al., 2022) | 4m | vaccinated at birth | fever,cough and papule | tuberculosis antibody(-) | infectious lesions | PCR | no | yes | H, R, Z | no | death |
| Shi B (Maron et al., 2021) | 4m | vaccinated at birth | fever and lymphadenopathy | acid fast bacilli stain (-) | not mentioned | PCR | no | no | H, R, E, Lfx | gene therapy | alive |
| Bacalhau S (Shi et al., 2020) | 4m | vaccinated at birth | fever and pustules | not mentioned | not mentioned | not mentioned | no | yes | Lzd, Mpm, H, R, E, Clr | HSCT | alive |
| Jaing TH (Bacalhau et al., 2011) | 5m | vaccinated at the third day of life | decreased appetite and tachypneic | not mentioned | infectious lesions | sputum culture | no | no | H, R, E, Cfx | UCBT | alive |
| Huang LH (Jaing et al., 2006) | 8m | vaccinated at the third day of life | fever and rash | not mentioned | infectious lesions and absence of athymus | PCR | no | no | H, R, Z | no | death |
| Culic S (Huang et al., 2006) | 6m | not mentioned | chronic diarrhea | not mentioned | not mentioned | not mentioned | no | no | not mentioned | no | death |





10.3389/fcimb.2024.1341236

transplantation (Kimura et al., 2020). In this case, both CD4 T cell count and CD8 T cell count returned to normal after UCBT, while the CD4/CD8 T cell ratio was inverted, suggesting that the risk of rejection was reduced in the process of immune reconstitution.

The patient's initial clinical symptoms were consistent with HPS. Such atypical clinical symptoms also hinder early identification of MTB infection. HPS is an aggressive and fatal syndrome of excessive inflammation due to abnormal immune activation caused by the absence of normal downregulation by activated macrophages and lymphocytes. The occurrence of HPS leads to the production of cytokine abundance and an imbalance of the host immune response, which can cause multiple organ damage or failure (Allen and McClain, 2015). TB-associated HPS has been shown to have a poor prognosis and high mortality rates (up to 50%) (Padhi et al., 2015). Several cases of HPS due to BCG have been reported in adults with bladder cancer after intravesical instillation of BCG, but rarely in children (Bacalhau et al., 2011). In this case, BCG induced the polarization of M0-type macrophages into M1-type macrophages (Liu et al., 2015). Key transcription factors such as NFkB, STAT1, STAT5 and IRF3 can mediate M1-type macrophages' release of numerous cytokines and chemokines (Bosco, 2019). The lack of normal feedback regulation leads to excessive macrophage polarization with highly elevated levels of cytokines, which triggers the predisposing factors of HPS. In addition, this patient had a severe congenital immune deficiency, which not only predisposed him to severe invasive infections but also induced a systemic cytokine storm. As soon as the diagnosis of HPS is suspected, treatment with etoposide and corticosteroids should be initiated (Henter et al., 2007). Nevertheless, these immunosuppressive treatments can exacerbate the course of the BCG infection. Thus, confirming the diagnosis of BCG infection and timely anti-tuberculosis treatment need to be synchronized and implemented as quickly as possible. Shi B has reported a pediatric case of HPS secondary to BCG disease in which the treatment ultimately failed due to a delayed diagnosis (Maron et al., 2021). Our strategy was to maintain a full-dose regimen of HPS and avoid neutropenia. In addition, we adopted the method of mNGS combined with tNGS to quickly identify *M.bovis* and its resistance genes, and developed a more targeted drug regimen as soon as possible. At the same time, levofloxacin, a second-line anti-tuberculosis drug, was added in time to achieve full coverage of MTB.

We reviewed the literature reporting BCG infection in X-SCID patients over the past thirty years. All seven children presented with fever, cough, chronic diarrhea, or other nonspecific manifestations. All reported cases began with clinical symptoms in infancy. Screening results for latent TB infection were negative. Two of the children had male relatives who died from infantile infections. After anti-tuberculosis therapy and undergoing transplantation or gene therapy, four patients had achieved immunologic reconstitution. Unlike our case, which used mNGS combined with tNGS to detect *M.bovis* and found its drug resistance, most of the MTB reported in these literatures were detected by PCR, and no drug resistance genes were found (Culic et al., 2004; Huang et al., 2006; Jaing et al., 2006; Bacalhau et al., 2011; Shi et al., 2020; Maron et al., 2021; Liu et al., 2022). The conventional etiological culture-based detection is time-consuming (generally 2–6 weeks) (Assemie et al., 2020). Positive culture specimens need a further drug sensitivity test, but it takes another 2 to 4 weeks. In recent years, the use of mNGS for the detection of MTB and NTM has received considerable attention because of its shortened turnaround time and significant sensitivity (Gu et al., 2019). mNGS does not need to refer to the culture results to analyze the results, nor does it require specific amplification. It can directly conduct non-differential and non-selective high-throughput sequencing of nucleic acids in clinical samples and detect pathogenic microorganisms (including viruses, bacteria, fungi, and parasites) promptly and objectively. Compared with traditional methods, the sensitivity and positive predictive value of mNGS in the diagnosis of MTB infection were significantly higher than those of smear, culture and GeneXpert MTB/RIF. MTB is an intracellular growth bacterium that releases fewer extracellular nucleic acids. MTB was considered positive by mNGS when at least 1 read was mapped, due to the difficulty of DNA extraction and low possibility of contamination (Miao et al., 2018). Due to the completely conserved DNA sequences in the MTB complex and insufficient nucleic acid sequence information, it is difficult for mNGS to determine the species within the MTB complex and detect multiple resistant mutations (Chae and Shin, 2018). tNGS technology is a targeted enrichment sequencing technology based on NGS, which combines multiple PCR amplifications with sequencing technology. It has a clear range of target pathogens. The subspecies and subtypes can be further differentiated. This method contributes to the early diagnosis of BCG disease. In addition, it can capture numerous different drug resistance gene sequences, which enable molecular diagnosis of drug resistance (Murphy et al., 2023). Compared with mNGS, tNGS has the remarkable advantages of a clear pathogen spectrum and a low sequencing cost. We adopted mNGS combined with tNGS for species identification. This combination of the two methods provides rapid, sensitive, and specific identification of *M. bovis* and detects resistance to antituberculosis drugs. Furthermore, the combination of mNGS and tNGS is also useful for evolutionary tracking, strain typing, and virulence prediction of MTB.

It is extremely rare for HPS to be triggered by BCG. Both HPS and BCGosis have a high mortality rate, so early diagnosis of the cause and treatment are critical. The favorable sensitivity and specificity of mNGS combined with tNGS in the identification of *M.bovis* can help us identify specific pathogens and drug resistance. However, in the detection of susceptibility of MTB to various drugs, the genotype method cannot completely replace traditional phenotyping because the genetics of drug resistance in MTB are not completely clear (Ko et al., 2019). At present, there is relatively little literature on the value of high-throughput sequencing technology in the diagnosis of BCGosis. In this case report, we only show the experience of a single patient. While it provides in-depth information about this particular case, it's not generalizable to a larger population. We also did not map a phylogenetic tree analysis of *M.bovis* based on existing NGS information. Relevant case-control studies and cohort studies can be designed to further explore the clinical significance of mNGS combined with tNGS in the future. The clinical significance of mNGS combined with tNGS needs to be further explored in the future. With the continuous optimization of high-throughput sequencing technology, the recognition of the value of high-throughput sequencing technology in the diagnosis of latent MTB infection, BCGosis and drug resistance will gradually deepen.

## Conclusion

In conclusion, we report the successful treatment of a BCGosis case. For PID, BCG vaccine injection may lead to BCGosis. This disease can lead to some non-specific complications, such as HPS. The mNGS combined with the tNGS may be a good choice to identify the *M.bovis* and its drug resistance in children. Their combined application significantly improved the sensitivity and specificity of the detection of *M.bovis*.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Ethics statement

The studies involving humans were approved by West China Second University Hospital of Sichuan University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

## Author contributions

HZ: Conceptualization, Data curation, Investigation, Resources, Writing – original draft. YL: Data curation, Methodology, Writing – review & editing. ZZ: Validation, Visualization, Writing – review & editing. HL: Formal Analysis, Project administration, Supervision, Writing – review & editing. DL: Data curation, Supervision, Writing – review & editing. SW: Conceptualization, Data curation, Methodology, Resources, Software, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## Acknowledgments

The authors thank all the clinical staff who contributed to the study.

## Conflict of interest

Author ZZ is employed by Hugobiotech Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1341236/full#supplementary-material

## References

Allen, C. E., and McClain, K. L. (2015). Pathophysiology and epidemiology of hemophagocytic lymphohistiocytosis. *Hematol. Am. Soc. Hematol. Educ. Program* 2015, 177–182. doi: 10.1182/asheducation-2015.1.177

Amanati, A., Pouladfar, G., Kadivar, M. R., Sanaei Dashti, A., Jafarpour, Z., Haghpanah, S., et al. (2017). A 25-year surveillance of disseminated Bacillus Calmette-Guérin disease treatment in children in Southern Iran. *Medicine* 96 (52), e9035. doi: 10.1097/MD.0000000000009035

Assemie, M. A., Alene, M., Petrucka, P., Leshargie, C. T., and Ketema, D. B. (2020). Time to sputum culture conversion and its associated factors among multidrug-resistant tuberculosis patients in Eastern Africa: A systematic review and meta-analysis. *Int. J. Infect. Dis.* 98, 230–236. doi: 10.1016/j.ijid.2020.06.029

Bacalhau, S., Freitas, C., Valente, R., Barata, D., Neves, C., Schäfer, K., et al. (2011). Successful handling of disseminated BCG disease in a child with severe combined immunodeficiency. *Case Rep. Med.* 2011, 527569. doi: 10.1155/2011/527569

Beviere, M., Reissier, S., Penven, M., Dejoies, L., Guerin, F., Cattoir, V., et al. (2023). The role of next-generation sequencing (NGS) in the management of tuberculosis: practical review for implementation in routine. *Pathogens* 12 (8), 978. doi: 10.3390/pathogens12080978

Bosco, M. C. (2019). Macrophage polarization: Reaching across the aisle? *J. Allergy Clin. Immunol.* 143 (4), 1348–1350. doi: 10.1016/j.jaci.2018.12.995

Chae, H., and Shin, S. J. (2018). Importance of differential identification of Mycobacterium tuberculosis strains for understanding differences in their prevalence, treatment efficacy, and vaccine development. *J. Microbiol.* 56 (5), 300–311. doi: 10.1007/s12275-018-8041-3

Culic, S., Kuzmic, I., Culic, V., Martinic, R., Kuljis, D., Pranic-Kragic, A., et al. (2004). Disseminated BCG infection resembling langerhans cell histiocytosis in an infant with severe combined immunodeficiency: a case report. *Pediatr. Hematol. Oncol.* 21 (6), 563–572. doi: 10.1080/08880010490477257

de la Morena, M. T., Leonard, D., Torgerson, T. R., Cabral-Marques, O., Slatter, M., Aghamohammadi, A., et al. (2017). Long-term outcomes of 176 patients with X-linked hyper-IgM syndrome treated with or without hematopoietic cell transplantation. *J. Allergy Clin. Immunol.* 139 (4), 1282–1292. doi: 10.1016/j.jaci.2016.07.039

Dvorak, C. C., Haddad, E., Heimall, J., Dunn, E., Cowan, M. J., Pai, S. Y., et al. (2023). The diagnosis of severe combined immunodeficiency: Implementation of the PIDTC 2022 Definitions. *J. Allergy Clin. Immunol.* 151 (2), 547–555.e5. doi: 10.1016/j.jaci.2022.10.021

Gu, W., Miller, S., and Chiu, C. Y. (2019). Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol.* 14, 319–338. doi: 10.1146/annurev-pathmechdis-012418-012751

Hardin, O., Lokhnygina, Y., and Buckley, R. H. (2022). Long-term clinical outcomes of severe combined immunodeficiency patients given nonablative marrow transplants. *J. Allergy Clin. Immunol. Pract.* 10 (4), 1077–1083. doi: 10.1016/j.jaip.2021.11.032

Henter, J. I., Aricò, M., Egeler, R. M., Elinder, G., Favara, B. E., Filipovich, A. H., et al. (1997). HLH-94: a treatment protocol for hemophagocytic lymphohistiocytosis. *HLH study Group Histiocyte Society. Med. Pediatr. Oncol.* 28 (5), 342–347. doi: 10.1002/(sici)1096-911x(199705)28:5<342::aid-mpo3>3.0.co;2-h

Henter, J. I., Horne, A., Aricó, M., Egeler, R. M., Filipovich, A. H., Imashuku, S., et al. (2007). HLH-2004: Diagnostic and therapeutic guidelines for hemophagocytic lymphohistiocytosis. *Pediatr. Blood Cancer* 48 (2), 124–131. doi: 10.1002/pbc.21039

Huang, L. H., Shyur, S. D., Weng, J. D., Huang, F. H., and Tzen, C. Y. (2006). Disseminated cutaneous bacille calmette-guérin infection identified by polymerase chain reaction in a patient with X-linked severe combined immunodeficiency. *Pediatr. Dermatol.* 23 (6), 560–563. doi: 10.1111/j.1525-1470.2006.00309.x

Jaing, T. H., Lee, W. I., Lin, T. Y., Huang, J. L., Chen, S. H., and Chow, R. (2006). Successful unrelated mismatched cord blood transplantation in an infant with severe combined immunodeficiency and Mycobacterium bovis bacillus Calmette-Guerin disease. *Pediatr. Transplant.* 10 (4), 501–504. doi: 10.1111/j.1399-3046.2006.00490.x

Kimura, S. I., Sato, M., Misaki, Y., Yoshimura, K., Gomyo, A., Hayakawa, J., et al. (2020). Prospective validation of the L-index reflecting both the intensity and duration of lymphopenia and its detailed evaluation using a lymphocyte subset analysis after allogeneic hematopoietic stem cell transplantation. *Transpl. Immunol.* 58, 101262. doi: 10.1016/j.trim.2019.101262

Ko, D. H., Lee, E. J., Lee, S. K., Kim, H. S., Shin, S. Y., Hyun, J., et al. (2019). Application of next-generation sequencing to detect variants of drug-resistant Mycobacterium tuberculosis: genotype-phenotype correlation. *Ann. Clin. Microbiol. Antimicrob.* 18 (1), 2. doi: 10.1186/s12941-018-0300-y

Lange, C., Aaby, P., Behr, M. A., Donald, P. R., Kaufmann, S. H. E., Netea, M. G., et al. (2022). 100 years of Mycobacterium bovis bacille Calmette-Guérin. *Lancet Infect. Dis.* 22 (1), e2–e12. doi: 10.1016/S1473-3099(21)00403-5

Liu, S., Huo, F., Dai, G., Wu, J., Qin, M., Mao, H., et al. (2022). Case Report: Immune reconstitution inflammatory syndrome after hematopoietic stem cell transplantation for severe combined immunodeficiency. *Front. Immunol.* 13. doi: 10.3389/fimmu.2022.960749

Liu, Q., Tian, Y., Zhao, X., Jing, H., Xie, Q., Li, P., et al. (2015). NMAAP1 expressed in BCG-activated macrophage promotes M1 macrophage polarization. *Mol. Cells* 38 (10), 886–894. doi: 10.14348/molcells.2015.0125

Locatelli, F., Kabbara, N., Ruggeri, A., Ghavamzadeh, A., Roberts, I., Li, C. K., et al. (2013). Outcome of patients with hemoglobinopathies given either cord blood or bone marrow transplantation from an HLA-identical sibling. *Blood* 122 (6), 1072–1078. doi: 10.1182/blood-2013-03-489112

Marciano, B. E., Huang, C. Y., Joshi, G., Rezaei, N., Carvalho, B. C., Allwood, Z., et al. (2014). BCG vaccination in patients with severe combined immunodeficiency: Complications, risks, and vaccination policies. *J. Allergy Clin. Immunol.* 133 (4), 1134–1141. doi: 10.1016/j.jaci.2014.02.028

Maron, G., Kaste, S., Bahrami, A., Neel, M., Malech, H. L., Puck, J. M., et al. (2021). Successful SCID gene therapy in infant with disseminated BCG. *J. Allergy Clin. Immunol. Pract.* 9 (2), 993–995.e1. doi: 10.1016/j.jaip.2020.09.004

Miao, Q., Ma, Y., Wang, Q., Pan, J., Zhang, Y., Jin, W., et al. (2018). Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin. Infect. Dis.* 67 (suppl_2), S231–S240. doi: 10.1093/cid/ciy693

Murphy, S. G., Smith, C., Lapierre, P., Shea, J., Patel, K., Halse, T. A., et al. (2023). Direct detection of drug-resistant Mycobacterium tuberculosis using targeted next generation sequencing. *Front. Public Health* 11. doi: 10.3389/fpubh.2023.1206056

Norouzi, S., Aghamohammadi, A., Mamishi, S., Rosenzweig, S. D., and Rezaei, N. (2012). Bacillus Calmette-Guérin (BCG) complications associated with primary immunodeficiency diseases. *J. Infect.* 64 (6), 543–554. doi: 10.1016/j.jinf.2012.03.012

Ong, R. Y. L., Chan, S. B., Chew, S. J., Liew, W. K., Thoon, K. C., Chong, C. Y., et al. (2020). Disseminated bacillus-calmette-guerin infections and primary immunodeficiency disorders in Singapore: A single center 15-year retrospective review. *Int. J. Infect. Dis.* 97, 117–125. doi: 10.1016/j.ijid.2020.05.117

Padhi, S., Ravichandran, K., Sahoo, J., Varghese, R. G., and Basheer, A. (2015). Hemophagocytic lymphohistiocytosis: An unusual complication in disseminated Mycobacterium tuberculosis. *Lung India* 32 (6), 593–601. doi: 10.4103/0970-2113.168100

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17 (5), 405–424. doi: 10.1038/gim.2015.30

Shi, B., Chen, M., Xia, Z., Xiao, S., Tang, W., Qin, C., et al. (2020). Hemophagocytic syndrome associated with Mycobacterium bovis in a patient with X-SCID: a case report. *BMC Infect. Dis.* 20 (1), 711. doi: 10.1186/s12879-020-05421-9

Spolski, R., Li, P., and Leonard, W. J. (2018). Biology and regulation of IL-2: from molecular mechanisms to human therapy. *Nat. Rev. Immunol.* 18 (10), 648–659. doi: 10.1038/s41577-018-0046-y

Tangye, S. G., Al-Herz, W., Bousfiha, A., Chatila, T., Cunningham-Rundles, C., Etzioni, A., et al. (2020). Human inborn errors of immunity: 2019 update on the classification from the international union of immunological societies expert committee. *J. Clin. Immunol.* 40 (1), 24–64. doi: 10.1007/s10875-019-00737-x

Thomas, L., Verghese, V. P., Chacko, A., Michael, J. S., and Jeyaseelan, V. (2022). Accuracy and agreement of the Tuberculin Skin Test (TST) and the QuantiFERON-TB Gold In-tube test (QFT) in the diagnosis of tuberculosis in Indian children. *Indian J. Med. Microbiol.* 40 (1), 109–112. doi: 10.1016/j.ijmmb.2021.05.022

Toskas, A., Karamanakos, G., Bountzona, I., Maria, A., and Manganas, K. (2022). Secondary haemophagocytic lymphohistiocytosis syndrome (HLH) after intravesical instillation of bacillus calmette-guérin (BCG): A case report and review of the literature. *Eur. J. Case Rep. Intern. Med.* 9 (10), 3395. doi: 10.12890/2022_003395

Zhang, C., Zhang, Z.-Y., Wu, J.-F., Tang, X.-M., Yang, X.-Q., Jiang, L.-P., et al. (2011). Clinical characteristics and mutation analysis of X-linked severe combined immunodeficiency in China. *World J. Pediatr.* 9 (1), 42–47. doi: 10.1007/s12519-011-0330-4

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 04 March 2024
DOI 10.3389/fcimb.2024.1356804



OPEN ACCESS

EDITED BY
Qing Wei,
Genskey Co. Ltd, China

REVIEWED BY
Mohamed A. Abouelkhair,
The University of Tennessee, Knoxville, United States
Michael John Calcutt,
University of Missouri, United States

*CORRESPONDENCE
Jun Tan
Doctortan0718@126.com

RECEIVED 16 December 2023
ACCEPTED 19 February 2024
PUBLISHED 04 March 2024

CITATION
Tan J, Wu L, Zhan L, Sheng M, Tang Z, Xu J and Ma H (2024) Optimal selection of specimens for metagenomic next-generation sequencing in diagnosing periprosthetic joint infections. *Front. Cell. Infect. Microbiol.* 14:1356804. doi: 10.3389/fcimb.2024.1356804



# Optimal selection of specimens for metagenomic next-generation sequencing in diagnosing periprosthetic joint infections

Jun Tan[1]*, Lingxiao Wu[2], Lijuan Zhan[3], Minkui Sheng[1], Zhongxin Tang[1,2], Jianzhong Xu[2] and Haijun Ma[1]

[1]Department of Mini-invasive Spinal Surgery, The Third People's Hospital of Henan Province, Zhengzhou, Henan, China, [2]Department of Orthopedic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China, [3]Department of Neurology, People's Hospital of Zhengzhou, Zhengzhou, Henan, China

**Objective:** This study aimed to assess the diagnostic value of metagenomic next-generation sequencing (mNGS) across synovial fluid, prosthetic sonicate fluid, and periprosthetic tissues among patients with periprosthetic joint infection (PJI), intending to optimize specimen selection for mNGS in these patients.

**Methods:** This prospective study involved 61 patients undergoing revision arthroplasty between September 2021 and September 2022 at the First Affiliated Hospital of Zhengzhou University. Among them, 43 cases were diagnosed as PJI, and 18 as aseptic loosening (AL) based on the American Musculoskeletal Infection Society (MSIS) criteria. Preoperative or intraoperative synovial fluid, periprosthetic tissues, and prosthetic sonicate fluid were collected, each divided into two portions for mNGS and culture. Comparative analyses were conducted between the microbiological results and diagnostic efficacy derived from mNGS and culture tests. Furthermore, the variability in mNGS diagnostic efficacy for PJI across different specimen types was assessed.

**Results:** The sensitivity and specificity of mNGS diagnosis was 93% and 94.4% for all types of PJI specimens; the sensitivity and specificity of culture diagnosis was 72.1% and 100%, respectively. The diagnostic sensitivity of mNGS was significantly higher than that of culture ($X^2$ = 6.541, $P$=0.011), with no statistically significant difference in specificity ($X^2$ = 1.029, $P$=0.310). The sensitivity of the synovial fluid was 83.7% and the specificity was 94.4%; the sensitivity of the prosthetic sonicate fluid was 90.7% and the specificity was 94.4%; and the sensitivity of the periprosthetic tissue was 81.4% and the specificity was 100%. Notably, the mNGS of prosthetic sonicate fluid displayed a superior pathogen detection rate compared to other specimen types.

**Conclusion:** mNGS can function as a precise diagnostic tool for identifying pathogens in PJI patients using three types of specimens. Due to its superior ability in pathogen identification, prosthetic sonicate fluid can replace synovial fluid and periprosthetic tissue as the optimal sample choice for mNGS.

KEYWORDS

**metagenomic, prosthesis-related infections, next-generation sequencing - NGS, diagnosis, sonication**

# 1 Introduction

Periprosthetic joint infection (PJI) stands as one of the most catastrophic and challenging complications following total joint arthroplasty (TJA) (Gallo, 2020; Rietbergen et al., 2016). Rapid and accurate identification of pathogenic bacteria is crucial for determining the appropriate surgical approach and for selecting effective antibacterial drugs (Hersh et al., 2019; Young et al., 2023). Although widely used for diagnosing PJI, traditional blood tests like C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), and white blood cell counts (WBC) struggle to identify the causative organism (Nodzo et al., 2015). Microbial culture is extensively employed for diagnosing the pathogenesis of PJI; however, it exhibits a high false-negative rate. Approximately 40% of clinically diagnosed PJI cases yield negative culture results due to the existence of bacterial biofilms and prior antibiotic treatment (Rak et al., 2013; Yoon et al., 2017; Malekzadeh et al., 2010). Molecular diagnostic techniques based on polymerase chain reaction (PCR) are extensively employed in detecting pathogenic microorganisms (Hartley and Harris, 2014). However, multiplex PCR techniques exhibit low diagnostic sensitivity, and 16S rDNA/rRNA PCR encounters difficulty in identifying fungal or multi-pathogenic infections, displaying a poor ability in identifying contaminating bacteria (Villa et al., 2017; Huang et al., 2018).

Next-generation sequencing, particularly metagenomic next-generation sequencing (mNGS), is a swiftly developing and extensively employed technology in the clinical diagnosis of pathogenic microorganisms (Chiu and Miller, 2019). mNGS integrates high-throughput sequencing and bioinformatics analysis to discern sequence data from all nucleic acid fragments within samples, detecting microbial species and their respective abundances relying on the BLAST database (Gu et al., 2021). Wilson et al. (2014) initially utilized mNGS for clinically diagnosing neuroleptospirosis in a 14-year-old patient experiencing severe meningoencephalitis. The successful diagnosis enabled targeted antibiotic treatment, leading to the patient's eventual recovery and demonstrating, for the first time, the utility of mNGS in clinical diagnosis (Wilson et al., 2014). The progressive application of mNGS in diagnosing infectious diseases has demonstrated enhanced detection of pathogenic microorganisms, particularly in conditions like ocular, intracranial, and pulmonary infections (O'Flaherty et al., 2018; Brown et al., 2018; Doan et al., 2017). However, there is a scarcity of studies utilizing mNGS for detecting pathogenic microorganisms in PJI, primarily relying on a single specimen for pathogen detection, such as synovial fluid, prosthetic sonicate fluid, and periprosthetic tissue (Dudareva et al., 2018). mNGS results correlate with the tissue type, volume of the sampled specimen, and the proportions of human DNA and pathogenic bacterial DNA in the sample (Thoendel et al., 2016). Variations exist in mNGS test outcomes among distinct samples obtained from the same patient (Miao et al., 2018). Choosing the most suitable specimen is crucial for pathogen identification, not just for traditional culture methods but also for molecular diagnostic technologies (Huang et al., 2018; Larsen et al., 2018; Flurin et al., 2022).

This study aimed to prospectively assess the diagnostic efficacy of mNGS across three different specimen types obtained from patients with PJI. This approach should facilitate the selection of more suitable tests for detecting pathogens in these patients by employing multiple samples.

# 2 Materials and methods

## 2.1 Patient selection

This study received approval from our institution's Medical Ethics Committee (Approval No:2020-KY-821), and all patients provided informed consent by signing an informed consent form. We prospectively included 61 patients who underwent hip or knee revision procedures at our institution from September 2021 to September 2022 due to PJI or aseptic loosening (AL).

The diagnosis of PJI after TJA is based primarily on the American Musculoskeletal Infection Association (MSIS) criteria, confirmed by meeting one major criterion or 4 out of the 6 minor criteria (Parvizi et al., 2011). The major criteria include (1): Isolation of the same pathogen from two distinct tissue or fluid specimens sourced from infected joints (2); Presence of a sinus tract communicating with the prosthesis. Minor criteria consist of (1): Elevated ESR and CRP levels (2); Increased synovial fluid WBC (3); Elevated synovial fluid neutrophil percentage (4); Presence of pus in the affected joint (5); Isolation of a microorganism in either

periprosthetic tissue or fluid culture (6); The intraoperative microscopic examination of frozen sections of infected tissue revealed an observation of greater than 5 neutrophils in all 5 high-magnification fields.

The inclusion criteria comprised (1): Patients diagnosed with PJI or AL according to the MSIS diagnostic criteria, undergoing revision surgery at our institution (2); Patients undergoing mNGS testing and microbial culture analysis of synovial fluid, prosthetic sonication fluid from periprosthetic tissue, and periprosthetic tissue (3); Patients with complete medical records. The exclusion criteria were (1): Patients with samples suspected of contamination during collection, transportation, or processing (2); Patients with an unclear diagnosis or incomplete clinical information (3); Patients presenting other inflammatory lesions or malignancies potentially affecting the outcomes.

## 2.2 Specimen collection

The same surgeon conducted arthrocentesis to collect synovial fluid under ultrasound guidance either before or during the operation.

Ultrasound cleavage of the prosthesis removed intraoperatively was performed according to the method of Trampuz et al. (2007) for reference. The process included placing the extracted implants in a sterile, sealable container, submerging them in 500 ml of Ringer's solution, subjecting the samples to 30 seconds of vortexing and oscillation, and subsequently performing sonication for five min (40 ± 2 kHz, 0.22 ± 0.04 W/cm$^2$) in a sterile water bath. Subsequently, after vortexing and oscillation for an additional 30 seconds, the sonicated product was transferred to a centrifuge tube and centrifuged at 4000 rpm/min for 15 minutes. The supernatant was then removed to collect the prosthetic sonicated fluid.

Intraoperatively, tissue samples were collected from five different sites with the most obvious inflammation, subsequently sheared, ground, lysed, and centrifuged to produce tissue homogenates for further analysis.

The three types of specimens obtained as described earlier were split into two portions: one for culture, while the other was frozen at -80°C for mNGS.

## 2.3 Microbial culture

Synovial fluid and prosthetic sonication fluid direct smears underwent acid-fast staining and were then cultured aerobically and anaerobically (35-37°C, 5%-7% $CO_2$) on blood agar plates for 6 and 14 days, respectively. The residual synovial fluid and prosthetic sonication fluid were introduced into BACTEC Peds Plus/F bottles and then incubated for 5 days in a BACTEC 9240 instrument (Becton Dickinson, Cockeysville, MD, USA). Positive blood culture bottles were Gram stained and then inoculated. Periprosthetic tissue homogenates were cultured on blood agar plates and incubated aerobically, anaerobically, and for fungal growth for 14 days.

Species identification and drug sensitivity analyses were conducted utilizing the VITEK 2-Compact (bioMérieux, Lyon, France), a fully automated system for bacterial identification and drug sensitivity analysis.

## 2.4 Metagenomic next-generation sequencing

### 2.4.1 DNA extraction

Total genomic DNA was extracted from samples using the TIANamp Micro DNA Kit (DP316, TIANamp Biotech, Beijing, China) according to the manufacturer's instructions. The extracted DNA underwent sonication using the Covaris S220 (Covaris, Woburn, MA, USA) to produce fragments ranging from 200 bp to 300 bp.

### 2.4.2 DNA library preparation and sequencing

The construction of DNA library followed the protocol of the Nextera XT library construction Kit (Illumina, San Diego, CA, USA), where the extracted DNA was initially fragmented into ~300 bp fragments, followed by the addition of different index sequences. The library size and quantification were assessed using the Agilent 2100 bioanalyzer system (Agilent Technologies, CA, USA). Subsequent to this, accurate quantification was reaffirmed through qPCR (Bio-Rad CFX96, Hercules, CA, USA). Following the equal mixing of libraries, high-throughput sequencing was performed on the Illumina Nextseq 550 DX sequencing platform.

### 2.4.3 Bioinformatics analysis

Initial filtering of the raw data was conducted using FastQC software (version 0.11.7, http://bioinformatics.babraham.ac.uk/projects/fastqc/), including removing low-quality data, sequences shorter than 35 bp, repeated sequences, and adaptor contamination to produce high-quality sequencing data. Human host sequences were removed through mapping to the human reference genome (hg19) using the Burrows-Wheeler Alignment (version 0.7.13, http://bio-bwa.sourceforge.net) (Li and Durbin, 2009). The remaining data underwent analysis with Kraken2 (Wood et al., 2019) (version 2.1.1) and Bracken (Lu et al., 2017) (version 2.6.2) to identify and quantify pathogenic microorganisms. The resultant clinical test reports deliver accurate and reliable information on microbiological test results. The microbial reference whole-genome data were sourced from the National Center for Biotechnology Information (http://ncbi.nlm.nih.gov/genome) (Figure 1).

### 2.4.4 Definition of the mNGS threshold

Relative abundance in genus level (RAG) refers to the proportion of matched microbial genera among all microorganisms of the same type (bacteria, fungi, viruses, parasites) detected. The genome coverage rate (CR) primarily represents the proportion of the detected pathogenic nucleic acid sequence length in relation to the total length of the reference genome used for comparison.

In this study, saline served as a negative control. mNGS were conducted concurrently for the same batch of samples (both the negative control and clinical samples). Microorganisms identified in negative controls are commonly viewed as environmental

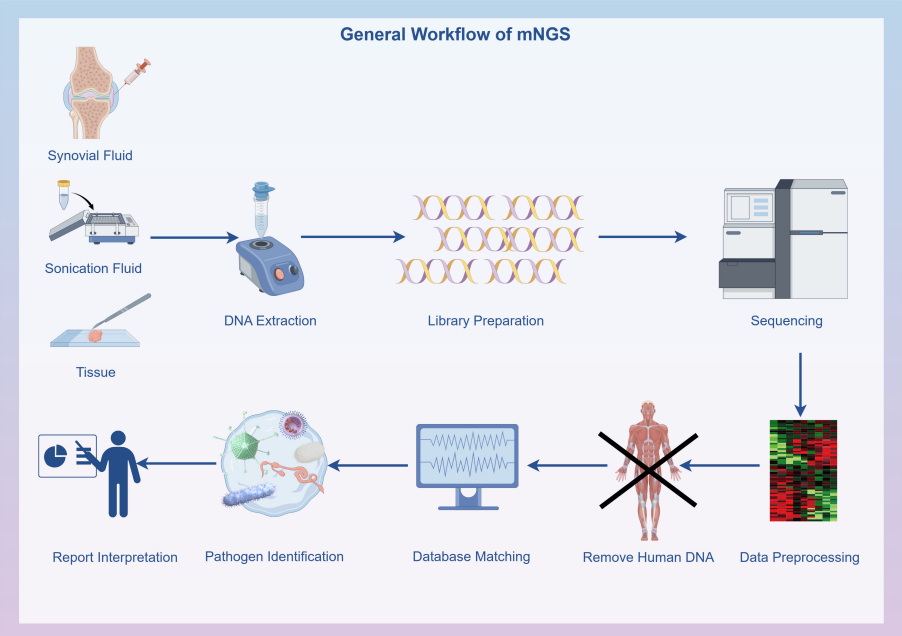



FIGURE 1
General workflow of metagenomic next-generation sequencing: Specimen collection, DNA extraction, DNA library preparation, Sequencing, Bioinformatics analysis and Result interpretation. The flowchart was drawn with Figdraw (www.figdraw.com).

contaminants. If these microorganisms are detected in the same batch of clinical samples, they are generally classified as "background microorganisms" rather than "pathogenic microorganism".

Referring to the proposal of Li et al. (2018) and considering our experimental findings, the thresholds were set as follows (1): As *Burkholderia, Aspergillus, Delftia, Sodaria, Sphingobium, Alternaria, Ralstonia, Albugo*, etc. can be detected in other types of samples from the same laboratory and are seldom associated with causing PJI (Zhang et al., 2019), we categorized them as background microorganisms. These microorganisms were recognized as causative pathogens only when RAG >80% (Ivy et al., 2018) (2). RAG >15% and the number of reads >100 were employed as the optimal thresholds for bacterial identification (3). Considering the limited quantity of nucleic acid obtained from a fungus, we established RAG >15% and the number of reads >50 as the optimal thresholds for fungi (Miao et al., 2018) (4). *Mycobacterium tuberculosis* was deemed positive when it met the criterion of having at least one read aligned to the reference genome at the species or genus level.

## 2.5 Statistical analysis

All clinical and laboratory data underwent analysis employing the non-parametric Mann-Whitney U test, chi-squared test, and independent-samples t-test. The sensitivities and specificities of microbiological tests were compared using the McNemar's test to assess related proportions. Data analyses were conducted using SPSS 23.0 software (SPSS 23.0 for Windows, SPSS Inc., Chicago, IL, USA). $P < 0.05$ were deemed statistically significant.

# 3 Results

## 3.1 Patient characteristics

This study encompassed 61 patients, with 43 diagnosed with PJI and classified in the PJI group, while 18 patients diagnosed with AL were classified in the AL group based on the MSIS criteria. The PJI group comprised 43 patients (26 knees and 17 hips): 24 males, 19 females, with an average age of 62.31 ± 10.15 years, and a Body Mass Index (BMI) of 24.3 ± 2.9 kg/m$^2$. There were 18 patients (11 knees and 7 hips) in the AL group, including 8 males and 10 females; with an average age of 63.57 ± 11.37 years, and BMI of 25.1 ± 2.15 kg/m$^2$. There were no statistically significant differences in gender, age, or BMI between the two groups ($P > 0.05$). The characteristics of all the included patients are showed in Table 1.

## 3.2 mNGS and microbial culture results

A median 24,332,918 raw reads (interquartile range (IQR) 19,126,072 to 27,035,628) were generated from each sample sequenced. The human host nucleic acid sequences averaged 95.28% (ranging from 90.63% to 98.71%). Differences in raw reads and proportions of human nucleic acid sequences between synovial fluid, prosthetic sonication fluid, and periprosthetic tissue samples from the PJI and AL groups did not show statistical significance ($P>0.05$, Figure 2).

In the PJI group, 31 patients tested positive in culture, while 40 patients exhibited positive mNGS results. Within the AL group, one

TABLE 1 The characteristics of the PJI and AL groups.

| Variable | Prosthetic joint infection (n = 43) | Aseptic failure (n = 18) | *P* |
|---|---|---|---|
| Sex, n (%) | | | 0.417* |
| Male | 24 (55.8%) | 8 (44.4%) | |
| Female | 19 (44.2%) | 10 (55.6%) | |
| Age, yrs, mean (SD) | 62.31 (10.15) | 63.57 (11.37) | 0.079† |
| BMI, kg/m$^2$ (SD) | 24.3 (2.91) | 25.1 (2.15) | 0.305† |
| Site of arthroplasty, n (%) | | | 0.962* |
| Knee | 26 (60.5%) | 11 (61.1%) | |
| Hip | 17 (39.5%) | 7 (38.9%) | |
| Sinus tract, n (%) | 4 (9.3%) | 0 (0%) | 0.012* |
| Antibiotics received two weeks prior to surgery, n (%) | 27 (62.8%) | 2 (11.1%) | <0.001* |
| Serum ESR, mm/h, mean (SD) | 65.28 (38.51) | 15.38 (12.07) | 0.022‡ |
| Serum CRP, mg/dl, mean (SD) | 45.57 (30.37) | 4.03 (1.38) | 0.015‡ |
| Synovial WBC, cells/ml mean (SD) | 17931 (8237.07) | 612.91 (533.76) | 0.004‡ |
| Synovial PMN, % (SD) | 73.55 (16.74) | 38.27 (13.58) | 0.031‡ |

†Independent-samples t-test. *Chi-squared test. ‡Mann-Whitney U test. PMN, polymorphonuclear neutrophils; SD, standard deviation; WBC, white blood cell. ESR, erythrocyte sedimentation rate;CRP, C-reactive protein.

patient exhibited positive mNGS results, whereas all patients tested negative in culture.

The pathogens detected by both culture and mNGS are presented in Supplementary Table 1. Among the 43 PJI cases analyzed with culture, pathogenic microorganisms were isolated in 31 cases (72.1%), with all cases attributed to a single microorganism. Pathogenic microorganisms were isolated in 40 out of the 43 (93.0%) PJI cases using mNGS, 33 cases of infections were caused by a single microorganism, and 7 were polymicrobial.

In the comparison of diagnostic performance between mNGS and culture, mNGS demonstrated higher sensitivity, both in the combined analysis of all specimen types and when considering each individual type of specimen. In summary, the sensitivity of mNGS was 93.0%, which was significantly higher than that of culture (72.1%, $X^2 = 6.541$, $P$=0.011, Figure 3).

The turnaround time for mNGS in our patients was 24-28 hours. In contrast, the mean durations for pathogen culture feedback for bacteria, fungi, and mycobacteria were ≥3, 7, and 45 days, respectively. Thus, mNGS exhibits a time advantage in detection.

## 3.3 Consistency of mNGS and culture results

Comparing the mNGS and culture results among the 43 patients in the PJI group, 31 were double positive and 3 were double negative, while 9 showed positivity in mNGS but negativity in culture.

Among the double positive cases, 23 exhibited complete matching (*Staphylococcus aureus* in 10, *Staphylococcus epidermidis* in 5, *Staphylococcus haemolyticus* in 2, *Enterococcus faecalis* in 3, and *Escherichia coli* in 3), 7 showed partial matching (one of pathogens is the same). However, only one case exhibited a discordant result. Although *Enterobacter cloacae* was isolated by culture, the mNGS result indicated *S. epidermidis* as the pathogenic microorganism (Figure 4). The microbial results of mNGS and culture in PJI group are listed in Supplementary Table 2.

mNGS identified pathogenic microorganisms in 9 culture-negative cases, including *S. aureus* (2), *E. faecalis* (1), *S. haemolyticus* (1), *S. epidermidis* (1), *Mycobacterium abscessus* (1), *E. coli* (1), *E. cloacae* (1), *Candida albicans* (1). In the AL group, all culture results were negative. However, in one case, *E. coli* was detected in the sonication fluid by mNGS.

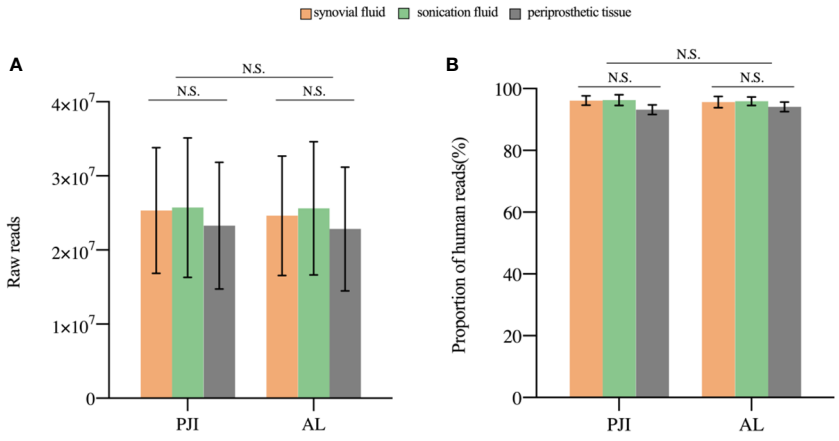



FIGURE 2
The sequencing characteristics of the different specimens. **(A)** The raw reads of the three samples in the PJI and AL groups was not statistically ($P$>0.05). **(B)** The difference in the proportion of human reads was not statistically significant ($P$>0.05). N.S., non-significant; PJI, periprosthetic joint infection; AL, aseptic failure.

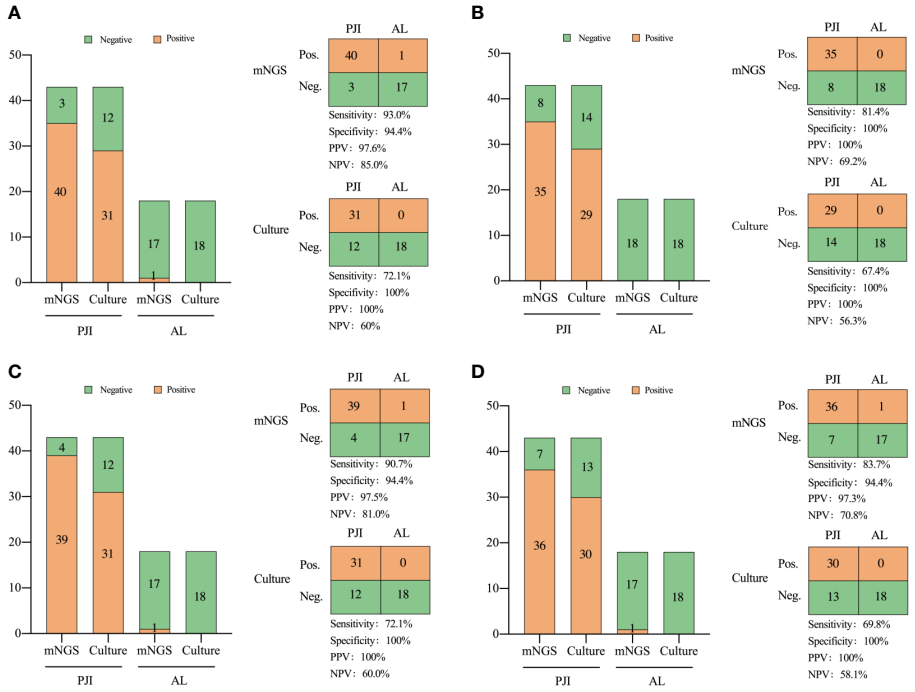



FIGURE 3
Comparison of the diagnostic efficiency of metagenomic next-generation sequencing and culture for PJI in: **(A)** all types of specimens, **(B)** periprosthetic tissue, **(C)** prosthetic sonication fluid, and **(D)** synovial fluid. NPV, negative predictive value; PPV, positive predictive value. Neg., negative; Pos., positive.

## 3.4 Consistency of mNGS for different types of specimens

The mNGS analysis of sonication fluid demonstrated a sensitivity of 90.7% and a specificity of 94.4%. The sensitivity of prosthetic sonication fluid mNGS was significantly higher than that of the synovial fluid and prosthetic tissues ($P < 0.05$).

Consistency in pathogen species detected by mNGS across the three specimen types demonstrates the reliability of mNGS and underscores its proficiency in detecting pathogens across varied specimen types. Among the results with pathogen matches, 26 cases exhibited identical positive sequencing, while 3 cases displayed identical negative sequencing. In the partially matched cases that most accurately reflected diagnostic capability, prosthetic

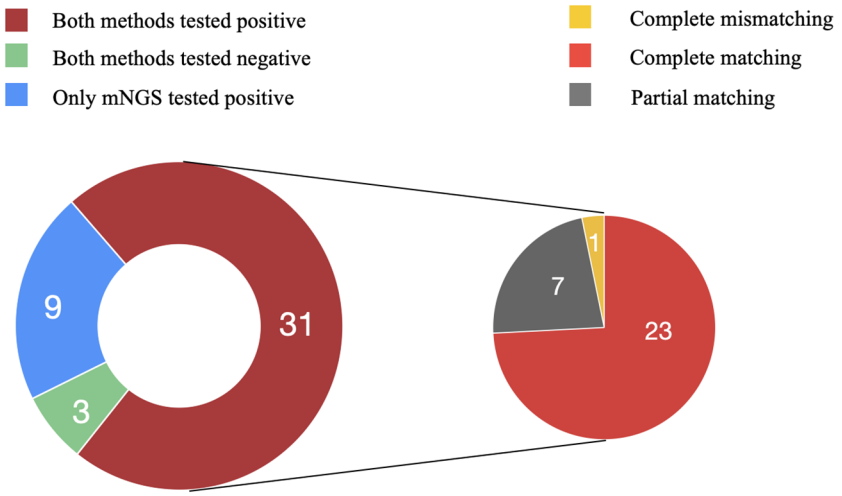



FIGURE 4
Consistency of metagenomic next-generation sequencing and culture results. Among the 43 patients in the PJI group, 31 were double positive and 3 were double negative, while 9 showed positivity in mNGS but negativity in culture. Among the double positive cases, 23 exhibited complete matching, 7 showed partial matching, and one case exhibited complete mismatching.

sonication fluid (92.9%, 13/14) demonstrated the highest pathogen detection rate (both single and multiple pathogens), followed by synovial fluid (71.4%, 10/14).

### 3.5 Effect of antibiotic exposure on test results

The results of mNGS and culture are affected by antibiotic exposure. In the PJI group, mNGS demonstrated a sensitivity of 92.3% in 27 patients (62.8%) who applied antibiotics within 2 weeks before the test, which was significantly better than that of culture in 62.3%, ($X^2$ = 6.857, $P$=0.009). Among the 16 patients not administered antibiotics, the sensitivity of the mNGS (93.8%) was not significantly different from that of the culture (87.5%) ($X^2$ = 0.368, $P$=0.544, Figure 5).

### 3.6 Combined application of mNGS and culture

The combined application of mNGS and culture may enhance pathogen detection capabilities. Therefore, as shown in Table 2, we explored the effectiveness of the combined application of mNGS and culture based on the results when using multiple samples. The combination of mNGS and culture showed the highest sensitivity of 93.0% (40/43) across all three specimens. Prosthetic sonication fluid exhibited superior mNGS performance compared to other two types of specimen, making it the optimal choice for mNGS specimens in clinical practice. The combined application of mNGS using prosthetic sonication fluid and a single specimen culture achieved the highest sensitivity of 93.0% (40/43). Nevertheless, the performance difference between combining mNGS of prosthetic sonication fluid with culture and using mNGS of prosthetic sonication fluid only was not significant (90.7% vs. 93.0%, $X^2$ = 0.156, $P$ = 0.693).

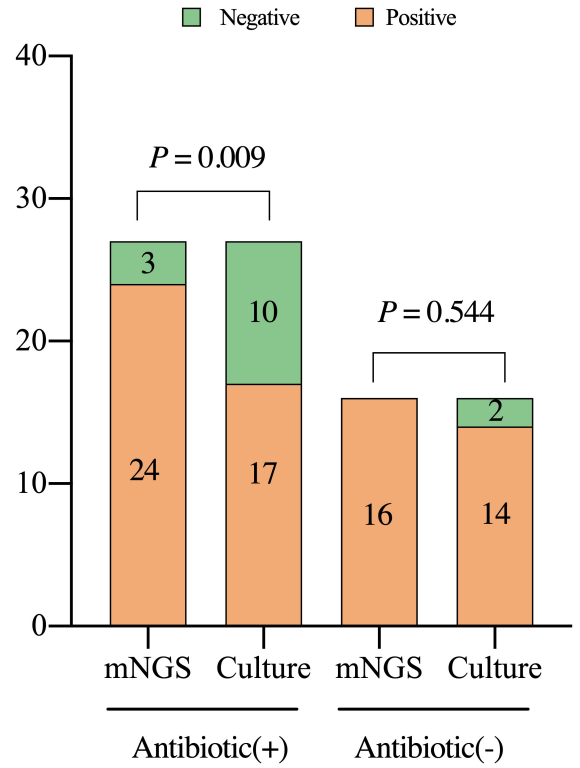



FIGURE 5
Detection characteristics of metagenomic next-generation sequencing and culture in periprosthetic joint infections among patients with or without previous treatment with antibiotics. Antibiotic (+), the patients were administered with antibiotics within two weeks prior to the test; mNGS, metagenomic next-generation sequencing.

## 4 Discussion

Accurate and rapid identification of infectious microorganisms in hip and knee PJI is crucial for guiding the application of antibiotics and improving infection control rate. At present, microbiological culture stands as the "gold standard" for diagnosing PJI. Nevertheless, the considerable false-negative rate of traditional culture presents several challenges in choosing antibiotics (Karbysheva et al., 2019). As a new molecular technology, mNGS sequences DNA fragments in clinical samples, extracting microbial sequence and species information through bioinformatics approaches (Adams et al., 2009). This study examined mNGS efficiency in detecting pathogens across three sample types using a multi-sample parallel control approach.

Our results show that the sensitivity of mNGS was higher than that of culture. These results align with our group's prior statistical analysis studies on mNGS (Tan et al., 2022) and are consistent with studies by Yu et al. (2023) and Huang et al. (2020) on mNGS. Moreover, for culture-negative patients, mNGS showed outstanding detection ability (75.0%, 9/12). Thoendel et al. (2016) demonstrated that mNGS yielded positive results in 94.8% of culture-positive PJI cases. Moreover, in culture-negative PJI cases, it remained positive in 43.9% of instances. Ivy et al. (2018) demonstrated that in cases of PJI with negative microbial cultures, mNGS detected pathogenic microorganisms in 84% of cases. This is mainly due to the fact that mNGS combines high-throughput sequencing with bioinformatics analysis to directly detect nucleic acids extracted from samples and detect pathogenic bacteria without enrichment. This fundamentally avoids the problem of fastidious growth conditions for some pathogenic bacteria that are difficult to culture. In addition, among the double positive cases, 96.8% (30/31) were either completely matched or partly matched, with only one case exhibiting complete mismatching. The proportion of matched and partly matched cases among the double positive cases was > 90%, indicating an acceptable concordance between mNGS and culture.

The mNGS of prosthetic sonication fluid can detect more pathogenic microorganisms than synovial fluid and periprosthetic tissue. Trampuz et al. (2007) initially employed sonication to treat 120 joint prostheses and observed an enhanced detection rate of bacteria. In our study, mNGS could detect pathogenic

TABLE 2 The combined application of mNGS and culture.

| Detection combination method | Sensitivity, n (%) | Specificity, n (%) | PPV, n (%) | NPV, n (%) | Youden Index |
|---|---|---|---|---|---|
| mNGS (PT+SF+PSF) + Culture (PT+SF+PSF) | 40/43 (93.0%) | 17/18 (94.4%) | 40/41 (97.6%) | 17/20 (85.0%) | 87.4 |
| mNGS (PT+SF+PSF) | 40/43 (93.0%) | 17/18 (94.4%) | 40/41 (97.6%) | 17/20 (85.0%) | 87.4 |
| Culture (PT+SF+PSF) | 31/43 (72.1%) | 18/18 (100%) | 31/31 (100%) | 18/30 (60.0%) | 72.1 |
| mNGS (PSF) + Culture (PT+SF+PSF) | 40/43 (93.0%) | 17/18 (94.4%) | 40/41 (97.6%) | 17/20 (85.0%) | 87.4 |
| mNGS (PSF) + Culture (PSF) | 40/43 (93.0%) | 17/18 (94.4%) | 40/41 (97.6%) | 17/20 (85.0%) | 87.4 |
| mNGS (PSF) + Culture (PT) | 39/43 (90.7%) | 17/18 (94.4%) | 39/40 (97.5%) | 17/21 (80.9%) | 85.1 |
| mNGS (PSF) + Culture (SF) | 39/43 (90.7%) | 17/18 (94.4%) | 39/40 (97.5%) | 17/21 (80.9%) | 85.1 |
| mNGS (PSF) | 39/43 (90.7%) | 17/18 (94.4%) | 39/40 (97.5%) | 17/21 (80.9%) | 85.1 |

mNGS, metagenomic next-generation sequencing; NPV, negative predictive value; PPV, positive predictive value; PSF, prosthetic sonicate fluid; PT, periprosthetic tissue; SF, synovial fluid; Youden Index = Sensitivity + Specificity − 1

microorganisms in the prosthetic sonication fluid of all 9 culture-negative PJI cases. Additionally, mNGS detected pathogens in 7 PJI cases involving multiple microorganisms. Thoendel et al. (2018) and Zhang et al. (2019) also showed that mNGS of prosthetic sonication fluid possesses high sensitivity in the detection of PJI pathogens. The detection rate of pathogenic microorganisms in the PJI group of our study surpassed that reported by Mei et al. (2023), who analyzed polymicrobial PJI cases using mNGS. Given its superior detection capabilities than synovial fluid and periprosthetic tissue, prosthetic sonication fluid stands out as the optimal choice for mNGS specimens. This is mainly attribute to the prosthetic sonication fluid can destroy the bacterial biofilm that colonizes the prosthesis, and enrich the concentration of pathogenic bacteria to be examined by centrifugation, which improves the detection efficiency of pathogenic bacteria. However, prosthetic sonication fluid also has its disadvantages, as it can only be obtained after intraoperative removal of the infected prosthesis, making it difficult to obtain a sample by puncture preoperatively to guide preoperative treatment (Huang et al., 2020).

The pathogenic bacteria detected by both methods in Sample 9 showed complete inconsistency. Several reasons could account for this inconsistency. Firstly, pitfalls in processing, establishment of background thresholds, and nucleic acid isolation can lead to failure to identify an organism and/or insufficient genome coverage to allow for accurate mapping to a species level. Secondly, the manifestation of PJI can be highly variable and the presence of other infecting organisms and the host microbiome can influence the relative abundances of various pathogens. Thirdly, insufficient DNA extraction can lead to a low number of reads or insufficient genome coverage in the setting of greater background reads. Weaver et al. (2019) similarly found discordance between genomical analysis and culture results was possible and reported positive cultures of organisms falling below their whole genome shotgun sequencing thresholds. Similarly, Ivy et al. (2018) observed that mNGS was incapable of identifying pathogens in 14 cases of culture-positive PJI.

While mNGS demonstrates distinct advantages in pathogen detection, its inherent limitations cannot be ignored. The mNGS results are susceptible to interference by background microorganisms, which mainly include exogenous bacterial and nucleic acid contamination caused by experimental manipulation or the experimental environment. Hence, rigorous adherence to sterilization and experimental protocols, implementation of blank controls, and rectification of false-positive results due to sample contamination are imperative. Moreover, the choice of the mNGS threshold significantly influences the outcomes. A threshold set too low could introduce numerous background microorganisms, causing false positives, whereas an excessively high threshold might overlook pathogenic bacteria, resulting in false negatives. The establishment of appropriate thresholds is a subject of controversy. Threshold settings must be carefully selected based on different sequencing methods and case types, ensuring a balance between diagnostic sensitivity and specificity. Generally, threshold setting relies on three indicators: the number of reads, RAG, and CR. In this study, we used a strict algorithm to obtain the results by reference to Street et al. (2017), employing a diagnostic threshold of RAG >15% and number of sequences reads > 100. In contrast to culture, mNGS presents challenges in conducting drug sensitivity tests for pathogenic bacteria, particularly for certain multi-drug resistant infections, thereby making it difficult to empirically administer antibiotics based on sensitivity test outcomes. In addition, the high cost of mNGS is a major barrier limiting its widespread use in the clinic practice.

There are some limitations in this study. Firstly, the variety of pathogenic bacteria detected in this study was limited, with some common pathogenic bacteria going undetected. This may be attributed to antibiotic use within the 2 weeks preceding testing, potentially hindering the growth of certain common microorganisms on the culture medium. Secondly, this study is a single-center study with a small sample size, which may also result in limitations in the types of pathogens detected and introduce bias into the results. To address these limitations, further support from larger samples, expanded datasets, and broader studies in subsequent stages is required. Finally, for multi-drug resistant infections, it is difficult for mNGS to conduct drug sensitivity testing and guide antibiotic selection. In a further study, we hope

to conduct targeted detection of drug resistance genes to accurately guide the use of antibiotics.

## 5 Conclusions

In conclusion, mNGS can function as a precise diagnostic tool for identifying pathogens in PJI patients using three types of specimens. Due to its superior ability in pathogen identification, prosthetic sonicate fluid can replace synovial fluid and periprosthetic tissue as the optimal sample choice for mNGS. Additionally, employing three types of specimens in this study could contribute to the enhanced processing and analysis of mNGS technology, facilitating its thorough application in identifying pathogens in PJI patients.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: China National GeneBank Database (CNGBdb) with accession number CNP0005321.

## Ethics statement

The studies involving humans were approved by Medical Ethics Committee of the First Affiliated Hospital of Zhengzhou University (Approval No:2020-KY-821). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

JT: Writing – original draft, Writing – review & editing. LW: Writing – review & editing. LZ: Writing – review & editing. MS: Writing – review & editing. ZT: Writing – review & editing. JX: Writing – review & editing. HM: Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Joint Project of Medical Science and Technology of Henan Province (LHGJ20190859) and the National Clinical Key Specialty Construction Project (Yu Wei Medical Letter 2023 [30]).

## Acknowledgments

We extend our appreciation to the diligent and dedicated efforts of all the staff members who carried out the intervention and evaluation components of the study. We would also like to express our gratitude to all the patients in this study for their selfless commitment to medical research.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1356804/full#supplementary-material

SUPPLEMENTARY TABLE 1
Detection of pathogens by mNGS and culture in the three types of specimen.

SUPPLEMENTARY TABLE 2
Microbial results of mNGS and culture in PJI group.

## References

Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., et al. (2009). Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Mol. Plant Pathol.* 10, 537–545. doi: 10.1111/j.1364-3703.2009.00545.x

Brown, J. R., Bharucha, T., and Breuer, J. (2018). Encephalitis diagnosis using metagenomics: application of next generation sequencing for undiagnosed cases. *J. Infect.* 76, 225–240. doi: 10.1016/j.jinf.2017.12.014

Chiu, C. Y., and Miller, S. A. (2019). Clinical metagenomics. *Nat. Rev. Genet.* 20, 341–355. doi: 10.1038/s41576-019-0113-7

Doan, T., Acharya, N. R., Pinsky, B. A., Sahoo, M. K., Chow, E. D., Banaei, N., et al. (2017). Metagenomic DNA sequencing for the diagnosis of intraocular infections. *Ophthalmology.* 124, 1247–1248. doi: 10.1016/j.ophtha.2017.03.045

Dudareva, M., Barrett, L., Figtree, M., Scarborough, M., Watanabe, M., Newnham, R., et al. (2018). Sonication versus tissue sampling for diagnosis of prosthetic joint and other orthopedic device-related infections. *J. Clin. Microbiol.* 56(12):e00688–18. doi: 10.1128/JCM.00688-18

Flurin, L., Hemenway, J. J., Fisher, C. R., Vaillant, J. J., Azad, M., Wolf, M. J., et al. (2022). Clinical use of a 16S ribosomal RNA gene-based sanger and/or next generation

sequencing assay to test preoperative synovial fluid for periprosthetic joint infection diagnosis. *mBio.* 13, e0132222. doi: 10.1128/mbio.01322-22

Gallo, J. (2020). Prosthetic joint infection: updates on prevention, diagnosis and therapy. *J. Clin. Med.* 9(12):3892. doi: 10.3390/jcm9123892

Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryke, D., et al. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* 27, 115–124. doi: 10.1038/s41591-020-1105-z

Hartley, J. C., and Harris, K. A. (2014). Molecular techniques for diagnosing prosthetic joint infections. *J. Antimicrob. Chemother.* 69 Suppl 1, i21–i24. doi: 10.1093/jac/dku249

Hersh, B. L., Shah, N. B., Rothenberger, S. D., Zlotnicki, J. P., Klatt, B. A., and Urish, K. L. (2019). Do culture negative periprosthetic joint infections remain culture negative? *J. Arthroplasty* 34, 2757–2762. doi: 10.1016/j.arth.2019.06.050

Huang, Z., Li, W., Lee, G. C., Fang, X., Xing, L., Yang, B., et al. (2020). Metagenomic next-generation sequencing of synovial fluid demonstrates high accuracy in prosthetic joint infection diagnostics: mNGS for diagnosing PJI. *Bone Joint Res.* 9, 440–449. doi: 10.1302/2046-3758.97.BJR-2019-0325.R2

Huang, Z., Wu, Q., Fang, X., Li, W., Zhang, C., Zeng, H., et al. (2018). Comparison of culture and broad-range polymerase chain reaction methods for diagnosing periprosthetic joint infection: analysis of joint fluid, periprosthetic tissue, and sonicated fluid. *Int. Orthop.* 42, 2035–2040. doi: 10.1007/s00264-018-3827-9

Ivy, M. I., Thoendel, M. J., Jeraldo, P. R., Greenwood-Quaintance, K. E., Hanssen, A. D., Abdel, M. P., et al. (2018). Direct detection and identification of prosthetic joint infection pathogens in synovial fluid by metagenomic shotgun sequencing. *J. Clin. Microbiol.* 56(9):e00402–18. doi: 10.1128/JCM.00402-18

Karbysheva, S., Grigoricheva, L., Golnik, V., Popov, S., Renz, N., and Trampuz, A. (2019). Influence of retrieved hip- and knee-prosthesis biomaterials on microbial detection by sonication. *Eur. Cell Mater.* 37, 16–22. doi: 10.22203/eCM

Larsen, L. H., Khalid, V., Xu, Y., Thomsen, T. R., and Schønheyder, H. C. (2018). Differential contributions of specimen types, culturing, and 16S rRNA sequencing in diagnosis of prosthetic joint infections. *J. Clin. Microbiol.* 56(5):e01351–17. doi: 10.1128/JCM.01351-17

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Gao, H., Meng, H., Wang, Q., Li, S., Chen, H., et al. (2018). Detection of pulmonary infectious pathogens from lung biopsy tissues by metagenomic next-generation sequencing. *Front. Cell Infect. Microbiol.* 8, 205. doi: 10.3389/fcimb.2018.00205

Lu, J., Breitwieser, F. P., Thielen, P., and Salzberg, S. L. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Science.* 3, e104. doi: 10.7717/peerj-cs.104

Malekzadeh, D., Osmon, D. R., Lahr, B. D., Hanssen, A. D., and Berbari, E. F. (2010). Prior use of antimicrobial therapy is a risk factor for culture-negative prosthetic joint infection. *Clin. Orthop Relat. Res.* 468, 2039–2045. doi: 10.1007/s11999-010-1338-0

Mei, J., Hu, H., Zhu, S., Ding, H., Huang, Z., Li, W., et al. (2023). Diagnostic role of mNGS in polymicrobial periprosthetic joint infection. *J. Clin. Med.* 12(5):1838. doi: 10.3390/jcm12051838

Miao, Q., Ma, Y., Wang, Q., Pan, J., Zhang, Y., Jin, W., et al. (2018). Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin. Infect. Dis.* 67, S231–Ss40. doi: 10.1093/cid/ciy693

Nodzo, S. R., Bauer, T., Pottinger, P. S., Garrigues, G. E., Bedair, H., Deirmengian, C. A., et al. (2015). Conventional diagnostic challenges in periprosthetic joint infection. *J. Am. Acad. Orthop Surg.* 23 Suppl, S18–S25. doi: 10.5435/JAAOS-D-14-00385

O'Flaherty, B. M., Li, Y., Tao, Y., Paden, C. R., Queen, K., Zhang, J., et al. (2018). Comprehensive viral enrichment enables sensitive respiratory virus genomic identification and analysis by next generation sequencing. *Genome Res.* 28, 869–877. doi: 10.1101/gr.226316.117

Parvizi, J., Zmistowski, B., Berbari, E. F., Bauer, T. W., Springer, B. D., Della Valle, C. J., et al. (2011). New definition for periprosthetic joint infection: from the Workgroup of the Musculoskeletal Infection Society. *Clin. Orthop Relat. Res.* 469, 2992–2994. doi: 10.1007/s11999-011-2102-9

Rak, M., Barlic-Maganja, D., Kavcic, M., Trebse, R., and Cor, A. (2013). Comparison of molecular and culture method in diagnosis of prosthetic joint infection. *FEMS Microbiol. Lett.* 343, 42–48. doi: 10.1111/femsle.2013.343.issue-1

Rietbergen, L., Kuiper, J. W., Walgrave, S., Hak, L., and Colen, S. (2016). Quality of life after staged revision for infected total hip arthroplasty: a systematic review. *Hip Int.* 26, 311–318. doi: 10.5301/hipint.5000416

Street, T. L., Sanderson, N. D., Atkins, B. L., Brent, A. J., Cole, K., Foster, D., et al. (2017). Molecular diagnosis of orthopedic-device-related infection directly from sonication fluid by metagenomic sequencing. *J. Clin. Microbiol.* 55, 2334–2347. doi: 10.1128/JCM.00462-17

Tan, J., Liu, Y., Ehnert, S., Nüssler, A. K., Yu, Y., Xu, J., et al. (2022). The effectiveness of metagenomic next-generation sequencing in the diagnosis of prosthetic joint infection: A systematic review and meta-analysis. *Front. Cell Infect. Microbiol.* 12, 875822. doi: 10.3389/fcimb.2022.875822

Thoendel, M., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., et al. (2016). Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J. Microbiol. Methods* 127, 141–145. doi: 10.1016/j.mimet.2016.05.022

Thoendel, M. J., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., et al. (2018). Identification of prosthetic joint infection pathogens using a shotgun metagenomics approach. *Clin. Infect. Dis.* 67, 1333–1338. doi: 10.1093/cid/ciy303

Trampuz, A., Piper, K. E., Jacobson, M. J., Hanssen, A. D., Unni, K. K., Osmon, D. R., et al. (2007). Sonication of removed hip and knee prostheses for diagnosis of infection. *N Engl. J. Med.* 357, 654–663. doi: 10.1056/NEJMoa061588

Villa, F., Toscano, M., De Vecchi, E., Bortolin, M., and Drago, L. (2017). Reliability of a multiplex PCR system for diagnosis of early and late prosthetic joint infections before and after broth enrichment. *Int. J. Med. Microbiol.* 307, 363–370. doi: 10.1016/j.ijmm.2017.07.005

Weaver, A. A., Hasan, N. A., Klaassen, M., Karathia, H., Colwell, R. R., and Shrout, J. D. (2019). Prosthetic joint infections present diverse and unique microbial communities using combined whole-genome shotgun sequencing and culturing methods. *J. Med. Microbiol.* 68, 1507–1516. doi: 10.1099/jmm.0.001068

Wilson, M. R., Naccache, S. N., Samayoa, E., Biagtan, M., Bashir, H., Yu, G., et al. (2014). Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N Engl. J. Med.* 370, 2408–2417. doi: 10.1056/NEJMoa1401268

Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* 20, 257. doi: 10.1186/s13059-019-1891-0

Yoon, H. K., Cho, S. H., Lee, D. Y., Kang, B. H., Lee, S. H., Moon, D. G., et al. (2017). A review of the literature on culture-negative periprosthetic joint infection: epidemiology, diagnosis and treatment. *Knee Surg. Relat. Res.* 29, 155–164. doi: 10.5792/ksrr.16.034

Young, B. C., Dudareva, M., Vicentine, M. P., Hotchen, A. J., Ferguson, J., and McNally, M. (2023). Microbial persistence, replacement and local antimicrobial therapy in recurrent bone and joint infection. *Antibiotics (Basel).* 12(4):708. doi: 10.3390/antibiotics12040708

Yu, Y., Wang, S., Dong, G., and Niu, Y. (2023). Diagnostic performance of metagenomic next-Generation sequencing in the diagnosis of prosthetic joint infection using tissue specimens. *Infect. Drug Resist.* 16, 1193–1201. doi: 10.2147/IDR.S397260

Zhang, C., Fang, X., Huang, Z., Li, W., Zhang, C. F., Yang, B., et al. (2019). Value of mNGS in sonication fluid for the diagnosis of periprosthetic joint infection. *Arthroplasty.* 1, 9. doi: 10.1186/s42836-019-0006-4

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 19 March 2024
DOI 10.3389/fcimb.2024.1378112



OPEN ACCESS

EDITED BY
Qing Wei,
Genskey Co. Ltd, China

REVIEWED BY
Zhuyun Qian,
Capital Medical University, China
Jing Wang,
Hugobiotech Co. Ltd., China

*CORRESPONDENCE
Yanfang Jiang
yanfangjiang@hotmail.com

†These authors have contributed equally to this work

RECEIVED 29 January 2024
ACCEPTED 06 March 2024
PUBLISHED 19 March 2024

CITATION
Zou H, Gao S, Liu X, Liu Y, Xiao Y, Li A and Jiang Y (2024) Combination of metagenomic next-generation sequencing and conventional tests unraveled pathogen profiles in infected patients undergoing allogeneic hematopoietic stem cell transplantation in Jilin Province of China. *Front. Cell. Infect. Microbiol.* 14:1378112. doi: 10.3389/fcimb.2024.1378112



# Combination of metagenomic next-generation sequencing and conventional tests unraveled pathogen profiles in infected patients undergoing allogeneic hematopoietic stem cell transplantation in Jilin Province of China

Hongyan Zou[1†], Sujun Gao[2†], Xiaoliang Liu[2], Yong Liu[1], Yunping Xiao[1], Ao Li[1] and Yanfang Jiang[1*]

[1]Key Laboratory of Organ Regeneration & Transplantation of the Ministry of Education, Genetic Diagnosis Center, The First Hospital of Jilin University, Changchun, China, [2]Department of Hematology, The First Hospital of Jilin University, Changchun, China

**Background:** Infection is the main cause of death for patients after allogeneic hematopoietic stem cell transplantation (HSCT). However, pathogen profiles still have not been reported in detail due to their heterogeneity caused by geographic region.

**Objective:** To evaluate the performance of metagenomic next-generation sequencing (mNGS) and summarize regional pathogen profiles of infected patients after HSCT.

**Methods:** From February 2021 to August 2022, 64 patients, admitted to the Department of Hematology of The First Hospital of Jilin University for HSCT and diagnosed as suspected infections, were retrospectively enrolled.

**Results:** A total of 38 patients were diagnosed as having infections, including bloodstream (*n* =17), pulmonary (*n* =16), central nervous system (CNS) (*n* =4), and chest (*n* =1) infections. Human betaherpesvirus 5 (CMV) was the most common pathogen in both bloodstream (*n* =10) and pulmonary (*n* =8) infections, while CNS (*n* =2) and chest (*n* =1) infections were mainly caused by Human gammaherpesvirus 4 (EBV). For bloodstream infection, *Mycobacterium tuberculosis* complex (*n* =3), *Staphylococcus epidermidis* (*n* =1), and *Candida tropicalis* (*n* =1) were also diagnosed as causative pathogens. Furthermore, mNGS combined with conventional tests can identify more causative pathogens with high sensitivity of 82.9% (95% CI 70.4-95.3%), and the total coincidence rate can reach up to 76.7% (95% CI 64.1-89.4%).

**Conclusions:** Our findings emphasized the importance of mNGS in diagnosing, managing, and ruling out infections, and an era of more rapid, independent, and impartial diagnosis of infections after HSCT can be expected.

KEYWORDS

allogeneic hematopoietic stem cell transplantation, pathogen profiles, MNGs, total coincidence rate, Jilin Province

# Introduction

Hematopoietic stem cell transplantation (HSCT) is a potential radical treatment for hematological and lymphatic malignancies (Garcia et al., 2013). The overall survival rate of HSCT has been significantly improved due to the improvement of patient care after transplantation (Li et al., 2021). Of all the organ-specific complications that can occur after HSCT, incidence rate of pneumonia, which is complex and difficult to treat, can reach up to 30-60% in HSCT recipients, followed by bloodstream (Zanella et al., 2021) and central nervous system (CNS) infections. Pneumonia remains the leading cause of non-recurrent mortality after transplantation. Accordingly, not only timely but accurate diagnosis are needed to improve the prognoses and decrease the mortality.

For infection diagnosis, ideal samples include blood, cerebrospinal fluid (CSF), and bronchoalveolar lavage fluid (BALF) (Panackal, 2021; Shane et al., 2017; Berhane et al., 2021; Herbozo et al., 2021; Dorresteijn et al., 2020), while more than 60% of pathogens cannot be detected by blood, CSF, or BALF culture (Chen et al., 2020; Zhang et al., 2020; Saha et al., 2018). However, some pathogens with negative blood culture, such as *Streptococcus pneumoniae*, may result in high mortality rates (Kumar et al., 2021). Limited number of pathogens could be detected in a single experiment using hypothesis-based PCR (Marcilla-Vazquez et al., 2018) or antibody methods (Ge et al., 2021). These disadvantages hindered extensive application of conventional tests in accurate diagnosis, not to mention HSCT patients with co-infections (Jiang et al., 2021). With the first successful application of metagenomics next-generation sequencing (mNGS) in clinical infection of CNS (Wilson et al., 2014), this unbiased technology has been extensively used for diagnosing various infections (Wu et al., 2020; Chen et al., 2021; Gu et al., 2021), including co-infections (Jiang et al., 2021).

Although more and more studies used mNGS for pathogen detection in patients after HSCT, pathogen profiles still have not been reported in detail due to their heterogeneity caused by geographic region. To obtain regional pathogen profiles for guiding clinical diagnosis and treatment, we retrospectively enrolled the infected patients after HSCT and summarized the detection results revealed by mNGS and conventional tests (CT). Besides, we also evaluated the performance of mNGS and CT against final clinical diagnosis, including sensitivity, specificity, and total coincidence rate (TCR).

# Methods

## Ethics statement

This study was reviewed and approved by the Ethical Review Committee of the First hospital of Jilin University (approval no. 2022-566). All procedures followed were in strict compliance with the Ethical Review of Biomedical Research Involving Human Subjects (2016), the Declaration of Helsinki, and the International Ethical Guidelines for Biomedical Research Involving Human Subjects.

## Study population

A total of 64 patients undergoing HSCT in the First hospital of Jilin University from February 2021 to August 2022 diagnosed as suspected infection were enrolled in the retrospective study (Figure 1). The inclusion criteria were as follows. The patients with suspected bloodstream infection were enrolled in reference to the diagnostic criteria of sepsis. Suspected CNS infection was based on acute fever (> 38.5°C) and one of the following signs: a) meningeal irritation, b) increased anterior fontanelle tension, c) stiff neck, and d) consciousness disorders. Suspected pulmonary infection was considered if the patient had new opacity on imaging

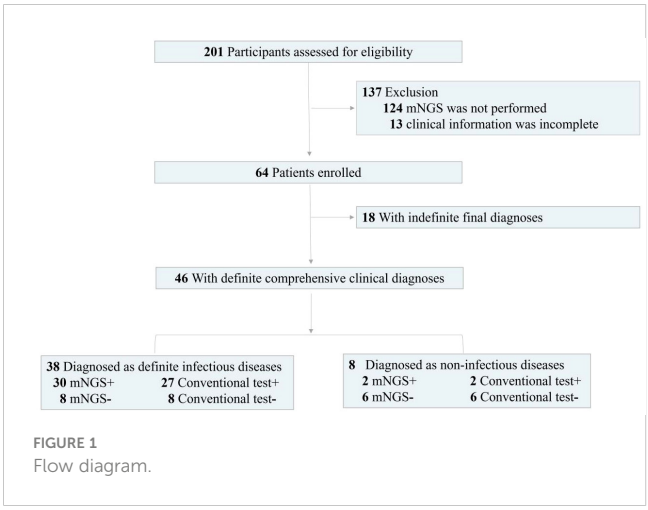



FIGURE 1
Flow diagram.

examination and at least one of the following symptoms: a) respiratory distress, b) fever, c) cough, and d) peripheral leukocytosis (> $10\times10^9$/L) or leucopenia (< $4\times10^9$/L). The exclusion criteria included 1) mNGS was not performed and 2) clinical information was incomplete. Physical information and clinical characterization were collected. After signing informed consents, CSF, blood, and BALF samples were respectively collected from patients with suspected CNS, bloodstream, and pulmonary infections, respectively. CT included routine bacterial and fungal smears and cultures, serum antigen (Galactomannan) and antibody (Human betaherpesvirus 5 (CMV) and Human gammaherpesvirus 4 (EBV)) tests, computerized tomography, and PCR. mNGS was performed on all of samples.

## CT assay

### Smear and culture

Gram staining and KOH test were used for bacterial and fungal smear, respectively. The bacterial and fungal cultivation were performed in the incubators at 35°C and 28°C, respectively. Culture media for bacterial culture included blood agar, chocolate agar and MacConkey agar. Besides, Lowenstein Jensen Medium was used for suspected infection of *Mycobacterium tuberculosis*. Sabauraud agar and CHROMagar Candida medium were used for fungal culture.

### Quantitative PCR

After total DNA extraction, CMV and EBV detection were performed using commercial PCR methods (DA0121 and DA0151, DaAn Gene Co., Ltd. of Sun Yat-sen University, Guangzhou, China) according to the manufacturer's instructions.

### mNGS pipeline

Conventional mNGS and hybridization capture-based targeted mNGS were used to detect pathogens. The 0.6 mL blood, 0.6 mL CSF, and 1.0 mL BALF samples were respectively used to extract DNA using TIANamp Micro DNA Kit DP316 (TIANGEN Biotech, Beijing, China). Qubit 4.0 (Thermo Fisher Scientific, MA, USA) was used to measure extracted DNA concentrations. QIAseq Ultralow Input Library Kit (QIAGEN, Hilden, Germany) was used to construct metagenomics libraries (Ji et al., 2020). While for targeted NGS, the constructed library from each sample was used for hybrid capture-based enrichment of microbial probes. Inspected and qualified library was sequenced on Nextseq 550 platform (Illumina, San Diego, USA).

To remove adapter and low-quality, low complexity, and short reads of < 35 bp, raw data generated by the sequencing were filtered (Chen et al., 2021). To exclude human DNA sequences, reads were mapped to the human reference genome hg38 using bowtie2 (Langmead and Salzberg, 2012) to obtain clean reads. And then, the clean reads were blasted against a microbial genome database constructed according to the published microbial genome databases, including reference sequence database at National Center for Biotechnology Information. Finally, microbial information at species level can be obtained.

As control, using the same procedure and bioinformatics analysis, negative and positive controls were also set during the mNGS tests of samples from patients. The reads per million (RPM) of each detected pathogen were calculated. For the detected bacteria (*Mycobacterium* excluded), fungi (*Cryptococcus* excluded), or parasites, a positive mNGS result was defined when the microorganism was not detected in the negative control ('No template' control, NTC) and genome coverage of detected sequences belonged to this microorganism ranked top10 among the microbes in the same genus or when its ratio of $RPM_{sample}$ to $RPM_{NTC}$ was ($RPM_{sample}/RPM_{NTC}$) > 10 if the $RPM_{NTC}\neq0$. For viruses, *Mycobacterium*, and *Cryptococcus*, a positive mNGS result was considered when the virus was not detected in NTC and at least 1 specific read was mapped to species or when $RPM_{sample}/RPM_{NTC}$ was > 5 if the $RPM_{NTC}\neq0$. The interpretation of mNGS results was performed by 2-3 clinical adjudicators. Positive mNGS results were defined according to whether the pathogens were the most commonly reported pathogens or the infections by the pathogens were in accordance with clinical features of patients (Zhang et al., 2020; Zhang et al., 2023).

## Diagnostic assessment

According to mNGS results (including both positive and negative mNGS results), complete laboratory examinations, the treatment response of the patients, and clinical experiences, the final clinical diagnoses and causative pathogens were independently made by 2-3 clinicians with expertise in infectious diseases. According to final comprehensive clinical diagnoses, we divided enrolled patients into definite and indefinite clinical diagnosis groups. Definite clinical diagnosis group included patients with infectious and non-infectious diagnoses. Indefinite clinical diagnosis group included patients whose clinical characteristics or laboratory examinations were not adequate for diagnoses and patients lost during follow-up duration.

Given semi-quantitative characteristics of mNGS, dynamic surveillance of infections can be conducted (Zhang et al., 2020). To evaluate the efficacy of medication in reference to mNGS results, comparison was performed as follows. Firstly, some microbes detected by the 1$^{st}$ mNGS were diagnosed as causative pathogens of patient. Subsequently, if specific reads of the pathogen detected by the 2$^{nd}$ mNGS decreased after effective anti-microbial treatment in reference to mNGS results and the infection symptoms were partially improved, we defined that mNGS can provide positive reference for treatment strategies. Otherwise, if continuous two or more mNGS results were negative at short interval and patient was recovered without anti-microbial treatment, we defined that negative mNGS results can be used to rule out infection.

## Statistical analysis

Counts and percentages were presented for independent variables. Mean ± standard error (SE) was calculated for continuous variables with normal distributions, while medians and interquartile ranges (IQRs) were used for abnormal distributions. Confidence intervals were calculated according to the formula: CI=Average ± 1.96SE. Spearman correlation analysis was conducted, and $P$ value of $< 0.05$ were considered statistically significant. The data were analyzed using IBM SPSS 25.0 and R 4.1.1.

## Data availability

Sequencing data were deposited to the National Genomics Data Center under accession numbers PRJCA014245 and CRA009447. The authors declare that the main data supporting the findings are available within this article. The other data generated and analyzed for this study are available from the corresponding author upon reasonable request.

# Results

## Clinical characteristics of the patients undergoing HSCT

A total of 64 patients undergoing HSCT diagnosed as having suspected infections were enrolled in the retrospective study (Figure 1), including 33 males and 31 females. The patients underwent transplantation due to different kinds of leukemias, including acute myeloid leukemia ($n$ =20), acute lymphoblastic leukemia ($n$ = 14), acute leukemia (myeloid + lymphoblastic leukemia) ($n$ = 1), chronic myeloid leukemia ($n$ = 1), myelodysplastic syndrome ($n$ = 9), aplastic anemia ($n$ = 1), and lymphoma ($n$ = 2) (Table 1). For further diagnoses, samples were collected from 64 patients for mNGS and conventional tests, including 31 blood, 22 BALF, 7 CSF, 2 drainage liquid, 1 hydropericardium, and 1 hydrothorax samples (Figure 2A). According to final comprehensive clinical diagnoses, we divided enrolled patients into definite ($n$ = 46) and indefinite ($n$ = 18) clinical diagnosis groups. Among the definite group, a total of 38 patients were diagnosed as having infections, including bloodstream ($n$ = 17), pulmonary ($n$ = 16), CNS (Viral encephalitis, $n$ = 4), and chest ($n$ = 1) infections (Figure 2B).

## Regional pathogen profiles in different systems

There are quite differences in pathogen profiles among different systems (Figure 3). More than 75% of infections were caused by viruses and only 1 patient had bacterial-fungal co-infection. Bacteria, fungi, and viruses were diagnosed as causative pathogens in bloodstream and pulmonary infections, while pathogens causing CNS and chest infections were found to be viruses. For viruses, CMV was the most common pathogen in both bloodstream ($n$ = 10, detection rate of 58.8%) and pulmonary ($n$ = 8) infections, while CNS ($n$ = 2) and chest ($n$ = 1) infections were mainly caused by EBV. Besides, EBV can also cause bloodstream ($n$ = 4) and pulmonary ($n$ = 5) infections with high detection rates.

Apart from viruses, *Pneumocystis jirovecii* ($n$ =2), *Aspergillus* ($n$ =1), *Enterobacter cloacae* ($n$ =1), *Finegoldia magna* ($n$ =1), and *Mycoplasma pneumoniae* ($n$ =1) were detected in pulmonary infection. For bloodstream infection, *Mycobacterium tuberculosis* complex ($n$ =3), *Staphylococcus epidermidis* ($n$ =1), and *Candida tropicalis* ($n$ =1) were diagnosed as causative pathogens. Identification of regional pathogen profiles can help guide clinical empirical treatment before accurate diagnoses to improve prognosis of patients.

## Comparison between mNGS and CT

Differences in positive rate between mNGS and CT for different samples were found (Figure 4 and Supplementary Table S1). For BALF sample, the positive rate of mNGS (77%) was significantly higher than that of CT (59.1%). For blood sample, mNGS can detect microbes from 15 out of 31 samples (48.4%), while CT can detect microbes from 17 out of 25 samples (68.0%). Besides, positive rates of both mNGS and CT for CSF sample were low. Overall, the

TABLE 1 Baseline characteristics of patients after HSCT.

| | Cases ($n$ = 64) | |
|---|---|---|
| Sex ($n$, %) | Male | 33 (51.6%) |
| | Female | 31 (48.4%) |
| Age (years) | Median (min,max) | 40.4 (0.9, 75) |
| WBC ($\times 10^9$/L) | Average | 3.4 ± 3.2 |
| PCT (ng/mL) | Average | 7.9 ± 26.7 |
| CRP (mg/L) | Average | 69.9 ± 83.1 |
| Underlying disease ($n$, %) | Total | 6 (9.4%) |
| | Hypertension | 2 (3.1%) |
| | Post-operation of lung cancer | 1 (1.6%) |
| | Hepatitis B | 1 (1.6%) |
| | Food allergy | 1 (1.6%) |
| | Cholecystectomy | 1 (1.6%) |
| Clinical diagnosis (n, %) | Aplastic anemia | 1 (1.6%) |
| | Acute lymphoblastic leukemia | 14 (21.9%) |
| | Acute myeloid leukemia | 20 (31.3%) |
| | Myelodysplastic syndrome | 9 (14.1%) |
| | Chronic myeloid leukemia | 1 (1.6%) |
| | Lymphoma | 2 (3.1%) |

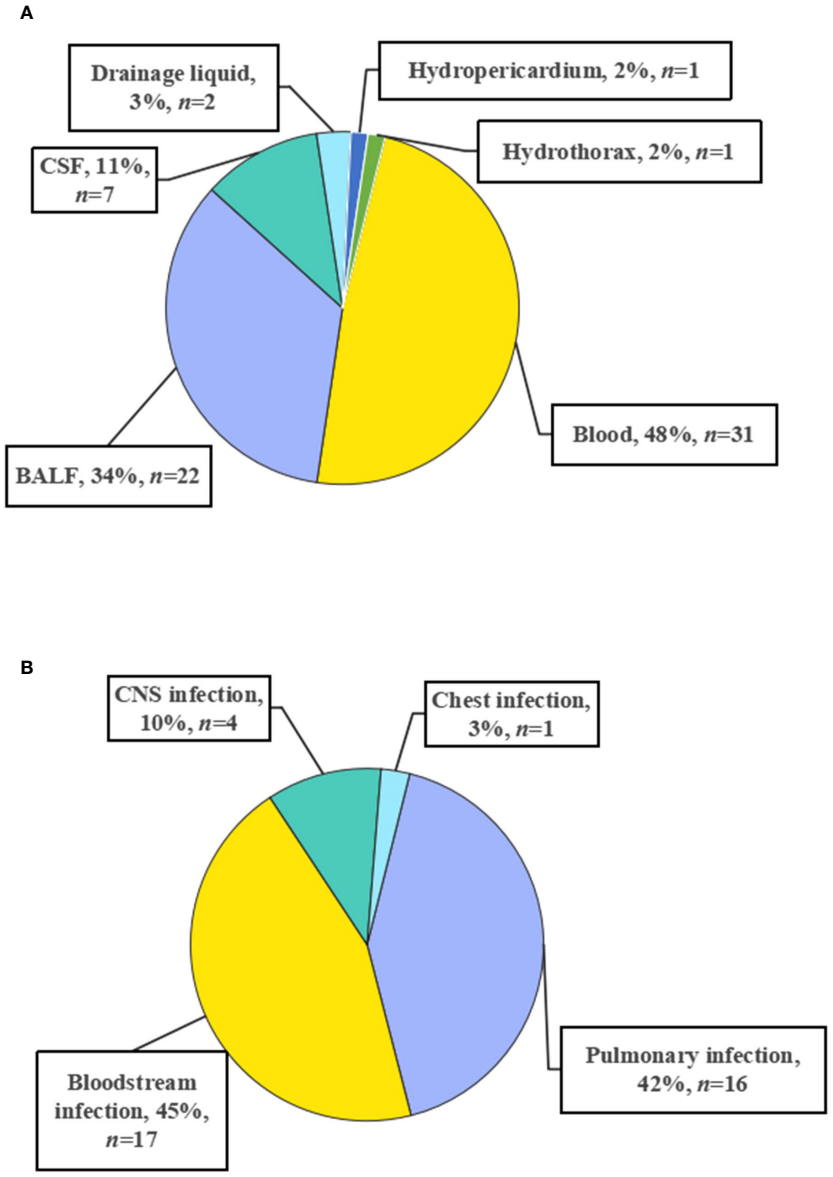



FIGURE 2
Sample type **(A)** and infection site **(B)**.

positive rate of mNGS (56.3%) was close to that of CT (60.3%). Accordingly, we proposed that mNGS might be an alternative or supplemental examination for diagnosing infections when CT fails.

In our study, mNGS was performed on the whole patients enrolled, while CT was performed on 58 patients (Figure 5). There were 24 patients with both mNGS and CT positive, and 14 patients with both mNGS and CT negative, while inconsistent results between mNGS and CT were found in more than 34% of 58 patients. For the patients with mNGS positive and CT negative ($n$ =9), about half of patients (4 out of 9) can be directly diagnosed only by mNGS results with positive coincidence rate of 44.4% against final clinical diagnoses. A same trend was found in the patients with mNGS negative and CT positive. The above results indicate that there were no significant differences in detecting microbes between mNGS and CT.

Taking final clinical diagnoses as gold standard, we further compared the performance of mNGS and CT in diagnosing infections (Figure 5). The sensitivity of mNGS was slightly lower [52.6% (95% confidence interval (CI) 36.8-68.5%)] than that of CT [65.7% (95% CI 50.0-81.4%)]. The specificities of both mNGS and CT were 75%. Positive predictive values (PPV) of both mNGS [90.9% (95% CI 78.9-100.0%)] and CT [92.0% (95% CI 81.4-100.0%)] were >90%, while negative predictive values (NPV) of both mNGS (25.0% [95% CI 7.7-42.3%)] and CT [33.3% (95% CI 11.6-55.1%)] were low. Besides, more than half of mNGS (TCR, [56.5% (95% CI 42.2-70.8%)] and CT [67.4% (95% CI 53.4-81.4%)] results were consistent with the final clinical diagnoses. Most importantly, compared with mNGS or CT, mNGS combined with CT can identify more causative pathogens with higher sensitivity of 82.9% (95% CI 70.4-95.3%), and the TCR can reach up to 76.7%

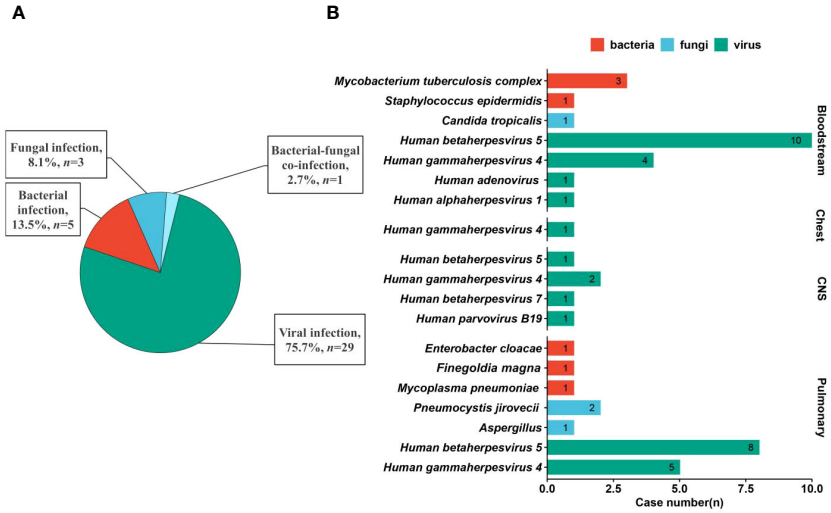



FIGURE 3
Causative pathogens in different systems. **(A)** Infection type in patients after HSCT. **(B)** Pathogen profiles causing bloodstream, chest, CNS, and pulmonary infections based on final clinical diagnoses.

(95% CI 64.1-89.4%). The above results unravel that mNGS can be considered as a supplement to CT for infection diagnosis in patients after HSCT mainly infected by viruses.

## Semi-quantitative value of mNGS

Given semi-quantitative characteristics of mNGS, we performed mNGS on some patients at different intervals to further verify the efficacy of medication in reference to mNGS results (Figure 6). In cases 1-3, CMV, EBV, or *M.tuberculosis* were detected by the 1st mNGS and diagnosed as causative pathogens, respectively. After effective anti-microbial treatment, specific reads detected by the 2nd mNGS decreased to low level or even to 0, and the infection symptoms of these cases were partially improved. We also use mNGS to rule out infection in some case. Case 4 was finally diagnosed as non-infection disease by the two negative mNGS results at 28 d interval. The above results further emphasize the importance of mNGS in providing reference for treatment strategies.

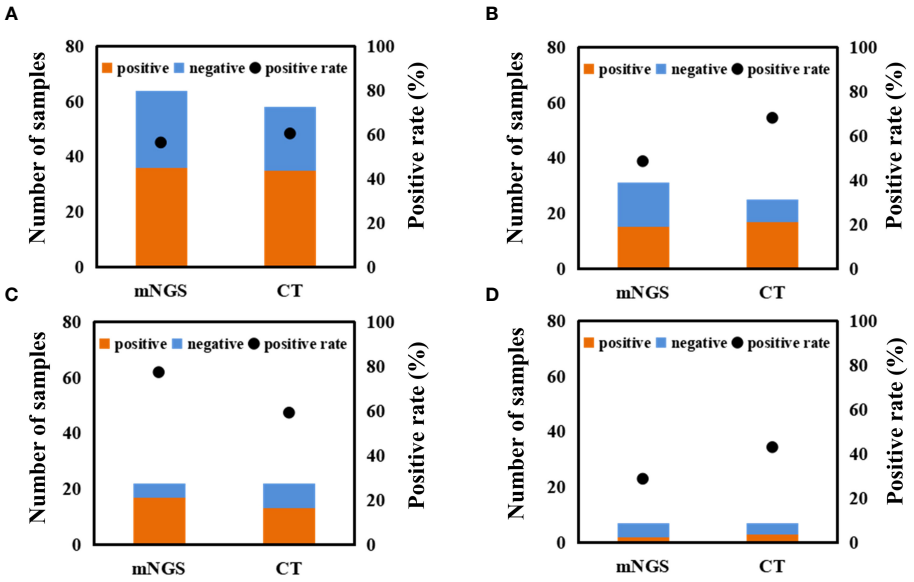



FIGURE 4
Comparison between mNGS and CT in microbial detection using different samples. Differences in microbial detection between mNGS and CT using different samples. **(A)** all of samples. **(B)** blood sample. **(C)** BALF sample. **(D)** CSF sample. 'CT' represents conventional tests.

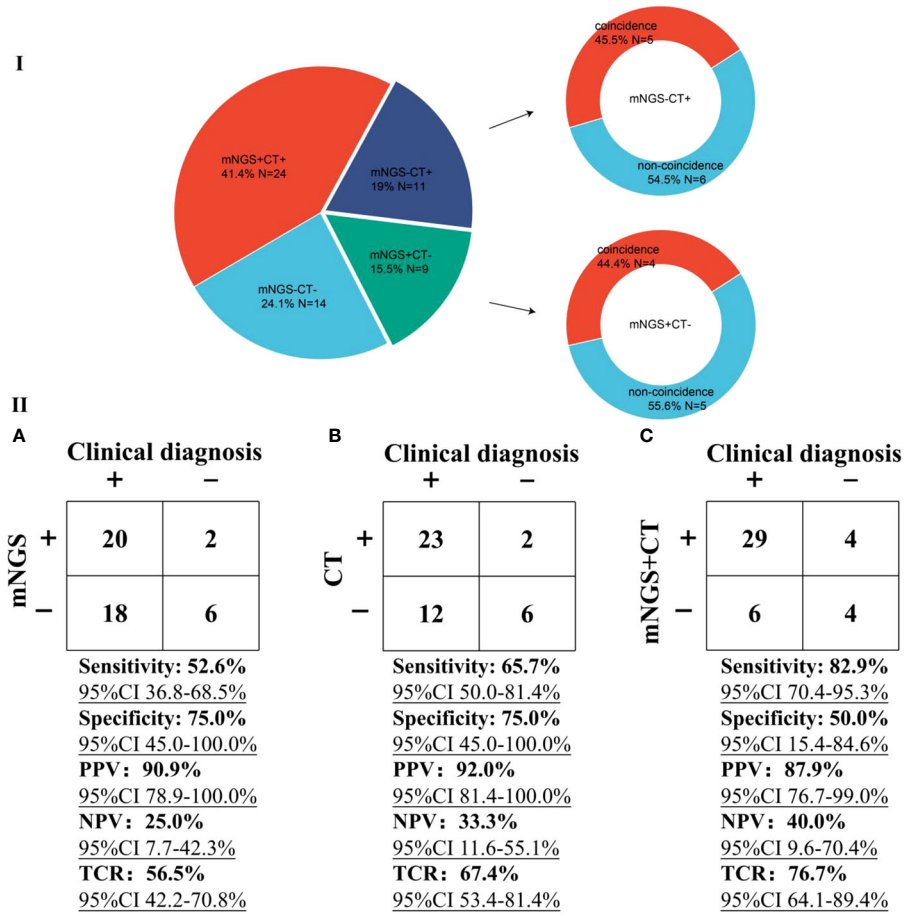



FIGURE 5
Performance comparison between mNGS and CT. I, Comparison of differences in detection between mNGS and CT. 'N' was the number of cases. '+' and '-' represent positive and negative results, respectively. II, Performance of mNGS **(A)**, CT **(B)**, and mNGS combined with CT **(C)** against final clinical diagnoses.

## Discussion

To the best of our knowledge, this is the first report on summarizing regional pathogen profiles of infected patients after HSCT in Jilin Province of China revealed by mNGS and CT. Bloodstream infection was the most common infection in patients after HSCT, followed by pulmonary infection, CNS infection, and chest infection. Taking final clinical diagnoses as gold standard, we found that more than 75% of infections were caused by viruses, and CMV and EBV were found to be the most common pathogens. Besides, mNGS can be considered as a supplement to CT for infection diagnosis in patients after HSCT infected mainly by viruses.

We found that the performance of mNGS in diagnosing infection was not better than that of CT, which was contrary to the results of previous studies (Qu et al., 2022; Gu et al., 2021; Chen et al., 2020; Yu et al., 2022). It was reported that the sensitivity of mNGS was determined by the pathogen DNA ratio in sample (Ebinger et al., 2021). Host depletion methods, such as differential lysis method, can filter human DNA (Ji et al., 2020), increasing pathogen DNA ratio (Ji et al., 2020; Gu et al., 2021; Thoendel et al., 2018) at the expense of some viruses, parasites, and bacteria (Ji et al., 2020). Besides, host depletion methods may bring in contamination of engineered strains from reagents (Gu et al., 2021), decreasing the detection accuracy (Han et al., 2021). In our study, viral infection accounted for more than 75%, and host depletion method was also included in mNGS pipeline. However, there was no need to deplete host DNA for PCR in CT used in our study, avoiding the loss of viral genes. The above might be the reasons why the performance of mNGS was slightly lower than that of CT. Furthermore, we found that mNGS combined with CT can identify more causative pathogens and the TCR can reach up to 76.7%. Accordingly, we propose that mNGS and CT should be simultaneously performed on patients mainly infected by viruses to improve diagnostic accuracy.

In our study, CMV and EBV were found to be the most common pathogens causing infections in patients after HSCT. EBV is responsible for 2-5% of viral encephalitis and meningitis patients, and a review covering a 10-year period revealed that EBV was occasionally detected in CSF using PCR (Lee et al., 2021). Latent infections in B lymphocytes and myeloid cells are essential for persistence and transmission of EBV and CMV (Fragkou et al., 2021), respectively, and EBV is estimated to infect >90% of adults worldwide (Hong et al., 2021). Besides, treatment-related immunosuppression provides conditions for reactivation and infection of latent viruses (Fragkou et al., 2021; Arruti et al.,

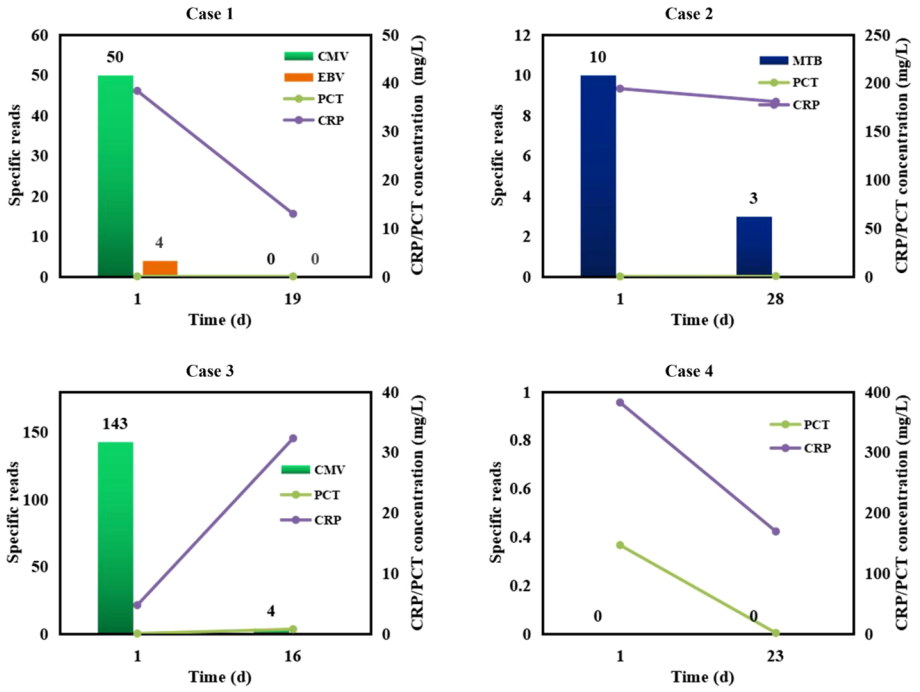



FIGURE 6
Semi-quantitative value of mNGS in the dynamic surveillance of infections. Case 1, CMV and EBV infections. Case 2, *M.tuberculosis* (MTB) infection. Case 3, CMV infection. Case 4, non-infection.

2017), and viral reactivation in patients after HSCT is associated with significant morbidity and mortality (Fragkou et al., 2021). Previous studies also found that EBV was the leading cause of CNS viral infection in patients after HSCT (Qu et al., 2022; Liu et al., 2019), which was consistent with our study (Figure 3). Meanwhile, CMV was found to be one of the most significant pathogens causing morbidity and mortality in patients after HSCT (Imlay and Kaul, 2021; Haidar et al., 2020), and we also found that the most common pathogen was CMV in patients after HSCT. Given the high risk of viral reactivation, we propose that clinicians should strictly perform seropositivity screening for viruses on both donors and candidates according to pre-transplantation evaluation guidelines (Fragkou et al., 2021; Tomblyn et al., 2009; U.S. Food and Drug Administration, 2019), to judge whether pre-emptive treatment should be conducted.

Furthermore, reads of semi-quantitative mNGS can reflect disease progression and treatment efficacy (Ge et al., 2021; Zhang et al., 2020; Ai et al., 2018). Besides, the value of mNGS as being a "rule-out" role (O'Grady, 2019; Chiu and Miller, 2019) is beneficial to minimizing the abuse of anti-microbial drugs (Ge et al., 2021). Zhang et al. performed several mNGS tests at different intervals on nine patients to evaluate the role of mNGS during the treatments of CNS infections, and found that for the patients with effective antimicrobial treatment in reference to mNGS results, mNGS semi-quantitative sequencing reads correlated with CSF WBC and glucose ratio level (Zhang et al., 2020). Chen et al. also exhibited the important role of repeated mNGS tests in diagnosing and managing *Escherichia coli* infection of neonates (Chen et al., 2022). The findings of our study provide more evidences to further emphasize the importance of mNGS in diagnosing, managing, and ruling out infections, and an era of more rapid, independent, and impartial diagnosis of infections in patients after HSCT can be expected. Given the high cost of mNGS, it needs skilled staff to be correctly performed.

## Limitations

RNA mNGS was not performed on the patients in this study, and more samples from multiple hospitals are needed to further evaluate the performance of both DNA and RNA mNGS in diagnosing infections in patients after HSCT. More negative controls from non-infectious cases should be included to further detect the potential of mNGS in ruling out infection. In addition, the potential of mNGS to help with the timely adjustment of treatments should also be evaluated.

## Conclusions

Our findings highlight the importance of mNGS combined with CT in diagnosing infections of patients after HSCT, especially for viral infection. Bloodstream infection was the most common infection in patients after HSCT, followed by pulmonary infection, CNS infection, and chest infection. Most of infections were caused by viruses, including CMV and EBV. Based on our findings, extensive application of mNGS in diagnosing infections after HSCT could be expected.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://ngdc.cncb.ac.cn/search/?dbId=bioproject&q=PRJCA014245, PRJCA014245.

## Ethics statement

This study was reviewed and approved by the Ethical Review Committee of the First hospital of Jilin University (approval no. 2022-566).

## Author contributions

HZ: Writing – original draft, Writing – review & editing. SG: Writing – original draft, Writing – review & editing. XL: Writing – original draft. YL: Formal analysis, Writing – original draft. YX: Writing – original draft. AL: Writing – original draft. YJ: Funding acquisition, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (nos. 22174137), Jilin Province Science and Technology Agency (nos. JJKH20211210KJ, JJKH20211164KJ, JLSWSRCZX2020-009, 20200901025SF, and 20200403084SF), and Beijing Medical Award Foundation (YXJL-2021-1097-0645.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1378112/full#supplementary-material

## References

Ai, J. W., Zhang, H. C., Cui, P., Xu, B., Gao, Y., Cheng, Q., et al. (2018). Dynamic and direct pathogen load surveillance to monitor disease progression and therapeutic efficacy in central nervous system infection using a novel semi-quantitive sequencing platform. *J. Infect.* 76, 307–310. doi: 10.1016/j.jinf.2017.11.002

Arruti, M., Piñeiro, L., Salicio, Y., Cilla, G., Goenaga, M., and de Munain, A. L. (2017). Incidence of varicella zoster virus infections of the central nervous system in the elderly: a large tertiary hospital-based series (2007–2014). *J. Neurovirol.* 23, 451–459. doi: 10.1007/s13365-017-0519-y

Berhane, M., Gidi, N. W., Eshetu, B., Gashaw, M., Tesfaw, G., Wieser, A., et al. (2021). Clinical profile of neonates admitted with sepsis to neonatal intensive care unit of Jimma Medical Center, a tertiary hospital in Ethiopia. *Ethiopian J. Health Sci.* 31, 485–494. doi: 10.4314/ejhs.v31i3.5

Chen, Y., Feng, W., Ye, K., Guo, L., Xia, H., Guan, Y., et al. (2021). Application of metagenomic next-generation sequencing in the diagnosis of pulmonary infectious pathogens from bronchoalveolar lavage samples. *Front. Cell Infect. Microbiol.* 11, 168. doi: 10.3389/fcimb.2021.541092

Chen, H., Yin, Y., Gao, H., Guo, Y., Dong, Z., Wang, X., et al. (2020). 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.* 71, S416–S426. doi: 10.1093/cid/ciaa1516

Chen, L., Zhao, Y., Wei, J., Huang, W., Ma, Y., Yang, X., et al. (2022). Metagenomic next-generation sequencing for the diagnosis of neonatal infectious diseases. *Microbiol. Spectr.* 10, e0119522. doi: 10.1128/spectrum.01195-22

Chiu, C. Y., and Miller, S. A. (2019). Clinical metagenomics. *Nat. Rev. Genet.* 20, 341–355. doi: 10.1038/s41576-019-0113-7

Dorresteijn, K. R., Brouwer, M. C., Jellema, K., and van de Beek, D. (2020). Bacterial external ventricular catheter-associated infection. *Expert Rev. Anti. Infect. Ther.* 18, 219–229. doi: 10.1080/14787210.2020.1717949

Ebinger, A., Fischer, S., and Höper, D. (2021). A theoretical and generalized approach for the assessment of the sample-specific limit of detection for clinical metagenomics. *Comput. Struct. Biotechnol. J.* 19, 732–742. doi: 10.1016/j.csbj.2020.12.040

Fragkou, P. C., Moschopoulos, C. D., Karofylakis, E., Kelesidis, T., and Tsiodras, S. (2021). Update in viral infections in the intensive care unit. *Front. Med.* 8, 575580. doi: 10.3389/fmed.2021.575580

Garcia, J. B., Lei, X., Wierda, W., Cortes, J. E., Dickey, B. F., Evans, S. E., et al. (2013). Pneumonia during remission induction chemotherapy in patients with acute leukemia. *Ann. Am. Thorac. Soc.* 10, 432–440. doi: 10.1513/AnnalsATS.201304-097OC

Ge, M., Gan, M., Yan, K., Xiao, F., Yang, L., Wu, B., et al. (2021). Combining metagenomic sequencing with whole exome sequencing to optimize clinical strategies in neonates with a suspected central nervous system infection. *Front. Cell Infect. Microbiol.* 11. doi: 10.3389/fcimb.2021.671109

Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryke, D., et al. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* 27, 115–124. doi: 10.1038/s41591-020-1105-z

Haidar, G., Boeckh, M., and Singh, N. (2020). Cytomegalovirus infection in solid organ and hematopoietic cell transplantation: state of the evidence. *J. Infect. Dis.* 221, S23–S31. doi: 10.1093/infdis/jiz454

Han, D., Diao, Z., Lai, H., Han, Y., Xie, J., Zhang, R., et al. (2021). Multilaboratory assessment of metagenomic next-generation sequencing for unbiased microbe detection. *J. Adv. Res.* 38, 213–222. doi: 10.1016/j.jare.2021.09.011

Herbozo, C., Julca, I., Flores, F., Hernandez, R., and Zegarra, J. (2021). Incidence and microbiological characteristics of neonatal late onset sepsis in a neonatal intensive care unit in Peru. *Int. J. Infect. Dis.* 108, 171–175. doi: 10.1016/j.ijid.2021.05.012

Hong, J., Wei, D., Wu, Q., Zhong, L., Chen, K., Huang, Y., et al. (2021). Antibody generation and immunogenicity analysis of EBV gp42 N-terminal region. *Viruses* 13, 2380. doi: 10.3390/v13122380

Imlay, H. N., and Kaul, D. R. (2021). Letermovir and maribavir for the treatment and prevention of cytomegalovirus infection in solid organ and stem cell transplant recipients. *Clin. Infect. Dis.* 73, 156–160. doi: 10.1093/cid/ciaa1713

Ji, X.-C., Zhou, L.-F., Li, C.-Y., Shi, Y.-J., Wu, M.-L., Zhang, Y., et al. (2020). Reduction of human DNA contamination in clinical cerebrospinal fluid specimens improves the sensitivity of metagenomic next-generation sequencing. *J. Mol. Neurosci.* 70, 659–666. doi: 10.1007/s12031-019-01472-z

Jiang, J., Bai, L., Yang, W., Peng, W., An, J., Wu, Y., et al. (2021). Metagenomic next-generation sequencing for the diagnosis of pneumocystis jirovecii pneumonia in non-HIV-infected patients: A retrospective study. *Infect. Dis. Ther.* 10, 1733–1745. doi: 10.1007/s40121-021-00482-y

Kumar, C. S., Subramanian, S., Murki, S., Rao, J., Bai, M., Penagaram, S., et al. (2021). Predictors of mortality in neonatal pneumonia: an INCLEN childhood pneumonia study. *Indian Pediatr.* 58, 1040–1045. doi: 10.1007/s13312-021-2370-8

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lee, G.-H., Kim, J., Kim, H.-W., and Cho, J. W. (2021). Clinical significance of Epstein-Barr virus in the cerebrospinal fluid of immunocompetent patients. *Clin. Neurol. Neurosurg.* 202, 106507. doi: 10.1016/j.clineuro.2021.106507

Li, N., Cai, Q., Miao, Q., Song, Z., Fang, Y., and Hu, B. (2021). High-throughput metagenomics for identification of pathogens in the clinical settings. *Small Methods* 5, 2000792. doi: 10.1002/smtd.202000792

Liu, N., Kan, J., Cao, W., Cao, J., Jiang, E., Zhou, Y., et al. (2019). Metagenomic next-generation sequencing diagnosis of peripheral pulmonary infectious lesions through virtual navigation, radial EBUS, ultrathin bronchoscopy, and ROSE. *J. Int. Med. Res.* 47, 4878–4885. doi: 10.1177/0300060519866953

Marcilla-Vazquez, C., Martinez-Gutierrez, A., Carrascosa-Romero, M., Baquero-Cano, M., and Alfaro-Ponce, B. (2018). Neonatal viral meningitis. The importance of the polymerase chain reaction in their diagnosis. *Rev. Neurol.* 67, 484–490. doi: 10.33588/rn.6712.2018203

O'Grady, J. (2019). A powerful, non-invasive test to rule out infection. *Nat. Microbiol.* 4, 554–555. doi: 10.1038/s41564-019-0424-7

Panackal, A. V. (2021). Role of Biomarkers in Neonatal Sepsis. Are we in search of the Holy Grail? *Fetus Newborn* 1, 43–49. doi: 10.52314/fnb.2021.v1i2.19

Qu, Y., Ding, W., Liu, S., Wang, X., Wang, P., Liu, H., et al. (2022). Metagenomic next-generation sequencing vs. Traditional pathogen detection in the diagnosis of infection after allogeneic hematopoietic stem cell transplantation in children. *Front. Microbiol.* 13, 868160. doi: 10.3389/fmicb.2022.868160

Saha, S. K., Schrag, S. J., El Arifeen, S., Mullany, L. C., Islam, M. S., Shang, N., et al. (2018). Causes and incidence of community-acquired serious infections among young children in south Asia (ANISA): an observational cohort study. *Lancet* 392, 145–159. doi: 10.1016/S0140-6736(18)31127-9

Shane, A. L., Sánchez, P. J., and Stoll, B. J. (2017). Neonatal sepsis. *Lancet* 390, 1770–1780. doi: 10.1016/S0140-6736(17)31002-4

Thoendel, M. J., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., et al. (2018). Identification of prosthetic joint infection pathogens using a shotgun metagenomics approach. *Clin. Infect. Dis.* 67, 1333–1338. doi: 10.1093/cid/ciy303

Tomblyn, M., Chiller, T., Einsele, H., Gress, R., Sepkowitz, K., Storek, J., et al. (2009). Guidelines for preventing infectious complications among hematopoietic cell transplantation recipients: a global perspective. *Biol. Blood Marrow Transplant.* 15, 1143–1238. doi: 10.1016/j.bbmt.2009.06.019

U.S. Food and Drug Administration. (2019). *Testing donors of human cells, tissues, and cellular and tissue-based products (HCT/P): specific requirements*. Available at: https://www.fda.gov/vaccines-blood-biologics/safety-availability-biologics/testing-donors-human-cells-tissues-and-cellular-and-tissue-based-products-hctp-specific-requirements.

Wilson, M. R., Naccache, S. N., Samayoa, E., Biagtan, M., Bashir, H., Yu, G., et al. (2014). Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *New Engl. J. Med.* 370, 2408–2417. doi: 10.1056/NEJMoa1401268

Wu, M., Chen, Y., Xia, H., Wang, C., Tan, C. Y., Cai, X., et al. (2020). Transcriptional and proteomic insights into the host response in fatal COVID-19 cases. *Proc. Natl. Acad. Sci.* 117, 28336–28343. doi: 10.1073/pnas.2018030117

Yu, L., Zhang, Y., Zhou, J., Zhang, Y., Qi, X., Bai, K., et al. (2022). Metagenomic next-generation sequencing of cell-free and whole-cell DNA in diagnosing central nervous system infections. *Front. Cell Infect. Microbiol.* 12, 951703. doi: 10.3389/fcimb.2022.951703

Zanella, M. C., Cordey, S., Laubscher, F., Docquier, M., Vieille, G., Van Delden, C., et al. (2021). Unmasking viral sequences by metagenomic next-generation sequencing in adult human blood samples during steroid-refractory/dependent graft-versus-host disease. *Microbiome* 9, 28. doi: 10.1186/s40168-020-00953-3

Zhang, Y., Cui, P., Zhang, H.-C., Wu, H.-L., Ye, M.-Z., Zhu, Y.-M., et al. (2020). Clinical application and evaluation of metagenomic next-generation sequencing in suspected adult central nervous system infection. *J. Trans. Med.* 18, 1–13. doi: 10.1186/s12967-020-02360-6

Zhang, C., Li, Z., Wang, M., Zhou, J., Yu, W., Liu, H., et al. (2023). High specificity of metagenomic next-generation sequencing using protected bronchial brushing sample in diagnosing pneumonia in children. *Front. Cell Infect. Microbiol.* 13, 1165432. doi: 10.3389/fcimb.2023.1165432

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 28 March 2024
DOI 10.3389/fcimb.2024.1378379



OPEN ACCESS

EDITED BY
Jiemin Zhou,
Vision Medicals Co, Ltd., China

REVIEWED BY
Ying Liang,
Peking University Third Hospital, China
Ying Tang,
First Affiliated Hospital of Jilin University, China
Arturo Cortes-Telles,
IMSS-BIENESTAR, Mexico
Ping Fang,
The Second Affiliated Hospital of Xi'an Jiaotong University, China

*CORRESPONDENCE
Bing Sun
ricusunbing@126.com

†These authors have contributed equally to this work

RECEIVED 29 January 2024
ACCEPTED 21 March 2024
PUBLISHED 28 March 2024

CITATION
Tang X, Xu X-L, Wan N, Zhao Y, Wang R, Li X-Y, Li Y, Wang L, Li H-C, Gu Y, Zhang C-Y, Yang Q, Tong Z-H and Sun B (2024) Long-term outcomes of survivors with influenza A H1N1 virus-induced severe pneumonia and ARDS: a single-center prospective cohort study. *Front. Cell. Infect. Microbiol.* 14:1378379. doi: 10.3389/fcimb.2024.1378379



# Long-term outcomes of survivors with influenza A H1N1 virus-induced severe pneumonia and ARDS: a single-center prospective cohort study

Xiao Tang[1†], Xiao-Li Xu[2†], Na Wan[1], Yu Zhao[1], Rui Wang[1], Xu-Yan Li[1], Ying Li[1], Li Wang[1], Hai-Chao Li[1], Yue Gu[1], Chun-Yan Zhang[1], Qi Yang[2], Zhao-Hui Tong[1] and Bing Sun[1*]

[1]Department of Respiratory and Critical Care Medicine, Beijing Institute of Respiratory Medicine and Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China, [2]Department of Radiology, Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China

**Introduction:** Systematic evaluation of long-term outcomes in survivors of H1N1 is still lacking. This study aimed to characterize long-term outcomes of severe H1N1-induced pneumonia and acute respiratory distress syndrome (ARDS).

**Method:** This was a single-center, prospective, cohort study. Survivors were followed up for four times after discharge from intensive care unit (ICU) by lung high-resolution computed tomography (HRCT), pulmonary function assessment, 6-minute walk test (6MWT), and SF-36 instrument.

**Result:** A total of 60 survivors of H1N1-induced pneumonia and ARDS were followed up for four times. The carbon monoxide at single breath ($D_{LCO}$) of predicted values and the 6MWT results didn't continue improving after 3 months. Health-related quality of life didn't change during the 12 months after ICU discharge. Reticulation or interlobular septal thickening on HRCT did not begin to improve significantly until the 12-month follow-up. The $D_{LCO}$ of predicted values showed negative correlation with the severity degree of primary disease and reticulation or interlobular septal thickening, and a positive correlation with physical functioning. The $D_{LCO}$ of predicted values and reticulation or interlobular septal thickening both correlated with the highest tidal volume during mechanical ventilation. Levels of fibrogenic cytokines had a positive correlation with reticulation or interlobular septal thickening.

**Conclusion:** The improvements in pulmonary function and exercise capacity, imaging, and health-related quality of life had different time phase and impact on each other during 12 months of follow-up. Long-term outcomes of pulmonary fibrosis might be related to the lung injury and excessive lung fibroproliferation at the early stage during ICU admission.

KEYWORDS

**influenza A (H1N1) virus, severe community-acquired pneumonia (SCAP), acute respiratory distress syndrome (ARDS), pulmonary fibrosis, pulmonary function**

# Introduction

Influenza has threatened human health for decades (Taubenberger and Morens, 2006). It is estimated that global seasonal influenza-associated respiratory deaths account for 4.0–8.8 per 100,000 individuals annually (Iuliano et al., 2018). During the influenza A (H1N1) virus pandemic in the United States in 2009, 26% of patients with H1N1 pneumonia progressed to acute respiratory distress syndrome (ARDS), and 52% of them were admitted to intensive care unit (ICU) (Jain et al., 2012). The study from Mexico and Canada during the 2013–2014 influenza season showed that the 90-day mortality of critical illness in ICUs was 34.6% (Dominguez-Cherit et al., 2016). It is important to understand the long-term outcomes in survivors.

A study from Canada that focused on the one-year outcomes in survivors of ARDS showed that carbon monoxide diffusion capacity remained low throughout the 12-month follow-up (Herridge et al., 2003). At the 5-year follow-up, pulmonary function was normal to near-normal, but none of the survivors returned to normal predicted levels of physical function (Herridge et al., 2011). Another study from Spain on survivors of ARDS with 6-month follow-up showed a poorer health-related quality of life and mild to moderate pulmonary functional abnormalities compared with the healthy population (Masclans et al., 2011). Similar results were found in survivors among H1N1 patients during the 12-month follow-up (Luyt et al., 2012). However, most studies focused on long-term outcomes in the survivors of H1N1 during the 2009 pandemic. In addition, there are still fewer systematic evaluations of the physical function, pulmonary imaging, and quality of life in such patients. There have been many cases of critical illness of severe H1N1 pneumonia and ARDS in every flu season after 2009 (Iuliano et al., 2018; Li et al., 2019).

This study aimed to characterize long-term outcomes in survivors of severe H1N1 pneumonia and ARDS during 12 months of follow-up. And try to explore the relationship between the long-term pulmonary function and exercise capacity, health-related quality of life, pulmonary imaging, with clinical situation and serological biomarkers acquired in the early stage while ICU admission.

# Methods

## Study design and patients

This was a single-center, prospective, cohort study conducted in a 16-bed respiratory ICU. Patients admitted to the ICU from March 1, 2016, to December 31, 2020, were included in the study. Patients were included if they fulfilled the following criteria: (1) age above 18 years; (2) severe community acquired pneumonia (SCAP) (Cao et al., 2018); (3) meeting the Berlin definition of ARDS (Ferguson et al., 2012); and (4) detection of influenza A (H1N1) virus in sputum or bronchoalveolar lavage fluid (BALF) using real-time polymerase chain reaction (PCR). Exclusion criteria included expected ICU duration less than 48 hours and refuse to participate this study or follow-up.

This study was reviewed and approved by the Ethics Committee of Beijing Chao-Yang Hospital (2016-KE-61). Informed consent was obtained from the patients themselves or their legal guardians. All methods were carried out in accordance with relevant guidelines and regulations.

## Procedures and data collection

Demographic and clinical data of the patients during ICU stay were entered into an electronic case report form (eCRF) and included the following: demographic characteristics, underlying diseases, comorbidities, clinical symptoms, vital signs, laboratory tests, images of the lung, and microbiological findings. Antimicrobial therapy, respiratory support, complications, and outcomes were also recorded. The acute physiology and chronic health evaluation (APACHE) II, sequential organ failure assessment (SOFA), and acute lung injury score (Murray et al., 1988) was also collected while ICU admission.

The survivors were examined during four follow-up visits at outpatient department during 12 months, specifically at 1 month, at 3 months, at 6 months, and at 12 months after ICU discharge. Symptoms and vital signs were recorded. The patients underwent blood routine test, arterial blood gas (ABG) analysis, pulmonary high-resolution computed tomography (HRCT), pulmonary function assessment, and 6-minute walk test (6MWT) at every follow-up visit. Moreover, a 36-item short-form health survey (SF-36) (Skinner et al., 2015) was used to evaluate the physical and mental health function. All of the results were recorded in the eCRF. The diagnose of pulmonary function was referred to Chinese experts' consensus on the standardization of adult lung function diagnosis (Lei and Rong-Chang, 2022) and ERS/ATS technical standard on interpretive strategies for routine lung function tests (Stanojevic et al., 2022). 6MWT performed according to Chinese expert consensus on standardized clinical application of 6-minute walk test (Chinese Society of Cardiology CMA et al., 2022).

The primary outcome was the incidence of patients with abnormal pulmonary function at 12-month follow-up. The secondary outcome was the symptom, HRCT manifestation, pulmonary function, SF-36, and 6MWT at 4 times follow-up.

## Pulmonary HRCT

We performed lung segmentation, lesion extraction and labeling, and volume calculation using a dedicated multi-task deep learning algorithm developed for pulmonary pneumonia (Beijing Deepwise & League of PhD Technology Co.Ltd, China). One experienced radiologist with experience in pulmonary imaging interpretation reviewed the CT images. The CT images were evaluated and defined according to the Fleischner Society glossary of terms for thoracic imaging (Hansell et al., 2008) as the following radiologic patterns: (a) ground-glass opacification (GGO); (b) consolidation; (c) bronchiectasis; (d) reticulation or interlobular septal thickening; and (e) emphysema. The extent of disease at HRCT was evaluated as a CT score. Bilateral lungs were divided into

five lung zones, where each lung zone was assigned a score that was based on the following: score 0, 0% involvement; score 1, less than 25% involvement; score 2, 25% to less than 50% involvement; score 3, 50% to less than 75% involvement; and score 4, 75% or greater involvement. Summation of scores provided overall lung involvement (maximal CT score for both lungs was 20) (Ooi et al., 2004).

## Measurement of inflammatory and fibrogenic cytokines

Serum specimens were collected dynamically during the first nine days for the measurement of inflammatory cytokines and fibrogenic cytokines. A human cytokine panel (Procarta Plex™, Affmetrix Inc., CA, USA) consisting of IFN-γ, IL-12p70, IL-13, IL-1b, IL-2, IL-4, IL-5, IL-6, TNF-α, GM-CSF, IL-18, IL-10, IL-17A, IL-21, IL-22, IL-23, IL-27, IL-9, IFN-α, IL-15, IL-1a, IL-1RA, IL-7, TNF-β, and IL-31 was used to measure inflammatory cytokines. Fibrogenic cytokines consisted of hyaluronic acid, laminin, type IV collagen, type III procollagen (Bioscience, Tianjin, China), and Krebs Von den Lungen-6 (KL-6) (Fujirebio Inc., Tokyo, Japan).

## Statistical analysis

Data analysis was performed using SPSS 23.0 (IBM Corp., Armonk, NY, USA) software. Categorical variables were summarized using frequencies and percentages, and continuous data were presented as the medians (interquartile ranges [IQRs]). The Mann–Whitney *U* test was used for continuous variables, and the $\chi 2$ test or Fisher's exact test was used for categorical variables. Differences between groups were tested by one-way analysis of variance test. The overall time course for pulmonary function, HRCT manifestations, and eight dimensions of the SF-36 instrument was analyzed using two-way analysis of variance for repeated measures. Pearson correlation analysis was used to analyze the correlation between cytokine levels, pulmonary function, pulmonary manifestations on HRCT, and dimensions of the SF-36 instrument in four visits during the 12-month follow-up after ICU discharge. Stepwise regression and collinearity diagnostics were used to analyze the multicollinearity among the variables. *P* values lower than 0.05 were considered to be statistically significant.

# Results

From March 1, 2016, to December 31, 2019, a total of 345 patients with SCAP and ARDS were admitted to the respiratory ICU of Beijing Chao-Yang Hospital. In total, 92 (26.7%) of them were diagnosed with influenza A (H1N1) virus–induced SCAP, of which 70 (76.1%) patients were male, and the median age was 49 (41, 63) years. The median acute lung injury score was 3.25 (2.75, 3.74), and Acute Physiology and Chronic Health Evaluation (APACHE) II was 12 (9, 18) at the time of admission. The median $PaO_2/FiO_2$ was 107.5 (77.0, 137.8) mm Hg. All of the patients received mechanical ventilation during ICU therapy. Invasive mechanical ventilation (IMV) was used in 70 (76.1%) patients, including 38 (41.3%) patients who received extracorporeal membrane oxygenation (ECMO) support. The ICU mortality was 34.8%, and the median length of ICU stay was 16 (10, 27) days (Supplementary Table S1). Sixty survivors underwent four follow-up examinations at 1, 3, 6, and 12 months after ICU discharge. Six patients were lost to follow-up, and one patient died for cerebrovascular disease (Supplementary Figure S1).

## Symptoms, pulmonary function, and 6MWT

At the one-month follow-up, 39.3% of the survivors had cough. The number of patients with respiratory tract symptoms decreased gradually with time. Fourteen survivors (24.0%) had hypoxemia at the one-month follow-up, which decreased to four patients (7.0%) at the three-month follow-up. However, at the 12-month follow-up, there were still four patients with cough, and one patient had dyspnea and hypoxemia (Table 1).

Pulmonary function assessment at the one-month follow-up revealed gas transfer impairments in 32 patients (55.2%), of which

TABLE 1 Symptoms, pulmonary function, and 6MWT during 12 months of follow-up.

| Follow-up | 1 month (n = 58) | 3 months (n = 57) | 6 months (n = 54) | 12 months (n = 53) | *P* |
|---|---|---|---|---|---|
| Symptom (n, %) | | | | | |
| Cough | 24 (39.3) | 12 (21.0) | 7 (12.7) | 4 (7.5) | 0.416 |
| Sputum | 15 (24.6) | 8 (14.0) | 6 (10.9) | 3 (5.7) | 0.969 |
| Short breath | 4 (6.6) | 3 (5.3) | 0 | 1 (1.9) | 0.537 |
| Dyspnea | 8 (13.1) | 0 | 0 | 1 (1.9) | 0.094 |
| Arterial blood gas analysis | | | | | |
| pH | 7.41 (7.39, 7.43) | 7.40 (7.39, 7.42) | 7.41 (7.40, 7.41) | 7.41 (7.39, 7.42) | 0.820 |
| $PaCO_2$ (mm Hg) | 32.0 (31.2, 34.5) | 32.4 (30.7, 33.5) | 33.8 (30.6, 34.9) | 35.8 (33.7, 37.2) | 0.660 |

*(Continued)*

10.3389/fcimb.2024.1378379

TABLE 1 Continued

| Follow-up | 1 month (n = 58) | 3 months (n = 57) | 6 months (n = 54) | 12 months (n = 53) | *P* |
|---|---|---|---|---|---|
| **Arterial blood gas analysis** | | | | | |
| $PaO_2$ (mm Hg) | 89.2 (74.4, 95.9) | 91.9 (86.8, 99.2) | 88.1 (79.3, 93.5) | 89.7 (76.5, 93.4) | 0.489 |
| $PaO_2/FiO_2$ (mm Hg) | 406.8 (354.1, 455.2) | 439.7 (410.5, 474.4) | 421.5 (379.5, 447.4) | 428.2 (365.3, 446.4) | 0.431 |
| Hypoxemia (n, %) | 14 (24.0) | 4 (7.0) | 3 (5.4) | 1 (1.9) | 0.001 |
| **Pulmonary function** | | | | | |
| FVC (L) | 3.30 (2.51, 3.84) | 3.74 (2.87, 4.26) * | 3.87 (3.09, 4.22) * | 3.82 (3.17, 4.26) | 0.009 |
| Difference | | 0.560 (0.143, 0.977) | 0.185 (0.005, 0.364) | 0.075 (−0.151, 0.302) | |
| FVC% of predicted value | 81.8 (71.6, 89.4) | 97.2 (81.7, 101.3) * | 97.2 (87.6, 103.0) * | 93.9 (89.0, 112.4) | 0.011 |
| Difference | | 17.445 (4.533, 30.358) | 5.327 (0.356, 10.299) | 3.173 (−3.277, 9.622) | |
| FEV 1 (L) | 2.72 (2.15, 3.18) | 3.02 (2.42, 3.50) * | 2.96 (2.56, 3.45) | 3.00 (2.54, 3.28) | 0.038 |
| Difference | | 0.385 (0.058, 0.711) | 0.106 (−0.039, 0.252) | 0.045 (−0.109, 0.198) | |
| FEV 1% of predicted value | 82.0 (73.1, 90.4) | 93.1 (82.2, 101.2) * | 97.2 (83.9, 101.4) * | 92.9 (83.0. 108.8) | 0.024 |
| Difference | | 14.300 (0.570, 28.030) | 5.845 (1.100, 10.590) | 1.464 (−4.870, 7.797) | |
| FEV 1/FVC | 85.8 (81.2, 89.0) | 84.5 (80.9, 96.2) | 85.7 (82.4, 97.5) | 85.0 (79.9, 95.3) | 0.558 |
| Difference | | 1.609 (−7.952, 1.170) | −1.736 (−15.754, 12.282) | −1.855 (−7.188, 3.479) | |
| MEF75/25 | 2.90 (2.15, 4.00) | 3.33 (2.22, 3.90) | 2.76 (2.09, 3.27) | 2.48 (2.06, 3.75) | 0.787 |
| Difference | | 0.098 (−0.777, 0.974) | 0.020 (−0.620, 0.660) | −0.027 (−0.364, 0.311) | |
| MEF 75/25% of predicted value | 75.95 (57.1, 96.9) | 79.1 (62.5, 100.3) | 73.5 (61.2, 99.0) | 72.1 (55.3, 100.9) | 0.788 |
| Difference | | 4.063 (−24.678, 32.804) | 1.427 (−18.643, 21.516) | −0.850 (−11.580, 9.880) | |
| TLC (L) | 4.51 (3.53, 5.30) | 5.23 (3.88, 5.79) | 5.16 (4.39, 6.08) | 5.20 (4.25, 5.90) | 0.033 |
| Difference | | 1.267 (−0.098, 2.633) | 0.175 (−0.401, 0.752) | 0.088 (−0.172, 0.349) | |
| TLC% of predicted value | 73.65 (63.85, 86.15) | 84.1 (76.1, 92.7) * | 90.3 (76.9, 95.1) | 92.3 (84.6, 99.4) | <0.001 |
| Difference | | 14.320 (4.233, 24.407) | 5.440 (−0.210, 11.090) | 1.380 (−3.833, 6.593) | |
| $D_{LCO}$ (mL/min/mm Hg) | 5.29 (3.42, 6.99) | 6.15 (5.00, 7.90) | 6.44 (5.74, 7.86) | 6.78 (6.06, 8.52) | <0.001 |
| Difference | | 1.022 (−1.027, 3.070) | 0.994 (−0.511, 2.499) | 0.355 (−0.334, 1.043) | |
| $D_{LCO}$% of predicted value | 56.7 (45.7, 69.0) | 71.9 (59.0, 82.9) * | 73.7 (68.3, 86.8) | 81.2 (73.8, 90.8) | <0.001 |
| Difference | | 23.364 (8.953, 37.774) | 5.900 (−4.602, 16.402) | 3.882 (−5.510, 13.273) | |
| $D_{LCO}$/VA (mL/min/mm Hg/L) | 1.33 (1.13, 1.54) | 1.28 (1.11, 1.56) | 1.37 (1.22, 1.62) | 1.60 (1.22, 1.79) | 0.584 |
| Difference | | −0.189 (−1.349, 0.971) | 0.076 (−0.144, 0.296) | 11.326 (−26.684, 49.336) | |
| $D_{LCO}$/VA% of predicted value | 77.5 (67.1, 91.9) | 84.3 (75.2, 94.6) * | 88.3 (79.0, 104.9) | 95.2 (80.5, 111.7) | 0.001 |
| Difference | | 16.3 (5.2, 27.5) | 5.2 (−4.5, 14.9) | 2.9 (−5.2, 11.1) | |
| 6MWT | 450 (372, 519) | 540 (421, 597) * | 534 (493, 600) * | 538 (430, 604) | 0.005 |
| | | 135 (56, 214) | 76 (98, 250) | −5 (70, 61) | |

P: Two-way analysis of variance for repeated measures to analyze the difference between four visits during 12 months.
One-way analysis of variance was used to analyze differences between subgroups. Differences showed the comparison with the last follow-up visit. *: with significant difference compared with the result of the last follow-up visit.
$D_{LCO}$, diffusion capacity of the lung for carbon monoxide at single breath; $D_{LCO}$/VA, diffusion capacity of the lung for carbon monoxide corrected by alveolar volume; $FiO_2$, fraction of inspired oxygen; FEV 1, the first-second forced expiratory volume; FVC, forced vital capacity; MEF 25, maximum expiratory flow rate at 25% vital capacity; MEF 50, maximum expiratory flow rate at 50% vital capacity; MEF 75, maximum expiratory flow rate at 75% vital capacity; 6MWT, 6-minute walk test; $PaCO_2$, arterial partial pressure of carbon dioxide; $PaO_2$, arterial partial pressure of oxygen; and TLC, total lung capacity.

24 patients (41.4%) had limited ventilation, while eight patients had small-airway dysfunction. The dynamic evaluation showed that the pulmonary function significantly improved at the 3-month follow-up compared with the 1-month follow-up. Single breath diffusing capacity of lung for carbon monoxide ($D_{LCO}$) of predict values was 71.9% (59.0%, 82.9%) at the 3-month follow-up, significantly higher than 56.7% (45.7%, 69.0%) at the 1-month follow-up. However, the comparison between 3-month, 6-month, and 12-month follow-up showed no significant difference. At the 12-month follow-up, there were still 12 survivors (21.8%) with gas transfer impairments, in whom $D_{LCO}$ of predict values was 69.7% (66.1%, 75.2%). Seven (12.7%) patients had small-airway dysfunction, and four (7.3%) patients had limited ventilation. The distance at 6MWT at the 1-month follow-up was 450 m (372 m, 519 m), which increased significantly to 540 m (421 m, 597 m) at the 3-month follow-up but did not further improve at the 6-month and 12-month follow-up (Table 1 and Figure 1).

## Physical and mental health function

The SF-36 instrument was used to evaluate the physical and mental health function. At the 1-month follow-up, the SF-36 instrument showed poor manifestations in eight dimensions of physical and mental health function. Physical functioning and social functioning significantly improved at the 3-month follow-up. Then, at 6 months after ICU discharge, bodily pain and role–physical significantly improved. However, general health, vitality, mental health, and role–emotion did not change during 12 months after ICU discharge (Table 2 and Figure 2A).

## Pulmonary HRCT

The median patient's whole-lung volume by volumetric HRCT analysis was 3352 (2241, 3893) ml at the 1-month follow-up, which increased to 3769 (2696, 4643) ml at the 3-month follow-up ($P$ = 0.017). At the one-month visit, the main pulmonary HRCT manifestation was GGO, of which the median score was 14 (10, 18). GGO improved significantly at the next three follow-up visits ($P < 0.001$). Consolidation was fully disappeared at 3 months after ICU discharge. However, bronchiectasis, reticulation, or interlobular septal thickening did not begin to improve significantly until the 12-month follow-up (Table 3 and Figures 2B, Supplementary Figure S2).

## Correlation analysis

$D_{LCO}$ of predict values at four follow-up examinations showed a moderate negative correlation with reticulation or interlobular septal thickening manifested on HRCT at the corresponding time point (Figure 2B). Other indicators of pulmonary function and HRCT manifestations did not correlate. There were moderate to strong positive correlations between $D_{LCO}$ of predict values and physical functioning ($r$ = 0.683, $P < 0.001$), role–physical ($r$ = 0.622, $P < 0.001$), role–emotional ($r$ = 0.507, $P$ = 0.003), and 6MWT ($r$ = 0.522, $P$ = 0.004) at the 3-month follow-up (Supplementary Table S2). At the 12-month visit, $D_{LCO}$ of predict values still showed moderate positive correlations with role–physical ($r$ = 0.431, $P$ = 0.032) and general health ($r$ = 0.491, $P$ = 0.013) (Supplementary Table S2 and Figures 3A, B).

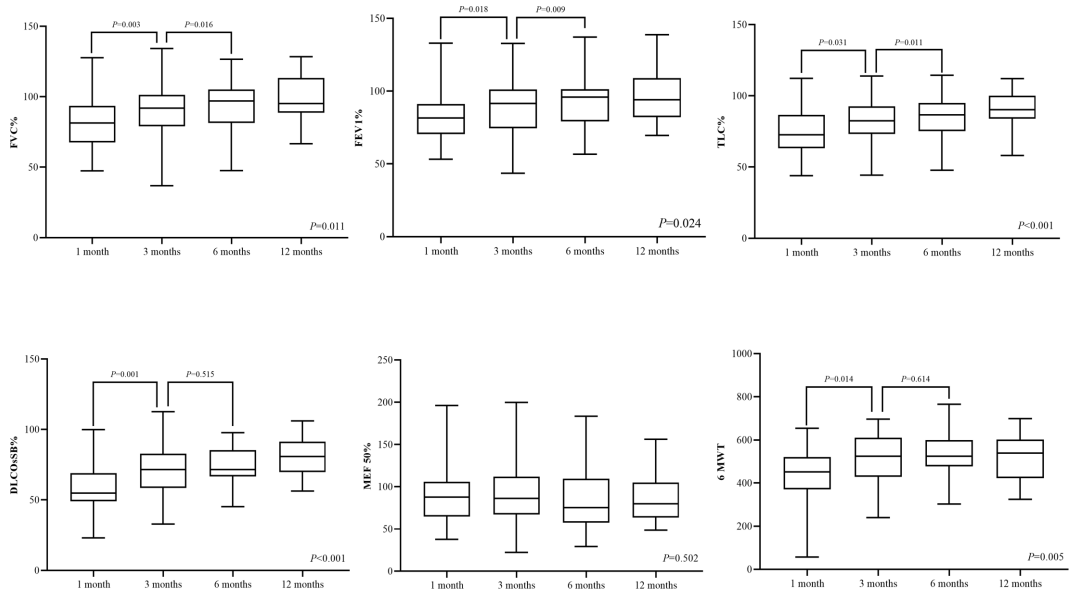



FIGURE 1
Pulmonary function assessment and 6MWT in four visits during 12 months of follow-up. $D_{LCO}$, diffusion capacity of the lung for carbon monoxide at single breath; TLC, total lung capacity; FEV 1, the first-second forced expiratory volume; FVC, forced vital capacity; MEF 50, maximum expiratory flow rate at 50% vital capacity; and 6MWT, 6-minute walk test.

10.3389/fcimb.2024.1378379

TABLE 2 The scores in the SF-36 instrument during 12 months of follow-up.

| Follow-up | 1 month (n = 58) | 3 months (n = 57) | 6 months (n = 54) | 12 months (n = 53) | *P* |
|---|---|---|---|---|---|
| Bodily pain | 74 (70, 100) | 92 (74, 100) | 92 (84, 100) * | 100 (82, 100)* | 0.030 |
| Difference | | 16 (−1, 32) | 4 (−4, 12) | −1 (−8, 5) | |
| Physical functioning | 70 (55, 90) | 92 (80, 95) * | 95 (85, 95) * | 95 (82, 98) * | 0.050 |
| Difference | | 23 (5, 41) | 3 (−2, 8) | 2 (−7, 12) | |
| Role–physical | 0 (0, 0) | 50 (0, 100) | 100 (25, 100)* | 100 (75, 100)* | 0.002 |
| Difference | | 32 (−12, 77) | 15 (−13, 43) | 25 (−4, 54) | |
| General health | 72 (50, 82) | 80 (57, 97) | 72 (52, 87) | 76 (57, 95) | 0.118 |
| Difference | | 5 (−6, 16) | 13 (−1, 27) | 2 (−13, 17) | |
| Vitality | 40 (25, 55) | 38 (20, 50) | 35 (25, 40) | 40 (35, 50) | 0.262 |
| Difference | | 6 (−12, 23) | −14 (−30, 2) | 3 (−8, 14) | |
| Social functioning | 55.6 (33.3, 77.8) | 88.9 (66.7, 100) * | 89 (78, 100) * | 100 (89, 100) * | 0.004 |
| Difference | | 21 (4, 38) | 9 (−8, 26) | 2 (−9, 13) | |
| Role–emotional | 83.3 (0, 100) | 100 (66.7, 100) | 78 (68, 88) | 100 (83, 100) | 0.258 |
| Difference | | 23 (−10, 57) | 17 (−6, 39) | −14 (−37, 9) | |
| Mental health | 80 (64, 88) | 87 (72, 92) | 78 (68, 88) | 80 (68, 92) | 0.382 |
| Difference | | 4 (−10, 17) | −1 (−7, 5) | −8 (−24, 8) | |
| Reported health transition | 25 (0, 25) | 37.5 (25, 75) | 50 (25, 100)* | 75 (25, 100)* | 0.003 |
| Difference | | 22 (−2, 47) | 27 (−6, 60) | −3 (−32, 27) | |

P: Two-way analysis of variance for repeated measures to analyze the difference between four visits during 12 months.
One-way analysis of variance was used to analyze differences between subgroups. Differences showed the comparison with the last follow-up visit. *: there was significant difference of the result compared with the 1-month follow-up.

Regarding the variables during ICU hospitalization, acute lung injury score ($r$ = −0.471, $P$ = 0.017) and APACHE II during ICU admission ($r$ = −0.691, $P$ < 0.001) had a negative correlation with $D_{LCO}$ of predict values at the 12-month follow-up. The highest tidal volume was the only monitoring parameter during mechanical ventilation that had a negative correlation with $D_{LCO}$ of predict values ($r$ = −0.542, $P$ = 0.025) and a positive correlation with reticulation or interlobular septal thickening ($r$ = 0.606, $P$ = 0.003) at the 12-month follow-up (Figures 3C, D). Regarding the fibrogenic cytokines at the early stage of ICU admission, KL-6

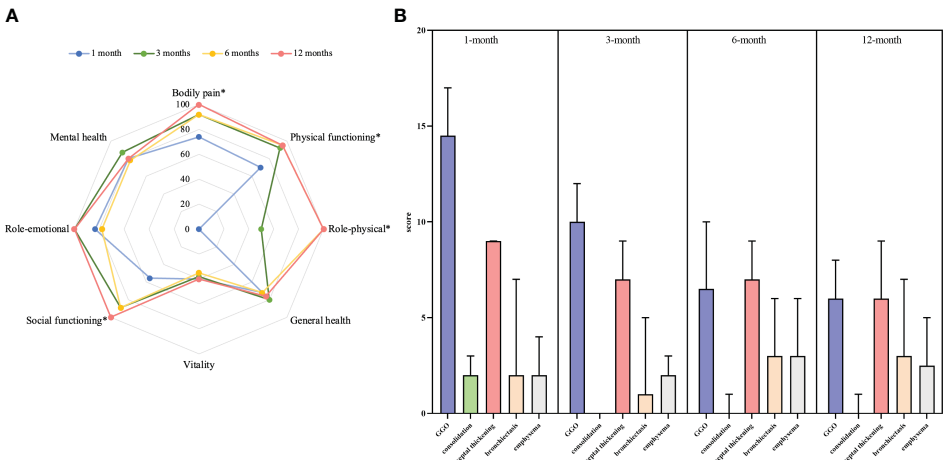



FIGURE 2
SF-36 and pulmonary HRCT scores in four visits during 12 months of follow-up of the survivors of H1N1 virus–induced SCAP and ARDS. **(A)** Eight dimensions of the SF-36 instrument. *: Two-way analysis of variance for repeated measures to analyze the difference between the four visits showed significant within-group differences. **(B)** Pulmonary HRCT scores of different manifestations. GGO, ground-glass opacity; HRCT, high-resolution computed tomography.

10.3389/fcimb.2024.1378379

TABLE 3 Pulmonary HRCT descriptions during 12 months of follow-up.

| HRCT manifestation | 1 month (n = 58) | 3 months (n = 57) | 6 months (n = 54) | 12 months (n = 53) | *P* |
|---|---|---|---|---|---|
| GGO | 14 (10, 18) | 10 (6, 16)* | 9 (4, 11)* | 6 (2, 10)* | <0.001 |
| Difference | | −3 (−5, −1) | −3 (−6, −1) | −2 (−5, 1) | |
| Consolidation | 3 (2, 5) | 0 (0, 4)* | 0 (0, 3)* | 0 (0, 2)* | 0.015 |
| Difference | | −2 (−4, 0) | 0 (−1, 0) | 0 (−1, 0) | |
| Reticulation/interlobular septal thickening | 9 (6, 11) | 9 (6, 10) | 8 (6, 10) | 6 (5, 9) ** | 0.031 |
| Difference | | 0 (−1, 1) | −1 (−2, 0) | −1 (−1, 0) | |
| Bronchiectasis | 6 (2, 11) | 5 (2, 9) | 5 (1, 8) | 4 (1, 8) ** | 0.018 |
| Difference | | −1 (−3, 0) | −1 (−1. 0) | 0 (−1, 0) | |
| Emphysema | 4 (1, 8) | 4 (2, 8) | 4 (2, 9) | 4 (2, 10) | 0.123 |
| Difference | | 0 (−1, 1) | 0 (−1, 1) | 0 (0, 1) | |
| Lung volume | 3352 (2241, 3893) | 3769 (2696, 4643)* | 3537 (2486, 4840)* | 4012 (2872, 4721)* | 0.004 |
| Difference | | 584 (173, 995) | −31 (−287, 225) | 174 (−105, 453) | |

P: Two-way analysis of variance for repeated measures to analyze the difference between four visits during 12 months.
One-way analysis of variance was used to analyze differences between subgroups. Differences showed the comparison with the last follow-up visit. *: the result of this follow-up with significant difference compared with the 1-month follow-up. **: the result of this follow-up with significant difference compared either with the 1-month follow-up, or 3-month follow-up.
GGO, ground-glass opacity; HRCT, high-resolution computed tomography.

level showed moderately positive correlations with reticulation or interlobular septal thickening at 1-month and 3-month follow-up, while type IV collagen, type III procollagen, hyaluronic acid, and laminin had moderate to strong positive correlations with reticulation or interlobular septal thickening at 6 months and 12 months after ICU discharge (Supplementary Table S3 and Figures 3E–H). There were no correlations between inflammatory cytokines and HRCT manifestations. Very weak multicollinearity among the above variables was shown in the collinearity diagnostics (Supplementary Table S4).

## Discussion

This study systematically assessed the long-term outcomes of pulmonary function, exercise capacity, HRCT manifestations, and

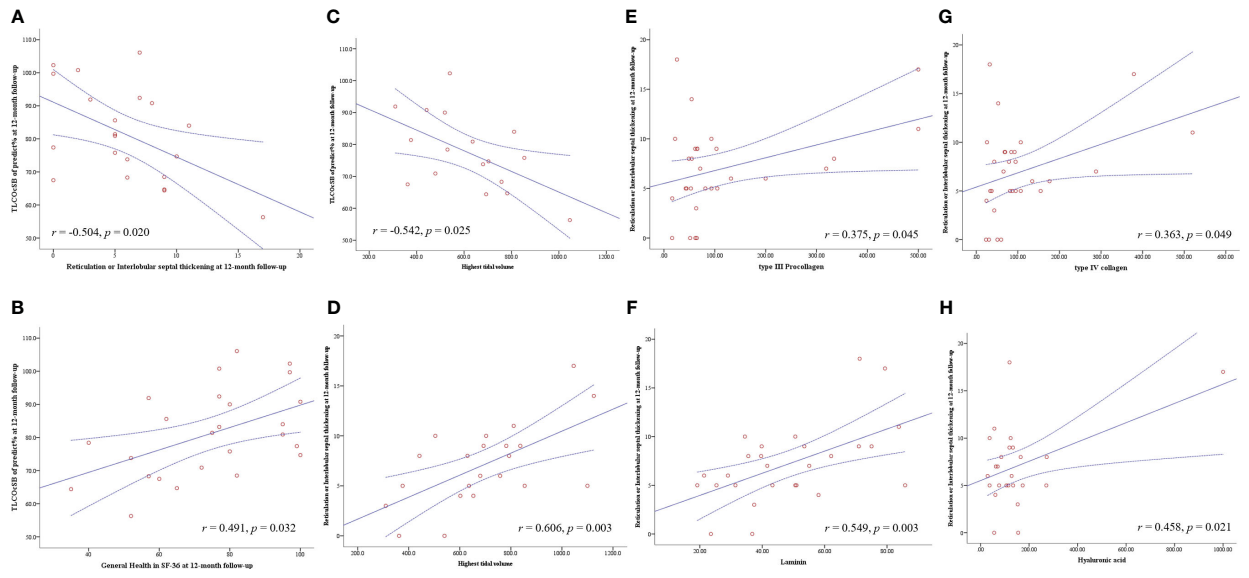



FIGURE 3
Pearson correlation analysis was used to analyze the correlation between the fibrogenic cytokines and variables during ICU therapy, pulmonary function, pulmonary HRCT, and SF-36 instrument in four visits during 12 months of follow-up after ICU discharge. The solid line is the fitted line of Pearson correlation analysis; the dotted line is the 95% confidence interval. $D_{LCO}$, diffusion capacity of the lung for carbon monoxide at single breath; HRCT, high-resolution computed tomography; and ICU, intensive care unit.

health-related quality of life in survivors of severe H1N1 pneumonia and ARDS who had been discharged from ICU. The improvements in pulmonary function and exercise capacity, HRCT, and health-related quality of life had different time phase and impact on each other. The decrease in $D_{LCO}$ was the main abnormality in pulmonary function, which improved at 3 months after ICU discharge. Although GGO and consolidation were fully absorbed at 3 months after ICU discharge, reticulation or interlobular septal thickening was the main long-term manifestation on HRCT, which did not begin to improve significantly until the 12-month follow-up, and had a negative correlation with $D_{LCO}$. Regarding health-related quality of life, general health and mental health were without any changes during one year after ICU discharge, even though there was a positive correlation between $D_{LCO}$ and general health. The severity of primary disease, tidal volume of mechanical ventilation, and levels of fibrogenic cytokines at the early stage of ICU admission affected the long-term outcomes of reticulation or interlobular septal thickening on HRCT.

A study from Greece evaluated 44 patients with H1N1 infection every three months, until 6 months after discharge. There was an improvement in pulmonary function tests at the second measurement, but there were no changes between the second and third measurements (Zarogoulidis et al., 2011). Huang's team in Taiwan followed up on nine survivors of ARDS due to severe H1N1 pneumonitis at 1, 3, and 6 months after hospital discharge. Pulmonary function and 6MWT results increased from 1 to 3 months after hospital discharge, but there was no further improvement from 3 to 6 months after discharge (Hsieh et al., 2018). Similar results were also found for ARDS of other etiologies. A 3-month follow-up study in survivors of ARDS secondary to coronavirus disease 2019 (COVID-19) showed that 82% of these patients had an impaired $D_{LCO}$ (Gonzalez et al., 2021). Similar to the previous literature, the present study showed that gas transfer impairment was the main manifestation of the abnormal pulmonary function, and in most survivors, it recovered to normal range at 3 months after ICU discharge. However, there were still 25% of patients in whom there was no improvement during 12 months.

Bilateral, peripheral GGO, and/or bilateral areas of consolidation are the predominant HRCT findings in the acute phase of H1N1-induced pneumonia (Marchiori et al., 2010). Short-term serial HRCT evaluation of patients with H1N1 infection showed that GGO and/or consolidation on initial CT scans tended to resolve fibrosis, which then resolved completely or displayed substantially reduced residual disease (Li et al., 2012). Another study showed that the most common HRCT finding in the convalescent stage in patients with H1N1 pneumonia was fibrosis (Feng et al., 2015). Nevertheless, few studies have focused on the long-term radiographic outcomes in survivors of H1N1 infection.

The present study revealed the dynamic HRCT evolution in survivors of H1N1-induced pneumonia and ARDS during 12 months after ICU discharge. GGO and consolidation were the main manifestations on HRCT at the early stage after ICU discharge, similar to that at the acute phase of H1N1 pneumonia. Especially, consolidation was mostly absorbed at the 3-month visit. However, reticulation or interlobular septal thickening did not begin to improve significantly until the 12-month follow-up. It means that reticulation or interlobular septal thickening is the main long-term HRCT finding in H1N1-induced pneumonia and ARDS. A similar result was found in the study of COVID-19 survivors (Gonzalez et al., 2021; Pan et al., 2022). We also discovered that the severity of reticulation or interlobular septal thickening correlated with diffusion capacity in such patients.

A study from Australia showed that health-related quality of life of survivors of severe H1N1 influenza was comparable to that of healthy population one year after ICU discharge (Skinner et al., 2015). However, researchers from Spain proposed a significant but temporary impact on health-related quality of life in the majority of patients with H1N1 infection (Hollmann et al., 2013). According to a review from Bein's team, surviving ARDS is associated with a long-term substantial reduction in health-related quality of life (Bein et al., 2018). Neufeld's team also reported that two-thirds of ARDS survivors had clinically significant fatigue symptoms in the first year of follow-up (Neufeld et al., 2020). In this study, we found that general health, vitality, mental health, and role–emotion in survivors of H1N1-induced pneumonia and ARDS did not change during 12 months after ICU discharge, which might be affected partially by impaired pulmonary function, and the experience in ICU might impact long-term mental health. Therefore, for the improvement in health-related quality of life, both psychological and physical recovery should be focused on.

Type III procollagen is recognized as a marker of ARDS-associated lung fibroproliferation (Forel et al., 2015). KL-6 is a marker of alveolar inflammation in ARDS (Nathani et al., 2008), which is also related to long-term outcomes (Kondo et al., 2011). In the present study, KL-6 was related to reticulation or interlobular septal thickening at the 1-month and 3-month visits, but not at the 12-month follow-up. In contrast, type IV collagen, type III procollagen, hyaluronic acid, and laminin showed correlations with reticulation or interlobular septal thickening at 6 months and 12 months after ICU discharge. We considered that the lung injury and excessive lung fibroproliferation at the early stage during ICU admission might be related to long-term pulmonary fibrosis in patients with H1N1-induced ARDS. It also influences the recovery of pulmonary function and quality of life.

Lung injury and excessive lung fibroproliferation at the early stage during ICU admission might not only be related to the severity of the primary disease (Burnham et al., 2014) but also to mechanical ventilation. Many studies have tried to explore the mechanism of ventilation-induced acute lung injury and fibroproliferation in patients with ARDS, which caused poor outcomes, and high tidal volume was identified as one of the important risk factors (Foda et al., 2001; Li et al., 2008; Cabrera-Benitez et al., 2014). Lung-protective mechanical ventilation with lower tidal volumes has been widely used and could decrease the mortality of patients with ARDS (Acute Respiratory Distress Syndrome N et al., 2000). In the present study, all of the patients received standard lung-protective mechanical ventilation; nevertheless, the relationship between the tidal volume and the long-term outcome was still observed, especially influencing the pulmonary function and fibrosis manifestation on HRCT at the 12-month visit. Therefore, strictly executing the mechanical ventilation

strategy with lower tidal volumes might help patients with ARDS in achieving better short-term or long-term outcomes. There is still controversy regarding the effectiveness of corticosteroids applied to reduce ventilation-induced lung fibrosis in severe H1N1 infection–induced ARDS (Burnham et al., 2014; Cabrera-Benitez et al., 2014; Moreno et al., 2018). Thus, effective pharmacological therapies to prevent the development of ventilator-induced and ARDS-related lung fibrosis should be the research priority in the future.

There are still some limitations to this study. First, this was a single-center cohort study, which may have induced the unavoidable selection bias. Second, the sample size was relatively small, which might have resulted in less statistical power. Third, this study focused on patients with H1N1 pneumonia–induced ARDS; therefore, extrapolation of the results to other etiologies of pneumonia and ARDS should be done with caution. Fourth, post-ICU life capability, workability, and delirium which was recognized as aspects related to long-term prognosis after ICU discharge was not collected in the present study. This study only describes the long-term prognosis from the perspectives of imaging, health-related quality of life, and pulmonary function. Lastly, inflammatory cytokines and fibrogenic cytokines were not measured in the bronchoalveolar lavage fluid, so we could only partly explain the injury degree of alveolar epithelial cells through the results from serum.

## Conclusions

The improvements in pulmonary function and exercise capacity, imaging, and health-related quality of life in survivors of critical H1N1-induced pneumonia and ARDS had different time phase and impact on each other during 12 months of follow-up after ICU discharge. Long-term pulmonary fibrosis outcome, which would influence the pulmonary function and quality of life, could be observed via HRCT longitudinal evaluation. It might be not only related to the severity of the primary disease but also to mechanical ventilation. These topics still need to be explored in multicenter studies, covering various etiologies and using larger sample sizes. Moreover, the mechanism also needs to be further investigated.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving humans were approved by the Ethics Committee of Beijing Chao-Yang Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

XT: Conceptualization, Data curation, Investigation, Writing – original draft. X-LX: Conceptualization, Data curation, Investigation, Writing – review & editing. NW: Investigation, Writing – review & editing. YZ: Investigation, Writing – review & editing. RW: Formal analysis, Investigation, Writing – review & editing. X-YL: Investigation, Writing – review & editing. YL: Investigation, Writing – review & editing. LW: Investigation, Writing – review & editing. H-CL: Investigation, Writing – review & editing. YG: Investigation, Writing – review & editing. C-YZ: Investigation, Writing – review & editing. QY: Investigation, Writing – review & editing. Z-HT: Funding acquisition, Writing – review & editing. BS: Conceptualization, Funding acquisition, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Clinical medicine development project of Beijing Hospital Authority (XMLX202105), clinical diagnosis and treatment technology and translational research project of Beijing (Z201100005520030), excellent talents development project of public health technical (XUEKEDAITOUREN-01-19), National Natural Science Foundation of China (82000103), and Reform and Development Program of Beijing Institute of Respiratory Medicine (Ggyfz202332).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1378379/full#supplementary-material

# References


Acute Respiratory Distress Syndrome N, Brower, R. G., Matthay, M. A., Morris, A., Schoenfeld, D., Thompson, B. T., et al. (2000). Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl. J. Med.* 342, 1301–1308. doi: 10.1056/NEJM200005043421801

Bein, T., Weber-Carstens, S., and Apfelbacher, C. (2018). Long-term outcome after the acute respiratory distress syndrome: different from general critical illness? *Curr. Opin. Crit. Care* 24, 35–40. doi: 10.1097/MCC.0000000000000476

Burnham, E. L., Janssen, W. J., Riches, D. W., Moss, M., and Downey, G. P. (2014). The fibroproliferative response in acute respiratory distress syndrome: mechanisms and clinical significance. *Eur. Respir. J.* 43, 276–285. doi: 10.1183/09031936.00196412

Cabrera-Benitez, N. E., Laffey, J. G., Parotto, M., Spieth, P. M., Villar, J., Zhang, H., et al. (2014). Mechanical ventilation-associated lung fibrosis in acute respiratory distress syndrome: a significant contributor to poor outcome. *Anesthesiology.* 121, 189–198. doi: 10.1097/ALN.0000000000000264

Cao, B., Huang, Y., She, D. Y., Cheng, Q. J., Fan, H., Tian, X. L., et al. (2018). Diagnosis and treatment of community-acquired pneumonia in adults: 2016 clinical practice guidelines by the Chinese Thoracic Society, Chinese Medical Association. *Clin. Respir. J.* 12, 1320–1360. doi: 10.1111/crj.12674

Chinese Society of Cardiology CMA, Professional Committee of Cardiopulmonary P, Rehabilitation of Chinese Rehabilitation Medical A and Editorial Board of Chinese Journal of C (2022). [Chinese expert consensus on standardized clinical application of 6-minute walk test]. *Zhonghua Xin Xue Guan Bing Za Zhi* 50, 432–442. doi: 10.3760/cma.j.cn112148-20211206-01054

Dominguez-Cherit, G., de la Torre, A., Rishu, A., Pinto, R., Namendys-Silva, S. A., Camacho-Ortiz, A., et al. (2016). (H1N1pdm09)-related critical illness and mortality in Mexico and Canada, 2014. *Crit. Care Med.* 44, 1861–1870. doi: 10.1097/CCM.0000000000001830

Feng, F., Xia, G., Shi, Y., and Zhang, Z. (2015). Longitudinal changes of pneumonia complicating novel influenza A (H1N1) by high-resolution computed tomography. *Radiol. Infect. Dis.* 2, 40–46. doi: 10.1016/j.jrid.2015.06.001

Ferguson, N. D., Fan, E., Camporota, L., Antonelli, M., Anzueto, A., Beale, R., et al. (2012). The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material. *Intensive Care Med.* 38, 1573–1582. doi: 10.1007/s00134-012-2682-1

Foda, H. D., Rollo, E. E., Drews, M., Conner, C., Appelt, K., Shalinsky, D. R., et al. (2001). Ventilator-induced lung injury upregulates and activates gelatinases and EMMPRIN: attenuation by the synthetic matrix metalloproteinase inhibitor, Prinomastat (AG3340). *Am. J. Respir. Cell Mol. Biol.* 25, 717–724. doi: 10.1165/ajrcmb.25.6.4558f

Forel, J. M., Guervilly, C., Hraiech, S., Voillet, F., Thomas, G., Somma, C., et al. (2015). Type III procollagen is a reliable marker of ARDS-associated lung fibroproliferation. *Intensive Care Med.* 41, 1–11. doi: 10.1007/s00134-014-3524-0

Gonzalez, J., Benitez, I. D., Carmona, P., Santisteve, S., Monge, A., Moncusi-Moix, A., et al. (2021). Pulmonary function and radiologic features in survivors of critical COVID-19: A 3-month prospective cohort. *Chest.* 160, 187–198. doi: 10.1016/j.chest.2021.02.062

Hansell, D. M., Bankier, A. A., MacMahon, H., McLoud, T. C., Muller, N. L., and Remy, J. (2008). Fleischner Society: glossary of terms for thoracic imaging. *Radiology.* 246, 697–722. doi: 10.1148/radiol.2462070712

Herridge, M. S., Cheung, A. M., Tansey, C. M., Matte-Martyn, A., Diaz-Granados, N., Al-Saidi, F., et al. (2003). One-year outcomes in survivors of the acute respiratory distress syndrome. *N Engl. J. Med.* 348, 683–693. doi: 10.1056/NEJMoa022450

Herridge, M. S., Tansey, C. M., Matte, A., Tomlinson, G., Diaz-Granados, N., Cooper, A., et al. (2011). Functional disability 5 years after acute respiratory distress syndrome. *N Engl. J. Med.* 364, 1293–1304. doi: 10.1056/NEJMoa1011802

Hollmann, M., Garin, O., Galante, M., Ferrer, M., Dominguez, A., and Alonso, J. (2013). Impact of influenza on health-related quality of life among confirmed (H1N1) 2009 patients. *PloS One* 8, e60477. doi: 10.1371/journal.pone.0060477

Hsieh, M. J., Lee, W. C., Cho, H. Y., Wu, M. F., Hu, H. C., Kao, K. C., et al. (2018). Recovery of pulmonary functions, exercise capacity, and quality of life after pulmonary rehabilitation in survivors of ARDS due to severe influenza A (H1N1) pneumonitis. *Influenza Other Respir. Viruses.* 12, 643–648. doi: 10.1111/irv.12566

Iuliano, A. D., Roguski, K. M., Chang, H. H., Muscatello, D. J., Palekar, R., Tempia, S., et al. (2018). Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet.* 391, 1285–1300. doi: 10.1016/S0140-6736(17)33293-2

Jain, S., Benoit, S. R., Skarbinski, J., Bramley, A. M., Finelli, L. Pandemic Influenza AVHIT (2012). Influenza-associated pneumonia among hospitalized patients with 2009 pandemic influenza A (H1N1) virus–United States, 2009. *Clin. Infect. Dis.* 54, 1221–1229. doi: 10.1093/cid/cis197

Kondo, T., Hattori, N., Ishikawa, N., Murai, H., Haruta, Y., Hirohashi, N., et al. (2011). KL-6 concentration in pulmonary epithelial lining fluid is a useful prognostic indicator in patients with acute respiratory distress syndrome. *Respir. Res.* 12, 32. doi: 10.1186/1465-9921-12-32

Lei, Z., and Rong-Chang, C. (2022). Chinese experts' consensus on the standardization of adult lung function diagnosis. *J. Clin. Pulmonary Med.* 27, 973–981.

Li, L. F., Liao, S. K., Huang, C. C., Hung, M. J., and Quinn, D. A. (2008). Serine/threonine kinase-protein kinase B and extracellular signal-regulated kinase regulate ventilator-induced pulmonary fibrosis after bleomycin-induced acute lung injury: a prospective, controlled animal experiment. *Crit. Care* 12, R103. doi: 10.1186/cc6983

Li, L., Liu, Y., Wu, P., Peng, Z., Wang, X., Chen, T., et al. (2019). Influenza-associated excess respiratory mortality in China, 2010-15: a population-based study. *Lancet Public Health* 4, e473–ee81. doi: 10.1016/S2468-2667(19)30163-X

Li, P., Zhang, J. F., Xia, X. D., Su, D. J., Liu, B. L., Zhao, D. L., et al. (2012). Serial evaluation of high-resolution CT findings in patients with pneumonia in novel swine-origin influenza A (H1N1) virus infection. *Br. J. Radiol.* 85, 729–735. doi: 10.1259/bjr/85580974

Luyt, C. E., Combes, A., Becquemin, M. H., Beigelman-Aubry, C., Hatem, S., Brun, A. L., et al. (2012). Long-term outcomes of pandemic 2009 influenza A(H1N1)-associated severe ARDS. *Chest.* 142, 583–592. doi: 10.1378/chest.11-2196

Marchiori, E., Zanetti, G., Hochhegger, B., Rodrigues, R. S., Fontes, C. A., Nobre, L. F., et al. (2010). High-resolution computed tomography findings from adult patients with Influenza A (H1N1) virus-associated pneumonia. *Eur. J. Radiol.* 74, 93–98. doi: 10.1016/j.ejrad.2009.11.005

Masclans, J. R., Roca, O., Munoz, X., Pallisa, E., Torres, F., Rello, J., et al. (2011). Quality of life, pulmonary function, and tomographic scan abnormalities after ARDS. *Chest.* 139, 1340–1346. doi: 10.1378/chest.10-2438

Moreno, G., Rodriguez, A., Reyes, L. F., Gomez, J., Sole-Violan, J., Diaz, E., et al. (2018). Corticosteroid treatment in critically ill patients with severe influenza pneumonia: a propensity score matching study. *Intensive Care Med.* 44, 1470–1482. doi: 10.1007/s00134-018-5332-4

Murray, J. F., Matthay, M. A., Luce, J. M., and Flick, M. R. (1988). An expanded definition of the adult respiratory distress syndrome. *Am. Rev. Respir. Dis.* 138, 720–723. doi: 10.1164/ajrccm/138.3.720

Nathani, N., Perkins, G. D., Tunnicliffe, W., Murphy, N., Manji, M., and Thickett, D. R. (2008). Kerbs von Lungren 6 antigen is a marker of alveolar inflammation but not of infection in patients with acute respiratory distress syndrome. *Crit. Care* 12, R12. doi: 10.1186/cc6785

Neufeld, K. J., Leoutsakos, J. S., Yan, H., Lin, S., Zabinski, J. S., Dinglas, V. D., et al. (2020). Fatigue symptoms during the first year following ARDS. *Chest.* 158, 999–1007. doi: 10.1016/j.chest.2020.03.059

Ooi, G. C., Khong, P. L., Muller, N. L., Yiu, W. C., Zhou, L. J., Ho, J. C., et al. (2004). Severe acute respiratory syndrome: temporal lung changes at thin-section CT in 30 patients. *Radiology.* 230, 836–844. doi: 10.1148/radiol.2303030853

Pan, F., Yang, L., Liang, B., Ye, T., Li, L., Li, L., et al. (2022). Chest CT patterns from diagnosis to 1 year of follow-up in patients with COVID-19. *Radiology.* 302, 709–719. doi: 10.1148/radiol.2021211199

Skinner, E. H., Haines, K. J., Howe, B., Hodgson, C. L., Denehy, L., McArthur, C. J., et al. (2015). Health-related quality of life in Australasian survivors of H1N1 influenza undergoing mechanical ventilation. *A Multicenter Cohort Study. Ann. Am. Thorac. Soc* 12, 895–903. doi: 10.1513/AnnalsATS.201412-568OC

Stanojevic, S., Kaminsky, D. A., Miller, M. R., Thompson, B., Aliverti, A., Barjaktarevic, I., et al. (2022). ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur. Respir. J.* 60 (1), 2101499. doi: 10.1183/13993003.01499-2021

Taubenberger, J. K., and Morens, D. M. (2006). 1918 Influenza: the mother of all pandemics. *Emerg. Infect. Dis.* 12, 15–22. doi: 10.3201/eid1209.05-0979

Zarogoulidis, P., Kouliatsis, G., Papanas, N., Spyratos, D., Constantinidis, T. C., Kouroumichakis, I., et al. (2011). Long-term respiratory follow-up of H1N1 infection. *Virol. J.* 8, 319. doi: 10.1186/1743-422X-8-319

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 10 April 2024
DOI 10.3389/fcimb.2024.1386506



OPEN ACCESS

EDITED BY
Jiemin Zhou,
Vision Medicals Co, Ltd, China

REVIEWED BY
Huaqiang Zhou,
Sun Yat-sen University Cancer Center (SYSUCC), China
Yi Xu,
Fifth Medical Center of the PLA General Hospital, China
Lingwei Wang,
Jinan University, China

*CORRESPONDENCE
Jinping Zheng
jpzhenggy@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 15 February 2024
ACCEPTED 22 March 2024
PUBLISHED 10 April 2024

CITATION
Wang Z, Li S, Cai G, Gao Y, Yang H, Li Y, Liang J, Zhang S, Hu J and Zheng J (2024) Mendelian randomization analysis identifies druggable genes and drugs repurposing for chronic obstructive pulmonary disease. *Front. Cell. Infect. Microbiol.* 14:1386506. doi: 10.3389/fcimb.2024.1386506



# Mendelian randomization analysis identifies druggable genes and drugs repurposing for chronic obstructive pulmonary disease

Zihui Wang†, Shaoqiang Li†, Guannan Cai†, Yuan Gao, Huajing Yang, Yun Li, Juncheng Liang, Shiyu Zhang, Jieying Hu and Jinping Zheng*

National Center for Respiratory Medicine, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou Medical University, Guangzhou, China

**Background:** Chronic obstructive pulmonary disease (COPD) is a prevalent condition that significantly impacts public health. Unfortunately, there are few effective treatment options available. Mendelian randomization (MR) has been utilized to repurpose existing drugs and identify new therapeutic targets. The objective of this study is to identify novel therapeutic targets for COPD.

**Methods:** Cis-expression quantitative trait loci (cis-eQTL) were extracted for 4,317 identified druggable genes from genomics and proteomics data of whole blood (eQTLGen) and lung tissue (GTEx Consortium). Genome-wide association studies (GWAS) data for doctor-diagnosed COPD, spirometry-defined COPD (Forced Expiratory Volume in one second [FEV1]/Forced Vital Capacity [FVC] <0.7), and FEV1 were obtained from the cohort of FinnGen, UK Biobank and SpiroMeta consortium. We employed Summary-data-based Mendelian Randomization (SMR), HEIDI test, and colocalization analysis to assess the causal effects of druggable gene expression on COPD and lung function. The reliability of these druggable genes was confirmed by eQTL two-sample MR and protein quantitative trait loci (pQTL) SMR, respectively. The potential effects of druggable genes were assessed through the phenome-wide association study (PheWAS). Information on drug repurposing for COPD was collected from multiple databases.

**Results:** A total of 31 potential druggable genes associated with doctor-diagnosed COPD, spirometry-defined COPD, and FEV1 were identified through SMR, HEIDI test, and colocalization analysis. Among them, 22 genes (e.g., MMP15, PSMA4, ERBB3, and LMCD1) were further confirmed by eQTL two-sample MR and protein SMR analyses. Gene-level PheWAS revealed that ERBB3 expression might reduce inflammation, while GP9 and MRC2 were associated with other traits. The drugs Montelukast (targeting the MMP15 gene) and MARIZOMIB (targeting the PSMA4 gene) may reduce the risk of spirometry-defined COPD. Additionally, an existing small molecule inhibitor of the APH1A gene has the potential to increase $FEV_1$.

**Conclusions:** Our findings identified 22 potential drug targets for COPD and lung function. Prioritizing clinical trials that target these identified druggable genes with existing drugs or novel medications will be beneficial for the development of COPD treatments.

KEYWORDS

Mendelian randomization, druggable genes, drugs repurposing, chronic obstructive pulmonary disease, lung function

# Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a complex lung condition characterized by persistent progressive airflow limitation, along with chronic respiratory symptoms (dyspnea, cough, sputum, etc.). Recent large-scale studies have revealed that the global prevalence of COPD among people aged 30-79 years was estimated to be 10.3% in 2019, with a staggering 391.9 million patients worldwide (Adeloye et al., 2022). The prevalence is expected to continue rising due to factors such as increasing smoking rates, air pollution, and an aging population. In addition, COPD ranks as the third leading cause of death globally, with approximately 2.5 deaths per minute worldwide, and a large proportion of deaths are caused by infections (S. European Observatory on Health, Policies et al., 2023). Given the high burden of COPD-related mortality and morbidity, it is necessary to implement interventions to reduce its prevalence and minimize the associated disease burden.

Despite extensive research, the treatment options for COPD remain limited. Currently, the main treatment involves bronchodilators, hormones, and mucolytic agents. However, these drugs only provide temporary relief of symptoms and modest improvements in lung function without significantly impacting the overall prevalence of the disease. Moreover, existing drugs are ineffective in controlling the symptoms and disease progression in some patients. Therefore, there is an urgent need to develop new therapeutic drugs.

Incorporating genetics into drug development shows promise in addressing this issue, as genetically-supported drugs are more likely to be successful in clinical trials (Trajanoska et al., 2023). For instance, the identification of the PCSK9 gene's association with coronary artery disease risk has led to the development and successful application of PCSK9 inhibitors (Stoekenbroek et al., 2015). Druggable genes refer to those that can be targeted by drugs, including small molecules and biotherapeutic agents, to modulate their activity (Finan et al., 2017). Genome-wide association studies (GWAS) have been effective in identifying single nucleotide polymorphisms (SNPs) associated with COPD (Silverman, 2020). However, GWAS may fall short in pinpointing causative genes for direct drug development. In this context, Mendelian randomization (MR) provides a more reliable approach by mimicking the effects of gene overexpression, similar to randomized controlled trials (Figure 1). As a result, MR informs drug development strategies (Schmidt et al., 2020) and has been employed to explore potential opportunities for drug repurposing (Folkersen et al., 2020; Gaziano et al., 2021; Ou et al., 2021; Storm et al., 2021).

In this study, we hypothesized that increased expression of the druggable gene in whole blood (exposure) could modify the risk of developing COPD (outcome). We extracted cis-expression quantitative trait loci (cis-eQTL) and protein quantitative trait loci (pQTL) for 4,317 druggable genes to investigate the causal effects of druggable gene expression on doctor-diagnosed COPD, spirometry-defined COPD (Forced Expiratory Volume in one second [FEV1]/Forced Vital Capacity [FVC]<0.7), and FEV1. Furthermore, the safety of the identified potential druggable genes and opportunities for repurposing licensed or clinical-stage drugs in COPD were assessed.

# Methods

## Overall design

We observed the hospitalization of 56 patients with COPD to observe the real-life disease burden of COPD on patients and to demonstrate the importance of drug development. The overall scheme of COPD drug target screening was shown in Figure 2. Briefly, we conducted a Summary-data-based Mendelian Randomization (SMR) analysis, using genetic variants that affect the expression of druggable genes in blood and lung tissue as instrumental variables to estimate the causal relationships between these genetic variations and doctor-diagnosed COPD, FEV1/FVC<0.7, and FEV1. The SMR method utilizes summary-level data from GWAS and QTL studies to test for associations between gene expression levels and complex traits of interest. Two-sample MR estimates the causal effects by separately estimating the associations between instrumental variables and both exposure factors and outcomes, requiring two sets of independent sample data. SMR can be used to prioritize potential genes for further research, while two-sample MR is suitable for causal inference. For

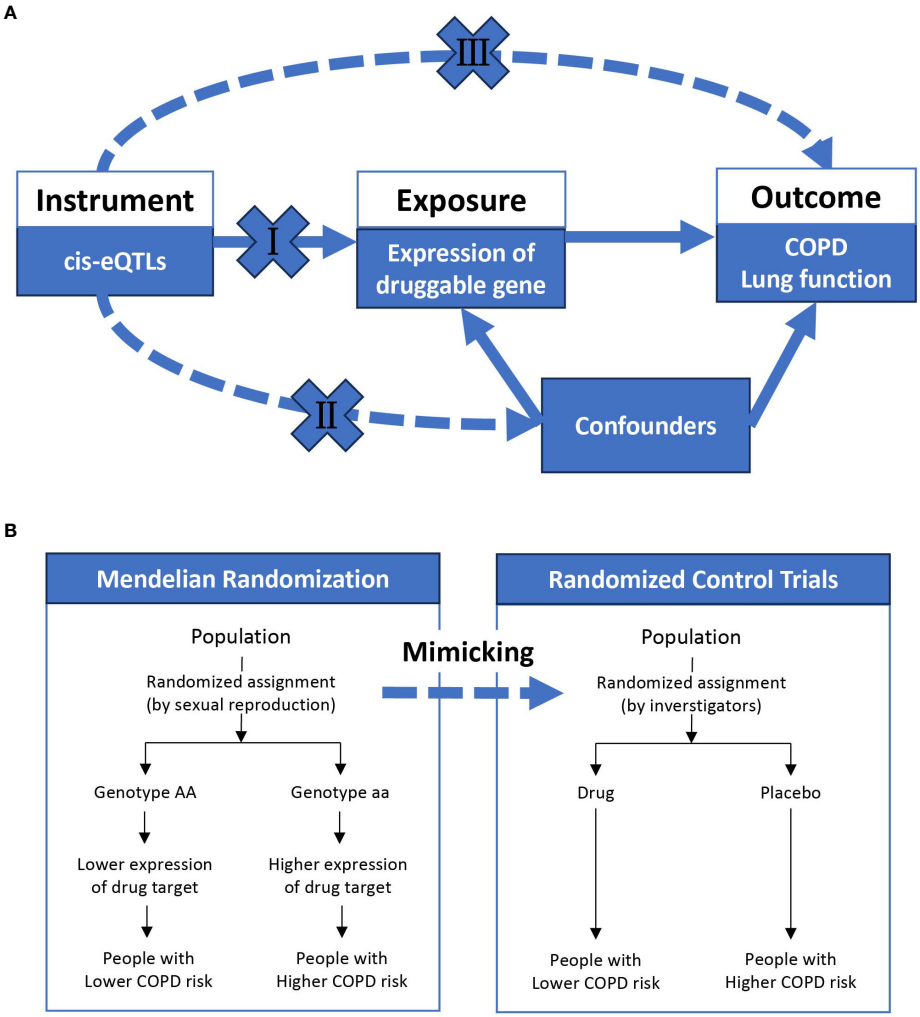



FIGURE 1
The overview of Mendelian randomization. **(A)** shows the basic assumptions underlying Mendelian randomization: (I) genetic variation must be associated with exposure; (II) genetic variation must not be associated with any confounding factors; (III) genetic variation must not be directly related to outcome. **(B)** indicates that Mendelian randomization is similar to a randomized controlled trial, where different genotypes (on autosomes) are randomly assigned from one generation to the next according to Mendel's law, and a population randomly assigned to a genotype with high expression of a particular gene will be exposed to high expression of a particular gene for life. COPD, chronic obstructive pulmonary disease; eQTL, expression quantitative trait loci.

SMR results meeting the significance threshold and passing the heterogeneity independent instrument (HEIDI) test, we conducted colocalization analyses to determine if the same SNPs were driving both the outcome and exposure. Furthermore, eQTL two-sample MR and protein SMR analyses were employed to ensure consistency in the results. Finally, we matched existing drugs to the identified druggable genes and conducted a PheWAS analysis to preliminarily evaluate the safety of these genes as potential drug targets (i.e., whether they may cause other diseases).

This study used the big data platform of the Clinical Medical Research Centre of the First Affiliated Hospital of Guangzhou Medical University, and the patient's clinical data were obliterated for privacy-related information before they were provided to the investigators. The observational study was approved by the Ethics Committee of the First Affiliated Hospital of Guangzhou Medical University (ES-2023-017-01).

## Druggable genes list

Druggable genes were obtained from a review by Finan et al. (Finan et al., 2017). A total of 4317 druggable genes on autosomes were utilized in this study.

## Exposure data

The eQTL data for whole blood were obtained from the eQTLGen consortia (www.eQTLgen.org), while the lung-specific eQTL data were obtained from the GTEx v.8 consortia (www.gtexportal.org). Detailed descriptions of the data can be found in the original publications (Aguet et al., 2020; Võsa et al., 2021). The eQTLGen consisted of a total of 25,482 whole blood samples from 37 datasets, with the majority of samples being of

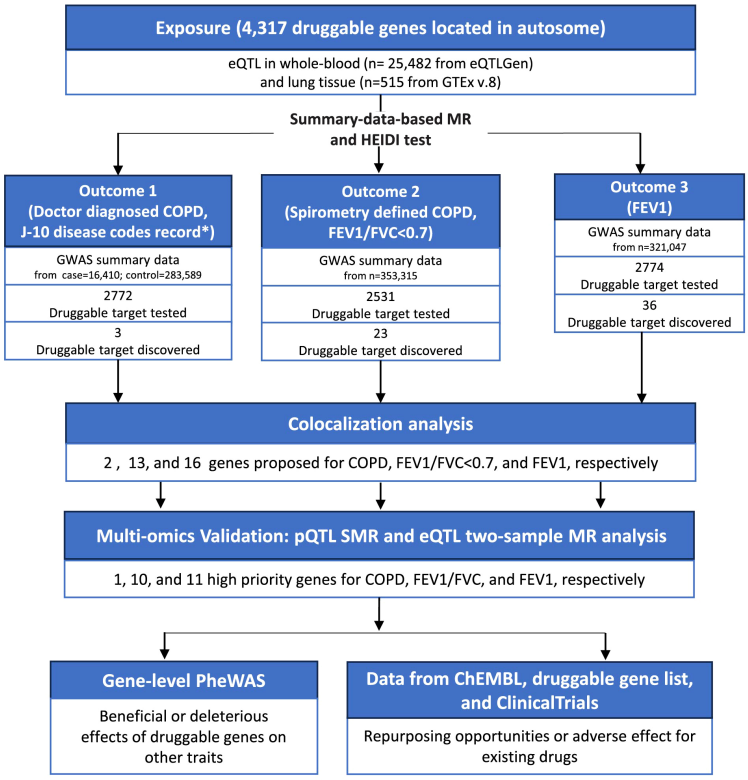



FIGURE 2
The flowchart of this study. COPD, chronic obstructive pulmonary disease; eQTL, expression quantitative trait loci; MR, Mendelian randomization; GWAS, genome-wide association studies; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; pQTL, protein quantitative trait loci; SMR, summary-data-based mendelian randomization; HEIDI, heterogeneity in dependent instruments; PheWAS, phenome-wide association studies.

European ancestry. Gene expression levels of the samples were analyzed by Illumina, Affymetrix expression arrays, and RNA-seq. The GTEx version 8 data contains lung tissue samples from 515 postmortem donors, the majority of whom were of European ethnicity. The mRNA from each tissue sample was sequenced at a median depth of 82.6 million reads. We Used eQTLGen and GTEx v.8 to identify all conditionally independent eQTLs in whole blood and lung tissue, which covered 2531 of the 4317 druggable genes in at least one tissue. The pQTL data for whole blood were obtained from Ferkingstad et al. study, which contained pQTLs for 4907 plasma proteins (Ferkingstad et al., 2021).

For the SMR analysis, we downloaded the cis-eQTL results from eQTLGen and GTEx v.8 consortia in SMR format in May 2023 (yanglab.westlake.edu.cn/software/smr). In the latest version, every SNP-gene combination with a distance <1Mb was included. To further validate the effect of druggable genes at the protein level, we extracted available cis-pQTLs from the study by Ferkingstad et al., considering those with a significance threshold of $P < 5e{-}8$.

For the two-sample MR analysis, the cis-eQTLs, which were defined as significant (5e-8) SNPs located within 5 kb upstream of the starting point or 5 kb downstream of the endpoint of a druggable gene, were extracted from the eQTLGen and GTEx v.8 consortia.

## Outcome data

FEV1 quantifies the maximal volume of air expelled in the first second of a forceful exhalation, serving as a critical parameter for evaluating pulmonary function. A reduced FEV1/FVC ratio (<0.7) typically indicates airflow obstruction (spirometry defined COPD). The doctor diagnosed COPD in this study comes from the J-10 disease codes record of the healthcare system. This diagnosis incorporates an assessment based on the patient's exposure history (such as smoking), symptoms (coughing, expectoration, or dyspnea), and post-bronchodilator pulmonary function tests. Conversely, spirometry-defined COPD (with a pre-bronchodilator FEV1/FVC<0.7) is identified during population health screenings. Participants meeting these conditions may not have any exposure history or symptoms (accounting for more than 70% of cases) (Fang et al., 2018) and did not undergo post-bronchodilator pulmonary function tests, making them a distinctly different cohort from those diagnosed with COPD by a doctor. Consequently, this study separately investigates the druggable genes associated with both doctor-diagnosed COPD and spirometry-defined COPD (pre-bronchodilator FEV1/FVC<0.7). All doctor-diagnosed COPD data were obtained from the FinnGen consortium (version R8, downloaded in May 2023), which comprises 16,410 patients with COPD and 283,589 controls (differentiated according to J-10 disease

codes in the medical electronic system). The cohorts for FEV1 and FEV1/FVC<0.7 were derived from the UK Biobank and SpiroMeta consortium. For further details regarding the data, please refer to the original literature and the official website (Supplementary Table 1) (Shrine et al., 2019; Higbee et al., 2021).

## Summary-data-based Mendelian randomization

The MR analysis is based on three fundamental assumptions: 1) genetic variation must be associated with the exposure; 2) genetic variation must not be associated with any confounding factors; and 3) genetic variation must not have a direct relationship with the outcome (Sanderson et al., 2022). As an extension of the MR concept, SMR employs two-step least squares (2SLS) to estimate the effect of the exposure on the outcome (e.g., doctor diagnosed COPD) using genetic variation that is significantly associated with the exposure (e.g., genes targetable for drug intervention at the transcriptional or protein expression level) (Sanderson et al., 2022). The causal associations were computed as follows:

$$\beta_{\text{druggable gene–COPD}} = \beta_{\text{SNP–COPD}} / \beta_{\text{SNP–druggable gene}}$$

The odds ratio (OR) estimates of the effect of druggable gene expression on COPD risk are calculated using the formula: $OR_{\text{druggable gene-COPD}}=\exp(\beta_{\text{druggable gene-COPD}})$, where OR represents the odds ratio estimate per 1-ln increase in the level of the druggable gene, and exp is the base of the natural logarithm.

SMR analyses were completed using the SMR software for Linux version 1.0.3 with default options (https://yanglab.westlake.edu.cn/software/smr). To account for multiple testing and control for genome-wide type I error rates, we used the Bonferroni correction to adjust the *P* values (Armstrong, 2014).

## HEIDI test

We conducted the HEIDI test to assess whether the observed associations were influenced by vertical pleiotropy rather than linkage disequilibrium (LD) of the dependent variable (Zhu et al., 2016). LD estimation was performed using the European pedigree genomes from the 1000 Genomes Project Consortium as a reference (Zheng-Bradley and Flicek, 2017). Associations with a significance level of $P < 0.01$ in the HEIDI test were excluded as they may be attributed to LD rather than pleiotropy.

## Colocalization

An alternative Bayesian test was performed to examine the colocalization of the two traits and estimate the posterior probability of shared variation (Giambartolomei et al., 2014). For each leading SNP in the outcome GWAS database, we retrieved all SNPs within a 100 kb range upstream and downstream. The posterior probability of H4 (PP.H4) was estimated, with the commonly used threshold of PP.H4 > 0.8 for GWAS and QTL associations (Navarro et al., 2021).

## Two-sample Mendelian randomization

To further validate the results of the SMR analysis, the two-sample MR approach was employed to estimate the effect on outcomes by using eQTL of druggable genes as instruments (Li et al., 2023). We assessed the strength of the correlation between instrumental variables and exposure using the F-statistic (>10) (Burgess and Thompson, 2011). In the presence of directional pleiotropy, we utilized the MR-egger estimated effect size as the outcome (Burgess and Thompson, 2017). The study proceeded with the two-sample MR analysis using MR-Egger, weighted median, inverse variance weighting, simple mode, and weighted mode methods. When multiple analytical methods yield similar results, we consider the MR results to be more robust. Cochran's Q test was used to assess the overall heterogeneity of Wald ratios (Higgins et al., 2003). In addition, the Steiger test was conducted to determine whether the exposure precedes the outcome (https://jean997.github.io/cause) (Hemani et al., 2017).

## Phenome-wide Mendelian randomization

To further assess the potential horizontal pleiotropy and any possible side effects of potential drug targets, we conducted a PheWAS analysis using the AstraZeneca portal (https://azphewas.com/). The original study utilized data from approximately 15,500 binary phenotypes and around 1,500 continuous phenotypes from about 450,000 participants who underwent exome sequencing, as published by the UK Biobank. The complete methodology can be found in the original article (Wang et al., 2021). We performed multiple corrections and set the threshold at 2E-9 to account for possible false positives.

All DNA positions are based on the human reference genome build hg19 (GRCh37). Data processing was carried out using R software version 4.2 with "TwosampeMR" and "easyMR" packages.

## Metagenomic next-generation sequencing (mNGS) pipeline

mNGS and hybridization capture-based targeted mNGS were used to detect pathogens. Briefly, about 1 mL of sample was centrifuged at 12,000 g for 5 min to collect the pathogens and human cells. Next, 50 µL of precipitate underwent depletion of host nucleic acid using 1 U of Benzonase (Sigma) and 0.5% Tween 20 (Sigma) and incubated at 37°C for 5 min. Subsequently, the DNA nucleic acid was extracted using a QIAamp UCP pathogen minikit. Qubit 4.0 (Thermo Fisher Scientific, MA, USA) was used to measure extracted DNA concentrations, which were used to construct metagenomics libraries. The inspected and qualified library was sequenced on the Nextseq 550 platform (Illumina, San

Diego, USA). For targeted NGS, the constructed library from each sample was used for hybrid capture-based enrichment of microbial probes. To remove adapter and low-quality, low complexity, and short reads of < 35 bp, raw data generated by the sequencing were filtered. The clean reads were blasted against a microbial database downloaded from the National Center for Biotechnology Information. Finally, microbial information at the species level can be obtained.

# Results

## Hospitalization and infections in patients with COPD

Recurrent acute exacerbations and hospitalization are crucial characteristics of severe COPD. We observed 56 patients with COPD admitted to the hospital from July 2020 to January 2024, most of whom were hospitalized with a combination of lung infections. During hospitalization, 94.6% of patients received antibiotics, 21.4% underwent mechanical ventilation, and 64.2% faced hospitalization costs over 20,000 RMB. Additionally, 51.8% were hospitalized multiple times during follow-up, with a mortality rate of 5.4% (Supplementary Table 2). The mNGS results suggested that *Pseudomonas aeruginosa* and Human gammaherpesvirus 4 were the most common bacterial and viral pathogens (Supplementary Figure 1).

## Instruments for druggable genes

We used eQTLs from whole blood and lung tissues to intersect with druggable genes to obtain druggable eQTLs. Finally, cis-eQTLs for 2529 druggable genes from eQTLGen's whole blood and 1042 druggable genes from GTEx v.8's lung tissues were extracted, both of which were used for SMR analyses with COPD and lung function as outcomes.

Among the genes identified by SMR/HEIDI and colocalization analysis, we obtained cis-pQTL data for 9 genes from the proteomics summary data of whole blood samples provided by Ferkingstad et al.

Our exposure and outcome samples were both predominantly European, so it can be assumed that the genetic variants used as instrumental variables (cis-eQTLs and cis-pQTLs) had a consistent role between the two different sample sets. The exposure and outcome samples were from completely independent studies, so there was no participant overlap.

## Druggable targets discovered by eQTL SMR and HEIDI test

We performed SMR analysis using doctor-diagnosed COPD GWAS summary data from the FinnGen consortium. Initially, we identified 3 genes (SLC22A5, C2, and GPC2) whose expression levels in whole blood were found to be correlated with doctor-diagnosed COPD (corrected for *P*-value by Bonferroni, *P*_HEIDI≥0.01, see Supplementary Table 3 for detailed data).

We utilized the FEV1/FVC<0.7 GWAS summary data from the UK biobank, and identified 23 genes whose expression levels in whole blood or lung tissues were associated with spirometry-defined COPD (corrected for *P*-value by Bonferroni, *P*_HEIDI≥0.01, see Supplementary Table 3 for detailed data).

The SMR analysis using FEV1 GWAS summary data identified 36 genes that were significantly associated with FEV1. Out of these genes, six were also found to be associated with spirometry-defined COPD.

## Druggable genes proposed after colocalization

Furthermore, colocalization analysis identified 2, 13, and 16 genes that shared the same SNPs (PP.H4 > 0.8 and/or PH4/PH3 > 2, Supplementary Table 4) in the results of doctor diagnosed COPD, FEV1/FVC<0.7 and FEV1, respectively (Figure 3).

## Druggable targets supported by eQTL two-sample MR

We performed eQTL two-sample MR (Table 1; Supplementary Table 5) on druggable genes that had passed the colocalization analysis to assess the consistency of direction and statistical significance.

## Druggable targets supported by protein SMR

We further performed protein SMR analysis (Table 1; Supplementary Table 5). Of the identified genes, 1 (GPC2) was associated with doctor-diagnosed COPD, 10 (TESK2, AKR1A1, LMCD1, ERBB3, PSMA4, MAST2, MMP15, MRC2, BMP4, and LTK) with FEV1/FVC <0.7, and 11 (APH1A, CHI3L2, PTK7, RASGRP3, HSP90AA1, TESK2, ADAM33, COL15A1, LINGO1, LTBR, and SULT1A1) with FEV1 in protein SMR results, without contradicting the expression SMR results. These genes were categorized as high priority.

In addition, we tested the underlying assumptions and sensitivity analyses for two-sample MR (Supplementary Table 6). The F-Statistic for the gene instrumental variables were all greater than 10. The Egger test indicated no evidence of horizontal pleiotropy for 16 genes and suggested that the IVW results could be used as the primary ones. The Steiger test confirmed the correct causal direction.

## Repurposing opportunities or adverse effects for existing drugs

Among the high-priority drug-targeted genes, MMP15 is associated with Montelukast and COL-3 (NSC-683551), while PSMA4 is linked to MARIZOMIB. These three drugs are small

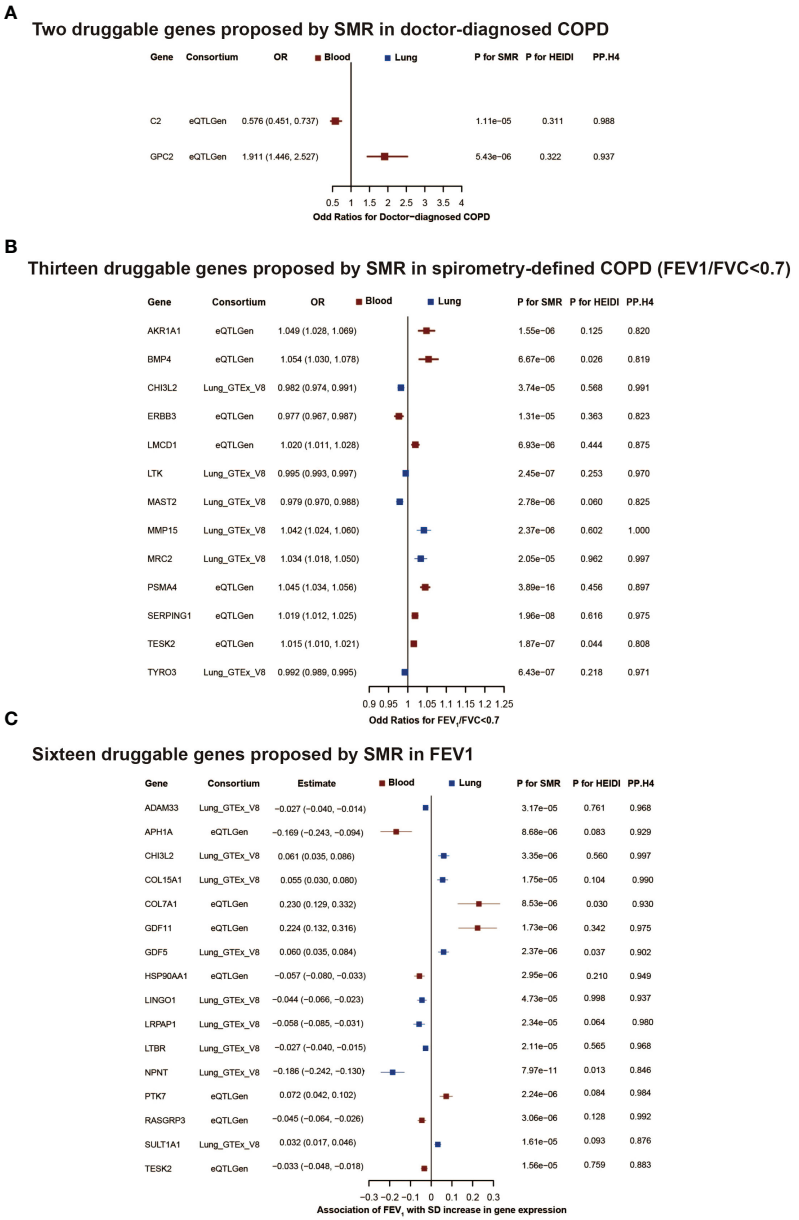



FIGURE 3
The druggable genes proposed by summary-data-based Mendelian randomization in COPD or lung function. The forest plot showed changes in doctor-diagnosed COPD risk, spirometry-defined COPD (FEV1/FVC<0.7) risk, and FEV1 level with an increase in gene expression per standard deviation. All genes passed eqtl SMR, HEIDI test, and colocalization analysis. The colors denote different tissues (red: blood, blue: brain tissue). In **(A, B)**, an OR<1 indicates a decrease in risk, while an OR>1 means an increase in risk. In **(C)**, a negative estimate corresponds to a decrease in FEV1, whereas a positive estimate denotes an increase in FEV1. COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; OR, odd ratios; SD, standard deviation; SMR, summary-data-based mendelian randomization.

molecule inhibitors, which may reduce the risk of spirometry-defined COPD. On the other hand, the APH1A gene, which codes for the gamma-secretase subunit APH-1A, appears to negatively impact FEV1 (β<0 in MR analysis). Several existing inhibitors against this target (e.g., TARENFLURBIL, BEGACESTAT, SEMAGACESTAT, NIROGACESTAT, RG-4733, and AVAGACESTAT) may counteract this harmful effect. Additionally, high expression of the ERBB3 gene may lower the risk of spirometry-defined COPD. However, various existing antitumor drugs, which are inhibitors of this target, may increase the risk of spirometry-defined COPD (Table 2).

## Phenome-wide MR evaluated side-effects of druggable genes

The PheWAS analysis revealed that increased GP9 expression may have adverse effects on platelets in addition to possibly

10.3389/fcimb.2024.1386506

TABLE 1 Multi-evidences supporting druggable genes whose expression was significantly associated with COPD and lung function.

| Gene | Tissue | SMR with HEIDI test | | Coloc analysis | Protein SMR with HEIDI test | | Two-sample MR | | Consortium | High priority |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Support | Direction | Support | Support | Direction | Support | Direction | | |
| Druggable genes whose expression was significantly associated with doctor-diagnosed COPD | | | | | | | | | | |
| C2 | Whole-blood | √ | – | √ | x | – | x | + | eQTLgen | x |
| GPC2 | Whole-blood | √ | + | √ | | | √ | + | eQTLgen | √ |
| Druggable genes whose expression was significantly associated with spirometry-defined COPD (FEV1/FVC<0.7) | | | | | | | | | | |
| TESK2 | Whole-blood | √ | + | √ | | | √ | + | eQTLGen | √ |
| AKR1A1 | Whole-blood | √ | + | √ | | | √ | + | eQTLGen | √ |
| LMCD1 | Whole-blood | √ | + | √ | | | √ | + | eQTLGen | √ |
| SERPING1 | Whole-blood | √ | + | √ | x | – | √ | + | eQTLGen | x |
| ERBB3 | Whole-blood | √ | – | √ | √ | – | √ | – | eQTLGen | √ |
| PSMA4 | Whole-blood | √ | + | √ | | | √ | + | eQTLGen | √ |
| MAST2 | Lung | √ | – | √ | | | √ | – | GTEx_V8 | √ |
| CHI3L2 | Lung | √ | – | √ | | | x | | GTEx_V8 | x |
| MMP15 | Lung | √ | + | √ | | | | | GTEx_V8 | √ |
| MRC2 | Lung | √ | + | √ | | | √ | + | GTEx_V8 | √ |
| BMP4 | Whole-blood | √ | + | √ | | | √ | + | eQTLGen | √ |
| LTK | Lung | √ | – | √ | | | √ | – | GTEx_V8 | √ |
| TYRO3 | Lung | √ | – | √ | x | + | √ | – | GTEx_V8 | x |
| Druggable genes whose expression was significantly associated with FEV1 | | | | | | | | | | |
| APH1A | Whole-blood | √ | – | √ | | | √ | – | eQTLgen | √ |
| CHI3L2 | Lung | √ | + | √ | | | √ | + | GTEx.v8 | √ |
| COL7A1 | Whole-blood | √ | + | √ | | | x | + | eQTLgen | x |
| GDF11 | Whole-blood | √ | + | √ | | | x | – | eQTLgen | x |
| GDF5 | Lung | √ | – | √ | | | x | + | GTEx.v8 | x |
| PTK7 | Whole-blood | √ | + | √ | √ | – | √ | + | eQTLgen | √ |
| RASGRP3 | Whole-blood | √ | – | √ | | | √ | – | eQTLgen | √ |
| HSP90AA1 | Whole-blood | √ | – | √ | | | √ | – | eQTLgen | √ |
| TESK2 | Whole-blood | √ | – | √ | | | √ | – | eQTLgen | √ |
| ADAM33 | Lung | √ | – | √ | | | √ | – | GTEx.v8 | √ |
| COL15A1 | Lung | √ | + | √ | √ | + | √ | + | GTEx.v8 | √ |
| LINGO1 | Lung | √ | – | √ | √ | – | √ | – | GTEx.v8 | √ |
| LRPAP1 | Lung | √ | – | √ | x | – | √ | – | GTEx.v8 | X |
| LTBR | Lung | √ | – | √ | | | √ | – | GTEx.v8 | √ |
| NPNT | Lung | √ | – | √ | x | + | √ | – | GTEx.v8 | X |
| SULT1A1 | Lung | √ | + | √ | | | √ | + | GTEx.v8 | √ |

COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; √, pass; x, fail; +, positive; -, negative; blank not possible to test; SMR, summary-data-based mendelian randomization; HEIDI test, heterogeneity in dependent instruments; Coloc, colocalization; MR, Mendelian randomization; High priority means there are no contradictory analysis results.

TABLE 2 Repurposing opportunities or adverse effects for existing drugs on COPD and lung function.

| Gene | Expression in tissue | Effect of gene | Outcome | Gene coding protein | Compound (drug) name | Action type | Molecule type | Current indication | Clinical phase | Evidence from |
|---|---|---|---|---|---|---|---|---|---|---|
| Repurposing opportunities for existing drugs | | | | | | | | | | |
| PSMA4 | Whole-blood | Increase risk | FEV1/FVC<0.7 | Proteasome subunit alpha type-4 | MARIZOMIB | Inhibitor | Small molecule | Cancer | Phase 3 | ChEMBL |
| MMP15 | Lung | Increase risk | FEV1/FVC<0.7 | Matrix metalloproteinase-15 | COL-3 (NSC-683551) | Inhibitor | Single protein | Cancer | Phase 1 | Clinicaltrials.gov |
| | | | | | Montelukast | Inhibitor | Small molecule | Asthma | Approved | Clinicaltrials.gov |
| APH1A | Whole-blood | Impair | FEV1 | Gamma-secretase subunit APH-1A | TARENFLURBIL | Inhibitor | Small molecule | Alzheimer Disease, Dementia | Phase 3 | ChEMBL |
| | | | | | BEGACESTAT | Inhibitor | Small molecule | Alzheimer Disease | Phase 1 | ChEMBL |
| | | | | | SEMAGACESTAT | Inhibitor | Small molecule | Alzheimer Disease | Phase 3 | ChEMBL |
| | | | | | NIROGACESTAT | Inhibitor | Small molecule | Cancer | Phase 3 | ChEMBL |
| | | | | | RG-4733 | Inhibitor | Small molecule | Cancer | Phase 2 | ChEMBL |
| | | | | | AVAGACESTAT | Inhibitor | Small molecule | Alzheimer Disease | Phase 2 | ChEMBL |
| LTBR | Lung | Impair | FEV1 | Tumor necrosis factor receptor superfamily member 3 | HCBE-11 | Inhibitor | Antibody | Cancer | Phase 1 | ChEMBL |
| Adverse effects of existing drugs | | | | | | | | | | |
| ERBB3 | Whole-blood | Decrease risk | FEV1/FVC<0.7 | Receptor tyrosine-protein kinase erbB-3 | CDX-3379 | Inhibitor | Antibody | Cancer | Phase 2 | ChEMBL |
| | | | | | MM-121 | Antagonist | Antibody | NSCLC and other cancer | Phase 2 | ChEMBL |
| | | | | | ELGEMTUMAB | Inhibitor | Antibody | Cancer | Phase 1 | ChEMBL |
| | | | | | DULIGOTUZUMAB | Inhibitor | Antibody | Cancer | Phase 2 | ChEMBL |
| | | | | | SAPITINIB | Inhibitor | Small molecule | NSCLC and other cancer | Phase 2 | ChEMBL |
| | | | | | LUMRETUZUMAB | Inhibitor | Antibody | NSCLC and other cancer | Phase 1 | ChEMBL |
| | | | | | AMG-888 | Inhibitor | Antibody | NSCLC and other cancer | Phase 2 | ChEMBL |

*(Continued)*





10.3389/fcimb.2024.1386506

10.3389/fcimb.2024.1386506

TABLE 2 Continued

| Gene | Expression in tissue | Effect of gene | Outcome | Gene coding protein | Compound (drug) name | Action type | Molecule type | Current indication | Clinical phase | Evidence from |
|---|---|---|---|---|---|---|---|---|---|---|
| Adverse effects of existing drugs | | | | | | | | | | |
| | | | | | SERIBANTUMAB | Inhibitor | Antibody | Cancer | Phase 2 | ChEMBL |
| | | | | | AV-203 | Inhibitor | Antibody | Cancer | Phase 1 | ChEMBL |
| | | | | | PATRITUMAB-DERUXTECAN | Binding-Agent | Antibody | NSCLC and other cancer | Phase 3 | ChEMBL |
| | | | | | PATRITUMAB | Inhibitor | Antibody | NSCLC and other cancer | Phase 3 | ChEMBL |
| | | | | | ISTIRATUMAB | Inhibitor | Antibody | Cancer | Phase 2 | ChEMBL |
| COL15A1 | Lung | improve | FEV1 | | OCRIPLASMIN | Inhibitor | Enzyme | Cancer | Approved | ChEMBL |
| | | | | | COLLAGENASE CLOSTRIDIUM HISTOLYTICUM | Inhibitor | Enzyme | Cancer | Approved | ChEMBL |

Repurposing opportunities or adverse effect data from druggable gene list, ClinicalTrials (https://www.ClinicalTrials.gov), or ChEMBL release.
COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity.

increasing the risk of spirometry-defined COPD. Conversely, increased ERBB3 expression is associated with a reduced risk of spirometry-defined COPD and inflammation suppression. Moreover, increased MRC2 expression affects cardio metabolism and the nervous system (Figure 4).

## Discussion

The real-world observational study emphasized the urgency and necessity of developing new drugs for COPD. In this study, we aimed to uncover druggable genes and explore drug repurposing possibilities for COPD and lung function. Through a comprehensive multi-omics approach, which encompassed SMR analysis with eQTL and pQTL data, colocalization analyses, and two-sample MR (MR Egger, Weighted Median, and Inverse Variance Weighted), we successfully identified and validated 22 potential druggable genes. Among these genes, MMP15, APH1A, LTBR, and PSMA4 emerged as potential targets for existing drugs. The potential effects of these targets were also investigated using PheWAS analysis. To our knowledge, this is the first study to use MR methods to identify drug targets for COPD and lung function, utilizing the largest publicly available QTL data and COPD GWAS data to date.

Montelukast and COL-3 (NSC-683551) have been identified as inhibitors of MMP expression, including MMP15, which plays a crucial role in extracellular matrix breakdown relevant to lung function. Previous studies have revealed the association of the MMP15 gene with lung function and its specific expression in multiple lung cell types (Soler Artigas et al., 2011; Chun et al., 2022). Moreover, several studies have linked various MMP family members, including MMP-9, MMP-12, and MMP-15, to COPD through the degradation of the alveolar extracellular matrix (Babusyte et al., 2007; Gharib et al., 2018; Wells et al., 2018; Zhou et al., 2020). Our findings suggest that inhibiting MMP15 could be an effective strategy to prevent spirometry-defined COPD. Montelukast, a common inhibitor of allergic reactions, has been widely prescribed for the treatment of asthma and chronic cough (von Mutius and Drazen, 2010; Wang et al., 2014). Since controlling asthma and airway hyperresponsiveness can help prevent spirometry-defined COPD (Couillard et al., 2023), it's likely that Montelukast could reduce the risk of spirometry-defined COPD by decreasing MMP levels in patients. On the other hand, COL-3, despite its broad inhibitory effects on MMPs, is limited in clinical use due to side effects like photosensitization. This highlights the need for maintaining a balance between efficacy and safety when considering strategies for lung function protection.

MARIZOMIB is a small molecule inhibitor that targets PSMA4. Our results suggest that PSMA4 expression increases the risk of spirometry-defined COPD. The significant association between PSMA and severe COPD has been confirmed by a genome-wide association study, which underscored the crucial role of PSMA4 in both whole blood and lung tissue (Sakornsakolpat et al., 2018). Additionally, extensive research investigating genetic variants on chromosome 15q25.1 in relation to COPD has consistently identified PSMA4 as a gene exhibiting substantial associations

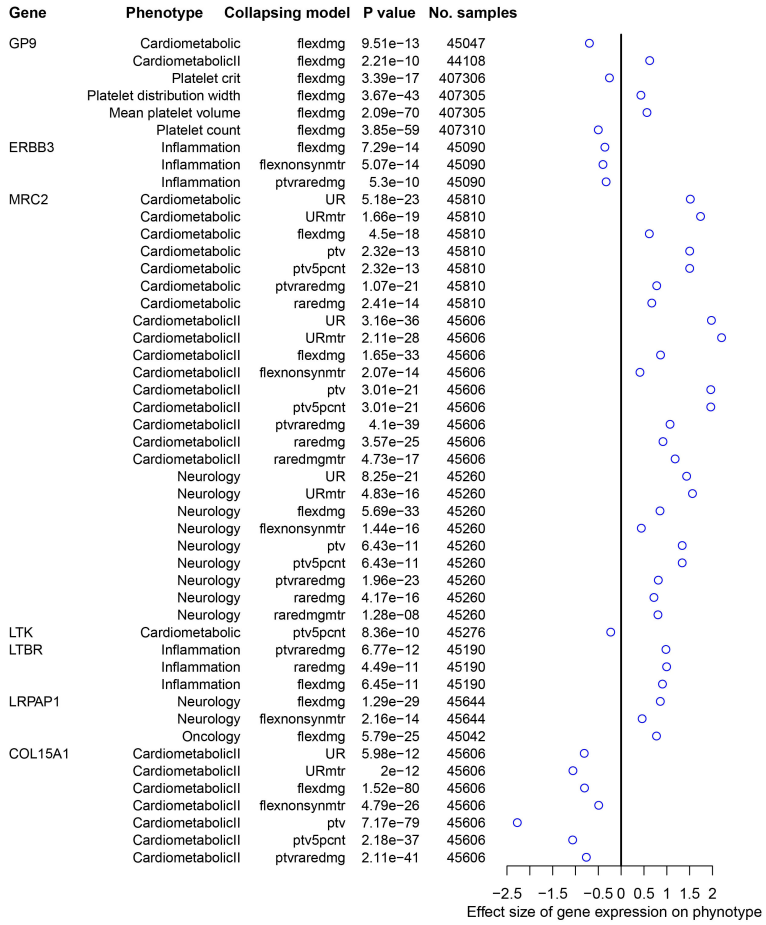



FIGURE 4
Traits significantly associated with druggable genes using AstraZeneca PheWAS portal. The effect size represents the extent of the impact of genetic variation on the phenotype, with a positive value indicating a positive correlation and a negative value indicating a negative correlation. PheWAS, phenome-wide association study.

with methylation differences in this genomic region (Nedeljkovic et al., 2018). Such findings are in alignment with our results. However, the potential clinical application of PSMA4 inhibitors, which are currently used in oncology, for COPD treatment remains to be explored.

Our analysis revealed a negative causal association between increased levels of ERBB3 expression in blood, indicating that the expression of this gene could act as a protective factor against spirometry-defined COPD. Although studies on the association of the ERBB3 gene with COPD are limited, one study has shown that mRNA levels of ERBB4 increase progressively from non-smokers to non-COPD smokers and then to COPD patients and is positively correlated with airflow obstruction severity (Anagnostis et al., 2013). This suggests that the ERBB receptors may contribute to the development or progression of COPD. However, an increase in ERBB3 expression does not necessarily indicate that the gene promotes the development of COPD. Instead, it may be a protective response of the organism to control the level of inflammation. Phenome-wide MR analysis in this study suggests that increased ERBB3 expression is associated with the suppression of inflammation. Evidence has shown that ERBB signaling can inhibit the production of TNF-α induced by LPS in immune activation associated with chronic systolic heart failure (Ryzhov et al., 2017). Therefore, the relationship between ERBB3 and COPD, as well as its mechanism of action, such as its inhibition of inflammation, requires further verification.

An association between increased levels of LMCD1 expression in the blood and an increased risk of spirometry-defined COPD was revealed in this study. It has been shown that LMCD1 plays an important role in the development of lung fibrosis and affects the properties of lung myofibroblasts (Jiang et al., 2021). In particular, in cases of systemic sclerosis-associated lung fibrosis, LMCD1 interacts with serum response factors in lung fibroblasts, which leads to increased contractile activity of lung myofibroblasts. This suggests that LMCD1 is a pro-fibrotic molecule that contributes to myofibroblast activation and sustained fibrotic proliferation (Bogatkevich et al., 2023). Lung fibrosis and lung myofibroblasts also play important roles in the pathophysiology of COPD, and it is reasonable to hypothesize that LMCD1 could potentially contribute to lung tissue fibrosis, thereby increasing the risk of spirometry-defined COPD.

Our study has also discovered several previously unreported genes associated with COPD or lung function. Specifically, we found that increased expression of APH1A and RASGRP3 is likely to lead to decreased FEV1, while increased expression of CHI3L2 is associated with increased FEV1. Besides, heightened

expression of GPC2 is linked to an increased risk of COPD, while increased expression of TESK2, AKR1A1, and MRC2 may result in an increased risk of spirometry-defined COPD. Increased expression of MAST2 appears to lower the risk of spirometry-defined COPD. Although no previous reports have been found on the direct association of these genes with COPD or lung function, they have well-defined roles in various other conditions, such as tumors, Alzheimer's disease, and inflammation. For example, GPC2 has been identified as a prognostic marker for several types of tumors (Liu et al., 2021; Chen et al., 2022), PTK7 shows potential as a target for CAR T-cell therapies in lung cancer treatment (Ma et al., 2023), APH1A (Todd et al., 2022) is involved in the development of Alzheimer's disease, and RasGRP3 acts to limit the inflammatory response during low-intensity pathogen infections (Lee et al., 2023). While our findings provide new insights into the potential involvement of these genes in COPD development, further studies are needed to determine their exact role in the disease.

Our study has several limitations. Firstly, while MR offers insights into causality, it presumes a linear relationship between exposure and outcome, which may not capture non-linear or U-shaped exposure-response relationships. Secondly, the QTLs used in our study may only show small differences in gene expression levels, which might fail to fully capture the gene's potential effects. Furthermore, the QTLs and FEV1 data included some individuals of non-European ancestry, whereas the Doctor-diagnosed COPD and FEV1/FVC<0.7 GWAS populations consisted of Europeans only. These differences in population may introduce potential bias in MR effect estimates due to differences in genetic background and chain imbalance patterns. Finally, this study mainly focused on European populations, which limits the generalizability of the findings to other ethnic groups. Therefore, further research and validation are necessary to generalize the results to other ethnicities.

Our study has identified potential therapeutic targets for COPD. In the future, well-established drugs like Montelukast, which targets the MMP15 gene, and Marizomib, targeting the PSMA4 gene, could be prioritized for clinical trials. However, the disease-modifying potential of many druggable genes requires further experimental validation.

## Conclusion

This study found and validated 22 potential druggable genes that show promise for COPD and lung function. Our findings provide genetic evidence supporting the potential therapeutic benefits of targeting these genes in the treatment of COPD. Clinical trials prioritizing existing drugs and novel medications targeting these identified druggable genes could potentially increase the likelihood of successful treatments.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Author contributions

ZW: Conceptualization, Data curation, Writing – original draft, Writing – review & editing, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization. SL: Conceptualization, Data curation, Writing – original draft, Writing – review & editing, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization. GC: Formal analysis, Writing – original draft, Writing – review & editing. YG: Formal analysis, Writing – original draft, Writing – review & editing. HY: Data curation, Writing – original draft, Writing – review & editing. YL: Data curation, Writing – original draft, Writing – review & editing. JL: Formal analysis, Writing – original draft, Writing – review & editing. SZ: Formal analysis, Writing – original draft, Writing – review & editing. JH: Formal analysis, Writing – original draft, Writing – review & editing, Funding acquisition. JZ: Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing, Resources, Supervision.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was sponsored by the Foundation of Guangzhou National Laboratory (SRPG22 018, SRPG22 016) and Guangzhou Science and Technology Plans (No. 202201020513). However, study funding had no influence on the study.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1386506/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Pathogen profiles of COPD patients with hospitalization.

# References

Adeloye, D., Song, P., Zhu, Y., Campbell, H., Sheikh, A., and Rudan, I. (2022). Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. *Lancet Respir. Med.* 10, 447–458. doi: 10.1016/S2213-2600(21)00511-7

Aguet, F., Anand, S., Ardlie, K. G., Gabriel, S., Getz, G. A., and Graubert, A. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 369, 1318–1330. doi: 10.1126/science.aaz1776

Anagnostis, A., Neofytou, E., Soulitzis, N., Kampas, D., Drositis, I., Dermitzaki, D., et al. (2013). Molecular profiling of EGFR family in chronic obstructive pulmonary disease: correlation with airway obstruction. *Eur. J. Clin. Invest.* 43, 1299–1306. doi: 10.1111/eci.12178

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic Physiol. Opt* 34, 502–508. doi: 10.1111/opo.12131

Babusyte, A., Stravinskaite, K., Jeroch, J., Lötvall, J., Sakalauskas, R., and Sitkauskiene, B. (2007). Patterns of airway inflammation and MMP-12 expression in smokers and ex-smokers with COPD. *Respir. Res.* 8, 81. doi: 10.1186/1465-9921-8-81

Bogatkevich, G. S., Atanelishvili, I., Bogatkevich, A. M., and Silver, R. M. (2023). Critical role of LMCD1 in promoting profibrotic characteristics of lung myofibroblasts in experimental and scleroderma-associated lung fibrosis. *Arthritis Rheumatol* 75, 438–448. doi: 10.1002/art.42344

Burgess, S., and Thompson, S. G. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* 40, 755–764. doi: 10.1093/ije/dyr036

Burgess, S., and Thompson, S. G. (2017). Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* 32, 377–389. doi: 10.1007/s10654-017-0255-x

Chen, G., Luo, D., Zhong, N., Li, D., Zheng, J., Liao, H., et al. (2022). GPC2 is a potential diagnostic, immunological, and prognostic biomarker in pan-cancer. *Front. Immunol.* 13, 857308. doi: 10.3389/fimmu.2022.857308

Chun, S., Akle, S., Teodosiadis, A., Cade, B. E., Wang, H., Sofer, T., et al. (2022). Leveraging pleiotropy to discover and interpret GWAS results for sleep-associated traits. *PloS Genet.* 18, e1010557. doi: 10.1371/journal.pgen.1010557

Couillard, S., Petousi, N., Smigiel, K. S., and Molfino, N. A. (2023). Toward a predict and prevent approach in obstructive airway diseases. *J. Allergy Clin. Immunol. Pract.* 11, 704–712. doi: 10.1016/j.jaip.2023.01.008

Fang, L., Gao, P., Bao, H., Tang, X., Wang, B., Feng, Y., et al. (2018). Chronic obstructive pulmonary disease in China: a nationwide prevalence study. *Lancet Respir. Med.* 6, 421–430. doi: 10.1016/S2213-2600(18)30103-6

Ferkingstad, E., Sulem, P., Atlason, B. A., Sveinbjornsson, G., Magnusson, M. I., Styrmisdottir, E. L., et al. (2021). Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* 53, 1712–1721. doi: 10.1038/s41588-021-00978-w

Finan, C., Gaulton, A., Kruger, F. A., Lumbers, R. T., Shah, T., Engmann, J., et al. (2017). The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* 9, 383. doi: 10.1126/scitranslmed.aag1166

Folkersen, L., Gustafsson, S., Wang, Q., Hansen, D. H., Hedman Å, K., Schork, A., et al. (2020). Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat. Metab.* 2, 1135–1148. doi: 10.1038/s42255-020-00287-2

Gaziano, L., Giambartolomei, C., Pereira, A. C., Gaulton, A., Posner, D. C., Swanson, S. A., et al. (2021). Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19. *Nat. Med.* 27, 668–676. doi: 10.1038/s41591-021-01310-z

Gharib, S. A., Manicone, A. M., and Parks, W. C. (2018). Matrix metalloproteinases in emphysema. *Matrix Biol.* 73, 34–51. doi: 10.1016/j.matbio.2018.01.018

Giambartolomei, C., Vukcevic, D., SChadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., et al. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PloS Genet.* 10, e1004383. doi: 10.1371/journal.pgen.1004383

Hemani, G., Tilling, K., and Davey Smith, G. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PloS Genet.* 13, e1007081. doi: 10.1371/journal.pgen.1007081

Higbee, D. H., Granell, R., Sanderson, E., Davey Smith, G., and Dodd, J. W. (2021). Lung function and cardiovascular disease: a two-sample Mendelian randomisation study. *Eur. Respir. J.* 58, 3. doi: 10.1183/13993003.03196-2020

Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj* 327, 557–560. doi: 10.1136/bmj.327.7414.557

Jiang, D., Dey, T., and Liu, G. (2021). Recent developments in the pathobiology of lung myofibroblasts. *Expert Rev. Respir. Med.* 15, 239–247. doi: 10.1080/17476348.2021.1829972

Lee, J. H., Kim, Y. S., and Leem, K. H. (2023). Citri reticulatae pericarpium limits TLR-4-triggered inflammatory response in raw264.7 macrophages by activating rasgrp3. *Int. J. Mol. Sci.* 24, 18. doi: 10.3390/ijms241813777

Li, Y., Sundquist, K., Zhang, N., Wang, X., Sundquist, J., and Memon, A. A. (2023). Mitochondrial related genome-wide Mendelian randomization identifies putatively causal genes for multiple cancer types. *EBioMedicine* 88, 104432. doi: 10.1016/j.ebiom.2022.104432

Liu, L., Yang, Y., Duan, H., He, J., Sun, L., Hu, W., et al. (2021). CHI3L2 is a novel prognostic biomarker and correlated with immune infiltrates in gliomas. *Front. Oncol.* 11, 611038. doi: 10.3389/fonc.2021.611038

Ma, H. Y., Das, J., Prendergast, C., De Jong, D., Braumuller, B., Paily, J., et al. (2023). Advances in CAR T cell therapy for non-small cell lung cancer. *Curr. Issues Mol. Biol.* 45, 9019–9038. doi: 10.3390/cimb45110566

Navarro, E., Udine, E., de Paiva Lopes, K., Parks, M., Riboldi, G., Schilder, B. M., et al. (2021). Dysregulation of mitochondrial and proteolysosomal genes in Parkinson's disease myeloid cells. *Nat. Aging* 1, 850–863. doi: 10.1038/s43587-021-00110-x

Nedeljkovic, I., Carnero-Montoro, E., Lahousse, L., van der Plaat, D. A., de Jong, K., Vonk, J. M., et al. (2018). Understanding the role of the chromosome 15q25.1 in COPD through epigenetics and transcriptomics. *Eur. J. Hum. Genet.* 26, 709–722. doi: 10.1038/s41431-017-0089-8

Ou, Y. N., Yang, Y. X., Deng, Y. T., Zhang, C., Hu, H., Wu, B. S., et al. (2021). Identification of novel drug targets for Alzheimer's disease by integrating genetics and proteomes from brain and blood. *Mol. Psychiatry* 26, 6065–6073. doi: 10.1038/s41380-021-01251-6

Ryzhov, S., Matafonov, A., Galindo, C. L., Zhang, Q., Tran, T. L., Lenihan, D. J., et al. (2017). ERBB signaling attenuates proinflammatory activation of nonclassical monocytes. *Am. J. Physiol. Heart Circ. Physiol.* 312, H907–h918. doi: 10.1152/ajpheart.00486.2016

Sakornsakolpat, P., Morrow, J. D., Castaldi, P. J., Hersh, C. P., Bossé, Y., Silverman, E. K., et al. (2018). Integrative genomics identifies new genes associated with severe COPD and emphysema. *Respir. Res.* 19, 46. doi: 10.1186/s12931-018-0744-9

Sanderson, E., Glymour, M. M., Holmes, M. V., Kang, H., Morrison, J., Munafò, M. R., et al. (2022). Mendelian randomization. *Nat. Rev. Methods Primers* 2, 6. doi: 10.1038/s43586-021-00092-5

Schmidt, A. F., Finan, C., Gordillo-Marañón, M., Asselbergs, F. W., Freitag, D. F., Patel, R. S., et al. (2020). Genetic drug target validation using Mendelian randomisation. *Nat. Commun.* 11, 3255. doi: 10.1038/s41467-020-16969-0

S. European Observatory on Health, Policies, Panteli, D., Polin, K., Webb, E., Allin, S., Barnes, A., et al. (2023). Health and care data: approaches to data linkage for evidence-informed policy. *Health Syst. Transit.* 2, 1–248.

Shrine, N., Guyatt, A. L., Erzurumluoglu, A. M., Jackson, V. E., Hobbs, B. D., Melbourne, C. A., et al. (2019). New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat. Genet.* 51, 481–493. doi: 10.1038/s41588-018-0321-7

Silverman, E. K. (2020). Genetics of COPD. *Annu. Rev. Physiol.* 82, 413–431. doi: 10.1146/annurev-physiol-021317-121224

Soler Artigas, M., Loth, D. W., Wain, L. V., Gharib, S. A., Obeidat, M., Tang, W., et al. (2011). Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat. Genet.* 43, 1082–1090. doi: 10.1038/ng.941

Stoekenbroek, R. M., Kastelein, J. J., and Huijgen, R. (2015). PCSK9 inhibition: the way forward in the treatment of dyslipidemia. *BMC Med.* 13, 258. doi: 10.1186/s12916-015-0503-4

Storm, C. S., Kia, D. A., Almramhi, M. M., Bandres-Ciga, S., Finan, C., Hingorani, A. D., et al. (2021). Finding genetically-supported drug targets for Parkinson's disease using Mendelian randomization of the druggable genome. *Nat. Commun.* 12, 7342. doi: 10.1038/s41467-021-26280-1

Todd, N. K., Huang, Y., Lee, J. Y., Doruker, P., Krieger, J. M., Salisbury, R., et al. (2022). GPCR kinases generate an APH1A phosphorylation barcode to regulate amyloid-β generation. *Cell Rep.* 40, 111110. doi: 10.1016/j.celrep.2022.111110

Trajanoska, K., Bhérer, C., Taliun, D., Zhou, S., Richards, J. B., and Mooser, V. (2023). From target discovery to clinical drug development with human genetics. *Nature* 620, 737–745. doi: 10.1038/s41586-023-06388-8

von Mutius, E., and Drazen, J. M. (2010). Choosing asthma step-up care. *N. Engl. J. Med.* 362, 1042–1043. doi: 10.1056/NEJMe1002058

Võsa, U., Claringbould, A., Westra, H. J., Bonder, M. J., Deelen, P., Zeng, B., et al. (2021). Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* 53, 1300–1310. doi: 10.1038/s41588-021-00913-z

Wang, K., Birring, S. S., Taylor, K., Fry, N. K., Hay, A. D., Moore, M., et al. (2014). Montelukast for postinfectious cough in adults: a double-blind randomised placebo-controlled trial. *Lancet Respir. Med.* 2, 35–43. doi: 10.1016/S2213-2600(13)70245-5

Wang, Q., Dhindsa, R. S., Carss, K., Harper, A. R., Nag, A., Tachmazidou, I., et al. (2021). Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature* 597, 527–532. doi: 10.1038/s41586-021-03855-y

Wells, J. M., Parker, M. M., Oster, R. A., Bowler, R. P., Dransfield, M. T., Bhatt, S. P., et al. (2018). Elevated circulating MMP-9 is linked to increased COPD exacerbation risk in SPIROMICS and COPDGene. *JCI Insight* 3, 22. doi: 10.1172/jci.insight.123614

Zheng-Bradley, X., and Flicek, P. (2017). Applications of the 1000 genomes project resources. *Briefings Funct. Genomics* 16, 163–170. doi: 10.1093/bfgp/elw027

Zhou, J. S., Li, Z. Y., Xu, X. C., Zhao, Y., Wang, Y., Chen, H. P., et al. (2020). Cigarette smoke-initiated autoimmunity facilitates sensitisation to elastin-induced COPD-like pathologies in mice. *Eur. Respir. J.* 56, 3. doi: 10.1183/13993003.00404-2020

Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., et al. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* 48, 481–487. doi: 10.1038/ng.3538

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 28 June 2024
DOI 10.3389/fcimb.2024.1409774



OPEN ACCESS

EDITED BY
Edwin Kamau,
Tripler Army Medical Center, United States

REVIEWED BY
António Machado,
Universidad San Francisco de Quito, Ecuador
Addy Cecilia Helguera-Repetto,
Instituto Nacional de Perinatología (INPER), Mexico

*CORRESPONDENCE
Ayodeji B. Oyenihi
aoyenihi@mdlab.com
Jason Trama
jtrama@ibr-genetics.com
John Osei Sekyere
joseisekyere@mdlab.com

†These authors have contributed equally to this work

RECEIVED 31 March 2024
ACCEPTED 07 June 2024
PUBLISHED 28 June 2024

CITATION
Oyenihi AB, Haines R, Trama J, Faro S, Mordechai E, Adelson ME and Osei Sekyere J (2024) Molecular characterization of vaginal microbiota using a new 22-species qRT-PCR test to achieve a relative-abundance and species-based diagnosis of bacterial vaginosis.
*Front. Cell. Infect. Microbiol.* 14:1409774.
doi: 10.3389/fcimb.2024.1409774



# Molecular characterization of vaginal microbiota using a new 22-species qRT-PCR test to achieve a relative-abundance and species-based diagnosis of bacterial vaginosis

Ayodeji B. Oyenihi[1*†], Ronald Haines[1†], Jason Trama[1*], Sebastian Faro[1,2], Eli Mordechai[1], Martin E. Adelson[1] and John Osei Sekyere[1*]

[1]Institute for Biomarker Research, Medical Diagnostic Laboratories, Genesis Biotechnology Group, Hamilton, NJ, United States, [2]Memorial Women's Care, Houston, TX, United States

**Background:** Numerous bacteria are involved in the etiology of bacterial vaginosis (BV). Yet, current tests only focus on a select few. We therefore designed a new test targeting 22 BV-relevant species.

**Methods:** Using 946 stored vaginal samples, a new qPCR test that quantitatively identifies 22 bacterial species was designed. The distribution and relative abundance of each species, α- and β-diversities, correlation, and species co-existence were determined per sample. A diagnostic index was modeled from the data, trained, and tested to classify samples into BV-positive, BV-negative, or transitional BV.

**Results:** The qPCR test identified all 22 targeted species with 95 – 100% sensitivity and specificity within 8 hours (from sample reception). Across most samples, *Lactobacillus iners, Lactobacillus crispatus, Lactobacillus jensenii, Gardnerella vaginalis, Fannyhessea (Atopobium) vaginae, Prevotella bivia*, and *Megasphaera* sp. type 1 were relatively abundant. BVAB-1 was more abundant and distributed than BVAB-2 and BVAB-3. No *Mycoplasma genitalium* was found. The inter-sample similarity was very low, and correlations existed between key species, which were used to model, train, and test a diagnostic index: *MDL-BV index*. The *MDL-BV index*, using both species and relative abundance markers, classified samples into three vaginal microbiome states. Testing this index on our samples, 491 were BV-positive, 318 were BV-negative, and 137 were transitional BV. Although important differences in BV status were observed between different age groups, races, and pregnancy status, they were statistically insignificant.

**Conclusion:** Using a diverse and large number of vaginal samples from different races and age groups, including pregnant women, the new qRT-PCR test and *MDL-BV index* efficiently diagnosed BV within 8 hours (from sample reception), using 22 BV-associated species.

KEYWORDS

**bacterial vaginosis (BV), *MDL-BV index*, vaginitis, qRT-PCR, BVAB, vaginal microbiome, machine learning**

## 1 Introduction

Bacterial Vaginosis (BV), a condition in which the normal *Lactobacillus*-rich vaginal microbiome becomes dominated by polymicrobial anaerobic bacterial species under non-acidic pH, remains the most common cause of abnormal vaginal discharge in reproductive-age women, with an estimated prevalence rate of 29% in North America (Abou Chacra et al., 2021; Elnaggar et al., 2023). The normal vaginal microbiome is dominated by three major *Lactobacillus* species, *L. crispatus, L. jensenii*, and *L. gasseri*, with low abundance of *L. acidophilus*, which protects the vagina by producing lactic acid, hydrogen peroxide, and bacteriocins that suppress bacterial growth (Pacha-Herrera et al., 2022; Wu et al., 2022). *Lactobacillus iners*, however, is quite enigmatic as it occurs in both healthy and unhealthy vaginal environments; *L. iners*, unlike *L. crispatus, L. jensenii*, and *L. gasseri*, produces the human non-functional L form of lactic acid (Basavaprabhu et al., 2020).

To produce lactic acid, which is necessary to maintain an acidic pH of < 4.5, *Lactobacillus* sp. uses glycogen deposited unto the vaginal walls from the epithelial cells (Shen et al., 2022; Navarro et al., 2023). The acidic pH keeps away most microbial species and allows *Lactobacillus* sp. to proliferate, keeping the species diversity in the normal vaginal microbiota low. Lactic acid is metabolized from glycogen, whose production and deposition are mediated by circulating estrogen in a dose-dependent manner. In the absence of glycogen or lactic-acid–producing species, the pH rises, allowing other microbes to proliferate, colonize the vagina, and increase the microbiota diversity. Hence, physiological factors disrupting or increasing estrogen levels indirectly affect *Lactobacillus* sp. and BV-associated bacterial species' prevalence (Shen et al., 2022; Navarro et al., 2023). The etiology and pathogenesis of BV are still not fully understood, making it important to study the involved microbes for timely diagnosis and treatment. Particularly, the interactions between species within the BV microbiome that cause pathologies, treatment failure and recurrence, or healing need further interrogations to enable a species biomarker-based diagnosis of BV.

The colonization of the vaginal microbiota by anaerobic bacterial and fungal species causes BV or vulvovaginal candidiasis, respectively, characterized by increased vaginal discharge with a fishy odor (Swidsinski et al., 2023). Dysbiosis also predisposes the vagina to sequelae of other sexually transmitted infections (STIs) and obstetric disorders. Although BV is asymptomatic in many women, it is associated with the development of common obstetric and gynecologic complications (Javed et al., 2019; Cheng et al., 2020; Gustin et al., 2021). It also increases the risk of contracting HIV and other STIs and pelvic inflammatory disease (PID) (Unemo et al., 2017). Notably, BV recurrence after treatment is common (estimated to affect 50% of women annually) (Hilbert et al., 2016; Coudray and Madhivanan, 2020; Coudray et al., 2020) and may be due to re-infections from sexual partners or failure of current treatment options (Coudray et al., 2020; Vodstrcil et al., 2021). Besides the health impacts of recurrence, BV treatment costs are increasing annually, specifically in the USA: this economic impact is particularly pronounced in BV-associated preterm births and other obstetric complications (Peebles et al., 2019; Watkins et al., 2024).

Species such as *Gardnerella vaginalis, Fannyhessea vaginae, Prevotella bivia, Megasphaera* sp., and BVAB are commonly implicated in the etiology of BV, with *G. vaginalis* and *F. vaginae* being common in most BV infections (Swidsinski et al., 2023) as sessiles. Swidsinksi et al. (2023) observed that vaginal epithelial cells from females with BV were covered with *G. vaginalis* biofilms that encapsulated other species, forming a polymicrobial biofilm, as well as by non-adherent planktonic bacterial cells. The biofilms associated with BV explain their ability to escape both the immune response and antimicrobial chemotherapy as biofilms prevent the immune cells and antimicrobial agents from reaching the pathogens (Cangui-Panchi et al., 2023; Swidsinski et al., 2023). This leads to high treatment failure and recurrence rates of about 50% (Muñoz-Barreno et al., 2021); indeed, the ability of these biofilm-coated polymicrobial species to spread into the uterus and fallopian tube explains their resistance to antimicrobials and the immune response (Swidsinski et al., 2023).

To diagnose BV, clinicians rely mostly on the classical clinical signs and symptoms outlined in Amsel's criteria (Amsel et al., 1983) or on the microscopically based Nugent score (Nugent et al., 1991). While these standard diagnostic methods have been effective over the years, they have been confronted with limitations such as being labor-intensive and time-consuming, subjective and unable to accurately identify pathogens (Muzny et al., 2023). As recently discussed by Muzny et al. (2023), nucleic acid tests ease the burden of laboriously going through the clinical testing criteria, saving

clinicians time and effort for other duties. The development of culture-independent molecular diagnostics has enabled the detection of non-cultivable bacterial species associated with BV and continues to revolutionize infectious disease diagnosis (Hilbert et al., 2016; van den Munckhof et al., 2019; Bhujel et al., 2021). Molecular tests such as real-time polymerase chain reaction (RT-PCR), combined with the traditional clinical diagnostic criteria, have greatly improved the sensitivity and specificity of detecting BV pathogens (Hilbert et al., 2016; van den Munckhof et al., 2019; Bhujel et al., 2021). They are also applicable in monitoring patient response to antibiotic therapy (Onderdonk et al., 2016; Elnaggar et al., 2023). This approach has proven more useful for identifying patients at risk for recurrent BV (Jones, 2019; Coudray and Madhivanan, 2020; Vodstrcil et al., 2021; Wu et al., 2022; Sobel and Vempati, 2024; Watkins et al., 2024).

Besides next-generation sequencing (NGS)-based metagenomics that target all genomes, current PCR diagnostics for BV mainly focus on a smaller spectrum of bacterial species. Hence, to increase the resolution and diagnostic power of PCR-based BV diagnostics (Balashov et al., 2014), we designed a new 22-species quantitative Real-Time PCR assay (qRT-PCR) that quantitatively detects 22 vaginal bacterial species within 8 hours (from sample reception). This newly designed proprietary qRT-PCR assay (Bacterial Vaginosis (with *Lactobacillus* Profiling) Panel®, Medical Diagnostic Laboratories, L.L.C. (MDL), New Jersey, USA) was used to screen 946 vaginal samples routinely obtained from different health centers across the United States. We further used the results, with assistance from machine-learning algorithms (Decision Trees and Random Forests) to design, train, and test an index (herein termed the *MDL-BV index*) that used a relative-abundance and species-based markers to classify the vaginal microbiome in three categories: BV-negative (healthy microbiome), transitional BV (between healthy and unhealthy vaginal microbiome), and BV-positive (abnormal microbiome).

# 2 Materials and methods

## 2.1 Specimen collection and processing

Clinical vaginal samples are routinely obtained from different healthcare centers across the United States for diagnostic processing at MDL. Historical vaginal specimens (n = 946) marked for disposal were received in *One*Swab® (Copan Diagnostics, CA, USA) or *Thin*Prep® (Hologic, MA, USA) transport media in a Clinical Laboratory Improvement Amendments (CLIA)-certified infectious disease laboratory facility between January and June 2023 and stored at −80°C were selected randomly for this study. This included specimens from symptomatic, asymptomatic, pregnant, or non-pregnant females and from whom vaginal profiling was requested. For each biological specimen that arrives at the laboratory facility, specimen accessioning occurs that assigns a random MDL number to ensure the de-identification of specimens. To further the de-identification of specimens during this study, samples that matched our collection criteria were randomized and the MDL numbers associated with each sample were not recorded.

## 2.2 Targeted bacterial species

The newly designed quantitative real-time PCR (qRT-PCR) assay (Bacterial Vaginosis (with *Lactobacillus* Profiling) Panel®) is a BV diagnostic assay designed by MDL to qualitatively and quantitatively detect 22 bacterial species that are found in the eubiotic and dysbiotic vaginal microbiome. These species are *L. crispatus, L. jensenii, L. gasseri, L. iners, L. acidophilus, Gardnerella vaginalis, Fannyhessea vaginae* (*Atopobium vaginae*), *Megasphaera* sp. types 1 and 2, *Prevotella bivia*, Bacterial Vaginosis-Associated Bacterium (BVAB) 1–3, *Ureaplasma urealyticum, Mycoplasma hominis, Mycoplasma genitalium, Mobiluncus curtisii, Mobiluncus mulieris, Sneathia sanguinegens, Bifidobacterium breve, Bacteroides fragilis*, and *Streptococcus anginosus* (Onderdonk et al., 2016; Muzny et al., 2020; Mondal et al., 2023; Powell et al., 2023).

## 2.3 DNA preparation

DNA from the vaginal specimens was extracted according to validated in-house laboratory protocols using QIAamp® DNA Mini Kit (QIAGEN, Maryland, USA) and the X-tractor Gene® DNA workstation (QIAGEN, Maryland, USA) with slight modifications. Canine Herpes Virus DNA was spiked into the samples and subsequently detected as an internal extraction control.

To serve as amplification standards, plasmids for each of the 22 bacterial species were generated by whole-gene synthesis (Genewiz®, Azenta Life Sciences, Waltham, USA). Briefly, a species-specific region of each species was amplified, synthesized, and cloned into the pUC-GW-AMP plasmid vector with an ampicillin-resistance marker. These plasmids were transformed into chemically competent *Escherichia coli* cells (One Shot™ TOP10, ThermoFisher Scientific, New Jersey, USA) and grown in liquid Lysogeny Broth (LB) containing 50 µg/mL Ampicillin. Plasmid DNA was isolated from overnight cultures using Wizard® Plus SV Minipreps DNA Purification Systems (Promega, Wisconsin, USA) according to the manufacturer's protocol. Extracted DNA was subsequently quantified spectrophotometrically using NanoDrop 1000 equipment (ThermoFisher Scientific, New Jersey, USA). Standard 10-fold serial dilutions of $10^8$ to $10^1$ DNA copies/µL were prepared for each bacterial species.

## 2.4 Identification of bacterial species by multiplex qRT-PCR

Each vaginal and plasmid control DNA sample was amplified in triplicates using species-specific primers and fluorescent probes on a CFX384 Touch Real-Time PCR Detection System (Bio-Rad, California, USA). The specific primer and probe sequences, as well as the multiplex RT-PCR conditions in the Bacterial Vaginosis (with *Lactobacillus* Profiling) Panel by Real-Time PCR assay are proprietary to MDL. An aliquot of 0.5µL of DNA was used as the template in a 4µL total PCR reaction mixture containing four primer sets, four probes, and the in-house—prepared mastermix of

Taq DNA polymerase and dNTPs (MDL, New Jersey, USA) in multiplex PCR reactions. Human β-globulin DNA served as the internal control while no-template control was included to account for potential extraneous nucleic acid contamination. After amplification, a standard curve (fluorescence vs cycle number) was generated for each species and the target DNA copy for each sample was extrapolated. A $10^2$ copies/μL concentration in each sample was established as the positive cutoff for a given bacterium. The relative percentage concentration of all the bacterial species identified in each vaginal specimen was then computed and tabulated (Supplementary Table S1).

## 2.5 Relative abundance- and species-based classification of BV

The mean, median, standard deviation, and concentration distribution of each species across samples as well as alpha-diversity, Shannon index, beta-diversity (Bray-Curtis Dissimilarity matrix), and relative abundance were calculated from the gDNA concentration data (obtained from the qRT-PCR amplification results of each species from the samples) and represented as box and Whisker plots, histograms, heatmaps, and charts. Using the distribution patterns in the data, a relative abundance or concentration-based cut-off criteria was established and used to define BV-positive (abnormal vaginal microbiota with bacterial vaginosis and depleted *Lactobacillus* sp.: ≤ 40% relative abundance), BV-negative (i.e., normal vaginal microbiota with ≥ 70% *Lactobacillus* sp. relative abundance), and transitional BV (transition between negative and positive BV microbiota with 30–50% *Lactobacillus* sp. relative abundance) status (see Table 1 for full definitions). Using the species-species correlation and co-existence data as well as previous research from literature, a species-based criteria was also established for determining BV-positive, BV-negative, and transitional BV status. A two-tier system was then created by synthesizing the species-based and concentration-based BV diagnosis criteria, which were used to interpret the PCR results. This criterion was called the *MDL-BV index*, which was further trained and tested on the qRT-PCR and demographics data using machine-learning algorithms (Decision Trees and Random Forests) in Python. The datasets were divided into two: one for training and one for testing. Python libraries such as numpy, pandas, scikitlearn, XGBoost, and matplotlib, were used for analysis and visualization.

## 2.6 Demographics

The age, race, and pregnancy status of the females from whom the samples were taken were collected and used to stratify the data to compare the different rates, distributions, and relative abundance of the various species, BV-positive, BV-negative, and transitional BV according to age, race, and pregnancy status.

## 2.7 Statistical analysis

Statistical analyses were performed with Prism version 10.1.2 (Graph-Pad, California USA) or the R statistical package. Quantitative data obtained in this study were expressed as the mean ± standard error of the mean (SEM). To detect bacterial associations and correlations, Pearson's correlation coefficient was chosen. The Kruskal–Wallis test was employed to determine any statistically significant differences in age distributions among BV classification groups. The data were parsed through Python (and Biopython) to calculate the significance of the coexistence between any two species using Chi-square and T-test.

Three types of groupings were used in this (Chi-square and T-test) analysis: (1) all *Lactobacillus* species were grouped into one and their absence/presence was compared with the absence/presence of each non-*Lactobacillus* species; (2) all *Lactobacillus* species except *L. iners* were grouped into one and their absence/presence was compared with the absence/presence of each non-*Lactobacillus* species; (3) the presence/absence of each *Lactobacillus* species was compared with that of each non-*Lactobacillus* species. The resulting data were tabulated in Supplementary Table S2, and

TABLE 1 *MDL-BV index* designed to interpret the new 22 species qRT-PCR results and diagnose vaginal samples as BV-positive, BV-negative, or transitional BV.

| BV Status# | Relative abundance (%) | | | |
|---|---|---|---|---|
| | *Lactobacillus* sp. | BVAB (1, -2, -3), *S. anginosus*, *B. fragilis* | *G. vaginalis*, *P. bivia*, *A. vaginalis*, *Megasphaera* sp. *1* | Megasphaera sp. 2, *U. urealyticum*, *Mycoplasma* sp., *Mobiluncus* sp., *S. sanguinegens*, *B. breve* |
| **Healthy microbiome (BV negative)** | ≥ 70 | ≤3 | ≤ 20 | ≤ 20 |
| **Transitional BV** | 30 – 60 | 0 - 10 | ≤ 30 | ≤ 50 |
| **BV positive** | ≤ 40 | 0 - 100 | 20 – 100 | 0 – 100 |

# (1) When a sample falls within both BV-negative and transitional BV, assign it to transitional BV. (2) When a sample falls within both BV-positive and transitional BV, assign it to BV-positive. (3) When at least three of a sample's four groups/biomarkers' relative abundances fall within a single BV state and only one falls in another BV state, override the single disagreement, and assign or classify the sample to the BV state with which the three relative abundances agree or fall into. (4) When biomarkers 2 (BVAB group) and 4 (Megasphaera sp. 2 group) have a relative abundance of zero, then a relative abundance of 0 – 30% for Group 1 and 70 – 100% for Group 3 is BV-positive; a relative abundance of 30 – 50% (Group 1) is transitional BV; and 60–100% (Group 1) is BV negative. (5) When every group is zero and only group 4 is 100%, define it as BV positive. (6) When all groups are zero, filter out that sample as there are no results to report

the significant results were filtered out and used to generate heat maps. The results were only considered significant if $P < 0.05$.

# 3 Results

## 3.1 Demographics

The vaginal samples were obtained from 946 women whose average age was 34.8 years, with a standard deviation of 13.3 years. The ages ranged from 18 to 84 years. The median age was 32 years; hence, the age distribution is somewhat skewed towards younger women [Figure 1; Supplementary Table S1 (Demographics)]. Fifty-three women reported as pregnant while the rest were either not pregnant or unknown. Race data was obtained for 245 samples: White (n = 145 samples), Black (n = 79), "Other race" (n = 15), Asian (n = 4), and Native American (n = 2) [Figure 1; Supplementary Table S1 (Demographics)]. The "Other race" category refers to those not falling within any of the four races above: Hispanic/Latino, Native Hawaiian or Pacific Islander, and mixed races.

The age distribution of the races [Figure 1; Supplementary Table S1 (Demographics)] shows a broader distribution of ages within the White population (with outliers) than that of the other races, albeit the median ages across all races fell within a narrow range of 28 – 35 years. The age distribution of White people ranged from 18 – 83 years, with 50% falling within 28 – 40 years. Whereas the ages of Black people ranged from 18 to 55 years, 50% were within the same 28 – 40 years range. The "Other races", which includes Hispanics/Latinos, had ages between 28 and 32 years and the median age was different among all the races [Figure 1; Supplementary Table S1 (Demographics)].

## 3.2 Identification and distribution of species across samples

The BV qRT-PCR assay efficiently identified the targeted 22 species with 95 – 100% sensitivity and specificity. Except for *M. genitalium* which was not detected in any sample, all the other species were present in at least one of the 946 samples. The percentage of the count of the identified species across all the samples are as follows: *L. iners* (75%), *G. vaginalis* (65%), *F. vaginae* (53%), *Megasphaera* sp. type 1 (44%), and *P. bivia* (41%) (Figure 2). These are followed by *L. crispatus* (29%), *L. jensenii* (23%), BVAB-1 (17%), *U. urealyticum* (13%), *L. gasseri* (12%), *M. hominis* (12%), *M. curtisii* (11%), and BVAB-3 (10%). The remaining species had less than 10% relative abundance across all samples combined (Figure 2).

To understand the distribution (spread) of the species and their concentrations across the samples, we used bar charts, histograms, and box and Whisker plots. *L. iners*, followed by *G. vaginalis, F. vaginae, Megasphaera* sp. type 1, *P. bivia, and L. crispatus* had better concentration distributions across the samples than the other species, with their median concentrations ($10^2$ and $10^6$ gDNA copies/μL) being found in more than 100 samples. Indeed, other species were found in low concentrations across all the samples (Supplementary Figures S1–S22). The relative concentration for

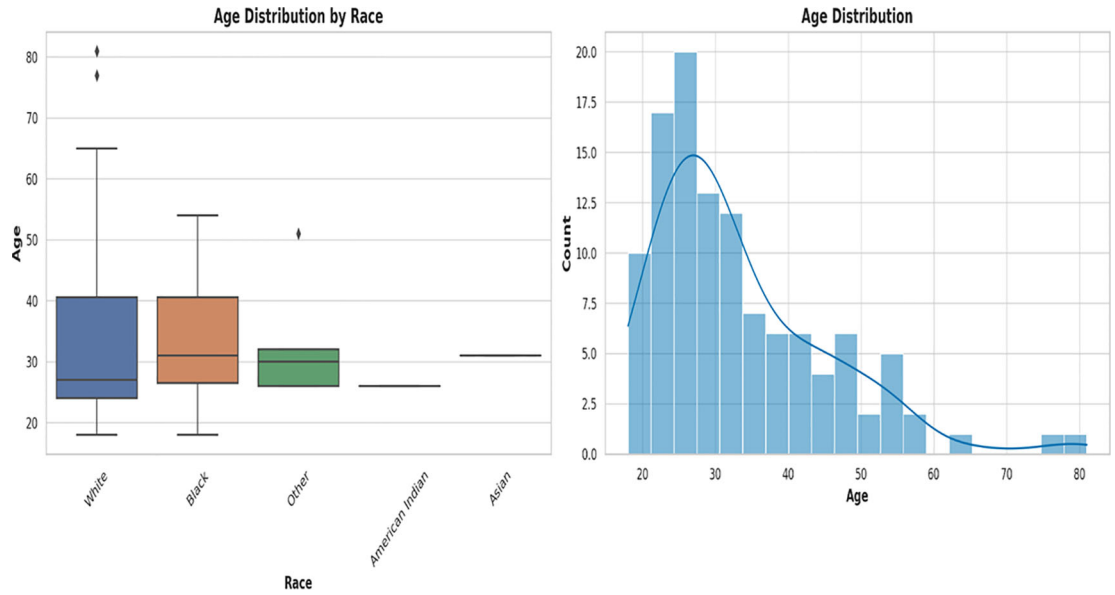



FIGURE 1
Box plot and histogram showing the age distribution among the races from which the vaginal samples were obtained. The box plot on the left, showing the age distribution among White people, Black people, Other (races), Native American or Alaska Natives, and Asians. White people were the largest population, followed by Black people. Other races are those who belong to none of the four races. The median ages among the races were very close within 28 and 35. An age distribution histogram on the right, shows the overall age distribution of participants in the study. The histogram includes a kernel density estimate (KDE) to show the smooth distribution of ages. The data is somewhat right-skewed, indicating a younger population with fewer older women. Most women fell into the 20–40 age range, with a peak around the late 20s to early 30s.

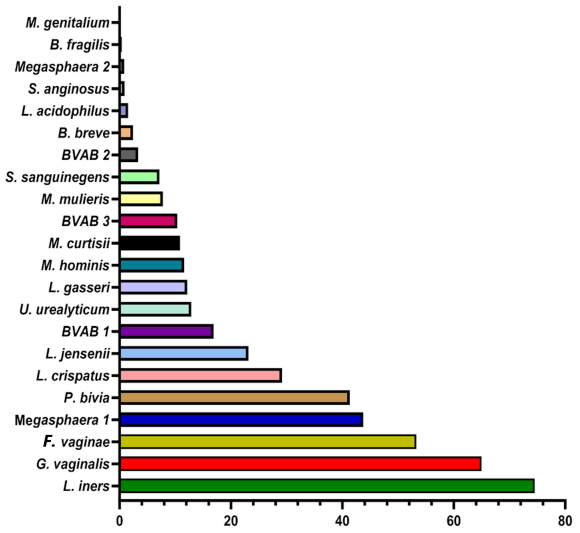



FIGURE 2
Percentage of the count of each species across all vaginal specimens (n = 946) as detected by qPCR. Among the 946 samples, it is observed that the most dominant species are *Lactobacillus iners, Gardnerella vaginalis, Atopobium (Fannyhessea) vaginae, Megasphaera* sp. type 1, *Prevotella bivia, Lactobacillus crispatus, and Lactobacillus jensenii*. The horizontal axis is the percentage of each species across all samples. No *Mycoplasma genitalium* was detected in any of all the samples.

each identified species ranged between $10^2$ and $10^8$ genomic DNA copies/μL and the median range was between $10^2$ and $10^6$ copies/μL (Figure 3).

The concentrations of the *Lactobacillus* species detected in this study were in the order, *L. iners* > *L. crispatus* > *L. jensenii* > *L. gasseri* > *L. acidophilus* (Figure 2). Additionally, *F. vaginae* and *G. vaginalis* occurred in similar amounts (~$10^5$ copies/μL) and were slightly more than the values obtained for *P. bivia* ($10^4$ copies/μL). BVAB-1 was the most predominant of the BVABs, and its relative concentration was 44% and 31% more than BVAB-3 and BVAB-2, respectively. Notably, over 98% of the *Megasphaera* species identified in this study belonged to type 1 with a mean concentration of ~$10^6$ copies/μL. Of the Mycoplasmas, *U. urealyticum* was found in higher concentrations than *M. hominis*. Although *M. curtisii* was 39% more abundant, it occurred in lower concentrations in vaginal swabs than *M. mulieris*.

## 3.3 Relative abundance, α- and β-diversities

We further investigated the per-sample species relative abundance and richness as well as inter-sample species diversity using alpha and beta diversities. The relative abundance of the species perfectly mirrored their concentration distribution across

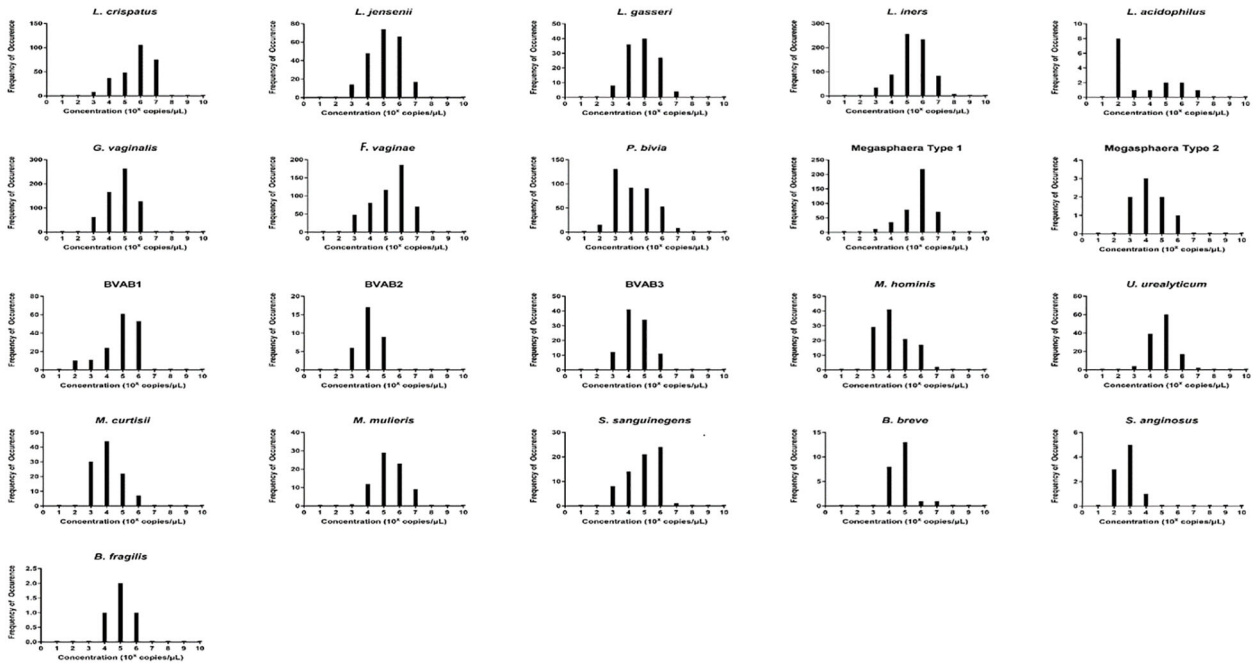



FIGURE 3
The relative concentration distribution of the identified bacterial species in vaginal swabs (n = 946). Each chart represents the concentration distribution per species; *M. genitalium* is not shown as was not detected in any of the samples. The vertical axis shows the count (frequency) of samples while the concentrations are shown on the horizontal axis. The median concentrations ($10^2$ and $10^6$ gDNA copies/μL) of *Lactobacillus iners* and *L. crispatus, Gardnerella vaginalis, Atopobium (Fannyhessea) vaginae, Prevotella bivia*, and *Megasphaera* sp. type 1 were found in more than 100 samples. The charts also do not show the count of samples with zero concentrations but only counts of samples with higher concentrations.

the samples (Figure 3; Supplementary Figures S1–S22), with *L. iners, L. jensenii, P. bivia, G. vaginalis, Megasphaera* sp. type 1, and *F. vaginae* being more abundant in many samples than the other species. The relative abundance, shown as a heatmap and a box plot (Figure 4), represents the abundance of each species across samples, normalized relative to the total concentration of microbial species within each sample. However, the other less abundant species also presented greater variability in relative abundance across the samples, as shown in the outliers (black stars) and absence of boxes (Figure 4B).

The alpha diversity of the species within each sample was determined using species richness and Shannon diversity index (Figure 5). The species richness shows that 50% of the samples had 2 – 6 species, with 8 – 12 species occurring in 100 – 300 samples (Figures 5A, C). The Shannon index, which ranges from 0 (no diversity) to 5 (practically 3.5, i.e., most diverse), showed that 50% of the samples' diversity index was between 0.7 and 1.8, with 200 – 280 samples having a diversity of > 2.0 (Figures 5B, D).

Inter-sample (β-) diversity, using the Bray-Curtis dissimilarity index and associated principal component analysis (PCoA), showed that a few of the samples shared strong similarities (Figure 6). Samples with close similarity in species diversity are shown as blue (purple) while those with little similarity are shown as yellow. The abundance of yellow in the matrix shows how different most of the samples are from one another (Figure 6A). The PCoA chart also reflected this, with a few samples clustering together but most of the samples were spatially separated (Figure 6B).

## 3.4 Correlation and co-existence of species

We performed correlation, Chi-square and T-test analyses of the data to determine the significance of the species-species co-existence in the samples (Supplementary Table S2). A heat map of the significant results was generated for easy visualization of the data (Figure 7; Supplementary Table S2). Unlike the the Chi-square test (Figures 7A–C), the T-test provided a significant association between the non-*Lactobacillus* species (BVAB-1 was significantly affected by the presence/absence of *L. iners* only but not by the other *Lactobacillus* species)–Figure 7D and the presence/absence of the *Lactobacillus* species (without *L. iners*–Figure 7E or as a group–Figure 7F). In the absence of *L. iners*, the *Lactobacillus* species were not significantly associated with *B. fragilis* (which was only significantly associated with *L. gasseri*) but rather, with *Megasphaera* sp. type 2 (Figure 7B). Hence, the Chi-square test provided a more stringent cut-off than the T-test: individually, the *Lactobacillus* species were significantly associated with 14 species (Figure 7C).

Therefore, the presence of BVAB-1 and *B. fragilis* were independent of *L. iners*. Furthermore, *L. crispatus* and *L. jensenii*

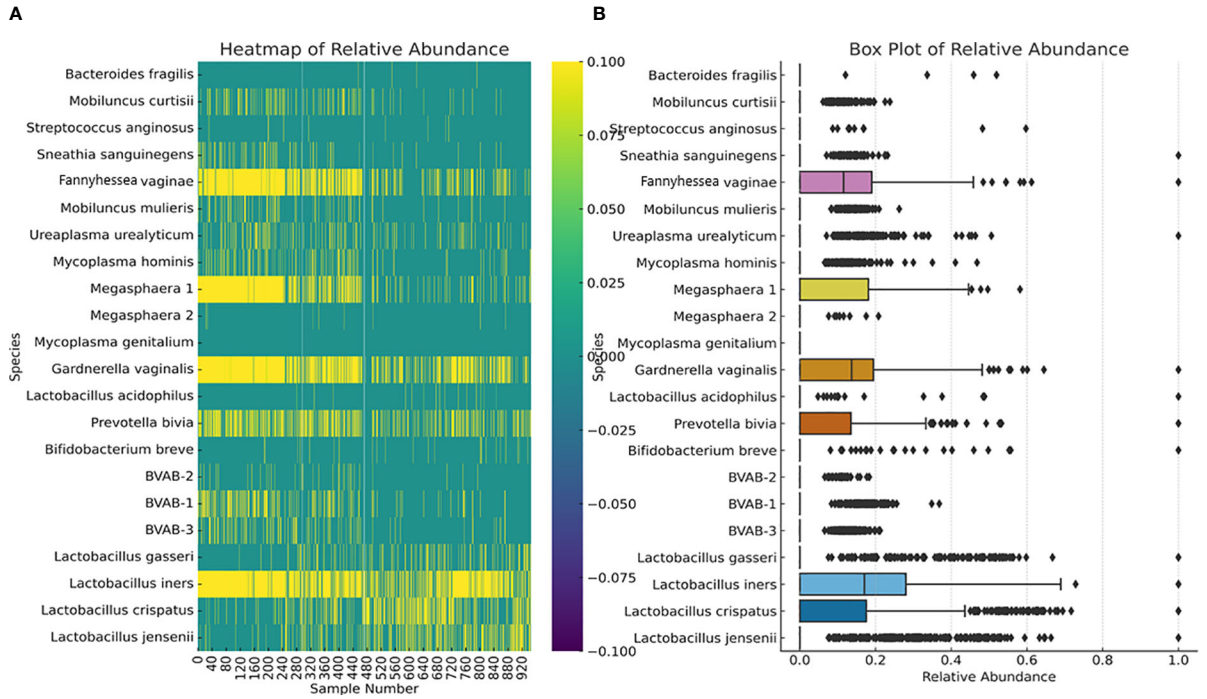



FIGURE 4
Relative abundance of the 22 species across all the 946 vaginal samples. The heatmap **(A)** highlights the distribution and intensity of species' relative abundance across all samples, offering a color-coded representation that easily identifies the most prevalent species. The box plot **(B)** provides a statistical summary of each species' relative abundance, including the median, interquartile range, and any outliers, offering insight into the variability and distribution of abundance for each species. This plot provides insights into the central tendency, spread, and outliers for the relative abundance of each microbial species within the dataset. Median values are represented by the line within each box. Interquartile range (IQR), indicating the middle 50% of the data, is shown by the box itself. Whisker extend to show the range of the data, i.e., 1.5 * IQR from the quartiles. Outliers are points outside the Whisker and are indicated as individual points.

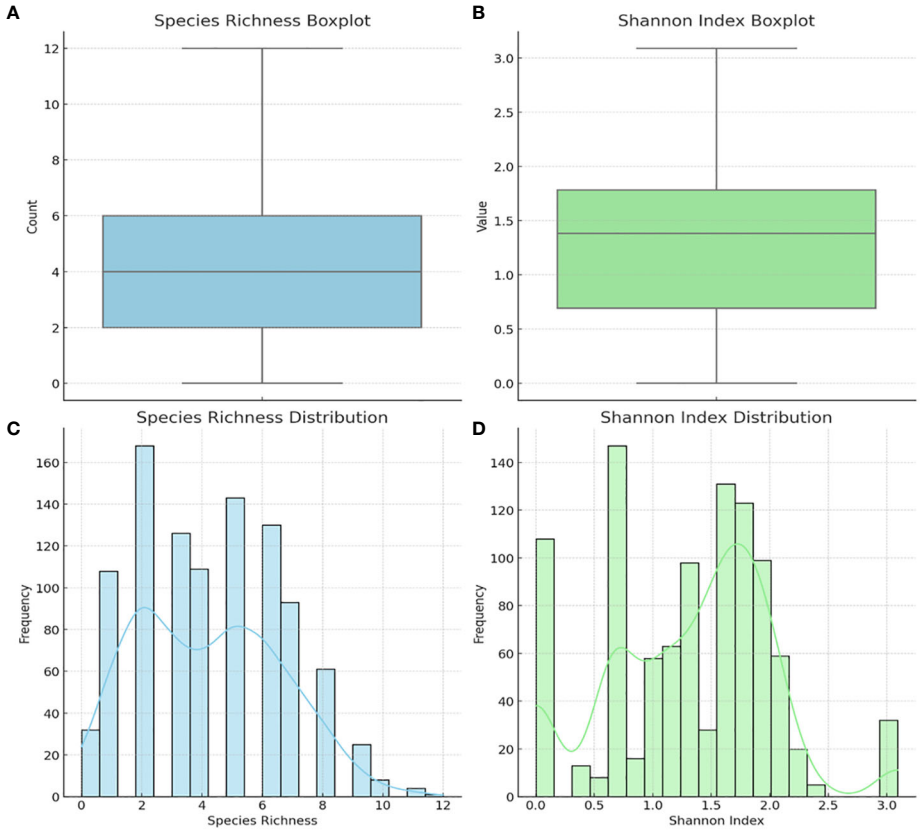



FIGURE 5
Alpha diversity of the 946 vaginal samples: species richness and Shannon index. Boxplot of Species Richness **(A)**, showing the count of species present across samples. This boxplot provides a clear visualization of the central tendency and variability in species richness. Boxplot of the Shannon Index **(B)**, indicating the diversity value that considers both abundance and evenness of species. Histogram of Species Richness Distribution **(C)**, providing a view of the frequency distribution of species counts across samples. Histogram of Shannon Index **(D)** visualizes the distribution of the Shannon index across samples, reflecting both the abundance and evenness of species. The Shannon index is a more comprehensive measure of diversity, considering not just the presence of species but also their relative abundances.

were significantly associated with the presence/absence of the same non-*Lactobacillus* species except *B. fragilis, B. breve, P. bivia*, and *U. urealyticum* (Figure 7). Although *L. acidophilus* was included in the Chi-square pairwise association test, it did not yield any significant results with any of the species (Supplementary Table S2).

Using Pearson's correlation coefficient, a clear pattern was observed regarding the co-existence of the 21 species within the vaginal microbiota: all the *Lactobacillus species*, except *L. iners*, inversely correlated with most of the other non-*Lactobacillus* species while species such as BVAB (1–3), *Megasphaera* sp. type 1, *P. bivia, G. vaginalis, F. vaginae, M. hominis, M. mulieris, M. curtisii*, and *S. sanguinegens* were positively correlated with each other. Notably, some species had very little or no inverse correlation with the *Lactobacillus* sp. (except *L. iners*): *S. anginosus, B. breve, U. urealyticum, M. hominis, Megasphaera* sp. type 2, and *B. fragilis. S. anginosus, B. breve*, and *B. fragilis* were the only non-*Lactobacillus* species with an inverse correlation with *L. iners;* the other species had a positive correlation with *L. iners*. The BVAB species did not correlate with each other as BVAB -2 was less correlated with both BVAB-1 and BVAB-3; these latter two species, however, had a strong positive correlation (Supplementary Figure S23).

## 3.5 Relative abundance- and species-based diagnostic criteria

All the *Lactobacillus* sp. were bundled together as a marker of a normal vaginal microbiome (Group 1). Owing to the absence of a Nugent score or Amsel data for the samples, we used a species-based criteria to select species that are not found in normal vaginal microbiota: BVAB-1, -2, -3, *B. fragilis*, and *S. anginosus*. These five species were grouped into a species marker (Group 2) to identify BV-positive samples (Table 1). Owing to the strong co-existence association between *G. vaginalis, P. bivia, Megasphaera* sp. type 1, and *A. vaginalis* (Supplementary Figures S10, S23), they were tied together to serve as a marker (Group 3) to fine-tune our criteria in distinguishing between transitional BV, BV-positive, and BV-negative samples. Finally, the remaining eight species were also grouped into a marker

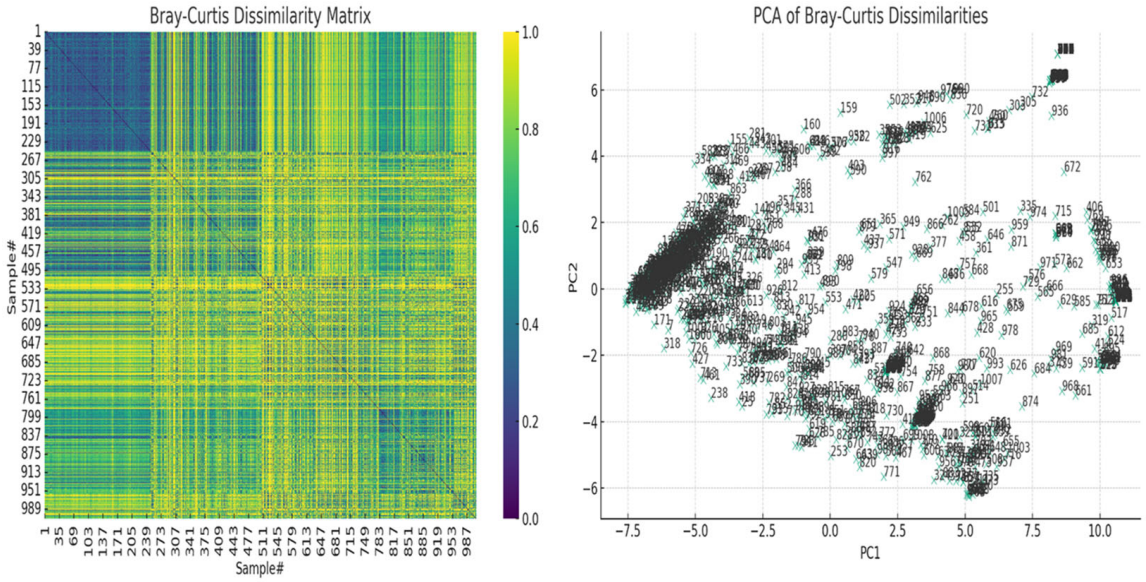



FIGURE 6
The Bray-Curtis dissimilarity matrix heatmap and associated principal component analysis (PCoA) for beta diversity. On the left, the Bray-Curtis Dissimilarity Matrix heatmap shows the pairwise dissimilarities between the samples. Higher values (closer to yellow, 1.0) indicate greater dissimilarity between samples, while lower values (closer to purple, 0.0) indicate greater similarity. On the right, the PCoA of Bray-Curtis Dissimilarities scatter plot visualizes the samples in a reduced two-dimensional space based on their dissimilarities. Each point represents a sample, and their positions reflect patterns of variation across the samples. Labels on the plot correspond to the sample numbers, helping to identify specific samples within the context of the PCoA. These show how dissimilar the samples were from each other.

(Group 4) to further distinguish between BV-positive, BV-negative, and transitional BV (Table 1).

The relative abundance distribution of the species within each of the four biomarker groups above was then used to select cut-off ranges for each group, incorporating a relative abundance-based criteria into the species-based criteria above. This resulted in a four-marker criterion, called the *MDL-BV index*, for diagnosing BV. We trained and tested this two-tier diagnostic *MDL-BV index* on our samples using machine-learning codes (Decision Trees and Random Forests) in Python, adjusting the ranges of each (species group) biomarker's relative abundance until an optimal range per (species group) biomarker was found. The results of the *MDL-BV index's* classifications were manually verified to ensure its veracity.

Observations made during the training and testing process made us include the following four instructions into the *MDL-BV index* code to enable categorization of all types of vaginal samples into their respective BV states: 1. When a sample falls within both BV-negative and transitional BV, assign it to transitional BV; 2. When a sample falls within both BV-positive and transitional BV, assign it to BV-positive; 3. When at least three of a sample's four groups/biomarkers' relative abundances fall within a single BV state and only one falls in another BV state, override the single disagreement, and assign or classify the sample to the BV state with which the three relative abundances agree; 4. When biomarkers 2 and 4 have a relative abundance of zero, then a relative abundance of 0 – 30% for Group 1 and 70 – 100% for Group 3 is BV-positive; a relative abundance of 30 – 50% (Group 1) is transitional BV and 60 – 100% (Group 1) is BV negative; 5. When every group is zero and only group 4 is 100%, define it as BV positive; 6. When all groups are zero, filter out that sample as there are no results to report (Table 1).

The final *MDL-BV index* was then tested on the data used in this study and 490 samples were classified as BV positive, 335 samples were classified as BV negative, and 151 samples were classified as Transitional BV. A manual verification of these classifications found them to be accurate, based on the MDL-BV index ranges. The relative abundance of each of the four biomarkers/groups and the final BV status classification based on these relative abundances, produced by the Python code, is shown in Supplementary Table S3 and summarized in Table 1.

## 3.6 Demographics affect BV

The effect of age, pregnancy status, and race on BV status was analyzed using a correlation heatmap (Supplementary Figure S24) and Box and Whisker distribution plots (Figure 8). Notably, almost all BV-negative cases were found within the White population. Transitional BV was found only among White and Black people, with BV-positive cases being widely distributed among the Black population aged 28 – 40; most White women who had BV were aged between 25 and 35. The age groups of the BV-positive and transitional BV populations were almost the same among Black people but very different among White people. The 'Other" race (including Latinos/Hispanics) also had substantial BV, with their ages between a tight window of 28 to 32 years. Although the number

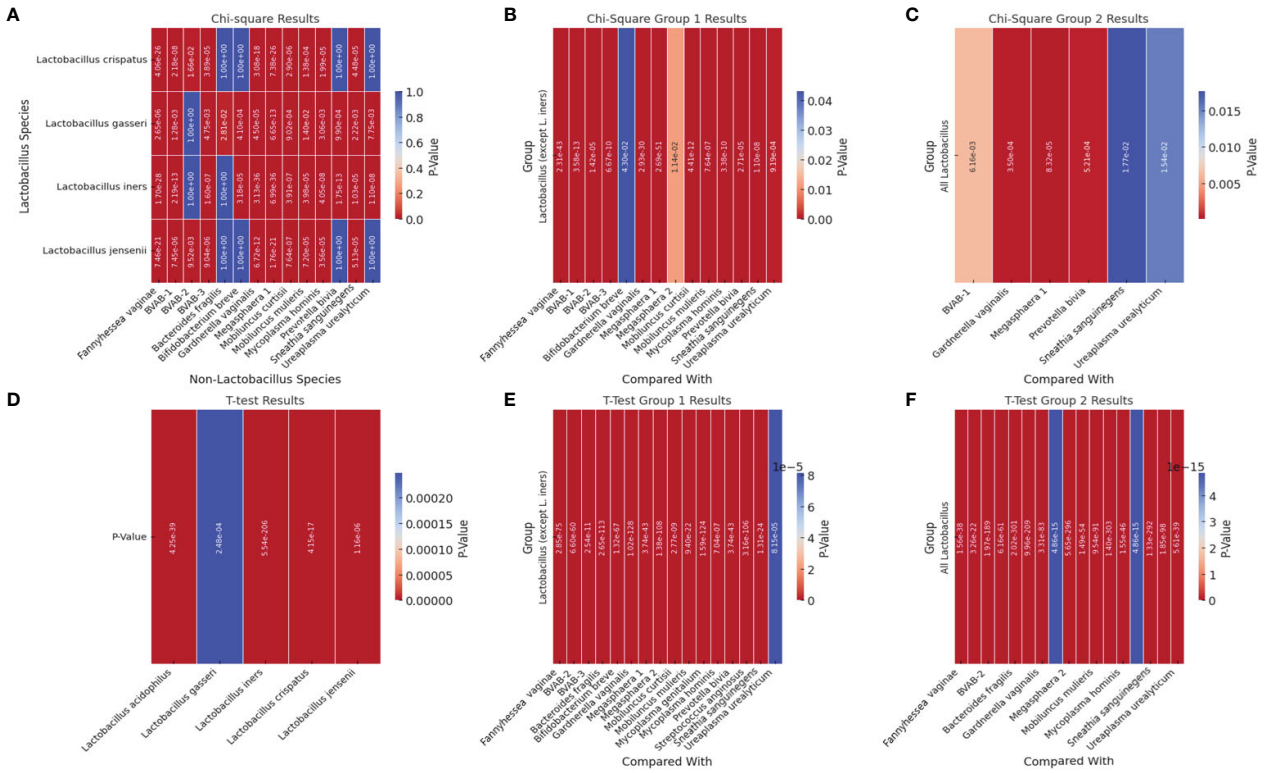



FIGURE 7
Heatmap of significant p-square values obtained from Chi-square and T-test analysis of pairwise comparisons between *Lactobacillus* and other non-*Lactobacillus* species within the samples. **(A)** visualizes the significant Chi-square p-values between *Lactobacillus* species and non-*Lactobacillus* species. **(B)** (Group 1) visualizes the significant Chi-square p-values of all *Lactobacillus* species (except *L. iners*) against each non-*Lactobacillus* species. **(C)** visualizes the significant Chi-square p-values from grouped comparisons of all *Lactobacillus* species against each non-*Lactobacillus* species. **(D)** shows the significant T-test p-values of all *Lactobacillus* species against all the other non-*Lactobacillus* species grouped together (T-test could not get individual species comparisons because of the structure of the data). **(E)** visualizes the significant T-test p-values from T-Test Group 1, which involves grouped comparisons of all *Lactobacillus* species (except *L. iners*) against each non-*Lactobacillus* species. **(F)** visualizes the significant T-test p-values from the "T-Test Group 2" sheet, which involves grouped comparisons of all *Lactobacillus* species against each non-*Lactobacillus* species. Rows represent different *Lactobacillus* species. Columns represent different non-*Lactobacillus* species. The color intensity indicates the level of significance, with cooler colors (towards blue) indicating lower p-values (higher significance) and warmer colors (towards red) indicating higher p-values (lower significance). Cells filled with a p-value of 1.0 represent non-significant results that were not included in the original significant results table and are shown for completeness.

of Asian and American Indian samples was relatively few, they were all BV-positive and fell within the median age range.

It is notable that most women who were pregnant were also BV-positive, with no transitional and BV-negative status being found among them. Further, the age differences between BV-positive pregnant women (24 – 30 years) and non-pregnant women (28 – 38 years) were wide. The Box plot in 8C shows a similar occurrence of BV among women aged 21 – 60, with very little incidence being found among women aged above 60 and below 20. Notably, transitional BV cases were more common among women aged between 41 and 50 than among the other age groups. BV-negative cases were mostly found among women aged between 21 – 30 years.

Chi-square tests indicated that there was no statistically significant association between race and BV status ($P$ = 0.217), and between pregnancy status and BV status ($P$ = 0.527) at significance levels ($P$ < 0.05). The Kruskal-Wallis test, which was used to assess whether there are statistically significant differences in age distributions between the BV positive and BV negative groups, found no statistically significant differences in the age distributions between the BV positive and BV negative groups ($P$ = 0.820) at significance levels ($P$ < 0.05). Therefore, age by itself, may not be a distinguishing factor between these two groups within the dataset analyzed.

## 3.7 Demographics affect variations in relative abundance

### 3.7.1 Variance of Lactobacillus shows lower abundance in BV-positive samples and higher abundance in BV-negative samples.

The relative abundance of *Lactobacillus* sp. showed that while *L. acidophilus, L. jensenii*, and *L. crispatus* had low mean abundance in BV-positive and transitional BV samples, and high mean abundance in BV-negative samples, *L. iners* and *L. gasseri* did not exhibit such phenomena. The mean abundance of *L. gasseri* was

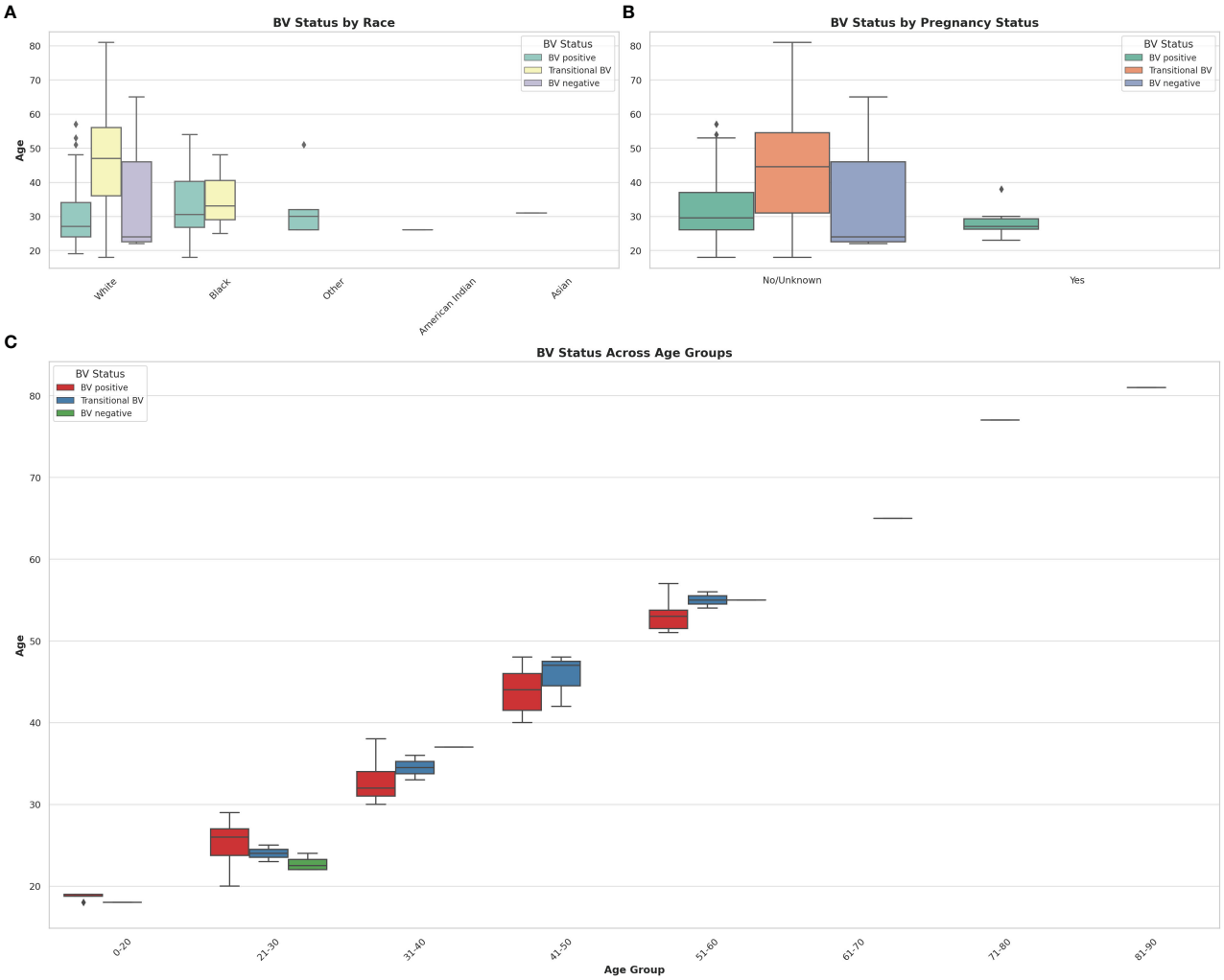



FIGURE 8
Box and Whisker plots showing the distribution of BV-positive, Transitional BV, and BV-negative samples across different races, ages, and pregnancy status. **(A)** shows the distribution of BV status across different races. **(B)** illustrates the distribution of BV status among pregnant and non-pregnant women. **(C)** visualizes the distribution of BV status across different age groups. The horizontal axes show the ages for the different classifications and distributions. The box colors are not the same for all the three plots and their respective keys in the upper right or left corners designate the correct interpretation of each color. The Whisker show the upper and lower quartiles while the boxes show the 25$^{th}$ and 75$^{th}$ percentile range (50% of the population) while the stars show outliers. Chi-square and Kruskal-Wallis tests respectively showed that there were no significant effect/association between race, pregnancy status, and age groups with BV.

similar across all three BV states (marginal variance). *L. iners* was more prevalent in transitional BV samples and less abundant in both BV-negative and BV-positive samples; its mean relative abundance in both BV-positive and –negative species was similar. The greatest variance was observed in *L. acidophilus*' relative abundance levels between BV-positive and BV-negative samples (Figure 9).

### 3.7.2 Variance of Lactobacillus and vaginal anaerobes show differences in pregnancy

The mean abundance of *L. acidophilus, L. gasseri*, and *L. iners* were slightly higher in pregnant women than in non-pregnant women. *L. crispatus* and *L. iners* had the highest relative abundance in non-pregnant and pregnant women, respectively. While *L. jensenii's* mean relative abundance in both pregnant and non-pregnant women was similar, *L. crispatus* stood out as the only *Lactobacillus* with a slightly higher relative abundance in non-pregnant women than in pregnant women. Notably, the variance between the mean relative abundance of *L. acidophilus* in pregnant and non-pregnant women was markedly wider than the other species (Figure 10; Supplementary Figure S25). Except for *L. crispatus*, the mean relative abundance of the *Lactobacillus* sp. was marginally higher in non-pregnant women than in pregnant women (Supplementary Figure S25).

The variance in mean relative abundance of the facultative and obligate anaerobes in pregnant and non-pregnant women differed per species (Figure 11; Supplementary Figure S26). Instructively, *B. fragilis, Megasphaera* sp. type 2, and *B. breve* were only present in non-pregnant women and absent in pregnant women. BVAB-3, *P. bivia, M. curtisii, F. vaginae and M. mulieris* had similar levels in

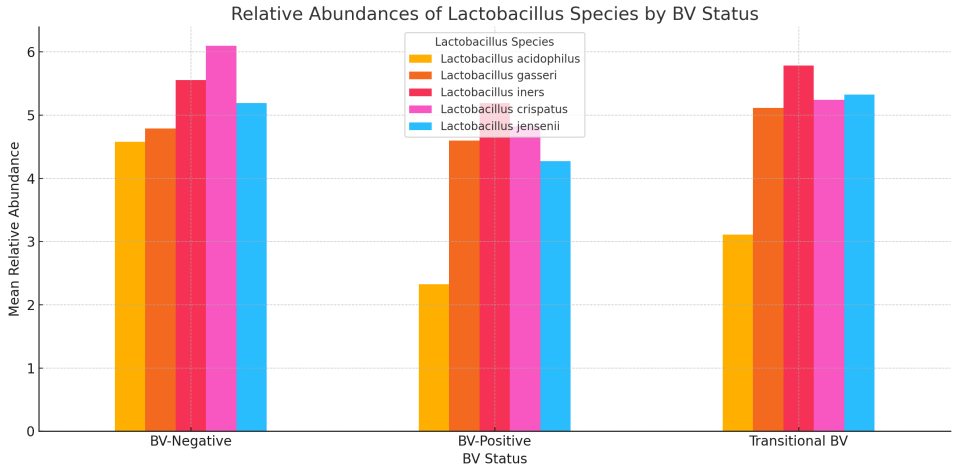



FIGURE 9
A bar chart displaying the relative abundances of five *Lactobacillus* species across three BV-Negative, BV-Positive, and Transitional BV statuses. Overall, Lactobacillus species tend to have higher mean abundances in BV-Negative samples, indicating their potential protective role against BV.

both cohorts, with a slight decrease in four out of the five species for pregnant women; only *M. curtisii* was slightly higher in pregnant women. The differential abundance of BVAB-1, BVAB-2, *G. vaginalis, Megasphaera* sp. type 1, *S. anginosus*, and *U. urealyticum* was higher in pregnant women. *S. sanguinegens* alone was higher in non-pregnant women (Figure 11; Supplementary Figure S26). Non-pregnant women had higher diversity of anaerobes (n = 16 species) and generally lower mean relative abundance than pregnant women (n = 13 species) (Supplementary Figure S27).

### 3.7.3 Variance of Lactobacillus and vaginal anaerobes show differences by age group

*L. acidophilus* was absent in females aged 61+ years and highest in those aged 41–50 years while *L. crispatus* was highest in those aged 61+ years. *L. jensenii* was most dominant in the 0–20 years' cohort while *L. gaserri* had similar abundance levels across all age groups, except the 51–60 age group where it was slightly lower. *L. iners* was also most abundant in the 0–20 years group (Figure 12; Supplementary Figure S28).

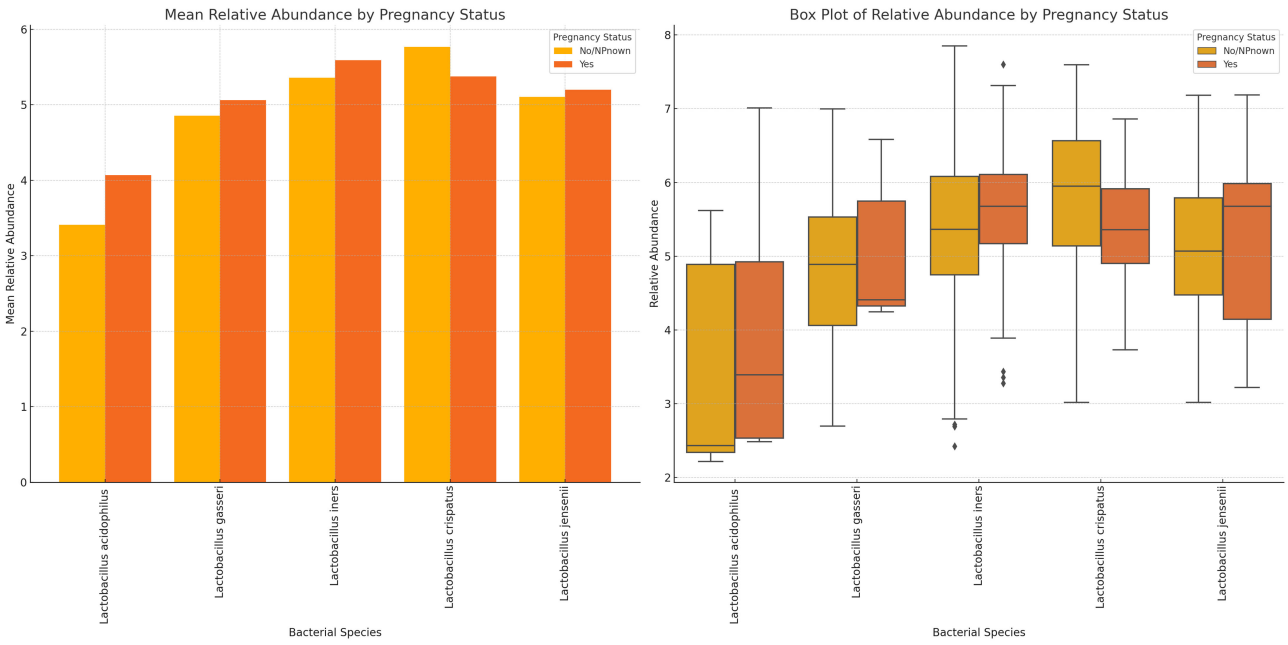



FIGURE 10
Plots displaying the relative abundances of five *Lactobacillus* species in vaginal samples collected from pregnant and non-pregnant women. *L. crispatus* was the highest in non-pregnant women while *L. iners* was highest in pregnant women. In all, the variance between these two cohorts were not so wide. *L. acidophilus* had a wider spread/distribution of relative abundances across the samples for the two cohorts.

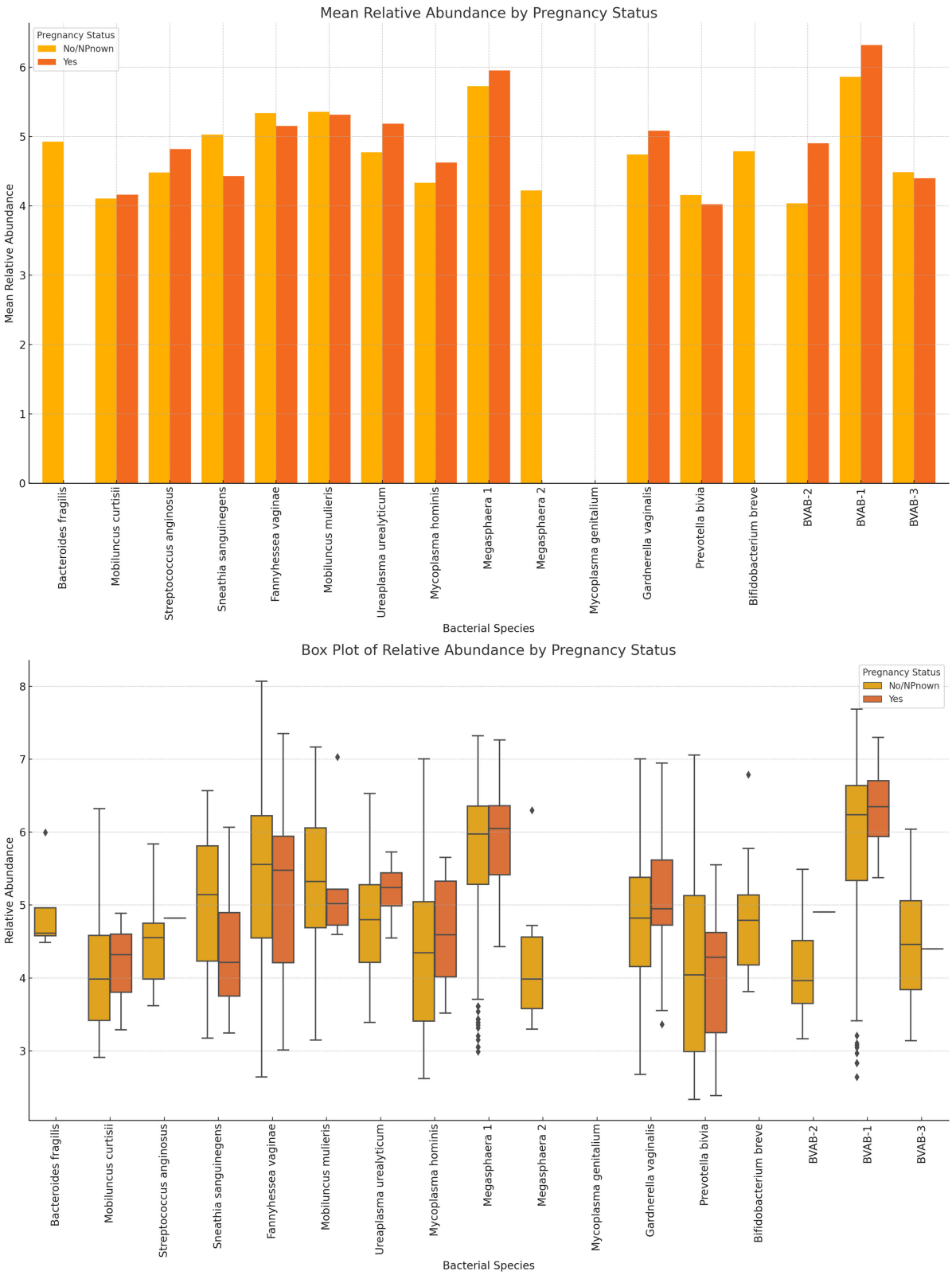



FIGURE 11
The mean and absolute relative abundances of various non-*Lactobacillus* bacterial species across pregnancy statuses (Pregnant and Not Pregnant). Species such as *B. fragilis, Megasphaera* sp. type 2, and *B. breve* were only present in non-pregnant women while the *S. sanguinegens* variance between pregnant and non-pregnant women was higher (in non-pregnant women). BVAB-3, *P. bivia, F. vaginae*, and *M. mulieris* had slightly higher mean abundances in non-pregnant women than in pregnant women. The rest had higher or slightly higher mean abundances in pregnant women than in non-pregnant women.

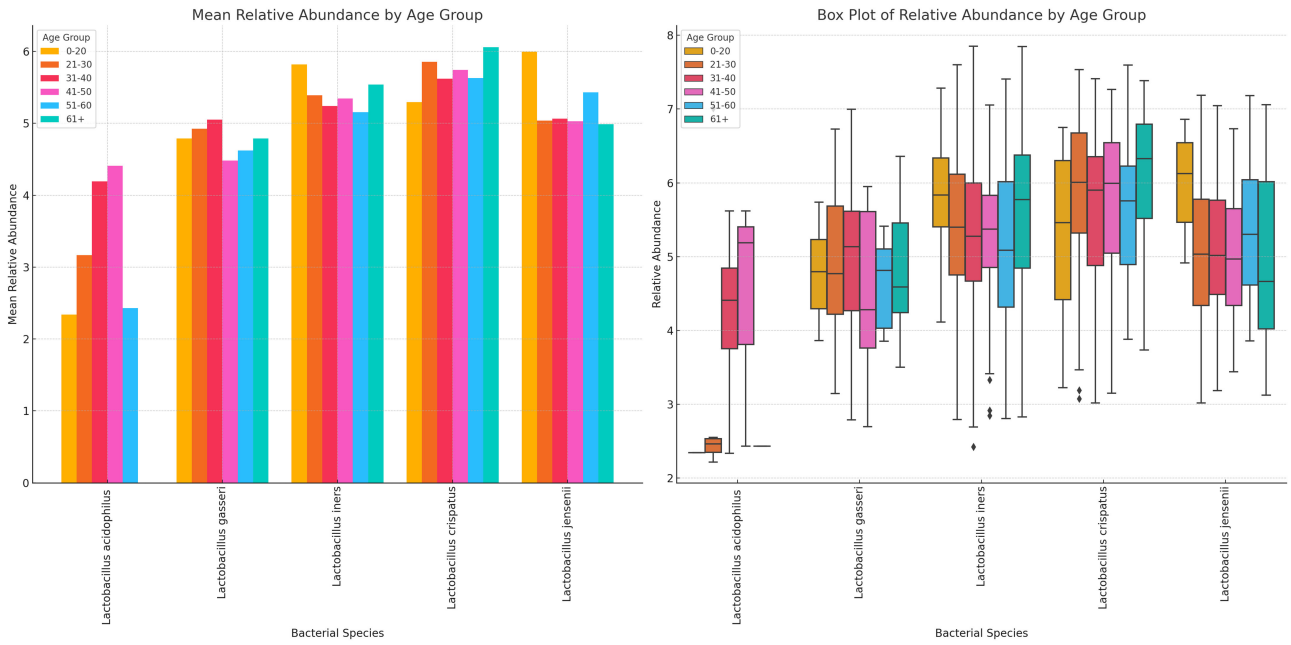



FIGURE 12
Charts displaying the mean and absolute relative abundances of five *Lactobacillus* species across different age groups (0–20, 21–30, 31–40, 41–50, 51–60, 61+). *L. acidophilus* was more abundant in age 41–50 group, followed by age 31–40 and 21–30 groups. *L. crispatus* was higher than the other species in all age groups except 0–20 group, where *L. iners* and *L. jensenii* were higher. Except. *L. acidophilus*, the other species had a wider distribution of relative abundances for the different age groups.

Likewise, the anaerobic bacterial species differed between age groups, with *S. sanguinegens, F. vaginae, M. mulieris, Megasphaera* sp. type 1, *G. vaginalis, B. breve*, and BVAB-3 being very dominant in females aged 61+ years. *B. breve* was very abundant in the 41–50 years group while *Megasphaera* sp. type 2 was also very prevalent in the 0–20-year group. *B. fragilis, S. anginosus, M. mulieris*, and *Megasphaera* sp. type 2 were also highly prevalent within the 31–40-year cohort. Notably, within the 21–30, 31–40, and 41–50 age groups, most of the species were mostly of similar or slightly higher relative abundance; the outliers are described above (Figure 13; Supplementary Figure S29).

There were 13 species in the 0–20 group, 15 in the 21–30, 41–50, and 51–60 groups, and 13 in the 61+ group, showing that the diversity increases after age 21, plateaus until age 60 and decreases after age 61. The highest relative abundances were mainly within the age 21–50 bracket for majority of the species (Figure 13; Supplementary Figure S29).

### 3.7.4 Variance of Lactobacillus and vaginal anaerobes show differences by age group

Variance of Lactobacillus and vaginal anaerobes show diversity and abundance vary by race *Lactobacillus* sp. and non-*Lactobacillus* sp. across the different racial groups. Specifically, *L. iners, L. crispatus*, and *L. jensenii* were present in most of the races, albeit *L. crispatus* was higher in most races and *L. acidophilus* was consistently lower than the other *Lactobacilli. L. iners* was very common in Asian people compared with the other *Lactobacilli; L. acidophilus* was higher in Others (Hispanics and Pacific Islanders) races and lower in the remaining races. *L. gasseri* was high in Asian, Black, and Others (Hispanics and Pacific Islanders) races across all races. Notably, *L. crispatus* was the highest *Lactobacilli* in White people (Figure 14; Supplementary Figures S30, S31).

Racial variations in relative abundance were most obvious in *B. fragilis, S. anginosus, M. hominis, Megasphaera* sp. type 2, *P. bivia, B. breve*, and BVAB-2. For many of the species, there were a higher relative abundance among Black people than White people: *Mobilincus* sp., *S. sanguinegens, U. urealyticum, G. vaginalis, Megasphaera* sp., and BVAB-2. *B. fragilis M. mulieris, G. vaginalis*, BVAB-1 *and* BVAB-3 were highly abundant in Asian people. Among White people, *F. vaginalis, M. hominis, S. sanguinegens, Megasphaera* sp. type 1, *G. vaginalis, P. bivia, B. breve*, BVAB-2 and BVAB- 3 were dominant (Figure 15; Supplementary Figures S32, S33).

## 4 Discussion

Unlike other infections with single etiological agents, BV is a polymicrobial infection with no clear consensus on the specific bacterial species responsible for the altered vaginal microbiota state (Mondal et al., 2023; Sobel and Vempati, 2024). The exact contributions of BV-associated microbes to the pathogenesis of BV, which may be relevant for accurate diagnosis and therapeutics, remain unresolved. Nevertheless, the occurrence of normal, transitional, and abnormal vaginal microbiota states is dependent on the composition and interactions of the different *Lactobacillus* and anaerobic species present (Onderdonk et al., 2016; Lynch et al., 2019; Chee et al., 2020; Drew et al., 2020). To address this gap between the vaginal microbiome dynamics and the current BV diagnostic tests'

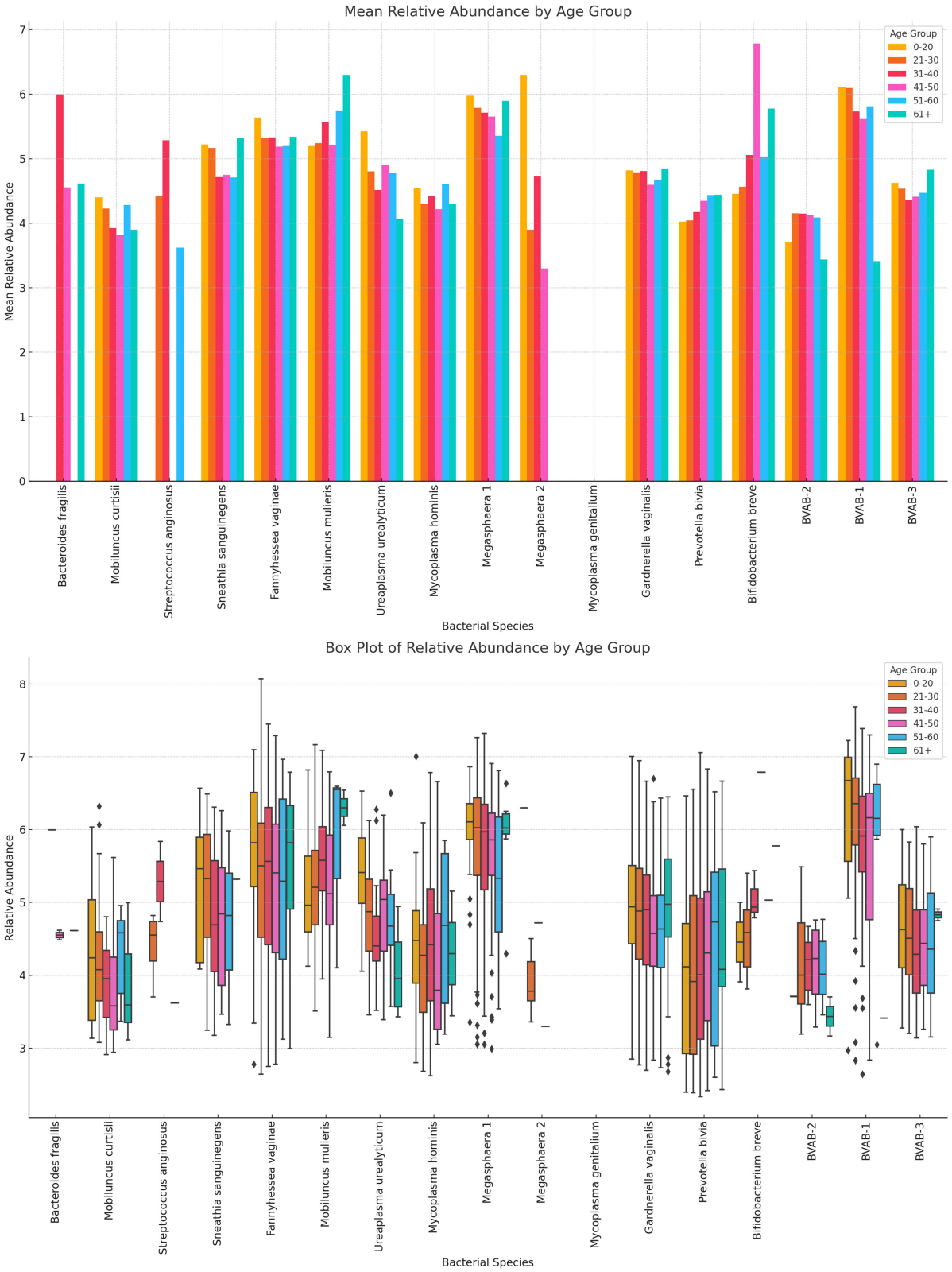



FIGURE 13
The bar chart displays the mean and absolute relative abundances of various non-*Lactobacillus* bacterial species across different age groups. Fourteen species were present in age 0–20 group, 15 were found in age 21–30, 41–50, and 51–60 groups, 16 was found in age 31–40 group, and 13 were found in 61+ age group. Species such as *P. bivia, F. vaginae* and *S. sanguinegens* had broader spread of relative abundances across the age groups. Overall, non-*Lactobacillus* bacterial species are primarily observed in the 21–30 age group, with little to no data available for other age groups. This indicates a higher microbial diversity in this age group compared to others.

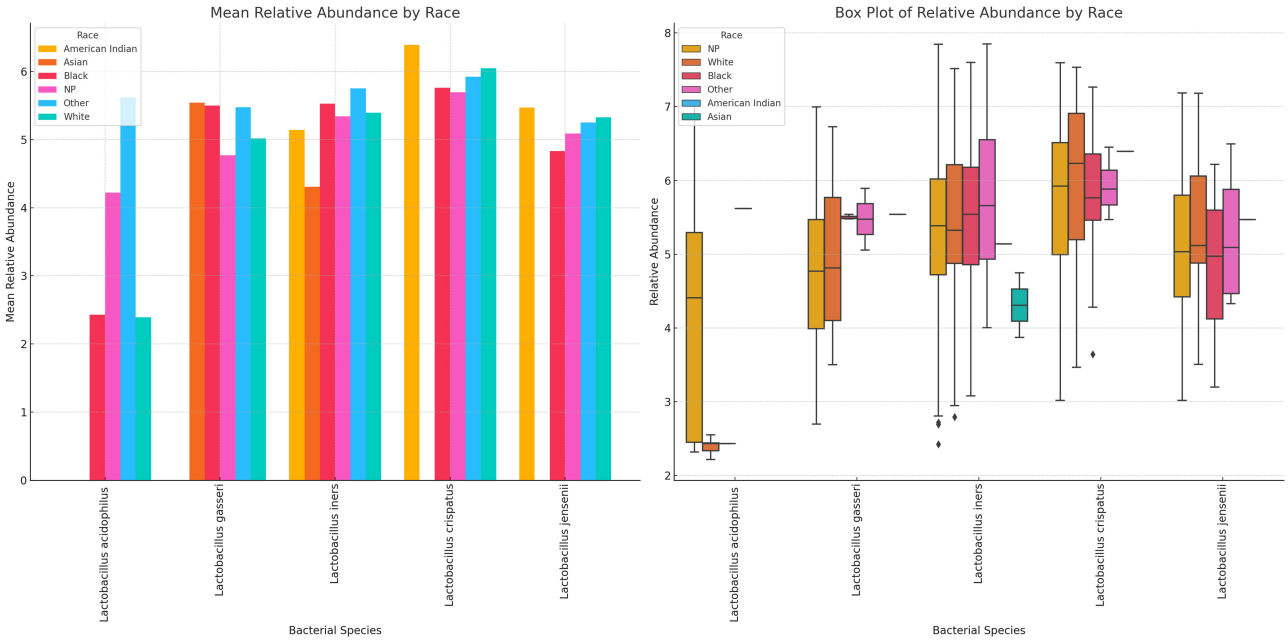



FIGURE 14
Mean and absolute relative abundance distribution of *Lactobacillus* species in different racial/ethnic groups. White, Black, and other races had more diversity of *Lactobacillus* sp. while Asian people had very little diversity. *L. acidophilus* was higher in other (Hispanics and Pacific Islanders) races than White and Black people. *L. crispatus* was higher in American Indians and White people while *L. gasseri* and *L. iners* were higher in Black and Other races. *L. gasseri* was also higher in Asian people. The relative abundance spread of the species differed per race. NP means not provided. The Whisker show the upper and lower quartiles while the boxes show the 25$^{th}$ and 75$^{th}$ percentile range (50% of the population) while the stars show outliers.

limitations, a new qRT-PCR test that detects and measures the gDNA concentrations of 22 species found in all the various conditions of the vaginal microbiome was designed. The relatively shorter turnaround time of this test and the ease of adoption in a routine diagnostic laboratory makes it possibly more valuable than the current classical tests (Muzny et al., 2023; Swidsinski et al., 2023).

With such a broad spectrum of bacterial species, the resolution of the test is enhanced to efficiently distinguish between normal, abnormal, and transitional microbiota. The test has been validated with 95 – 100% sensitivity and specificity with a short turnaround time of 8 hours (from sample reception). In studies where results from PCR or qPCR studies targeting 2 – 13 species have been compared with Amsel's criteria or Nugent's score, the results have been highly sensitive and specific (Menard et al., 2008; Malaguti et al., 2015; Lynch et al., 2019; Drew et al., 2020; Vodstrcil et al., 2021). We are therefore confident that subsequent studies with our test, using Nugent score or Amsel's criteria-diagnosed samples, may equally yield similar or better results with its larger bacteria target spectrum. Although this study is limited by the inability to compare the current results with clinical presentation data, its sensitivity and specificity in detecting the controls used confirm its efficiency.

In fact, the test's diagnostic capability is further appreciated when the make-up and sample size of our vaginal specimens are considered: 946 vaginal samples from a wide spectrum of ages (18 – 83), races (White, Black, Asian, American Indian, and "Other (Hispanics and Pacific Islanders)", not to mention the races of most women who failed to state their race), and pregnancy status. The Bray-Curtis dissimilarity matrix and the principal component analysis (PCoA) showed how dissimilar the samples were as few of the samples clustered together (Figure 6). The sample differences were further highlighted by the species richness (α-diversity) of the samples in which the number of species per sample ranged from 2 – 12, with 50% of the samples having 2 – 6 species per sample. Instructively, the Shannon index further clarified this observation by showing that the species diversity per sample was low, with most samples having an index between 0.8 and 1.8 and the most diverse sample having an index of 3.2 (out of 3.5).

Thus, although the samples were dissimilar in terms of composition, their diversity was relatively low, which could be characteristic of a healthy vaginal microbiome if the relative abundance of *Lactobacillus* sp. is high (Balashov et al., 2014). However, it is worth noting that the Shannon diversity is affected by the number of species sampled and the vaginal microbiome is naturally not as diverse as the gut microbiome (Javed et al., 2019; Muzny et al., 2020; Abou Chacra and Fenollar, 2021). Hence, although the 22 species used are more than any PCR test, it may not represent all the species in the vaginal microbiota. For instance, the alpha diversity of the samples was calculated through a direct count of the number of species observed in each sample without using any estimators or indices. Hence, it does not account for undetected species or attempt to estimate the total species richness in the samples, which methods like the Chao1 index aim to do. Notwithstanding, the representative nature of the samples used in this study is self-evident.

To our knowledge, no qPCR assay has the same broad-spectrum species target, making this assay an important

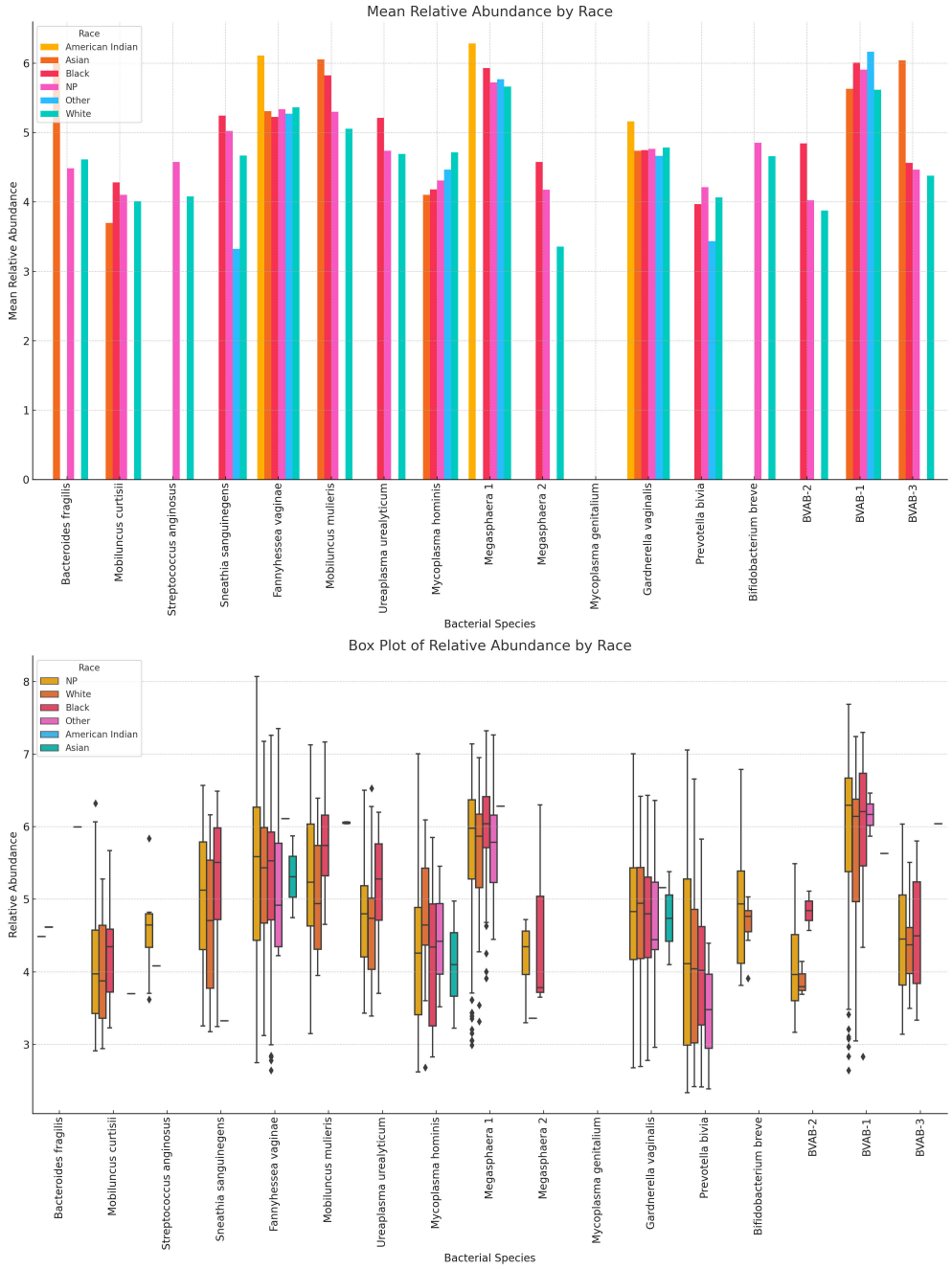



FIGURE 15
Relative abundance distribution of non-*Lactobacillus* species in different racial/ethnic groups. For many of the species, there were a higher relative abundance among Black people than White people, explaining the higher prevalence of BV in Black people than White people. *F. vaginae*, *Megasphaera* sp. type 1, and *G. vaginalis* were highest in American Indians. Relative abundance distribution of the species differed per species. The Whisker show the upper and lower quartiles while the boxes show the 25$^{th}$ and 75$^{th}$ percentile range (50% of the population) while the stars show outliers. NP means not provided.

innovation and the first to do so. Whereas other assays have used only two or up to 13 species (Malaguti et al., 2015), not all of the species were quantitatively detected as was done in this assay (Brotman and Ravel, 2008; Menard et al., 2008; Malaguti et al., 2015; Drew et al., 2020). This is despite the calls for increasing the bacterial spectrum for molecular BV diagnosis, focusing on *Megasphaera, G. vaginalis, F. vaginae*, and other anaerobes (Brotman and Ravel, 2008; Menard et al., 2008; Lynch et al., 2019). Moreover, owing to the biofilm-forming nature of these BV-associated species, which leads to immune evasion, high treatment failure rates and recurrence, a shorter-turnaround diagnostic tool is merited (Cangui-Panchi et al., 2023; Muzny et al., 2023). This study answers that call.

By both detecting and quantifying the 22 species in vaginal samples, we were able to obtain a clearer picture of the vaginal microbiome's composition and make a better diagnosis of its

condition; a feat only achievable with whole-genome metagenomics (Asante and Osei Sekyere, 2019; Osei Sekyere et al., 2020; Osei Sekyere et al., 2023). For instance, although BVAB-1, -2, and -3 were detected together in the same study (Fredricks et al., 2005), BVAB-2 has been commonly tested as a marker of BV positivity without BVAB-1 and -3. Yet, this study showed that BVAB-2 is the least prevalent/abundant among these three species, with BVAB-1, followed by BVAB-3, being the most common. A Brazilian study using 223 BV-positive samples also reported the same findings regarding the relative abundance of BVAB-1, BVAB-2, and BVAB-3, affirming our observations (Malaguti et al., 2015). A positive association between BV and high-risk Human Papilloma Virus (HPV) genotypes has been reported, owing to the occurrence of BVAB 1 and 3, and other BV-associated- bacteria in women co-infected with HIV and HPV (Naidoo et al., 2022). The same study also showed that the presence of BVAB 1 and 3 had an elevated likelihood of increasing the severity of cervical neoplasia in this population (Naidoo et al., 2022). Nevertheless, these species are not as closely related as initially thought as a recent phylogenetic analysis identified BVAB 1, 2, and 3 to be "*Clostridiales genomosp.*", *Oscillospiraceae bacterium* strain CHIC02, and *Mageeibacillus indolicus*, respectively (Osei Sekyere et al., 2023).

Additionally, the relative abundance of the various species identified by this test mirrors what has been reported (Malaguti et al., 2015; Bayigga et al., 2019; Abou Chacra et al., 2021; Zhang et al., 2022; Powell et al., 2023), with *Lactobacillus* sp. (except *L. acidophilus*), *G. vaginalis, A. (F.) vaginae, P. bivia*, and *Megasphaera* sp. type 1 being the most dominant and widely distributed (Figures 2, 4; Supplementary Figures S1–S22). The absence of *M. genitalium* [which was found in low abundance in other studies using BV-positive samples (Malaguti et al., 2015)], and lower abundance and prevalence of *L. acidophilus, M. hominis, U. urealyticum, Megasphaera* sp. type 2, *Mobiluncus* sp., *S. sanguinegens, B. breve*, and *S. anginosus* are also consistent with other findings (Atassi et al., 2019; Chee et al., 2020; Argentini et al., 2022; Wu et al., 2022), further crediting the diagnostic efficiency of this test.

Overall, the *Lactobacillus* species had higher mean abundances in BV-negative samples, lower abundances in BV-positive samples, and in-between for transitional BV, indicating their potential protective role against BV (Figure 9). The higher relative abundance of *L. iners* in pregnant women and transitional BV samples could further suggest that *L. iners* play a crucial role in transitioning the vaginal microbiome from BV-negative to BV-positive. The transitioning effect of *L. iners* has been investigated and reported earlier and these observations further throws more light on the veracity of this conclusion (Zheng et al., 2021), albeit other studies provide conflicting findings and call for additional studies (Carter et al., 2023).

It was observed from the correlation, and co-existence analysis that the five *Lactobacillus* sp. mostly co-existed together while *G. vaginalis, A. (F.) vaginae, P. bivia*, and *Megasphaera* sp. type 1 also co-existed in the same samples (Figures 4, 7; Supplementary Figures S1–S23). The Chi-square and T-tests also largely agreed with the significant association between the *Lactobacillus* sp. and the non-*Lactobacillus* species, with few exceptions (Figure 7). This was used to form two separate groups of species-based biomarkers. Furthermore, the BVABs, *S. anginosus*, and *B. fragilis*, which are known to be mainly associated with BV microbiomes and absent in normal vaginal microbiomes (Weyers et al., 2009; Matu et al., 2010; Malaguti et al., 2015; Schmidt et al., 2015; Barrientos-Durán et al., 2020), were also teased from the remaining non-*Lactobacillus* species and grouped into another biomarker group. The remaining eight species were then also bundled together into another fourth group (Table 1). Using the relative abundance distributions of each species (Figure 4; Supplementary Figures S1–S22), the relative abundance range for each of the four species biomarker groups was set to form the *MDL-BV index*, which was then further trained and tested on the data using machine-learning algorithms (Decision Trees and Random Forests) to diagnose BV. A relative abundance-based approach was adopted over a nominal concentration value because concentrations vary from sample to sample, and the swabbing method of sample collection among clinicians is not standardized; hence, samples collected by each swab is not quantitative.

The *MDL-BV index* has a high cut-off range for negative BV status (≥70% *Lactobacillus* sp. relative abundance) to ensure that samples diagnosed as normal were BV-negative in verity. It also errs on the side of caution by classifying samples that meets the criteria of both BV-negative and transitional or transitional and BV-positive, as transitional BV or BV-positive respectively. The training of the model on the data further revealed certain intricacies and nuances in the distribution of the species and their relative abundances, which were used to refine it by including more details and rules. For instance, in situations where only two of the four biomarkers had a relative abundance score, the ratios were adjusted (Table 1). This strengthened the diagnostic efficiency of the *MDL-BV index*, resulting in a final classification of the samples as 491 BV-positive, 318 BV-negative, and 137 transitional BV. A manual verification of the results confirmed the veracity of the classifications (Supplementary Table S3).

Instead of using just two biomarkers involving all *Lactobacillus* sp. and all non-*Lactobacillus* species (Balashov et al., 2014), this index uses four biomarkers, which further refines the ability of the index to correctly distinguish between transitional BV or BV-positive microbiomes. Evidently, a correct molecular diagnosis of BV is critical to its treatment and monitoring, making this test and index very important in gynecology and obstetrics. The next stage of this research is to train the model on large datasets with Nugent scores/Amsel criteria as a means of further enhancing it and strengthening its proof of concept in BV diagnosis.

The distribution of BV among the different demographic variables viz., age, race, and pregnancy status showed important differences, albeit none of the differences were statistically significant (Figure 8). Specifically, although the number of vaginal samples from Black people was lower than that from White people, there were more BV-positive samples from the former than from the latter. While most of the vaginal samples from White people were BV-negative, those from Black people were either BV-positive or transitional. Notably, the samples from the "Other" races (which includes Hispanics/Latinos and Pacific Islanders), Asians, and Native American Indians, although fewer in number, were also BV-positive, with none being negative. Furthermore, the mean age and age range for White people with transitional BV were higher than those of all the other races, including the mean age of White

people with abnormal vaginal microbiomes (Figure 8). Hence, although these differences were not statistically significant, they concur with other studies that show a higher prevalence of BV among Black and Hispanics than among White people and Asians (Peebles et al., 2019). This agreement between our data and other studies further confirms the accuracy of our *MDL-BV index*.

A further analysis of the racial vaginal microbiome composition showed that Black women had a higher prevalence of many of the anaerobes implicated in the BV pathogenesis, including *G. vaginalis, F. vaginae*, and *Megasphaera* sp., followed by Hispanics and Asian women who also had higher abundance of these BV-associated species than Caucasians (Figure 15; Supplementary Figures S32, S33). These findings are not singular as other studies have also identified the same. Particularly, Borgdoff et al. (2017) found a higher abundance of *L. crispatus* among Caucasians than Black and Asian women and a higher *G. vaginalis* and *L. iners* in Ghanaian and Asian (and Mediterranean) women, respectively (Borgdorff et al., 2017). Contrarily however, Roachford et al. (2022) also found a higher abundance of *L. crispatus* and *L. iners* in African Americans than in other ethnicities (Roachford et al., 2022); nevertheless, most studies agree that Africans have lower *Lactobacillus*-rich microbiomes compared with White women. These racial and ethnic variations were examined recently by Xin Wei et al. (2024) and found that the nucleotide identities of these vaginal denizens found in different ethnicities varied from each other, underscoring an evolutionary mechanism that explains these observed variations (Wei et al., 2024).

In effect, the genomes of the same species had differences when isolated from different races. While they also observed similarly higher abundance of BV-associated bacteria in Black pregnant women than other races (higher alpha diversity and Shannon index), they established that the alpha diversity remained fairly constant over gestational time and across ethnicities (Wei et al., 2024). This is similar to our finding in which the differences between pregnant and non-pregnant cohorts' vaginal microbiome variations were marginal. Notwithstanding, this study brings out certain unique findings that merit further investigation: the absence of *B. fragilis, Megasphaera* sp. type 2, and *B. breve* in pregnant women; the higher differential abundance of *G. vaginalis*, BVAB-1 and -2, *U. urealyticum, Megasphaera* sp. type 1, and *S. anginosus* in pregnant women; and the higher abundance of *S. sanguinegens* in non-pregnant women compared with pregnant women.

Albeit pregnancy had no significant effect on BV, we posit that the observed marginal variations of higher diversity and higher relative abundance in pregnant women than in non-pregnant women can be due to the hormonal changes that occur in the former. Hormonal variations in pregnancy also explain the lower relative abundance of *Lactobacillus* sp. (except *L. iners*) and anaerobes among non-pregnant women. Notably, *L. acidophilus* presented the widest variations between cohorts, suggesting that it is more sensitive to these changes than the other species and could be adopted as a biomarker. Indeed, the higher abundance of *L. crispatus* among non-pregnant and menopausal women, the absence of *L. acidophilus* in menopausal women, and the higher abundance of *L. iners* among pregnant women further show how the variations in hormones affect these species' relative abundance through changes in the hosts estrogen-medicated metabolisms (Auriemma et al., 2021). The higher abundance of *L. crispatus* and *L. iners* was been reported previously (Borgdorff et al., 2017; Wei et al., 2024).

Although BV among non-pregnant women was more prevalent than among pregnant women, BV-negative or transitional BV samples were not obtained from pregnant women (Figure 8). This is a concerning observation as BV in pregnancy is associated with preterm labor, low birth weight, premature rupture of membranes, miscarriage, chorioamnionitis, and birth asphyxia (Bayigga et al., 2019; Basavaprabhu et al., 2020; Vodstrcil et al., 2021). Although the pregnancy stage of the women from whom the samples were obtained is unknown, it is known that the first-trimester vaginal microbiome is similar to a BV-positive microbiome while the 2$^{nd}$ and 3$^{rd}$ trimesters are similar to a normal vaginal microbiome (Kaur et al., 2020). Hence, the stage of the pregnancy and vaginal symptoms should be considered in making a final therapeutic decision regarding pregnant women when using molecular-based tests. Furthermore, the difference in the number of pregnant women (n = 53) vis-a-viz the number of non-pregnant (or unknown pregnancy status) women is large and can skew the data. This is a major limitation in our data set and should be considered when analyzing the pregnancy data.

Although BV-positive samples were found in all age groups (18 – 83 years), they were mostly prevalent among women who were between 20 and 60 years, with a higher number of BV-positive and transitional BV samples being found among women between 41 – 50 years. The species variations per age group further support this observation as the anaerobic species diversity and abundance increase from after 20 years, plateau until age 50, and begin to decline again (Figures 12, 13; Supplementary Figures S28, S29). Indeed, the median age found in this study's population is similar to what was found in Brazil (Malaguti et al., 2015) and in the USA (Watkins et al., 2024), with women above 21 years being more likely to suffer recurrence of BV (Coudray and Madhivanan, 2020; Coudray et al., 2020). Hence, the age groups found in our data concur with that of other studies, which confirms that BV is common among reproductive-age women and lower before puberty and after menopause (Auriemma et al., 2021). However, age was not significantly associated with BV occurrence in this study.

## 5 Conclusion

The new qRT-PCR assay developed by MDL is the first of its kind to quantitatively detect 22 species in the vaginal microbiome. We developed a novel classification system, the new *MDL-BV index*, which was trained on large sets of data and based on four species-based and relative abundance-based biomarkers. This multifaceted two-tier approach of using species and relative abundance provides a better diagnostic resolution for a polymicrobial infection such as BV. We are working to extend this approach to apply to other microbiome-based diagnostics to make disease diagnosis reflective of the clinical presentations. Our study is, however, limited by the absence of a clinical Nugent score or Amsel's criteria; however, we are working to include these in the next studies that will involve the *MDL-BV index* and the 22-species qRT-PCR BV test. The clinical and socio-economic importance of this novel proprietary BV diagnostic test in obstetrics and gynecology is notable.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

## Ethics statement

Ethics approval was not required for this study as the samples used were leftover, discarded samples provided by healthcare providers and were anonymized, posing no risk of identifying individual patients. According to the Common Rule (45 CFR 46) and HIPAA (45 CFR 164.514(b)(2)), research involving anonymized data does not constitute human subjects research and thus does not require IRB approval. Written informed consent was not obtained as the study used leftover, discarded samples, and the study did not involve direct interaction with patients.

## Author contributions

AO: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. RH: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. JT: Investigation, Methodology, Supervision, Writing – review & editing. SF: Supervision, Validation, Writing – review & editing. EM: Funding acquisition, Resources, Supervision, Validation, Visualization, Writing – review & editing. MA: Funding acquisition, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing. JO: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by the Medical Diagnostics Laboratories, LLC.

## Acknowledgments

The authors are grateful to the technicians at Medical Diagnostic Laboratories for their direct and indirect assistance. The material and financial resources provided by Medical Diagnostic Laboratories towards this project are warmly acknowledged and deeply appreciated. We are also grateful to Annette Daughtry for assisting with the review of the initial draft.

## Conflict of interest

All the authors are scientists and staff of Medical Diagnostic Laboratories LLC.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1409774/full#supplementary-material

SUPPLEMENTARY TABLE 1
Raw demographic and qPCR data showing the demographics of the women from whom the samples were obtained and the qPCR gDNA concentrations of each species within the 946 samples.

SUPPLEMENTARY TABLE 2
Chi-square and T-test statistical data showing the significant and non-significant associations between *Lactobacillus* sp. and the other non-*Lactobacillus* species.

SUPPLEMENTARY TABLE 3
BV status classification data generated by the Python algorithm trained on the MDL-BV index ranges. The relative abundance of each group of biomarkers are summed up and shown under their respective groups. The ratio of these relative abundances are used to infer the BV status of the sample.

SUPPLEMENTARY FIGURES 1–22
Concentration distribution of each of the 22 bacterial species tested by the qPCR test across all the samples. Both boxplot and histograms charts are shown for each species.

SUPPLEMENTARY FIGURE 23
A correlation matrix showing the correlation between the individual species found within the samples. Different color codes are used to show positive and negative correlations.

SUPPLEMENTARY FIGURE 24
A correlation heatmap showing the correlation between age, pregnancy, race, and BV status.

# References


Abou Chacra, L., and Fenollar, F. (2021). Exploring the global vaginal microbiome and its impact on human health. *Microb. Pathog.* 160, 105172. doi: 10.1016/j.micpath.2021.105172

Abou Chacra, L., Fenollar, F., and Diop, K. (2021). Bacterial vaginosis: What do we currently know? *Front. Cell Infect. Microbiol.* 11. doi: 10.3389/FCIMB.2021.672429

Amsel, R., Totten, P. A., Spiegel, C. A., Chen, K. C. S., Eschenbach, D., and Holmes, K. K. (1983). Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. *Am. J. Med.* 74, 14–22. doi: 10.1016/0002-9343(83)91112-9

Argentini, C., Fontana, F., Alessandri, G., Lugli, G. A., Mancabelli, L., Ossiprandi, M. C., et al. (2022). Evaluation of modulatory activities of lactobacillus crispatus strains in the context of the vaginal microbiota. *Microbiol. Spectr.* 10, e02733–21. doi: 10.1128/spectrum.02733-21

Asante, J., and Osei Sekyere, J. (2019). Understanding antimicrobial discovery and resistance from a metagenomic and metatranscriptomic perspective: Advances and applications. *Environ. Microbiol. Rep.* 11, 62–86. doi: 10.1111/1758-2229.12735

Atassi, F., Pho Viet Ahn, D. L., and Lievin-Le Moal, V. (2019). Diverse Expression of Antimicrobial Activities Against Bacterial Vaginosis and Urinary Tract Infection Pathogens by Cervicovaginal Microbiota Strains of Lactobacillus gasseri and Lactobacillus crispatus. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.02900

Auriemma, R. S., Scairati, R., del Vecchio, G., Liccardi, A., Verde, N., Pirchio, R., et al. (2021). The vaginal microbiome: A long urogenital colonization throughout woman life. *Front. Cell Infect. Microbiol.* 11. doi: 10.3389/fcimb.2021.686167

Balashov, S. V., Mordechai, E., Adelson, M. E., Sobel, J. D., and Gygax, S. E. (2014). Multiplex quantitative polymerase chain reaction assay for the identification and quantitation of major vaginal lactobacilli. *Diagn. Microbiol. Infect. Dis.* 78, 321–327. doi: 10.1016/j.diagmicrobio.2013.08.004

Barrientos-Durán, A., Fuentes-López, A., de Salazar, A., Plaza-Díaz, J., and García, F. (2020). Reviewing the composition of vaginal microbiota: inclusion of nutrition and probiotic factors in the maintenance of eubiosis. *Nutrients* 12, 419. doi: 10.3390/nu12020419

Basavaprabhu, H. N., Sonu, K. S., and Prabha, R. (2020). Mechanistic insights into the action of probiotics against bacterial vaginosis and its mediated preterm birth: An overview. *Microb. Pathog.* 141, 104029. doi: 10.1016/j.micpath.2020.104029

Bayigga, L., Kateete, D. P., Anderson, D. J., Sekikubo, M., and Nakanjako, D. (2019). Diversity of vaginal microbiota in sub-Saharan Africa and its effects on HIV transmission and prevention. *Am. J. Obstet. Gynecol.* 220, 155–166. doi: 10.1016/j.ajog.2018.10.014

Bhujel, R., Mishra, S. K., Yadav, S. K., Bista, K. D., and Parajuli, K. (2021). Comparative study of Amsel's criteria and Nugent scoring for diagnosis of bacterial vaginosis in a tertiary care hospital, Nepal. *BMC Infect. Dis.* 21, 825. doi: 10.1186/s12879-021-06562-1

Borgdorff, H., van der Veer, C., Van Houdt, R., Alberts, C. J., De Vries, H. J., Bruisten, S. M., et al. (2017). The association between ethnicity and vaginal microbiota composition in Amsterdam, the Netherlands. *PLoS One* 12, e0181135. doi: 10.1371/journal.pone.0181135

Brotman, R. M., and Ravel, J. (2008). Ready or not: the molecular diagnosis of bacterial vaginosis. *Clin. Infect. Dis.* 47, 44–46. doi: 10.1086/588662

Cangui-Panchi, S. P., Ñacato-Toapanta, A. L., Enríquez-Martínez, L. J., Salinas-Delgado, G. A., Reyes, J., Garzon-Chavez, D., et al. (2023). Battle royale: Immune response on biofilms - host-pathogen interactions. *Curr. Res. Immunol.* 4, 100057. doi: 10.1016/j.crimmu.2023.100057

Carter, K. A., Fischer, M. D., Petrova, M. I., and Balkus, J. E. (2023). Epidemiologic evidence on the role of lactobacillus iners in sexually transmitted infections and bacterial vaginosis: A series of systematic reviews and meta-analyses. *Sex. Transm. Dis.* 50, 224–235. doi: 10.1097/OLQ.0000000000001744

Chee, W. J. Y., Chew, S. Y., and Than, L. T. L. (2020). Vaginal microbiota and the potential of Lactobacillus derivatives in maintaining vaginal health. *Microb. Cell Fact* 19, 203. doi: 10.1186/s12934-020-01464-4

Cheng, L., Norenhag, J., Hu, Y. O. O., Brusselaers, N., Fransson, E., Ährlund-Richter, A., et al. (2020). Vaginal microbiota and human papillomavirus infection among young Swedish women. *NPJ Biofilms Microbiomes* 6, 39. doi: 10.1038/s41522-020-00146-8

Coudray, M. S., and Madhivanan, P. (2020). Bacterial vaginosis-A brief synopsis of the literature. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 245, 143–148. doi: 10.1016/j.ejogrb.2019.12.035

Coudray, M. S., Sheehan, D. M., Li, T., Cook, R. L., Schwebke, J., and Madhivanan, P. (2020). Factors associated with the recurrence, persistence, and clearance of asymptomatic bacterial vaginosis among young African American women: A repeated-measures latent class analysis. *Sex. Transm. Dis.* 47, 832–839. doi: 10.1097/OLQ.0000000000001256

Drew, R. J., Murphy, T., Broderick, D., O'Gorman, J., and Eogan, M. (2020). An interpretation algorithm for molecular diagnosis of bacterial vaginosis in a maternity hospital using machine learning: proof-of-concept study. *Diagn. Microbiol. Infect. Dis.* 96, 114950. doi: 10.1016/j.diagmicrobio.2019.114950

Elnaggar, J. H., Ardizzone, C. M., Cerca, N., Toh, E., Łaniewski, P., Lillis, R. A., et al. (2023). A novel Gardnerella, Prevotella, and Lactobacillus standard that improves accuracy in quantifying bacterial burden in vaginal microbial communities. *Front. Cell Infect. Microbiol.* 13. doi: 10.3389/fcimb.2023.1198113

Fredricks, D. N., Fiedler, T. L., and Marrazzo, J. M. (2005). Molecular identification of bacteria associated with bacterial vaginosis. *N. Engl. J. Med.* 353, 1899–1911. doi: 10.1056/NEJMoa043802

Gustin, A., Cromarty, R., Schifanella, L., and Klatt, N. R. (2021). Microbial mismanagement: how inadequate treatments for vaginal dysbiosis drive the HIV epidemic in women. *Semin. Immunol.* 51, 101482. doi: 10.1016/j.smim.2021.101482

Hilbert, D. W., Smith, W. L., Paulish-Miller, T. E., Chadwick, S. G., Toner, G., Mordechai, E., et al. (2016). Utilization of molecular methods to identify prognostic markers for recurrent bacterial vaginosis. *Diagn. Microbiol. Infect. Dis.* 86, 231–242. doi: 10.1016/j.diagmicrobio.2016.07.003

Javed, A., Parvaiz, F., and Manzoor, S. (2019). Bacterial vaginosis: An insight into the prevalence, alternative treatments regimen and it's associated resistance patterns. *Microb. Pathog.* 127, 21–30. doi: 10.1016/j.micpath.2018.11.046

Jones, A. (2019). Bacterial vaginosis: A review of treatment, recurrence, and disparities. *J. Nurse Pract.* 15, 420–423. doi: 10.1016/j.nurpra.2019.03.010

Kaur, H., Merchant, M., Haque, M. M., and Mande, S. S. (2020). Crosstalk between female gonadal hormones and vaginal microbiota across various phases of women's gynecological lifecycle. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.00551

Lynch, T., Peirano, G., Lloyd, T., Read, R., Carter, J., Chu, A., et al. (2019). Molecular diagnosis of vaginitis: comparing quantitative PCR and microbiome profiling approaches to current microscopy scoring. *J. Clin. Microbiol.* 57, e00300–19. doi: 10.1128/JCM.00300-19

Malaguti, N., Bahls, L. D., Uchimura, N. S., Gimenes, F., and Consolaro, M. E. L. (2015). Sensitive detection of thirteen bacterial vaginosis-associated agents using multiplex polymerase chain reaction. *BioMed. Res. Int.* 2015, 645853. doi: 10.1155/2015/645853

Matu, M. N., Orinda, G. O., Njagi, E. N. M., Cohen, C. R., and Bukusi, E. A. (2010). *In vitro* inhibitory activity of human vaginal lactobacilli against pathogenic bacteria associated with bacterial vaginosis in Kenyan women. *Anaerobe* 16, 210–215. doi: 10.1016/j.anaerobe.2009.11.002

Menard, J. P., Fenollar, F., Henry, M., Bretelle, F., and Raoult, D. (2008). Molecular quantification of Gardnerella vaginalis and Atopobium vaginae loads to predict bacterial vaginosis. *Clin. Infect. Dis.* 47, 33–43. doi: 10.1086/588661

Mondal, A. S., Sharma, R., and Trivedi, N. (2023). Bacterial vaginosis: A state of microbial dysbiosis. *Med. Microecol.* 16, 100082. doi: 10.1016/j.medmic.2023.100082

Muñoz-Barreno, A., Cabezas-Mera, F., Tejera, E., and MaChado, A. (2021). Comparative effectiveness of treatments for bacterial vaginosis: A network meta-analysis. *Antibiot (Basel Switzerland)* 10, 978. doi: 10.3390/antibiotics10080978

Muzny, C. A., Cerca, N., Elnaggar, J. H., Taylor, C. M., Sobel, J. D., and van der Pol, B. (2023). State of the art for diagnosis of bacterial vaginosis. *J. Clin. Microbiol.* 61, e00837-22. doi: 10.1128/jcm.00837-22

Muzny, C. A., Łaniewski, P., Schwebke, J. R., and Herbst-Kralovetz, M. M. (2020). Host-vaginal microbiota interactions in the pathogenesis of bacterial vaginosis. *Curr. Opin. Infect. Dis.* 33, 59–65. doi: 10.1097/QCO.0000000000000620

Naidoo, K., Abbai, N., Tinarwo, P., and Sebitloane, M. (2022). BV associated bacteria specifically BVAB 1 and BVAB 3 as biomarkers for HPV risk and progression of cervical neoplasia. *Infect. Dis. Obstet. Gynecol.* 2022, 9562937. doi: 10.1155/2022/9562937

Navarro, S., Abla, H., Delgado, B., Colmer-Hamood, J. A., Ventolini, G., and Hamood, A. N. (2023). Glycogen availability and pH variation in a medium simulating vaginal fluid influence the growth of vaginal Lactobacillus species and Gardnerella vaginalis. *BMC Microbiol.* 23, 1–20. doi: 10.1186/s12866-023-02916-8

Nugent, R. P., Krohn, M. A., and Hillier, S. L. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J. Clin. Microbiol.* 29, 297–301. doi: 10.1128/jcm.29.2.297-301.1991

Onderdonk, A. B., Delaney, M. L., and Fichorova, R. N. (2016). The human microbiome during bacterial vaginosis. *Clin. Microbiol. Rev.* 29, 223–238. doi: 10.1128/CMR.00075-15

Osei Sekyere, J., Maningi, N. E., and Fourie, P. B. (2020). Mycobacterium tuberculosis, antimicrobials, immunity, and lung–gut microbiota crosstalk: current updates and emerging advances. *Ann. N. Y Acad. Sci.* 1467, 21–47. doi: 10.1111/nyas.14300

Osei Sekyere, J., Oyenihi, A. B., Trama, J., and Adelson, M. E. (2023). Species-specific analysis of bacterial vaginosis-associated bacteria. *Microbiol. Spectr.* 11, e04676-22. doi: 10.1128/spectrum.04676-22

Pacha-Herrera, D., Erazo-Garcia, M. P., Cueva, D. F., Orellana, M., Borja-Serrano, P., Arboleda, C., et al. (2022). Clustering analysis of the multi-microbial consortium by lactobacillus species against vaginal dysbiosis among Ecuadorian women. *Front. Cell Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.863208

Peebles, K., Velloza, J., Balkus, J. E., McClelland, R. S., and Barnabas, R. V. (2019). High global burden and costs of bacterial vaginosis: A systematic review and meta-analysis. *Sex. Transm. Dis.* 46, 304–311. doi: 10.1097/OLQ.0000000000000972

Powell, A. M., Sarria, I., and Goje, O. (2023). Microbiome and vulvovaginitis. *Obstet. Gynecol. Clin. North Am.* 50, 311–326. doi: 10.1016/j.ogc.2023.02.005

Roachford, O. S. E., Alleyne, A. T., and Nelson, K. E. (2022). Insights into the vaginal microbiome in a diverse group of women of African, Asian and European ancestries. *PeerJ* 10, e14449. doi: 10.7717/peerj.14449

Schmidt, K., Cybulski, Z., Roszak, A., Grabiec, A., Talaga, Z., Urbański, B., et al. (2015). Combination of microbiological culture and multiplex PCR increases the range of vaginal microorganisms identified in cervical cancer patients at high risk for bacterial vaginosis and vaginitis. *Ginekol Pol.* 86, 328–334. doi: 10.17772/gp/2417

Shen, L., Zhang, W., Yuan, Y., Zhu, W., and Shang, A. (2022). Vaginal microecological characteristics of women in different physiological and pathological period. *Front. Cell Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.959793

Sobel, J. D., and Vempati, Y. S. (2024). Bacterial vaginosis and vulvovaginal candidiasis pathophysiologic interrelationship. *Microorganisms* 12, 108. doi: 10.3390/microorganisms12010108

Swidsinski, S., Moll, W. M., and Swidsinski, A. (2023). Bacterial vaginosis—Vaginal polymicrobial biofilms and dysbiosis. *Dtsch Arztebl Int.* 120, 347. doi: 10.3238/arztebl.m2023.0090

Unemo, M., Bradshaw, C. S., Hocking, J. S., de Vries, H. J. C., Francis, S. C., Mabey, D., et al. (2017). Sexually transmitted infections: challenges ahead. *Lancet Infect. Dis.* 17, e235–e279. doi: 10.1016/S1473-3099(17)30310-9

van den Munckhof, E. H. A., van Sitter, R. L., Boers, K. E., Lamont, R. F., te Witt, R., le Cessie, S., et al. (2019). Comparison of Amsel criteria, Nugent score, culture and two CE-IVD marked quantitative real-time PCRs with microbiota analysis for the diagnosis of bacterial vaginosis. *Eur. J. Clin. Microbiol. Infect. Dis.* 38, 959–966. doi: 10.1007/s10096-019-03538-7

Vodstrcil, L. A., Muzny, C. A., Plummer, E. L., Sobel, J. D., and Bradshaw, C. S. (2021). Bacterial vaginosis: drivers of recurrence and challenges and opportunities in partner treatment. *BMC Med.* 19, 1–12. doi: 10.1186/s12916-021-02077-3

Watkins, E., Chow, C. M., Lingohr-Smith, M., Lin, J., Yong, C., Tangirala, K., et al. (2024). Treatment patterns and economic burden of bacterial vaginosis among commercially insured women in the USA. *J. Comp. Eff Res.* 13, e230079. doi: 10.57264/cer-2023-0079

Wei, X., Tsai, M.-S., Liang, L., Rand, L., Snyder, M., Jiang, C., et al. (2024). Vaginal microbiomes show ethnic evolutionary dynamics and positive selection of Lactobacillus adhesins driven by a long-term niche-specific process. *CellReports* 43, 114078. doi: 10.1016/j.celrep.2024.114078

Weyers, S., Verstraelen, H., Gerris, J., Monstrey, S., Lopes dos Santos Santiago, G., Saerens, B., et al. (2009). Microflora of the penile skin-lined neovagina of transsexual women. *BMC Microbiol.* 9, 102. doi: 10.1186/1471-2180-9-102

Wu, S., Hugerth, L. W., Schuppe-Koistinen, I., and Du, J. (2022). The right bug in the right place: opportunities for bacterial vaginosis treatment. *NPJ Biofilms Microbiomes* 81 2022, 8:1–811. doi: 10.1038/s41522-022-00295-y

Zhang, Q., Chen, R., Li, M., Zeng, Z., Zhang, L., and Liao, Q. (2022). The interplay between microbiota, metabolites, immunity during BV. *Med. Microecol* 11, 100049. doi: 10.1016/j.medmic.2021.100049

Zheng, N., Guo, R., Wang, J., Zhou, W., and Ling, Z. (2021). Contribution of lactobacillus iners to vaginal health and diseases: A systematic review. *Front. Cell Infect. Microbiol.* 11. doi: 10.3389/fcimb.2021.792787

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 11 July 2024
DOI 10.3389/fcimb.2024.1416884



OPEN ACCESS

EDITED BY
Jiemin Zhou,
Vision Medicals Co, Ltd, China

REVIEWED BY
Jiajia Zheng,
Peking University, China
Yang Jiao,
Second Military Medical University, China
Yong Liu,
First Affiliated Hospital of Jilin University, China

*CORRESPONDENCE
Hangyong He
yonghang2004@sina.com

†These authors have contributed equally to this work

RECEIVED 13 April 2024
ACCEPTED 26 June 2024
PUBLISHED 11 July 2024

CITATION
Zhang D, Fan B, Yang Y, Jiang C, An L, Wang X and He H (2024) Target next-generation sequencing for the accurate diagnosis of *Parvimonas micra* lung abscess: a case series and literature review. *Front. Cell. Infect. Microbiol.* 14:1416884. doi: 10.3389/fcimb.2024.1416884



# Target next-generation sequencing for the accurate diagnosis of *Parvimonas micra* lung abscess: a case series and literature review

Dongmei Zhang[1†], Boyang Fan[1,2†], Yuan Yang[1], Chunguo Jiang[1], Li An[1], Xue Wang[1] and Hangyong He[1,3*]

[1]Department of Respiratory and Critical Care Medicine, Beijing Institute of Respiratory Medicine and Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China, [2]Department of Respiratory, Beijing Huairou Hospital, Beijing, China, [3]Department of Pulmonary and Critical Care Medicine, Center of Respiratory Medicine, China-Japan Friendship Hospital, National Center for Respiratory Medicine, National Clinical Research Center for Respiratory Diseases, Beijing, China

**Background:** *Parvimonas micra* (*P. micra*) has been identified as a pathogen capable of causing lung abscesses; however, its identification poses challenges due to the specialized culture conditions for anaerobic bacterial isolation. Only a few cases of lung abscesses caused by *P. micra* infection have been reported. Therefore, we describe the clinical characteristics of lung abscesses due to *P. micra* based on our case series.

**Methods:** A retrospective analysis was conducted on eight patients who were diagnosed with lung abscesses attributed to *P. micra*. Detection of *P. micra* was accomplished through target next-generation sequencing (tNGS). A systematic search of the PubMed database using keywords "lung abscess" and "*Parvimonas micra/Peptostreptococcus micros*" was performed to review published literature pertaining to similar cases.

**Results:** Among the eight patients reviewed, all exhibited poor oral hygiene, with four presenting with comorbid diabetes. Chest computed tomography (CT) showed high-density mass shadows with necrosis and small cavities in the middle. Bronchoscopic examination revealed purulent sputum and bronchial mucosal inflammation. Thick secretions obstructed the airway, leading to the poor drainage of pus, and the formation of local abscesses leading to irresponsive to antibiotic therapy, which finally protracted recovery time. *P. micra* was successfully identified in bronchoalveolar lavage fluid (BALF) samples from all eight patients using tNGS; in contrast, sputum and BALF bacterial cultures yielded negative results, with *P. micra* cultured from only one empyema sample. Following appropriate antibiotic therapy, seven patients recovered. In previously documented cases, favorable outcomes were observed in 77.8% of individuals treated with antibiotics and 22.2% were cured after surgical interventions for *P. micra* lung abscesses.

**Conclusions:** This study enriches our understanding of the clinical characteristics associated with lung abscesses attributed to *P. micra*. Importantly, tNGS has emerged as a rapid and effective diagnostic test in scenarios where traditional

sputum cultures are negative. Encouragingly, patients with lung abscesses caused by *P. micra* infection exhibit a favorable prognosis with effective airway clearance and judicious anti-infective management.

KEYWORDS

*Parvimonas micra*, lung abscess, target next-generation sequencing, TNGS, diagnosis

## Introduction

Primary lung abscesses typically result from infections leading to necrosis and cavitation of the lung tissue. Aspiration of oral-derived anaerobic bacteria represents the predominant etiology of primary lung abscesses (Bansal et al., 2013). Individuals with periodontal disease are at an elevated risk of developing anaerobic lung abscesses. Since anaerobic organisms necessitate specific culture conditions, conventional sputum culture often yields negative results in the detection of anaerobic bacteria. Anaerobic organisms were isolated only in a limited number of lung abscess cases (Yazbeck et al., 2014).

*Parvimonas micra* (*P. micra*) is a gram-positive anaerobic coccus prevalent in the oral cavity, particularly in the gingival crevices. *P. micra* has been associated with aspiration pneumonia (Yu et al., 2021), lung abscesses (Zhang et al., 2021), and empyema (Vilcarromero et al., 2023). Nevertheless, clinical insights into *P. micra*-related lung abscesses remain rare.

Given the challenges in culturing and identifying anaerobic bacteria due to their vulnerability to atmospheric exposure, innovative approaches utilizing next-generation sequencing of metagenomes (NGS) have shown promise in enhancing the positive detection rate of anaerobic bacteria (Zhang et al., 2021).

In this study, we present findings from eight cases of lung abscess caused by *P. micra*, with the pathogenic bacterium identified through genomic sequencing. Additionally, a systematic review of prior cases detailed in the published literature was conducted to delineate the clinical features of lung abbesses attributed to this pathogen, which may be unfamiliar to pulmonary healthcare providers.

## Patients and methods

### Patients

A retrospective analysis was conducted on patients who were diagnosed with lung abscesses and admitted to Beijing Chao-Yang Hospital between July 2022 and December 2023. The results from bronchoalveolar lavage fluid (BALF) pathogen testing using target next-generation sequencing (tNGS) indicated *P. Micra* as the predominant bacterium, consistent with the clinical presentation. Two physicians independently confirmed the association of lung abscess with *P. micra*. Comprehensive patient medical records, laboratory tests, examinations, and treatment modalities were compiled for review.

### tNGS testing

Bronchoscopies were performed within 5 days of hospital admission, with BALF samples obtained from the segmental bronchus corresponding to the identified lesions. Immediately following collection, 5 ml of BALF was aseptically preserved at −20°C and transported to Beijing KingMed Diagnostics Laboratory within 4 h. Cellular material within the samples was enriched, lysed, and subjected to DNA extraction from 500 µl of fluid as per established protocols. Subsequent to polymerase chain reaction amplification and purification, sequencing was performed utilizing a genetic sequencing platform (KM MiniSeqDx-CN) and an automatic nucleic acid-protein analyzer (Qsep100, BiOptic, Taiwan). Sequence data interpretation was facilitated by the Pathogenic Microbial Data Analysis and Management System 1.0 (KingMed Diagnostics, Co., Ltd., China), which comprises a bacterial minimal genome database consisting of 198 respiratory pathogens.

### Literature review

A systematic search of the PubMed database was conducted to identify journal articles, employing the search terms “lung abscess” and “*Parvimonas micra/Peptostreptococcus micros*”. The inclusion criteria included papers published from 01/01/1980, to 31/12/2023. Five articles documenting pulmonary abscesses associated with *P. micra*, with a collective total of nine patients, were contained in the review.

## Results

### Characteristics

The general characteristics of the eight patients are outlined in Table 1. Their ages ranged from 28 to 83 years, with four patients

10.3389/fcimb.2024.1416884

TABLE 1 Clinical information of patients with *P. micra* lung abscess.

| Case | Age | Gender | BMI | Underlying disease | Smoking/drinking | Oral hygiene | Symptoms | Time* | WBC $10^9$/L | NEU $10^9$/L | Hgb g/L | PCT ng/ml | CRP mg/dl |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 72 | M | 20.28 | DM | N/N | Periodontal disease, dental calculus, false tooth | Cough sputum, | 30 days | 8.57 | 6.1 | 116 | <0.05 | 4.06 |
| 2 | 67 | F | 21.30 | – | N/N | dental calculus, teeth loss | Fever, cough, and sputum | 50 days | 5.3 | 3.41 | 117 | <0.05 | 0.34 |
| 3 | 83 | M | 27.22 | DM, HPT | Y/Y | Dental calculus, decayed teeth, teeth loss | Fever, cough and sputum, sore throat | 2 months | 10.56 | 7.6 | 109 | 0.06 | 11.5 |
| 4 | 54 | F | 24.43 | – | N/N | Decayed tooth | Fever, cough, and sputum | 2 months | 6.99 | 4.85 | 94 | <0.05 | 0.39 |
| 5 | 49 | M | 27.68 | DM, HPT | Y/N | Abscess tooth, dental calculus | Fever, cough and sputum, chest pain | 15 days | 16.61 | 13.86 | 116 | 0.34 | 35.8 |
| 6 | 57 | F | 26.03 | DM, HPT, CVD | N/N | Periodontal disease, false tooth, tooth loss | Fever, cough, and sputum | 20 days | 15.23 | 13.38 | 107 | 0.39 | 12.7 |
| 7 | 59 | M | 24.2 | – | N/Y | Periodontal disease, dental calculus | Cough and sputum, chest pain, hemoptysis | 70 days | 6.56 | 4.32 | 127 | <0.05 | 5.55 |
| 8 | 28 | M | 25.25 | – | N/Y | Decayed tooth | Cough, dyspnea | 14 days | 17.61 | 15.54 | 106 | 0.46 | 4.21 |

M, male; F, female; DM, diabetes mellitus; HPT, hypertension; CVD, cerebrovascular disease; N, no; Y, yes.
*Time from clinical onset to diagnosis.

having diabetes and none presenting HIV infection or receiving immunosuppressive treatment. The median duration from symptom onset to definitive diagnosis was approximately 40 days. Notably, all eight patients exhibited poor oral hygiene, with manifestations including periodontal disease and dental disease.

Primary symptoms were non-specific, yet common manifestations included fever (n = 5), cough (n = 8), and sputum production (n = 7). Additional complaints included sore throat, dyspnea, chest pain, and hemoptysis.

## Laboratory tests, chest CT, and bronchoscopy

Upon admission, routine blood tests revealed a median white blood cell count of $9.57 \times 10^9$/L, with half of the patients exhibiting normal procalcitonin (PCT) levels.

Chest CT findings demonstrated large-scale consolidation with necrosis and small cavities (Figure 1), devoid of classical air-fluid levels. The boundaries of the focusing area were unclear and blurred. The lesions were located in the right upper lobe (n = 2), left upper lobe (n = 2), right middle lobe (n = 1), and inferior lobe of the right lung (n = 3). One patient was complicated with liver abscesses and empyema.

Bronchoscopy revealed purulent sputum (Figure 1) and inflammatory changes in the bronchial mucosa at the affected sites.

CT-guided pulmonary biopsy was performed on three patients to differentiate from lung cancer. Pathology indicated inflammation of the lung tissue.

## Microbiological examinations

Although gram-positive cocci were identified in five cases via sputum or BALF smears, bacterial cultures yielded negative results for all eight patients. However, tNGS analysis of BALF samples confirmed *P. micra* as the predominant bacterium. Anaerobic culture of the pleural effusion identified *P. micra* in one patient complicated with concomitant empyema. Additionally, tNGS reports indicated the presence of other bacteria including *Fusbacterium nucleatum* (n = 5), *Klebsiella pneumoniae* (n = 2), *Escherichia coli* (n = 1), *Streptococcus pneumoniae* (n = 1), and *Acinetobacter baumannii* (n = 1).

## Treatment and outcomes

Following diagnosis, patients received tailored anti-infection therapies such as piperacillin–tazobactam (n = 3), ampicillin–sulbactam (n = 2), moxifloxacin (n = 1), meropenem (n = 1), and imipenem–cilastatin (n = 1). Mucolytic drugs and airway clearance techniques were also employed to facilitate sputum expectoration.

Their symptoms improved, and chest CT scans indicated resolution of the lung abscesses. Subsequent to initial therapy, oral antibiotics were administered for 2 to 5 months. Seven patients achieved satisfactory outcomes, with almost no

symptoms and almost completely absorbed lesions on chest CT. Two months later, one patient developed massive hemoptysis attributed to secondary *Enterococcus faecium* infection, leading to fatal pneumonia and sepsis.

## Literature review

A review of the literature identified other nine reported cases (Table 2), revealing cough (77.8%) and expectoration (77.8%) as the most common symptoms. Additional data indicated smoking history, alcohol consumption, periodontal conditions, and dental issues. The median time taken from clinical onset to make a diagnosis was 2 months. CT scan showed irregular mass shadows which had no obvious improvement after short-term initial empirical treatment. It is difficult to distinguish from malignant lesions. Thus, CT-guided percutaneous lung biopsies were performed in 55.6% of these patients. In published cases, metagenomics next-generation sequencing (mNGS) facilitated pathogen identification in more than half of the cases. The details of coinfections and treatment are shown in Table 3. A majority of *P. micra* lung abscess cases reported in the literature also displayed favorable outcomes, with 77.8% of patients cured via medication and 22.2% benefitting from surgical intervention.

# Discussion

The majority of lung abscesses are believed to be caused by the aspiration of anaerobic bacteria from the oral cavity, particularly from gingival crevices. *P. micra* is one of the most prevalent anaerobic microorganisms in the human oral cavity (Badri et al., 2019). However, due to challenges in laboratory culture, there are limited published data on the clinical characteristics of *P. micra* lung abscess, mainly in the last 5 years (Yun et al., 2019; Yang and Su, 2021; Zhang et al., 2021; Fukushima et al., 2023; Zhijun et al., 2023).

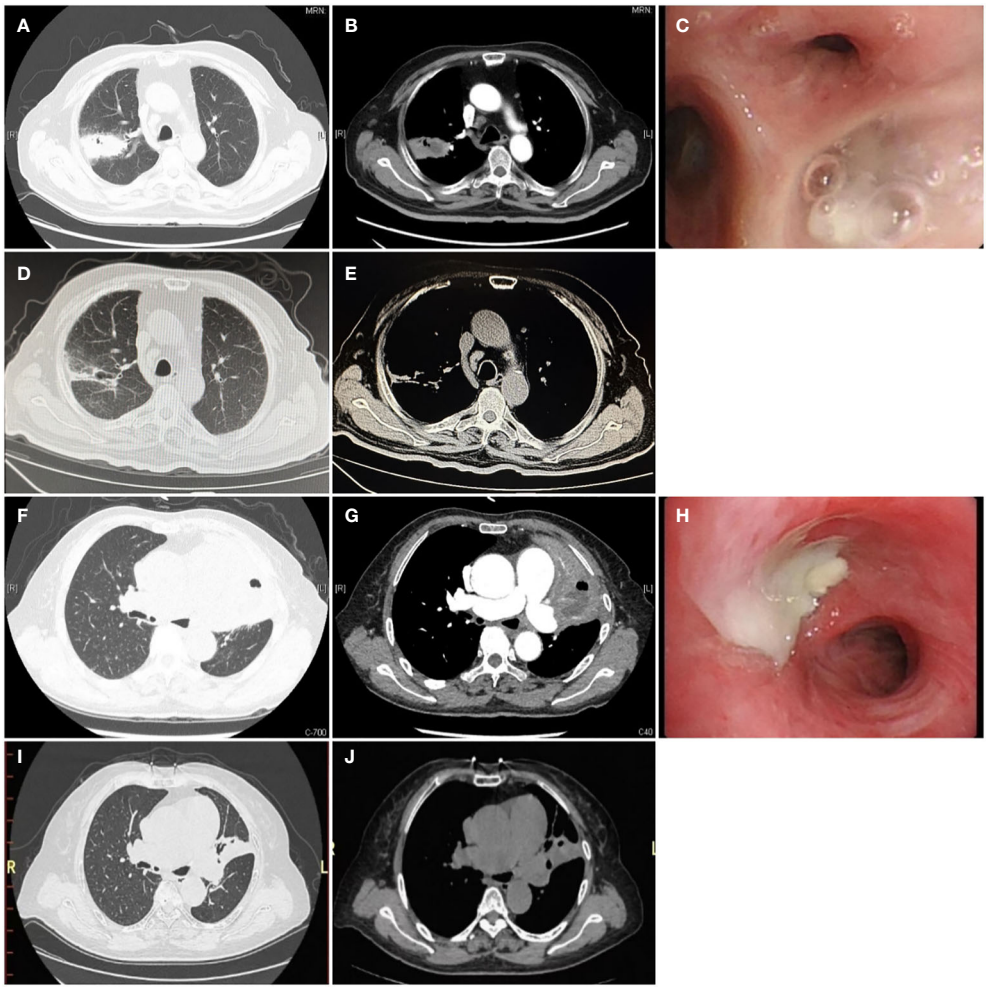



FIGURE 1
Chest computed tomography and bronchoscopy of *P. micra* lung abscess Computed tomography is displayed in the lung window **(A, D, F, I)**, soft tissue window **(B, E, G, J)**, and bronchoscopy **(C, H)**. **(A, B, F, G)** The initial CT scan showed large consolidations with necrosis and small cavities (white →) without a clear air-liquid level, which can be described as a "bubble levitation sign". **(C, H)** The bronchial segment was obstructed by purulent secretions (→). **(D, E, I, J)** After treatment for 3 months, the follow-up chest CT showed significant absorption of the lesions.

10.3389/fcimb.2024.1416884

TABLE 2 Clinical characteristics of *P. micra* in case series and in the literature.

| Characteristic | | Our case series (n = 8) | Literature cases (n = 9) |
|---|---|---|---|
| Age | Year median (range) | 55.5 (28-82) | 61.5 (46-82) |
| Gender | Male/female | 5/3 | 8/1 |
| Time from clinical onset to diagnosis (days) | | 40 (14-70) | 60 (4-260) |
| Symptom No. (%) | Fever | 5 (62.5) | 3 (33.3) |
| | Cough | 8 (100) | 7 (77.8) |
| | Expectoration | 7 (87.5) | 7 (77.8) |
| | Chest pain | 2 (25) | 3 (33.3) |
| | Dyspnea | 1 (12.5) | 2 (22.2) |
| | Hemoptysis | 1 (12.5) | 3 (33.3) |
| | Sore throat | 1 (12.5) | 0 |
| | Circulatory failure | 0 | 1 (11.1) |
| Method pathogen detection | BALF NGS | 8 (100) | 3 (33.3) |
| | Lung tissue NGS | 0 | 2 (22.2) |
| | BALF culture | 0 | 1 (11.1) |
| | Blood culture | 0 | 2 (22.2) |
| | Abscess culture | 1 (12.5) | 1 (11.1) |
| Underlying health status | Diabetes | 4 (50) | 1 (11.1) |
| | Hypertension | 3 (37.5) | 3 (33.3) |
| | Cerebrovascular disease | 1 (12.5) | 0 |
| | Bipolar disorder | 0 | 1 (11.1) |
| | Atrial fibrillation | 0 | 1 (11.1) |
| | Heart valve replacement | 0 | 1 (11.1) |
| | Healthy | 4 (50) | 3 (33.3) |
| Location of infection | Right upper lung lobe | 2 (25) | 3 (33.3) |
| | Right middle lung lobe | 1 (12.5) | 2 (22.2) |
| | Right lower lung lobe | 3 (37.5) | 0 |
| | Left upper lung lobe | 2 (25) | 3 (33.3) |
| | Left lower lung lobe | 0 | 2 (22.2) |
| Bronchoscopy | | 8 (100) | 7 (77.8) |
| CT-guided percutaneous lung puncture | | 3 (37.5) | 5 (55.6) |
| Prognosis | Cured with medicines | 7 (87.5) | 7 (77.8) |
| | Surgical resection | 0 | 2 (22.2) |
| | All-cause mortality | 1 (12.5) | 0 |

BALF, bronchoalveolar lavage fluid; NGS, next-generation sequencing.

Anaerobic lung infections are often associated with poor oral hygiene conditions and inadequate airway protection (Hata et al., 2020). Zhang et al (Zhang et al., 2021). reported that patients with a long smoking history and poor oral hygiene are susceptible to *P. micra* lung abscesses. However, in our study, smoking was not identified as a significant risk factor, as only two male patients had a long history of smoking, whereas the other patients were non-smokers. Lai et al. reported four cases of childhood pneumonia and abscesses caused by oral obligate anaerobes, primarily *P. micra*, where poor oral hygiene was a crucial risk factor (Zhijun et al., 2023). Therefore, poor oral health is an important risk factor for *P. micra*–induced lung abscesses. In addition, alcoholism should also be noted as a common predisposing condition for aspiration (Kuhajda et al., 2015).

10.3389/fcimb.2024.1416884

TABLE 3 Diagnosis, treatment, and prognosis of patients with *P. micra* lung abscess.

| Case | Age | Gender | Lesion site | Biopsy method | Diagnostic method | Coinfection | Treatment | Prognosis |
|---|---|---|---|---|---|---|---|---|
| 1 | 72 | M | Right middle lung lobe | N | BALF tNGS | *K. pneumoniae* | Ampicillin sulbactam for 10 days, and amoxicillin clavulanate 500/125 mg every 8 h for 3 months | Cured with medicines |
| 2 | 67 | F | Left upper lung lobe | N | BALF tNGS | N | Meropenem 1.0 g every 8 h for 11 days, piperacillin tazobactam 4.5 g every 8 h for 14 days, moxifloxacin 400 mg once a day for 4 months | Cured with medicines |
| 3 | 83 | M | Right upper lung lobe | CT-GPLP | BALF tNGS | *F. nucleatum* | Meropenem 1.0 g every 12 h for 8 days, piperacillin tazobactam 4.5 g every 8 h for 6 days, moxifloxacin 400 mg once a day for over 5 months | Cured with medicines |
| 4 | 54 | F | Right lower lung lobe | CT-GPLP | BALF tNGS Lung tissue NGS | *A. baumannii* | Piperacillin tazobactam 4.5 g every 8 h for 15 days, moxifloxacin 400 mg once a day for 15 days, amoxicillin and metronidazole for 4 months | Cured with medicines |
| 5 | 49 | M | Right lower lung lobe | N | BALF tNGS | *P. endodontalis* | Ampicillin intravenous for 12 days and amoxicillin clavulanate for 2.5 months | Cured with medicines |
| 6 | 57 | F | Left upper lung lobe | CT-GPLP EBUS-TBNA | BALF tNGS | N | Ceftazidime 2.0 g every 8 h for 17 days, moxifloxacin 400 mg once a day for 5 months | Cured with medicines |
| 7 | 59 | M | Right upper lung lobe | TBLB | BALF tNGS | *P. endodontalis* *F. nucleatum* | Meropenem 1.0 g every 8 h for 17 days, moxifloxacin 400 mg once a day for 1.5 months, Imipenem cilastatin 1.0 g every 8 h for 8 days. | Death from *Enterococcus fecal sepsis* |
| 8 | 28 | M | Right lower lung lobe | N | BALF tNGS, culture of pleural effusion | *F. nucleatum* | Meropenem 1.0 g every 8 h for 12 days, Piperacillin tazobactam 4.5 g every 8 h for 20 days, moxifloxacin for 3 months | Cured with medicines |
| 9 (Fukushima et al., 2023) | 57 | F | Right upper lung lobe | Surgical resection | Blood culture | N | Ceftriaxone, vancomycin, and azithromycin for 10 days | Cured with resection |
| 10 (Yang and Su, 2021) | 49 | M | Left lower lung lobe | N | Blood culture | *A. odontolyticus* | Ceftriaxone and clindamycin Ampicillin sulbactam Amoxicillin/clavulanate 6 months | Cured with medicines |
| 11 (Zhang et al., 2021) | 61 | M | Left upper lung lobe | CT-GPLP | Lung tissue NGS | N | Metronidazole 200 mg every 8 h for 3 months | Cured with medicines |
| 12 (Zhang et al., 2021) | 81 | M | Right middle lung lobe | EBUS-TBNA | BALF NGS | *F. nucleatum* | Amoxicillin clavulanate 375 mg every 8 h for 2 months | Cured with medicines |
| 13 (Zhang et al., 2021) | 46 | M | Left lower lung lobe | CT-GPLP | BALF NGS | *F. nucleatum*, *S. constellatus* | Moxifloxacin 400 mg once a day for 4 months | Cured with medicines |
| 14 (Zhang et al., 2021) | 82 | M | Right upper lung lobe | CT-GPLP | Lung tissue NGS | *P. intermedia*, *S. constellatus* | Metronidazole 200 mg every 8 h for 5 months | Cured with medicines |
| 15 (Zhang et al., 2021) | 62 | M | Left upper lung lobe | CT-GPLP | BALF NGS | *F. nucleatum* | Amoxicillin–clavulanate 375 mg every 8 h for 4 months | Cured with resection |

*(Continued)*

TABLE 3 Continued

| Case | Age | Gender | Lesion site | Biopsy method | Diagnostic method | Coinfection | Treatment | Prognosis |
|---|---|---|---|---|---|---|---|---|
| 16 (Yun et al., 2019) | 62 | M | Left upper lobe | N | BALF culture | N | Piperacillin–tazobactam 4.5 g every 6 h for 7 days, amoxicillin–clavulanic 875/125mg every 12 h for several weeks | Cured with medicines |
| 17 (Gorospe et al., 2014) | 67 | M | Right upper lobe and middle lobe | CT-GPLP | Chest wall abscess culture | N | Levofloxacin for 7 days Clindamycin 600 mg every 8 h for 8 weeks | Cured with medicines |

CT-GPLP, computed tomography-guided percutaneous lung puncture; EBUS-TBNA, endobronchial ultrasound-guided transbronchial needle aspiration; BALF, bronchoalveolar lavage fluid; NGS, next generation sequencing. N, no; M, male; F, female

Among our patients, three male patients had a history of regular alcohol consumption.

The symptoms observed in these patients were non-specific. All patients presented with a productive cough with sputum, and most experienced fever. Radiologically, chest CT scans revealed mass of lung consolidation with liquefactive necrosis and small cavities within the consolidation, without a distinct liquid–gas plane. Given the presence of small cavities with thick walls of consolidations on lung CT scan, along with non-specific symptoms, many patients underwent lung puncture biopsy for the differentiation of lung cancer. The lung biopsy pathology from these lung puncture biopsies revealed chronic inflammation of the lung tissue (Zhang et al., 2021). Gorospe et al. reported a case of a chest wall abscess of *P. micra* following CT-guided needle lung biopsy of the right lung consolidation (Gorospe et al., 2014).

In our case series, bronchoscopy revealed sticky secretion plugging in the bronchi as a feature of *P. micra* infection. For patients with lung consolidation and necrosis, bronchoscopy could be performed initially for pathogenic analysis. When empyema secretion plugging in the bronchial region was detected, without neoplasm, *P. micra* was isolated, and symptoms and radiological findings improved following targeted antibiotic therapy, percutaneous lung biopsy could be avoided.

The mNGS markedly improves pathogenic diagnosis, but it is expensive and involves procedural DNA and RNA sequencing, respectively. Target NGS is a more economical technology with a cost of approximately one-fifth of that of mNGS, with the similar advantages of rapid speed and high accuracy. The target detection of 198 pathogens include 80 bacteria, 79 DNA and RNA viruses, 32 fungi, and 7 mycoplasmas and chlamydia, which account for 95% respiratory infections (Li et al., 2022). Clinical data have shown that tNGS is effective and economical for diagnosis of respiratory infection (Li et al., 2022). However, NGS technology cannot achieve antibiotic sensitivity of the pathogen, and anaerobic culture remains important, particularly for determining antibiotic resistance. Additionally, as NGS often detects multiple bacterial species, clinicians must carefully ascertain the true pathogenic bacteria, by integrating patients' symptoms, laboratory tests, and radiological findings. Compared with traditional culture, tNGS had a shorter turnaround time for positive pathogen detection (1 day vs. 4 days). Finally, for detected pathogens with high NGS sequencing read numbers, especially in RNA sequencing tests, this approach could be helpful for the differentiation of colonization and infection.

tNGS is only a method for detecting microbes rather than a diagnostic method. However, tNGS is designed to target highly suspected microbes when they are commonly recognized as pathogens that cause lower respiratory tract infections. Furthermore, after its routine application in the clinical diagnostic process in clinical practice for pathogen identification in our hospital, its role as an examination for diagnostic purposes becomes increasingly important. Therefore, we could have a positive view of its promising diagnostic usage in the future. Finally, the results from tNGS could not be held as the only diagnostic evidence. To establish a precise diagnosis for an infection, tNGS should be combined with other clinical information such as medical history, symptoms, laboratory, and radiological findings.

In bronchoscopy, if bronchial blockage by purulent secretions is observed, which leads to local hypoxia and creates conditions for anaerobic bacterial growth, promoting sputum excretion and maintaining airway patency are critical treatment procedures on the base of drug therapy. Chest physiotherapy by respiratory therapists, including the use of devices such as the Acapella valve and high-frequency chest wall oscillation devices, are helpful in clearing airway secretions (Polverino et al., 2017)

*P. micra* is susceptible to many antibiotics, including penicillin G, ampicillin, cefepime, vancomycin, and metronidazole (Boattini et al., 2018; Fukushima et al., 2023). In the case of lung abscess, intravenous antibiotics were administered for 2 to 4 weeks, followed by oral antibiotic treatment for 2 to 3 months until the chest radiographic lesions resolved. Proper drainage helps with absorption and hastens recovery. Although lung abscesses associated with *P. micra* have rarely been reported, it has a benign prognosis, as most patients recover after antibiotic treatment.

In patients who do not respond to antibiotic therapy, catheter drainage or surgical resection should still be considered (Herth et al., 2005). Both percutaneous tube drainage (Lee et al., 2022) and endoscopic drainage (Unterman et al., 2017) have been shown to effectively reduce abscess size and improve clinical outcomes. However, catheter placement in *P. micra*-related lung abscesses with small cavities can be challenging (Herth et al., 2005). Surgical

resection may be necessary if antibiotic therapy is ineffective or in cases of life-threatening hemoptysis. In cases reported previously, two patients were cured after surgery.

This study has several limitations. First, this study has a small sample size, which may lead to bias. Furthermore, sputum or BALF cultures were negative for all patients. Finally, tNGS results do not cover all pathogens, which may be inaccurate in finding rare coinfected pathogens. To avoid missing important pathogens, retesting or conversion to mNGS would be considered, if the tNGS results not align with the clinical situation or if patients exhibited a poor response to treatment.

## Conclusion

In conclusion, our study systemically reports the characteristics of *P. micra*–related lung abscesses in adults based on the largest number of cases to date. Imaging features included mass consolidation with necrosis without a clear liquid–gas plane on chest CT, and sticky secretion plugging in the bronchial region on bronchoscopy. tNGS is an effective and cost-efficient tool for rapidly detecting pathogens. The lung abscesses caused by *P. micra* have a good prognosis with appropriate treatment. Improving oral health, promoting sputum excretion, and following an appropriate extended course of antibiotic treatment are crucial for recovery from *P. micra*-induced anaerobic lung abscesses.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

## Ethics statement

The studies involving humans were approved by Ethics Review Committee of the Beijing Chao-Yang Hospital, Capital Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

DZ: Validation, Writing – original draft, Conceptualization. BF: Data curation, Investigation, Writing – original draft. YY: Data curation, Investigation, Writing – original draft. CJ: Data curation, Investigation, Writing – review & editing. LA: Conceptualization, Supervision, Writing – review & editing. XW: Data curation, Investigation, Writing – review & editing. HH: Project administration, Supervision, Visualization, Writing – review & editing.

## Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Scientific Technology Support Project of the Xizang Autonomous Region (XZ202201ZY0037G).

## Conflict of interest

The authors declare that the research was conducted in the absence of commercial or financial relationships and without conflicts of interest.

## Publisher's note



## References

Badri, M., Nilson, B., Ragnarsson, S., Senneby, E., and Rasmussen, M. (2019). Clinical and microbiological features of bacteraemia with Gram-positive anaerobic cocci: a population-based retrospective study. *Clin.Microbiol. Infect.* 25, 760.e1–760.e6. doi: 10.1016/j.cmi.2018.09.001

Bansal, M., Khatri, M., and Taneja, V. (2013). Potential role of periodontal infection in respiratory diseases - a review. *J. Med. Life* 6, 244–248.

Boattini, M., Bianco, G., Cavallo, R., and Costa, C. (2018). *Parvimonasmicra* bacteremia following endoscopic retrograde cholangiopancreatography: A new route of infection. *Anaerobe* 54, 136–139. doi: 10.1016/j.anaerobe.2018.09.003

Fukushima, S., Hagiya, H., Naito, H., and Otsuka, F. (2023). Furious lung abscess due to *Parvimonasmicra*. *Respirol Case Rep.* 11, e01161. doi: 10.1002/rcr2.1161

Gorospe, L., Bermudez-Coronel-Prats, I., Gomez-Barbosa, C. F., Olmedo-Garcia, M. E., Ruedas-Lopez, A., and Gomez delOlmo, V. (2014). *Parvimonasmicra* chest wall abscess following transthoracic lung needle biopsy. *Korean J. Intern. Med.* 29, 834–837. doi: 10.3904/kjim.2014.29.6.834

Hata, R., Noguchi, S., Kawanami, T., Yamasaki, K., Akata, K., Ikegami, H., et al. (2020). Poor oral hygiene is associated with the detection of obligate anaerobes in pneumonia. *J. Periodontol.* 91, 65–73. doi: 10.1002/JPER.19-0043

Herth, F., Ernst, A., and Becker, H. D. (2005). Endoscopic drainage of lung abscesses: technique and outcome. *Chest* 127, 1378–1381. doi: 10.1378/chest.127.4.1378

Kuhajda, I., Zarogoulidis, K., Tsirgogianni, K., Tsavlis, D., Kioumis, I., Kosmidis, C., et al. (2015). Lung abscess-etiology, diagnostic and treatment options. *Ann. Transl. Med.* 3, 183. doi: 10.3978/j.issn.2305-5839.2015.07.08

Lee, J. H., Hong, H., Tamburrini, M., and Park, C. M. (2022). Percutaneous transthoracic catheter drainage for lung abscess: a systematic review and meta-analysis. *EurRadiol* 32, 1184–1194. doi: 10.1007/s00330-021-08149-5

Li, S., Tong, J., Liu, Y., Shen, W., and Hu, P. (2022). Targeted next generation sequencing is comparable with metagenomic next generation sequencing in adults with pneumonia for pathogenic microorganism detection. *J. Infect.* 85, e127–e129. doi: 10.1016/j.jinf.2022.08.022

Polverino, E., Goeminne, P. C., McDonnell, M. J., Aliberti, S., Marshall, S. E., Loebinger, M. R., et al. (2017). European Respiratory Society guidelines for the management of adult bronchiectasis. *Eur. Respir. J.* 50, 1700629. doi: 10.1183/13993003.00629-2017

Unterman, A., Fruchter, O., Rosengarten, D., Izhakian, S., Abdel-Rahman, N., and Kramer, M. R. (2017). Bronchoscopic drainage of lung abscesses using a pigtail catheter. *Respiration* 93, 99–105. doi: 10.1159/000453003

Vilcarromero, S., Small, M., Lizarzaburu, A., and Rivadeneyra-Rodriguez, A. (2023). Pleural empyema by *Parvimonasmicra* in an immunocompetent patient: a case report. *Rev. Peruana medicina Exp. y saludpublica* 40, 99–104. doi: 10.17843/rpmesp.2023.401.11956

Yang, H. W., and Su, Y. J. (2021). Cavitary lung mass presenting in an outpatient. *Am. J. Med.* 134, 1113–1114. doi: 10.1016/j.amjmed.2021.02.017

Yazbeck, M. F., Dahdel, M., Kalra, A., Browne, A. S., and Pratter, M. R. (2014). Lung abscess: update on microbiology and management. *Am. J. Ther.* 21, 217–221. doi: 10.1097/MJT.0b013e3182383c9b

Yu, Q., Sun, L., Xu, Z., Fan, L., and Du, Y. (2021). Severe pneumonia caused by Parvimonasmicra: a case report. *BMC Infect. Dis.* 21, 364. doi: 10.1186/s12879-021-06058-y

Yun, S. S., Cho, H. S., Heo, M., Jeong, J. H., Lee, H. R., Ju, S., et al. (2019). Lung abscess by Actinomycesodontolyticus and Parvimonasmicra co-infection presenting as acute respiratory failure: A case report. *Med. (Abingdon)* 98, e16911. doi: 10.1097/MD.0000000000016911

Zhang, Y., Song, P., Zhang, R., Yao, Y., Shen, L., Ma, Q., et al. (2021). Clinical characteristics of chronic lung abscess associated with *parvimonasmicra* diagnosed using metagenomic next-generation sequencing. *Infect. Drug Resist.* 14, 1191–1198. doi: 10.2147/IDR.S304569

Zhijun, L., Wenhai, Y., Peibin, Z., and Qingming, L. (2023). Pediatric pulmonary infection caused by oral obligate anaerobes: Case Series. *Front. Pediatr.* 11. doi: 10.3389/fped.2023.1226706

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 29 July 2024
DOI 10.3389/fcimb.2024.1398190



OPEN ACCESS

EDITED BY
Qing Wei,
Genskey Co. Ltd, China

REVIEWED BY
Feng Ye,
The First Affiliated Hospital of Guangzhou Medical University, China
Jia-Xin Shi,
The First People's Hospital of Lianyungang, China
Min Wang,
The First Affiliated Hospital of Henan University of Science and Technology, China

*CORRESPONDENCE
Xin Su
suxinjs@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 09 March 2024
ACCEPTED 09 July 2024
PUBLISHED 29 July 2024

CITATION
Cai X, Sun C, Zhong H, Cai Y, Cao M, Wang L, Sun W, Tao Y, Ma G, Huang B, Yan S, Zhong J, Wang J, Lu Y, Guan Y, Song M, Wang Y, Li Y and Su X (2024) The value of metagenomic next-generation sequencing with different nucleic acid extracting methods of cell-free DNA or whole-cell DNA in the diagnosis of non-neutropenic pulmonary aspergillosis.
*Front. Cell. Infect. Microbiol.* 14:1398190.
doi: 10.3389/fcimb.2024.1398190



# The value of metagenomic next-generation sequencing with different nucleic acid extracting methods of cell-free DNA or whole-cell DNA in the diagnosis of non-neutropenic pulmonary aspergillosis

Xiaomin Cai[1,2†], Chao Sun[2†], Huanhuan Zhong[2,3], Yuchen Cai[2], Min Cao[1], Li Wang[4], Wenkui Sun[5], Yujian Tao[6], Guoer Ma[7], Baoju Huang[4], Shengmei Yan[2], Jinjin Zhong[2], Jiamei Wang[8], Yajie Lu[2], Yuanlin Guan[9], Mengyue Song[3], Yujie Wang[2], Yuanyuan Li[2] and Xin Su[1,2*]

[1]Department of Respiratory and Critical Care Medicine, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China, [2]Department of Respiratory and Critical Medicine, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China, [3]Department of Respiratory and Critical Care Medicine, The Second Affiliated Hospital of Suzhou University, Suzhou, China, [4]Department of Respiratory and Critical Care Medicine, The Second Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China, [5]Department of Respiratory and Critical Care Medicine, Jiangsu Province Hospital, The First Affiliated of Nanjing Medical University, Nanjing, China, [6]Department of Respiratory and Critical Care Medicine, Affiliated Hospital of Yangzhou University, Yangzhou, China, [7]Department of Respiratory and Critical Care Medicine, Affiliated Hospital of Jiangsu University, Zhenjiang, China, [8]Department of Respiratory and Critical Care Medicine, Jinling Hospital, Nanjing Medical University, Nanjing, China, [9]Department of Research and Development, Hugobiotech Co., Ltd., Beijing, China

**Purpose:** Metagenomic next-generation sequencing(mNGS) is a novel molecular diagnostic technique. For nucleic acid extraction methods, both whole-cell DNA (wcDNA) and cell-free DNA (cfDNA) are widely applied with the sample of bronchoalveolar lavage fluid (BALF). We aim to evaluate the clinical value of mNGS with cfDNA and mNGS with wcDNA for the detection of BALF pathogens in non-neutropenic pulmonary aspergillosis.

**Methods:** mNGS with BALF-cfDNA, BALF-wcDNA and conventional microbiological tests (CMTs) were performed in suspected non-neutropenic pulmonary aspergillosis. The diagnostic value of different assays for pulmonary aspergillosis was compared.

**Results:** BALF-mNGS (cfDNA, wcDNA) outperformed CMTs in terms of microorganisms detection. Receiver operating characteristic (ROC) analysis indicated BALF-mNGS (cfDNA, wcDNA) was superior to culture and BALF-GM. Combination diagnosis of either positive for BALF-mNGS (cfDNA, wcDNA) or CMTs is more sensitive than CMTs alone in the diagnosis of pulmonary aspergillosis (BALF-cfDNA+CMTs/BALF-wcDNA+CMTs vs. CMTs: ROC analysis: 0.813 vs.0.66, P=0.0142/0.796 vs.0.66, P=0.0244; Sensitivity: 89.47% vs. 47.37%,

P=0.008/84.21% vs. 47.37%, P=0.016). BALF-cfDNA showed a significantly greater reads per million (RPM) than BALF-wcDNA. The area under the ROC curve (AUC) for RPM of *Aspergillus* detected by BALF-cfDNA, used to predict "True positive" pulmonary aspergillosis patients, was 0.779, with a cut-off value greater than 4.5.

**Conclusion:** We propose that the incorporation of BALF-mNGS (cfDNA, wcDNA) with CMTs improves diagnostic precision in the identification of non-neutropenic pulmonary aspergillosis when compared to CMTs alone. BALF-cfDNA outperforms BALF-wcDNA in clinical value.

KEYWORDS

**metagenomic next-generation sequencing (mNGS), cell-free DNA, whole-cell DNA, non-neutropenic pulmonary aspergillosis, pulmonary aspergillosis**

# 1 Introduction

Pulmonary Aspergillosis (PA) is a serious infectious fungal disease, commonly seen in immunocompromised patients (El-Baba et al., 2020). *Aspergillus fumigatus* is the predominant culprit, responsible for more than 70% of cases. Depending on the interaction between *Aspergillus* and individuals with varying immune status and underlying diseases, PA is classified as invasive pulmonary aspergillosis (IPA), allergic bronchopulmonary aspergillosis (ABPA) and chronic pulmonary aspergillosis (CPA). With the increasing use of corticosteroids or antimicrobials, an aging population, PA is not limited to immunosuppressed populations. High-risk hosts include individuals like those with chronic obstructive pulmonary disease (COPD) (Guinea et al., 2010) and critically ill patients (Taccone et al., 2015). In recent years, the morbidity of PA has been increasing. However, PA in non-neutropenic patients is hard to be recognized in the early stage due to the atypical clinical and radiological manifestations and limited sensitivity of traditional diagnostic methods. Thus, the mortality stays high. Current diagnostic tools, such as Aspergillus-specific IgG (Lu et al., 2023) and Pentraxin 3 (He et al., 2018) may have some value in the diagnosis of non-neutropenic PA, but we still need quicker and earlier detection methods for PA. Therefore, finding a sensitive, efficient, specific and less invasive method for early PA diagnosis is of great value of improving outcomes.

Since in 2014, when Charles Y. Chiu's team (Wilson et al., 2014) used metagenomic Next-Generation Sequencing (mNGS) to detect *Leptospira* for the first time to confirm the diagnosis of meningitis, mNGS has emerged as a novel molecular diagnostic technique. Over the past few years, mNGS has been recognized as a comprehensive, rapid and accurate method in detecting infectious pathogens in the nervous, respiratory and blood system infection offering advantages such as rapid detection, non-bias and broad spectrum. Commonly used samples include sputum, bronchoalveolar lavage fluid (BALF) and blood. However, few reports have been published about the application of mNGS in the study of PA in non-neutropenic patients. Basing on the methods of extracting nucleic acid, whole-cell DNA (wcDNA) and cell-free DNA (cfDNA) are both applied widely. mNGS of cfDNA may cause host DNA degradation, potentially leading to the loss of cfDNA in the supernatant. Conversely, fragmenting cells without degrading host DNA during the extraction of wcDNA from BALF samples can increase human DNA release. Determining the appropriate sample processing technique for clinical settings, but there are few reports on the clinical value of these two sample processing technologies.

Therefore, we performed this clinical study to analyze the pathogenicity of non-neutropenic patients with suspected PA in terms of sample type and nucleic acid extraction method. Our aims are to analyze the pathogenic profile in detail and to systematically evaluate the efficacy of mNGS in the diagnosis of PA.

# 2 Materials and methods

## 2.1 Study design and patient population

From March 2022 to October 2022, patients with suspected PA in non-neutropenic admitted to the Department of Respiratory and Critical Care Medicine of Nanjing Jinling Hospital, Nanjing Drum Tower Hospital, Nanjing First Hospital, Jiangsu Province Hospital, affiliated Hospital of Yangzhou university, Affiliated Hospital of Jiangsu university, Jiangsu Second Chinese Medicine Hospital.

Patients were eligible for enrollment if they were (1) age≥ 18 years; (2) including but not limited to: patients with underlying diseases, such as COPD, diabetes, and the use of corticosteroids; (3) have respiratory symptoms, like fever, cough, that have failed to respond to treatment with broad-spectrum antibacterial medication; (4) computed tomography (CT) showing lesions such as pulmonary nodules, infiltrative shadows or cavities. Patients were excluded from the study based on the following criteria: (1) age <18 years; (2) neutropenia (absolute neutrophil count $<0.5\times10^9$/L).

The patients were finally classified as the PA (including IPA and CPA) and non-PA groups. The clinical diagnosis of PA and whether the microorganism detected was pathogenic or colonizing was made by two senior pulmonologists based on host risk factors, clinical symptoms, chest computed tomography, laboratory findings, and response to treatment. The diagnostic criteria for IPA and CPA were mainly referred to the guidelines of the 2020 EORTC/MSGERC,2016 IDSA, 2017 ESCMID/ERS/ECMM (Denning et al., 2016; Patterson et al., 2016; Ullmann et al., 2018; Donnelly et al., 2020). Proven IPA requires a positive *Aspergillus* in sterile body fluids or tissues. The probable IPA needs the combination of (1) host factors like COPD, diabetes; and (2) clinical symptoms like fever, cough; and (3) CT showing lesions with or without a halo sign, infiltrative shadows, or cavities; and (4) microbiological evidence included positive results for *Aspergillus* culture or PCR, single Galactomannan (GM) test ≥1.0, or single serum/plasma GM ≥0.7 with BALF GM ≥0.8. The possible IPA needs at least one of the host factors and clinical features. CPA diagnosis relies on (1) clinical symptoms, like cough, sputum, or fever; (2) CT imaging like cavitation, fugal ball; and (3) *Aspergillus* culture positive or immunological response to *Aspergillus* (positive Aspergillus IgG or precipitin test). The disease has been present for at least 3 months.

BALF samples taken from patients with suspected PA underwent mNGS of cfDNA, mNGS of wcDNA, and conventional microbiological tests (CMTs). The CMTs included the GM test, culture, and smear microscopic for the bacteria and fungi, and smear microscopic for tuberculosis. Physical information and clinical details were investigated. The remaining BALF sample from each of the enrolled patients was collected into a 3 mL sterile tube and delivered to Hugobiotech (Hugobiotech, Beijing, China) immediately for mNGS of cfDNA and wcDNA(The minimum total volume of BALF required for the each experiment was 1.5 mL). The remaining 5 mL Blood samples from 8 enrolled patients were collected into a vacuum blood collective tube and delivered to Hugobiotech (Hugobiotech, Beijing, China) at room temperature and immediately for mNGS of cfDNA.

Concurrently, the remaining 3ml BALF from 62 of enrolled patients was collected into sterile tubes and immediately sent to KingMed (Guangzhou, China) for target next generation sequencing (tNGS) detection.

## 2.2 GM test and culture

The hospital laboratories performed the GM test using the double-sandwich enzyme-linked immunosorbent assay (Bio-Rad Laboratories). Samples of appropriately collected bronchoalveolar lavage fluid (BALF), comprising more than 10 mL, were cultured using CHROMagar and incubated at 35°C for three days.

## 2.3 mNGS detection

### 2.3.1 Nucleic acid extraction

Based on its manual, cfDNA and wcDNA were extracted from clinical samples using QIAamp DNA Micro Kit (QIAGEN, Hilden, Germany). For cfDNA extraction, the supernatant of the sample is taken after centrifugation. For wcDNA, the sample is extracted directly without centrifugation. Using Qubit 3.0 Fluoremeter (Invitrogen, Q33216) and agarose gel electrophoresis (Major Science, UVC1–1100) check DNA concentration and quality.

### 2.3.2 Library generation and sequencing

DNA library construction was carried out in line with the guidelines specified in the Qiagen library construction kit (QIAseq Ultralow Input Library Kit). Quality control of the library was conducted using both the Qubit Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA) and the 3.0 Fluoremeter (Invitrogen, Q33216). Eligible DNA libraries, labeled with different barcodes, were combined and sequenced using the SE75bp sequencing strategy and Illumina Nextseq 550 sequencing platform (Illumina, San Diego, USA).

### 2.3.3 Bioinformation pipeline

After gaining the sequencing data, we filtered out splice sequences, low quality, low complexity, and shorter sequences to obtain high-quality data. Then, we use SNAP software to remove human-derived sequences that match the human reference database (hg38). Next, we aligned the remaining data to the microbial genome database using BWA-MEM (processing time is approximately 20 minutes, and the memory requirement is around 20G). Finally, we compared the remaining data with the microbial genome database using Burrow Wheeler Alignment. This database contains an extensive collection of microbial genomes from NCBI having more than 30,000 microorganisms, including 17,748 species of bacteria, 11,058 species of viruses, 1,134 species of fungi, and 308 species of parasites. Finally, the microbial composition in the sample was determined. The positive criteria for the mNGS result were set as follows (Gu et al., 2021):

(1) To detect bacteria, fungi, and parasites, the sequencing coverage should be in the top 10 of all pathogens detected and not detected in the negative control (NTC), or the sample/NTC should have an RPM (reads per million mapped reads) ratio greater than 10.
(2) For viruses, tuberculosis, and cryptococci, at least one specific sequence should be detected and not detected in the NTC, or the sample/NTC RPM ratio should be greater than 5.

## 2.4 tNGS detection

### 2.4.1 Nucleic acid extraction and library preparation

The magnetic bead method is employed for the extraction of nucleic acids from samples.

The samples were subjected to amplification using ultra-multiplex PCR primers (a total of 153 respiratory pathogens (Supplementary Table 1) were analyzed with the aim of identifying highly conserved regions). This was achieved through the design of specific primers. The amplified PCR products were purified by magnetic beads and mixed with specific sequencing junction tags and a library amplifying

enzyme for the second round of amplification. The products of the second round of amplification were purified by magnetic beads for the second time to obtain the libraries.

### 2.4.2 Sequencing and bioinformatic analysis

Sequencing was conducted using the gene sequencer KM MiniSeqDx-CN. Following a comparison and analysis of the data from the sequencing machine with the database, the pathogenic situation in the samples was judged.

## 2.5 Statistical analysis

We used SPSS software (version 26, IBM Corp, Armonk, NY, USA), MedCalc (version 20.1), and Prism (version 9.5.1) for statistical analysis and drawing. Continuous variables were presented as mean ± SD. The t-test and Wilcoxon test for two group samples were used to compare the normal or abnormal distribution. We employed the Pearson chi-squared test and McNemar test (for paired data) or the Fisher's exact test for categorical data. The specificity and sensitivity of detection methods in diagnosing PA were calculated (percentage with 95% confidence interval [CI]) and compared (chi-squared). Spearman's r values were utilized to analyze their correlation. Furthermore, a receiver operating characteristic (ROC) curve was employed to determine the best test for identifying specific pathogens with true-positive results. The study implemented the Yoden index to establish the cut-off values for RPM in the ROC curve. A two-tailed P-value of less than 0.05 was considered statistically significant.

# 3 Results

## 3.1 Patient characteristics

A total of 71 suspected PA patients including 48 male and 23 female were enrolled in this study, Table 1 displays their clinical features. Most patients have underlying diseases (97.2%, 69/71), such as lung cancer (7), COPD (9), and diabetes (18). The clinicians finally diagnosed 19 cases of PA (12 cases of IPA, and 7 cases of CPA) and 52 cases of non-PA (26 cases of bacterial infection, 4 cases of non-infectious diseases, 8 cases of Mycobacterium tuberculosis (MTB), 6 cases of other fungal diseases, 3 cases of non-tuberculous mycobacteria (NTM), and 1 case of Chlamydia psittaci pneumonia,4 cases of bacterial co-infections with other fungal).

## 3.2 Species distribution and consistency of microorganisms detected by BALF-cfDNA, BALF-wcDNA, and CMTs for suspected non-neutropenic pulmonary aspergillosis

BALF-cfDNA detected 43 species (14 fungi, 19 bacteria, 6 viruses, 3 mycobacteria, 1 chlamydia), BALF-wcDNA detected 44

TABLE 1 Clinical characteristics of PA and non-PA in non-neutropenic patients on admission.

| Characteristic, n (%) | N=71 | PA (n=19) | non-PA (n=52) | P-value |
|---|---|---|---|---|
| Male/Female | 48/23 | 14/5 | 34/18 | 0.51 |
| Age, mean (SD),years | 61.17 ± 13.236 | 62.63 ± 7.974 | 60.63 ± 14.729 | 0.47 |
| Smoking history | 30 (42.25) | 9 (47.37) | 21 (40.38) | 0.60 |
| Drinking history | 12 (16.90) | 2 (10.53) | 10 (19.23) | 0.61 |
| Admitted to ICU | 18 (25.35) | 5 (26.32) | 13 (25.00) | 1.00 |
| Use of hormones for more than 3weeks within 60days | | | | 0.79 |
| Vein/Oral | 13 (18.31) | 4 (21.05) | 9 (17.31) | |
| Inhale | 1 (1.41) | 0 | 1 (1.92) | |
| Use of immunosuppressive agents within 30 days | 8 (11.27) | 1 (5.26) | 7 (13.46) | 0.59 |
| Underlying diseases | | | | |
| Hypertension | 25 (35.21) | 7 (36.84) | 18 (34.62) | 0.86 |
| Diabetes | 18 (25.35) | 7 (36.84) | 11 (21.15) | 0.30 |
| Bronchiectasis | 16 (22.54) | 7 (36.84) | 9 (17.31) | 0.16 |
| Cerebrovascular disease | 17 (23.94) | 4 (21.05) | 13 (25.00) | 0.97 |
| Pulmonary emphysema | 15 (21.13) | 4 (21.05) | 11 (21.15) | 1.00 |
| Pulmonary tuberculosis | 12 (16.90) | 5 (26.32) | 7 (13.46) | 0.36 |
| Cardiovascular disease | 12 (16.90) | 2 (10.53) | 10 (19.23) | 0.49 |

*(Continued)*

10.3389/fcimb.2024.1398190

TABLE 1 Continued

| Characteristic, n (%) | N=71 | PA (n=19) | non-PA (n=52) | P-value |
|---|---|---|---|---|
| **Underlying diseases** | | | | |
| Other solid organ tumor (except lung cancer) | 11 (15.49) | 5 (26.32) | 6 (11.54) | 0.25 |
| Chronic obstructive pulmonary disease | 9 (12.68) | 7 (36.84) | 2 (3.85) | 1.00 |
| Interstitial Lung Disease | 8 (11.27) | 0 | 8 (15.38) | 0.16 |
| Chronic kidney diseases | 7 (9.86) | 2 (10.53) | 5 (9.61) | 1.00 |
| Lung cancer | 7 (9.86) | 2 (10.53) | 5 (9.61) | 1.00 |
| Hepatopathy | 4 (5.63) | 1 (5.26) | 3 (5.77) | 0.81 |
| Congestive heart failure | 3 (4.22) | 0 | 3 (5.77) | 0.56 |
| Hematologic tumor | 2 (2.82) | 1 (5.26) | 1 (1.92) | 0.47 |
| Organ transplantation | 1 (1.41) | 0 | 1 (1.92) | 1.00 |
| **Clinical symptoms** | | | | |
| Fever | 35 (49.30) | 8 (42.11) | 27 (51.92) | 0.46 |
| Cough | 63 (88.73) | 17 (89.47) | 46 (88.46) | 1.00 |
| Shiver | 5 (7.04) | 1 (5.26) | 4 (7.69) | 1.00 |
| Eexpectoration | 61 (85.92) | 17 (89.47) | 44 (84.62) | 0.89 |
| Hemoptysis | 13 (18.31) | 5 (29.41) | 8 (15.38) | 0.56 |
| Chest distress | 33 (46.48) | 11 (57.89) | 22 (42.31) | 0.24 |
| Asthma/dyspnea | 35 (49.30) | 10 (52.63) | 25 (48.08) | 0.73 |
| Chest pain | 11 (15.49) | 4 (21.05) | 7 (13.46) | 0.68 |
| **Chest computed tomography images** | | | | |
| Infiltration or exudation | 45 (63.38) | 12 (63.16) | 33 (63.46) | 0.98 |
| Small nodule | 23 (32.39) | 8 (42.11) | 15 (28.85) | 0.32 |
| Wedge-shaped and segmental or lobar consolidation | 21 (29.58) | 4 (21.05) | 17 (32.69) | 0.32 |
| Cavitation sign | 15 (21.13) | 9 (47.37) | 6 (11.54) | 0.004 |
| Multiple clump-like infiltrates or consolidations along the bronchovascular bundle | 15 (21.13) | 5 (26.32) | 10 (19.23) | 0.78 |
| Tubercle | 14 (19.72) | 4 (21.05) | 10 (19.23) | 1.00 |
| Tree bud sign | 5 (7.04) | 3 (15.79) | 2 (3.85) | 0.07 |
| Mass | 4 (5.63) | 1 (5.26) | 3 (5.77) | 1.00 |
| Air crescent sign | 3 (4.22) | 2 (10.53) | 1 (1.92) | 0.05 |
| Halo sign | 1 (1.41) | 0 | 1 (1.92) | 1.00 |
| Pleural thickening | | | | 0.33 |
| Unilateral | 8 (11.27) | 1 (5.26) | 7 (13.46) | |
| Bilateral | 18 (25.35) | 7 (36.84) | 11 (21.15) | |
| Pleural effusion | | | | 0.25 |
| Unilateral | 15 (21.13) | 2 (10.53) | 13 (25.00) | |
| Bilateral | 11 (15.49) | 2 (10.53) | 9 (17.31) | |

PA, Pulmonary aspergillosis; non-PA, non- Pulmonary aspergillosis.

species (17 fungi, 18 bacteria, 5 viruses, 3 mycobacteria, 1 chlamydia), and CMTs detected 14 species (4 fungi, 9 bacteria,1 mycobacteria). As shown, BALF-mNGS (cfDNA, wcDNA) detected more species than CMTs. Five (*Rhizopus delemar, Klebsiella aerogenes, Serratia marcescens, Human mastadenovirus B, Human mastadenovirus C*) microorganisms were detected only by BALF-cfDNA. Six microorganisms (*Alternaria alternata, Candida parapsilosis Chaetomium globosum, Candida intermedia, Aspergillus glaucus, Human betaherpesvirus 6B*) were detected only by BALF-wcDNA. *Elizabethkingia meningosepticum* was identified by CMTs alone. *Escherichia coli* was seen by both CMTs and BALF-wcDNA, and the remaining species detected by CMTs were those that both mNGS methods could detect (Figure 1). *Aspergillus fumigatus* was the most frequently reported fungus in all three methods.

34 patients had positive responsible pathogens according to CMTs (34/71, 47.89%), 58 and 51 patients were tested positive by BALF-cfDNA and BALF-wcDNA (BALF-cfDNA:58/71, 81.69%; BALF-wcDNA:51/71, 71.83%). The positive rate of BALF-mNGS was higher than CMTs (BALF-cfDNA vs. CMTs: 81.69% vs. 47.89%, P<0.001; BALF-wcDNA vs. CMTs: 71.83% vs. 47.89%, P=0.002; BALF-cfDNA vs. BALF-wcDNA: 81.69% vs. 71.83%, P=0.016).

31 patients were positive for the pathogens tested using both three methods (BALF-cfDNA, BALF-wcDNA, and CMTs), the consistency between the three methods was as follows: (1) matched in 7 (7/31, 22.58%) patients (perfect agreement in pathogens detection across all three methods), (2) partially matched in 17 (17/31, 54.84%) patients (at least one microorganism overlapped between three methods), (3) wholly mismatched in 7 (7/31, 22.58%) patients (no overlap of the pathogen between the three methods). 4 patients were not detected by 3 methods and 1 patient was detected only through CMTs.

## 3.3 Differences in numbers for RPM detected by BALF-cfDNA, BALF-wcDNA

The RPM range observed was 1–86419 by BALF-cfDNA and 1–207274 by BALF-wcDNA. Generally, RPM detected by BALF-

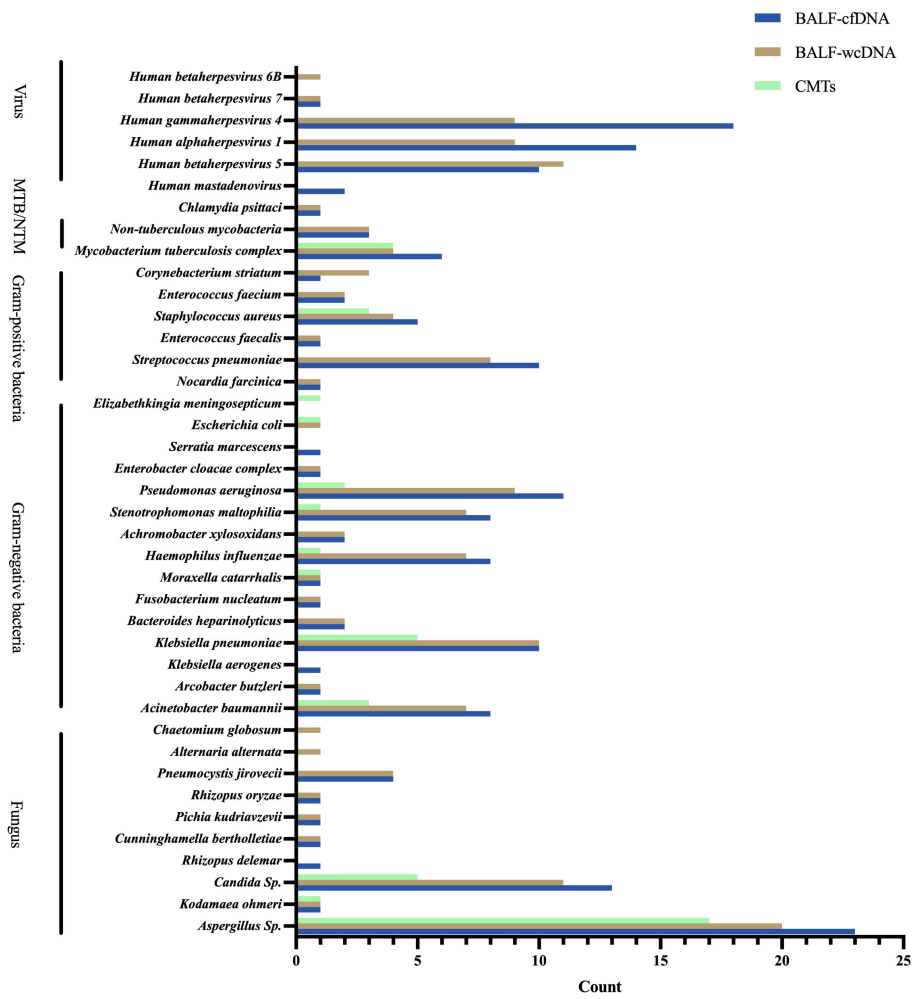



FIGURE 1
Species distribution of gram-positive bacteria, gram-negative bacteria, fungus, viruses, and chlamydia pasittaci detected by BALF-cfDNA, BALF-wcDNA, and CMTs. MTB, Mycobacterium tuberculosis; NTM, Non-tuberculous mycobacteria.

cfDNA was higher than those detected by BALF-wcDNA (29.5 vs. 19.5, P<0.001). For the Gram-positive bacteria, the RPM tested by BALF-cfDNA was greater than BALF-wcDNA (347.5 vs.119.5, P=0.008). Besides, for the detection of Gram-negative bacteria (188 vs. 97, P=0.071), fungus (17 vs. 9, P=0.467), MTB/NTM (29 vs.43, P=0.575), and virus (2 vs.2, P=0.163), there was no significant difference between BALF-cfDNA and BALF-wcDNA (Figure 2). These results reveal that mNGS of BALF-cfDNA captures more reads of microorganisms than mNGS of BALF-wcDNA.

## 3.4 Microbial distribution for pulmonary aspergillosis detected by BALF-cfDNA, BALF-wcDNA, and CMTs

For PA patients, BALF-cfDNA identified 21 species (8 fungi, 10 bacteria, 3 viruses), BALF-wcDNA detected 21 species (8 fungi, 9 bacteria, 4 viruses), and CMTs detected 8 species (3 fungi, 5 bacteria) (Figure 3).

The positive rate for *Aspergillus* was 78.95% (15/19) for BALF-cfDNA and 73.68% (14/19) for BALF-wcDNA, there was no diversity in the positive rate (P=1.00). 9 patients were positive for *Aspergillus* by CMTs, with a positive rate of 47.37% (9/19), which was no statistical difference with BALF-cfDNA and BALF-wcDNA (P=0.109; 0.18). The number of *Aspergillus* identified by CMTs only and BALF-mNGS (cfDNA and wcDNA both) only were 2 and 8 patients, respectively, 1 patient detected *Aspergillus* by BALF-cfDNA only, 7 patients detected *Aspergillus* through three methods both.

*Aspergillus fumigatus*, closely followed by *Aspergillus flavus*, is the most common causative of *Aspergillus* in patients with PA. A higher number of RPM indicated *Aspergillus* detection through BALF-cfDNA instead of BALF-wcDNA (48 vs. 6, P=0.001).

## 3.5 Comparison of diagnostic performance among BALF-cfDNA, BALF-wcDNA, and CMTs in non-neutropenic pulmonary aspergillosis

Using a clinical diagnosis as the gold standard, we compared the diagnostic accuracy of detection methods in non-neutropenic PA (Table 2). BALF-cfDNA showed a sensitivity of 78.95% and a specificity of 84.62%, with PPV and NPV of 65.22% and 91.67%, respectively. The sensitivity, specificity, PPV, and NPV of BALF-wcDNA were 73.68%, 88.46%, 70.00%, and 90.20%. The sensitivity and specificity of BALF-cfDNA were similar to BALF-wcDNA. The sensitivity and specificity of CMTs in diagnosing PA were 47.37% and 84.62%, whereas the PPV and NPV were 52.94% and 81.48%.

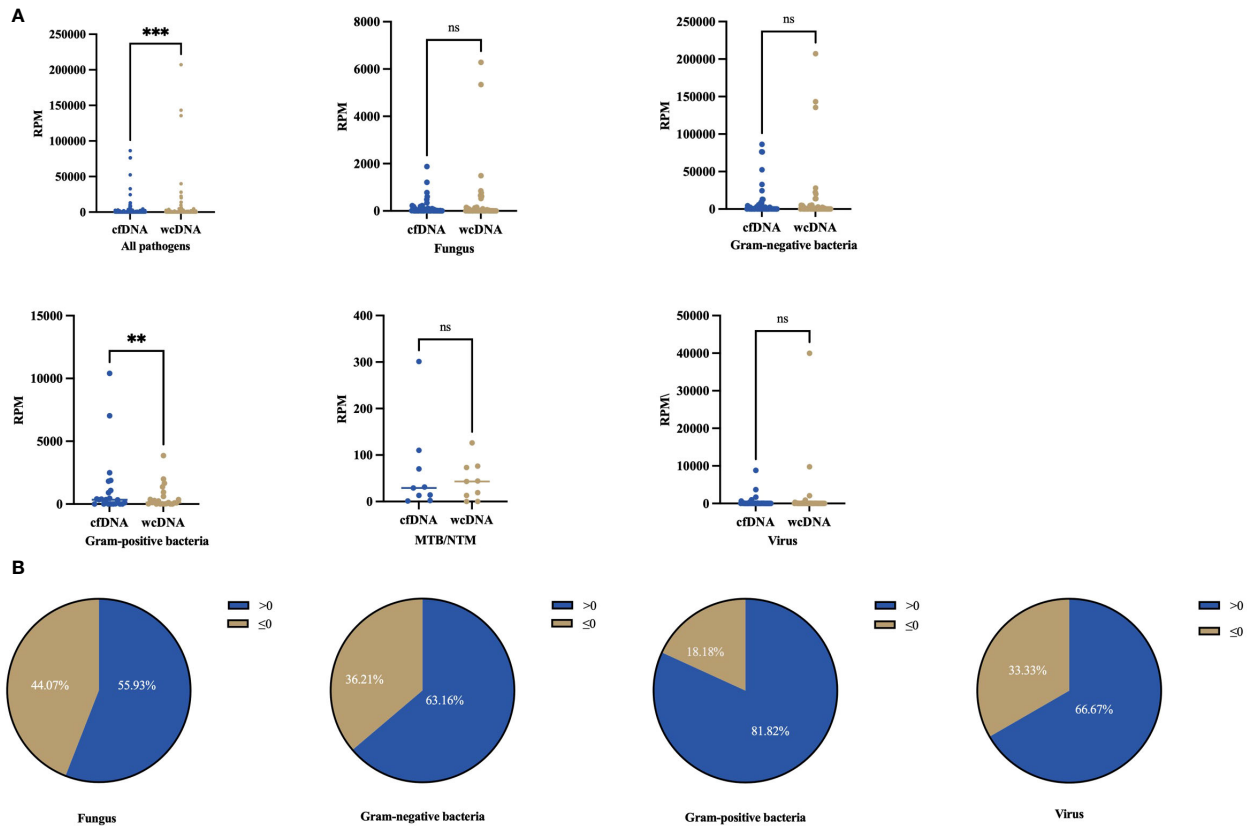



FIGURE 2
Differences in numbers for RPM detected by BALF-cfDNA, BALF-wcDNA. **(A)** Comparison of the number of RPM detected by BALF-cfDNA and BALF-wcDNA for all pathogens, Fungus, Gram-positive bacteria, Gram-negative bacteria, MTB/NTM, and virus. **(B)** >0 represented that the numbers of RPM detected by BALF-cfDNA were higher than that detected by BALF-wcDNA.≤0 represented that the numbers of RPM by BALF-wcDNA were higher or equal to BALF-cfDNA. **P<0.01; ***P<0.001; ns, no significant.

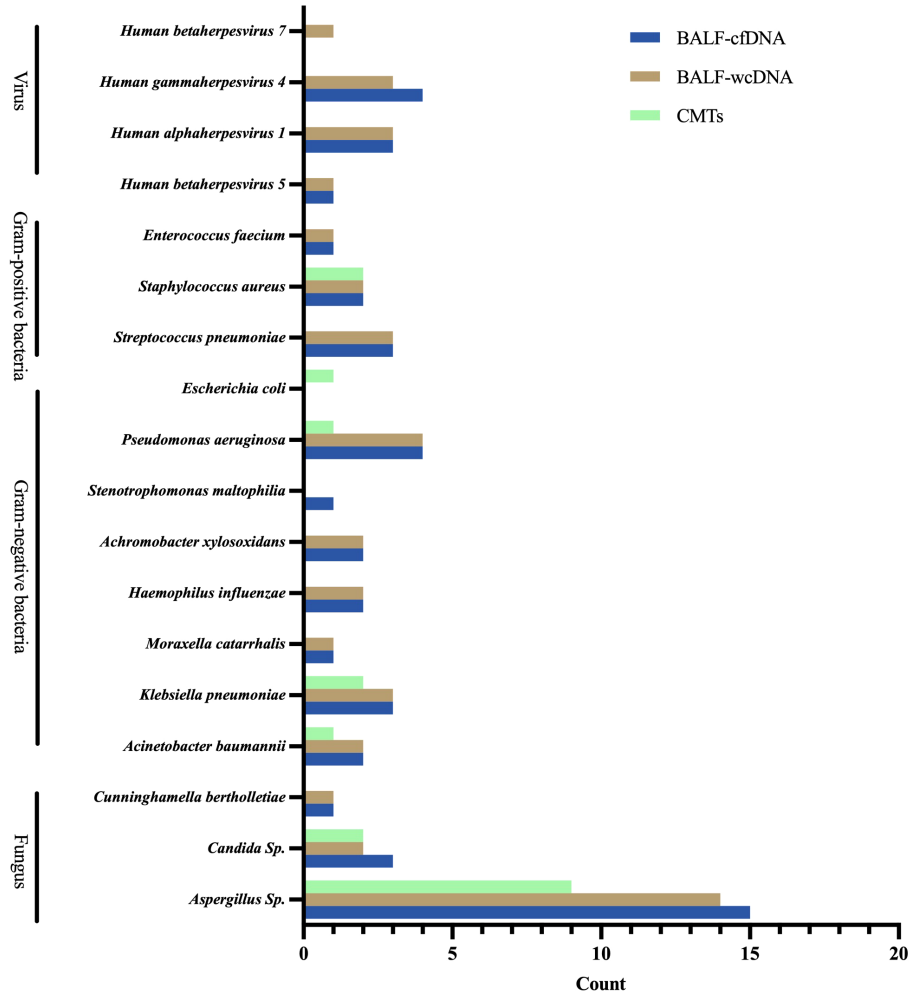



FIGURE 3
Microbial distribution for Pulmonary Aspergillus detected by BALF-cfDNA, BALF-wcDNA, and CMTs.

BALF-mNGS (cfDNA, wcDNA) perform better than culture or BALF-GM in sensitivity. No significant difference was observed in sensitivity and specificity between BALF-mNGS (cfDNA, wcDNA) and CMTs. Combination diagnosis of either positive for CMTs or BALF-mNGS (cfDNA, wcDNA) had significantly higher sensitivity, but significantly lower specificity than those of CMTs alone. (Tables 2, 3).

ROC analysis of BALF-cfDNA, BALF-wcDNA, and CMTs for the diagnosis of PA yielded an AUC of 0.818, 0.811, and 0.66. BALF-mNGS (cfDNA, wcDNA) and CMTs exhibited comparable diagnostic abilities, while BALF-mNGS (cfDNA, wcDNA) outperformed culture or BALF-GM. The combination of BALF-mNGS (cfDNA, wcDNA) and CMTs is more effective than CMTs alone in the diagnosis of PA (Figure 4; Table 3).

14 out of 71 patients had taken antifungal agents (Voriconazole, Caspofungin) before providing a sample. For these patients, Sensitivity and specificity results indicate no significant differences when comparing BALF-mNGS (cfDNA, wcDNA) and CMTs for diagnosing PA (Sensitivity: BALF-cfDNA/BALF-wcDNA, 75.00% vs. 37.50%, P=0.25/75.00% vs. 37.50%, P=0.25; Specificity: BALF-cfDNA/BALF-wcDNA, 83.33% vs. 66.67%, P=1.00/100.00% vs. 66.67%). There is no notable difference in the diagnostic efficacy of BALF-mNGS (cfDNA, wcDNA) between this patient group and those who did not receive antifungal medication (BALF-cfDNA: sensitivity, 75.00% vs. 83.33%, P=1.00/specificity, 83.33% vs. 84.78%, P=1.00; BALF-wcDNA: sensitivity, 75.00% vs.75.00%, P=1.00/specificity,100.00% vs. 86.96%, P=1.00).

## 3.6 "True positive", "False positive", "False negative" by BALF-mNGS

We observed that the diversity of RPM for *Aspergillus* detected by BALF-cfDNA in "True positive" and "False positive" patients (61 vs. 2.5, P=0.03), whereas there was no significant difference in BALF-wcDNA (14 vs. 6, P=0.231). (Figure 5) We utilized the ROC curve to assess the diagnostic performance of BALF-cfDNA in "True positive" PA patients. The area under the ROC curve (AUC) for the RPM was 0.779 (P=0.031) (Figure 4D), and the cut-off value calculated according to the Yoden index was greater than 4.5, the sensitivity and specificity were 100.00% and 62.50%. For BALF-wcDNA, the AUC for the RPM was 0.673 (P=0.232).

TABLE 2 Diagnostic performance of BALF-cfDNA, BALF-wcDNA, and CMTs in PA.

| Methods | Sensitivity (95%CI) | Specificity (95%CI) | PPV (95%CI) | NPV (95%CI) | AUC (95%CI) |
|---|---|---|---|---|---|
| BALF-cfDNA | 78.95% | 84.62% | 65.22% | 91.67% | 0.818 |
| | 0.5667–0.9149 | 0.7248–0.9199 | 0.4489–0.8119 | 0.8045–0.9671 | 0.708–0.899 |
| BALF-wcDNA | 73.68% | 88.46% | 70.00% | 90.20% | 0.811 |
| | 0.5121–0.8819 | 0.7703–0.946 | 0.481–0.8545 | 0.79–2-0.9574 | 0.7–0.894 |
| CMTs | 47.37% | 84.62% | 52.94% | 81.48% | 0.66 |
| | 0.2733–0.6829 | 0.7284–0.9199 | 0.3096–0.7383 | 0.6916–0.8962 | 0.538–0.768 |
| Culture | 21.05% | 96.15% | 66.67% | 76.92% | 0.586 |
| | 0.0851–0.433 | 0.8702–0.9932 | 0.3000–0.9408 | 0.6536–0.8549 | 0.463–0.702 |
| BALF-GM | 31.58% | 90.38% | 54.55% | 78.33% | 0.61 |
| | 0.1536–0.5399 | 0.7939–0.9582 | 0.2801–0.7873 | 0.6638–0.8688 | 0.487–0.723 |
| cfDNA+CMTs | 89.47% | 73.08% | 54.84% | 95.00% | 0.813 |
| | 0.6861–0.9813 | 0.5975–0.8323 | 0.3777–0.7084 | 0.8350–0.9911 | 0.703–0.896 |
| wcDNA+CMTs | 84.21% | 75.00% | 55.17% | 92.86% | 0.796 |
| | 0.6243–0.9448 | 0.6179–0.8477 | 0.3755–0.7159 | 0.8099–0.9754 | 0.684–0.882 |

cfDNA, cell-free DNA metagenomic next generation sequencing; wcDNA, whole-cell DNA metagenomic next generation sequencing; CMTs, conventional microbiological tests; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval; AUC, the area under the ROC curve.

The RPM of *Aspergillus* detected by BALF-cfDNA showed a positive correlation with BALF-GM (Spearman's r values:0.481, P=0.037), while no correlation was observed with serum-GM. Next, we compared the levels of serum-GM and BALF-GM in patients categorized as "True positive" or "False positive" based on BALF-cfDNA and BALF-wcDNA separately, we found a difference in BALF-GM (BALF-cfDNA:0.8 vs.0.16, P=0.014; BALF-wcDNA:0.8 vs.0.17, P=0.002), but no significant difference in serum-GM (BALF-cfDNA:0.25 vs.0.15, P=0.057, BALF-wcDNA:0.25 vs. 0.11, P=0.053).

There was variety in serum-GM between PA patients with true-positive and false-negative by BALF-cfDNA (0.25 vs. 0.1, P=0.008) or BALF-wcDNA (0.25 vs. 0.1, P=0.034) separately, and no differences in BALF-GM (BALF-cfDNA:0.84 vs. 0.32, P=0.305; BALF-wcDNA:0.88 vs. 0.33, P=0.246) (Figure 6).

## 3.7 Antifungal agents' application before enrollment

In our investigation, 8 out of 19 PA patients had received antifungal agents (Voriconazole, Caspofungin) before sample collection. 1 case (1/8, 12.50%) and 6 cases (6/8, 75.00%) were identified to have Aspergillus by CMTs and BALF-mNGS(cfDNA, wcDNA). We found no significant difference in the detection of *Aspergillus* among culture, CMTs, BALF-cfDNA, and BALF-wcDNA (P=0.344, 1.00, 0.065, 0.109) when treating PA patients with antifungal agents before sampling. In addition, there was also no diversity in serum GM (0.32 vs. 0.15, P=0.119), BALF-GM (0.54 vs. 0.75, P=0.354), RPM of *Aspergillus* detected by BALF-cfDNA (42 vs. 96, P=0.346) and BALF-wcDNA (6 vs. 32.5, P=0.154) (Figure 7).

TABLE 3 Comparison of the sensitivity, specificity, and ROC curve (AUC) among different diagnostic methods for suspected PA.

| sensitivity/ specificity/AUC | BALF-cfDNA | BALF-wcDNA | CMTs | Culture |
|---|---|---|---|---|
| BALF-wcDNA | 1.00/0.687/0.8412 | / | / | / |
| CMTs | 0.109/1.00/0.0602 | 0.18/0.774/0.0662 | / | / |
| Culture | 0.003/0.07/0.0018 | 0.006/0.219/0.0024 | 0.063/0.031/0.1911 | / |
| BALF-GM | 0.012/0.549/0.007 | 0.021/1.00/0.0088 | 0.25/0.25/0.2758 | 0.727/0.453/0.7651 |
| cfDNA+CMTs | 0.5/0.031/0.9053 | / | 0.008/0.031/0.0142 | / |
| wcDNA+CMTs | / | 0.5/0.016/0.7349 | 0.016/0.063/0.0244 | / |

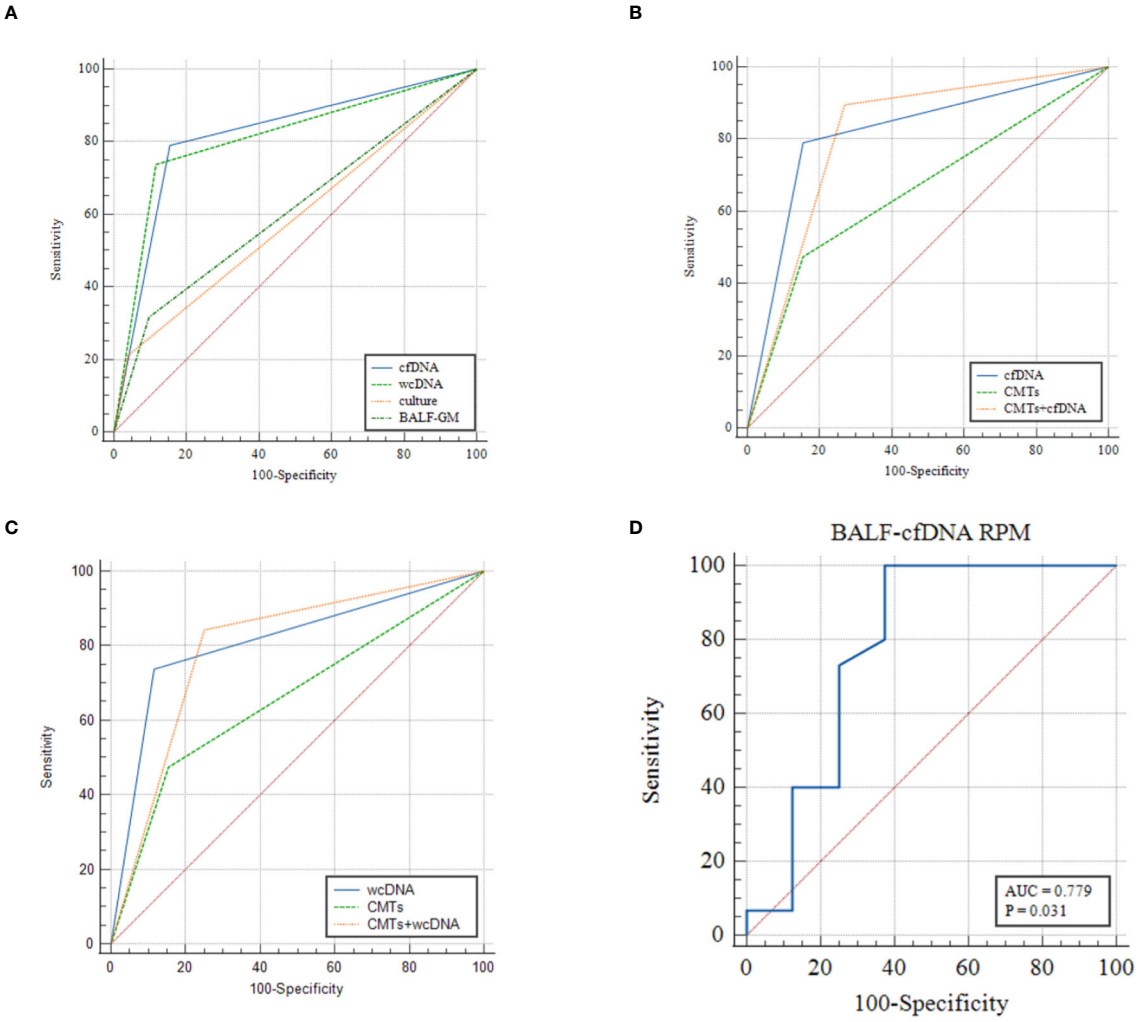



FIGURE 4
**(A–C)** ROC analysis of BALF-cfDNA, BALF-wcDNA, BALF-GM, culture and CMTs for the diagnosis of PA. **(D)** ROC analysis of RPM of Aspergillus detected by BALF-cfDNA in "True positive" PA.

## 3.8 Blood-cfDNA for suspected pulmonary aspergillosis

8 patients submitted both blood and BALF cfDNA, of whom 5 were diagnosed with PA (2 cases of CPA and 3 cases of IPA). However,4 of them did not detect any pathogens through blood-cfDNA. 3 patients (2 cases of IPA) detected *Aspergillus* through the blood cfDNA (Figure 8), and they also found Aspergills by BALF-mNGS (cfDNA, wcDNA).

A significant difference was observed in the serum-GM (0.78 vs.0.15, P=0.034) between the patients with positive and negative *Aspergillus* detection in blood-cfDNA, but there was no notable difference in BALF-GM (1.23 vs.0.62, P=1.00).

## 3.9 tNGS for suspected pulmonary aspergillosis

A total of 62 patients with suspected PA (18 PA and 44 non-PA) were subjected to tNGS. The pathogens detected are shown in Supplementary Figure 1.

The sensitivity and specificity of the diagnosis of PA by tNGS was 52.94% (0.3096–0.7383), 84.44% (0.7122–0.9225), and PPV and NPV were 56.25% (0.3318–0.769), and 82.61% (0.6928–0.9091), respectively.

The AUC values of tNGS, BALF-cfDNA, BALF-wcDNA and CMTs for PA in the 62 patients were re-analyzed using ROC curves, which yielded values of 0.716, 0.838, 0.808 and 0.658, respectively. The results demonstrated that tNGS did not exhibit a significant difference in comparison with the other methods (tNGS vs. BALF-cfDNA: 0.716 vs. 0.838, P=0.148; tNGS vs. BALF-wcDNA: 0.716 vs. 0.808, P=0.083; tNGS vs. CMTs: 0.716 vs. 0.658, P=0.521).

# 4 Discussion

This is the first report to assess the diagnostic efficacy of mNGS using cfDNA and wcDNA on BALF samples from patients with suspected PA. Our findings show a significantly higher detection rate of microorganisms in patients using BALF-mNGS (cfDNA, wcDNA) compared to CMTs. We provide evidence that

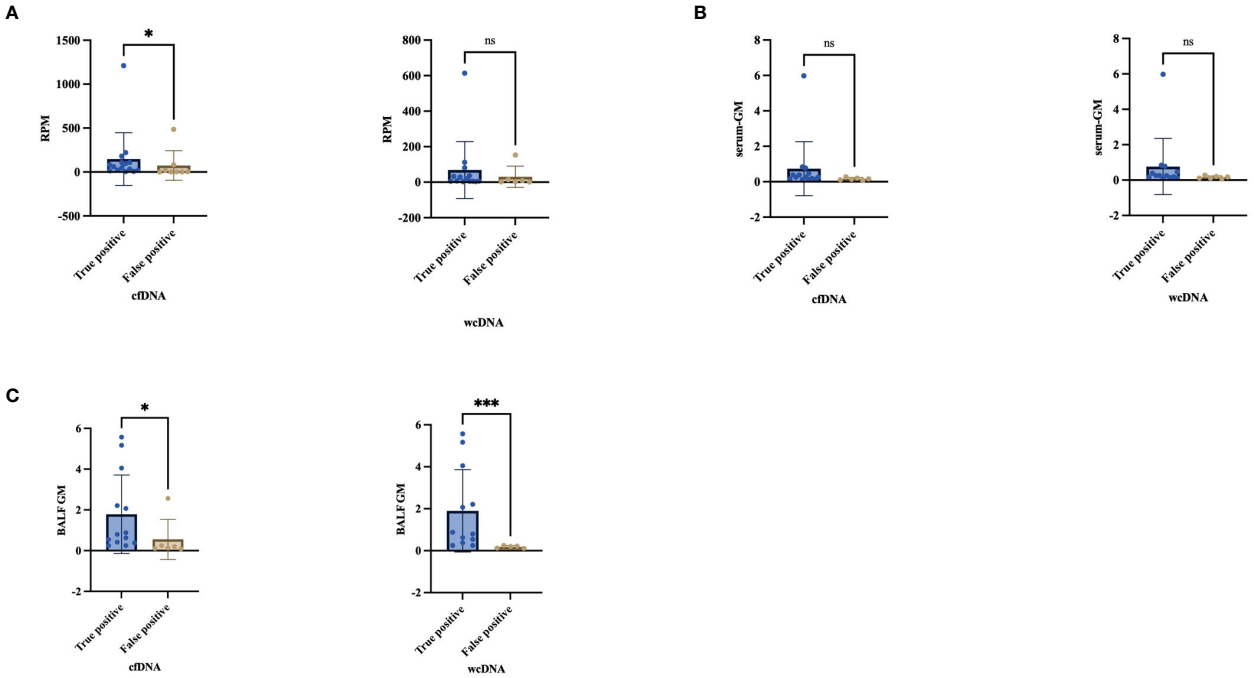



FIGURE 5
Comparison the difference between "True positive" and "False positive" by cfDNA, wcDNA. **(A)** The difference in the number of detected RPM between True positive and False positive. **(B, C)** Difference between serum-GM and BALF-GM in True positive and False positive. *P<0.05; ***P<0.001; ns, no significant.

BALF-mNGS might be the supplementary choice for the diagnosis of PA.

Rapid and precise identification of pathogens is hindered by the low sensitivity and time-consuming of conventional culture. Prior research indicates that mNGS permits impartial pathogen detection across diverse samples (Wei et al., 2022) and has outperformed culture in the investigation of infectious diseases (Zhang et al., 2020; Liang et al., 2022; Zuo et al., 2023). Our study demonstrates that BALF-mNGS (cfDNA, wcDNA) offers distinct advantages over culture in the diagnosis of PA with a high sensitivity. Conventional fungal and bacterial culture may overlook some pathogens due to differences in culture conditions, pathogen-specific incubation periods, and antimicrobial agents. Additionally, findings indicate that mNGS can promptly identify elusive pathogens that traditional culture methods may miss, possibly contributing to its superior sensitivity. The specificity of culture in diagnosing PA is not 100% and may be viewed as colonization or contamination. DNA extraction from certain PA

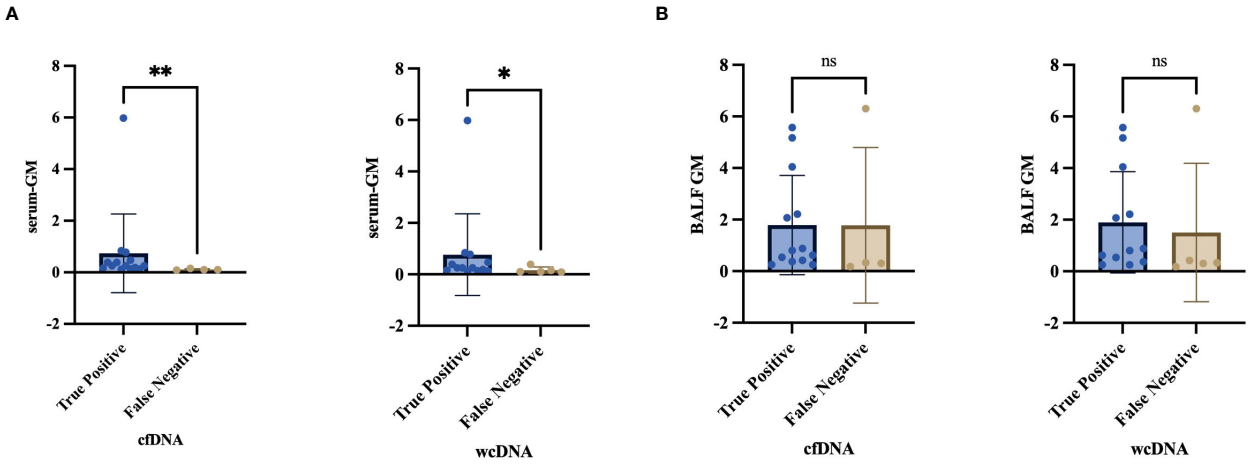



FIGURE 6
**(A, B)** Differential analysis of serum-GM, BALF-GM in the "True Positive" and "False Negative" PA by BALF-cfDNA, BALF-wcDNA. *P<0.05; **P<0.01; ns, no significant.

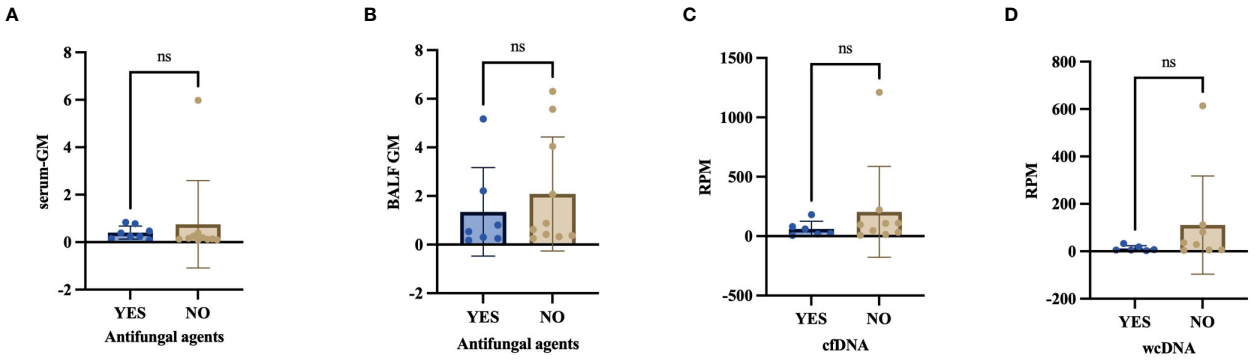



FIGURE 7
**(A–D)** Effect of antifungal agents on serum-GM, BALF-GM, RPM of Aspergillus detected by BALF-cfDNA and BALF-wcDNA. ns, no significant.

patients was hindered by *Aspergillus*' thick polysaccharide cell wall leading to negative mNGS results.

The diagnostic criteria for PA also include the GM test, which is frequently used clinically. GM test is an important mycologic evidence for diagnosing PA. Our findings indicate that BALF-mNGS has superior performance compared to BALF-GM for diagnosing PA. There have been limited studies comparing the diagnostic efficacy of BALF-mNGS with CMT in cases of PA. Thus, we conducted a comparison between the diagnostic performances of CMTs which includes culture and GM test to BALF-mNGS in diagnosing PA patients. While the sensitivity and specificity for CMTs were inferior to BALF-mNGS (cfDNA, wcDNA), there was no significant statistical difference between them. Combining a positive BALF-mNGS (cfDNA, wcDNA) or CMTs leads to a more effective diagnosis of PA than relying on CMTs alone, with higher sensitivity observed. This improvement can be attributed to the mNGS technique's capacity to identify microorganisms that are challenging to cultivate. However, the specificity of this combination of diagnostics is significantly reduced for a number of reasons. These include contamination during sample processing and the detection of non-pathogenic microbial DNA by macrogenomic sequencing techniques. These factors may result in the generation of false-positive results, thereby reducing the specificity of the diagnosis. Considering the cost-effectiveness and clinical value, it is recommended that CMTs be the primary option for suspected PA. BALF-mNGS testing could serve as a complementary approach in cases where CMTs are negative despite strong suspicion of *Aspergillus* infection. In contrast, the value of BALF-mNGS combined with CMTs in patients with neutropenic PA remains to be evaluated. Previous studies have indicated that in immunocompromised IPA patients, the diagnostic performance of mNGS was significantly superior to that of CMTs (Shi et al., 2023). Furthermore, the diagnostic efficacy of mNGS in combination with CT was superior to that of immunocompetent patients (Zhan et al., 2023). It is therefore hypothesized that in patients with neutropenic, the performance of mNGS in combination with CMTs for the diagnosis of PA should remain superior to that of CMTs and perform better than in non-neutropenic patients. However, this needs to be confirmed by further studies.

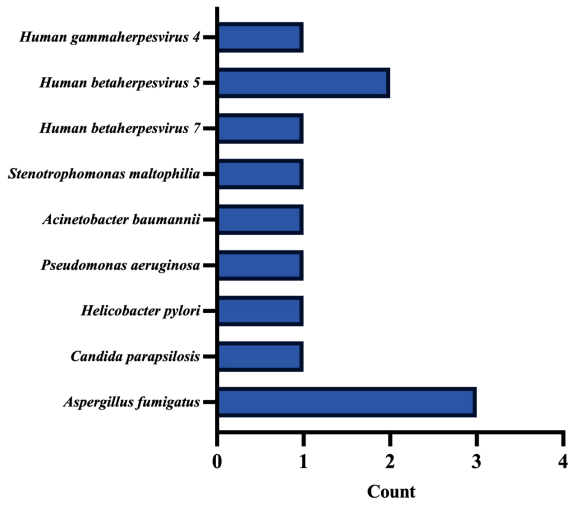



FIGURE 8
Pathogens distribution by blood cfDNA.

cfDNA and wcDNA have different performances in pathogen recognition due to different nucleic acid extraction methods. mNGS of cfDNA uses host DNA degradation methods, which can lead to the loss of cfDNA from the supernatant and the potential introduction of reagent contamination (Ji et al., 2020). However, the extraction of wcDNA from BALF sample via cell fragmentation without host DNA degradation increases the release of human DNA. Previous study has reported that cfDNA outperformed wcDNA in patients with pulmonary or central nervous system infections (He et al., 2022; Yu et al., 2022). However, there was no statistical difference between BALF-cfDNA and BALF-wcDNA in diagnosing PA. Our study illustrates that BALF-cfDNA is superior to BALF-wcDNA in the detection rate of pathogens and assessing *Aspergillus* infection, with a higher RPM for the microorganism. The ratio of DNA in the sample determines the sensitivity of mNGS (Ebinger et al., 2021). mNGS of cfDNA directly extract DNA from the supernatant of BALF (Gu et al., 2021). Conversely, wcDNA does not filter human DNA from the supernatant, potentially resulting in a greater ratio of pathogenic DNA for cfDNA than wcDNA derived

from the same BALF specimen. Besides, WcDNA may be better for the detection of the type of fungi. In this study, six microorganisms were exclusively detected by wcDNA, mostly fungi (5/6), and none had a high number of RPM detected. This occurrence may coincide with the fact that wcDNA extraction necessitates cell wall lysis and consequently detects more fungal species than cfDNA, or it may relate to sequencing contamination.

Plasma cell-free DNA sequencing has been widely used in clinical infectious diseases (Burnham et al., 2017; Armstrong et al., 2019). Consistent with the higher sensitivity of BALF-mNGS than blood-mNGS in patients with pneumonia as demonstrated by Chen et al. (Chen et al., 2020), our study similarly found that blood-mNGS was less sensitive than BALF-mNGS in detecting PA. Due to the limited number of blood samples in our study, we were unable to establish statistical significance in this comparison. We recruited the patients with CPA and IPA using blood-cfDNA. All patients with *Aspergillus* detected by blood-cfDNA were diagnosed with IPA except for one considered to be *Aspergillus* colonization. This could suggest that the blood-cfDNA is more suitable for patients with bloodstream rather than local infections, but further research is needed to confirm this. In contrast, the specificity of blood-mNGS in diagnosing PA was also lower than BALF-mNGS and CMTs in this study. This may be due to the small number of cases where blood-mNGS was used, and because *Aspergillus fumigatus* was detected in one patient by blood-mNGS who was later found to be colonized with *Aspergillus* following extensive clinical evaluation.

The most frequently detected pathogen was *Aspergillus fumigatus*, in agreement with previous research (Latgé and Chamilos, 2019). The identification of *Aspergillus* in respiratory samples does not necessarily indicate the presence of PA, as it is possible that the respiratory *Aspergillus* was colonized or contaminated during sequencing or other procedures. We try to distinguish the *Aspergillus* colonization by mNGS. Thus, patients with a confirmed PA diagnosis were identified as "True-positive" if *Aspergillus* was detected by the mNGS. Conversely, if the final diagnosis was non-PA but *Aspergillus* was detected by the mNGS, patients were classified as "False-positive". Additionally, a patient was labeled as "False-negative" if *Aspergillus* was not detected by the mNGS while they had a PA diagnosis.

The GM test identifies a polysaccharide antigen present in the cell wall of Aspergillus. The antigen can be released early during *Aspergillus* tissue invasion from the outer layer of the cell wall into the bloodstream, and can thus be detected in bodily fluids. The degree of fungal growth is reflected by the amount of antigen released, which in turn indicates the severity of the infection. In the investigation of patients with central nervous system infections, pathogen reads from mNGS cohered with modifications in cerebrospinal fluid WBC, exhibiting a connection between the number of pathogen readings and the extent of the disease's infection (Zhang et al., 2020). In this study, BALF-GM levels showed a positive correlation with the RPM of *Aspergillus*, true-positive patients had higher serum-GM levels than false-negative patients. Furthermore, serum-GM levels were found to be higher in individuals with positive blood-cfDNA for *Aspergillus* as compared to those with negative blood-cfDNA. This indirectly suggests that the RPM of Aspergillus in this study reflects the fungal loads. In several studies (Zhou et al., 2017; Sehgal et al., 2019; Dai et al., 2021), BALF-GM proved more effective than serum-GM in diagnosing non-neutropenic PA. Our study found BALF-GM to be more relevant than serum-GM in differentiating true-positive from false-positive patients. Wang et al (Wang et al., 2022), also discovered that true-positive patients with microorganisms detected by mNGS had more reads than false-positive patients in invasive fungal disease. Our previous study (Liu et al., 2021) demonstrates the capacity to distinguish Pneumocystis jirovecii Pneumonia and Pneumocystis jirovecii Colonization through pathogen reads using mNGS. In our study, true-positive patients displayed a greater RPM of *Aspergillus* than false-positive patients when utilizing BALF-cfDNA and the cut-off value was 4.5, whereas BALF-wcDNA was not significantly different. These findings suggest that BALF-cfDNA may be a more appropriate test for patients suspected to have PA than BALF-wcDNA. It appears that *Aspergillus* infection status can be inferred from the high or low RPM count.

Previous studies (Chen et al., 2020) have highlighted that using antimicrobial drugs prior to specimen collection can lead to inaccurate results in false-negative results in microbial cultures. Besides, Qing Miao et al. (Miao et al., 2018) demonstrated that mNGS is less susceptible to the effects of prior antibiotic exposure. Our study discovered a negligible impact of antifungal drug exposure on culture outcomes, CMTs, and the detection of *Aspergillus* by BALF-mNGS (cfDNA, wcDNA). This may be the reason for the shorter duration of antifungal drug use. It could also be due to the lower sensitivity of fungal cultures compared to bacterial cultures, resulting in antifungal drugs having a weaker effect. However, larger samples are needed to confirm this.

In this study, we also employed tNGS to assess its efficacy in diagnosing PA. tNGS represents a novel approach that combines the advantages of PCR and NGS. This technology amplifies target pathogen sequences by PCR, thereby reducing host nucleic acid interference and enhancing detection sensitivity. The results demonstrated that tNGS exhibited lower sensitivity values than BALF-mNGS in the diagnosis of PA. However, tNGS demonstrated comparable sensitivity to CMTs. This may be attributed to the fact that the tNGS technique employed in this study is limited to the detection of *Aspergillus fumigatus* and is unable to identify other types of *Aspergillus*, including *Aspergillus flavus* and *Aspergillus niger*, among others. The ROC curves also demonstrated superior diagnostic performance of the BALF-mNGS compared to tNGS, although the difference was not statistically significant. To date, no studies have been conducted to assess the applicability of tNGS in the context of non-neutropenic PA. Consequently, further studies with larger sample sizes are required to investigate the potential of tNGS in this area.

Our study has several limitations. Firstly, the sample size in this study was limited. In future studies, the sample size will be expanded in order to further stratify IPA and CPA, thereby providing more valuable information for the diagnosis and treatment of both diseases. Secondly, we were unable to retain blood from all enrolled patients at the same time as the BALF samples for simultaneous mNGS test, which prevented us from better assessing its diagnostic efficacy. Thirdly, due to COVID-19, the BALF sample could not be transported to the business lab for mNGS testing in time, which prevented the results from serving as a recommendation for using antimicrobial drugs.

## 5 Conclusion

In summary, the results indicate that BALF-mNGS (cfDNA, wcDNA) performs better than CMTs in detecting pathogens. Furthermore, when diagnosing non-neutropenic PA, BALF-mNGS (cfDNA, wcDNA) shows similarity to CMTs but superiority to culture and BALF-GM. Combining BALF-mNGS (cfDNA, wcDNA) with CMTs shows the potential to enhance the sensitivity of diagnostic performance. Additionally, the RPM of *Aspergillus* serves as an indicator of fungal loads. An additional advantage of BALF-cfDNA over BALF-wcDNA is the ability to differentiate between "True positive" and "False positive" patients with PA. Therefore, mNGS of BALF-cfDNA may present a novel diagnostic technology for PA.

## Data availability statement

The data presented in the study are deposited in the National Genomics Data Center (https://ngdc.cncb.ac.cn/?lang=zh) repository, accession number PRJCA028215.

## Ethics statement

The studies involving humans were approved by Ethics Committee of the Nanjing Jingling Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

XC: Data curation, Formal analysis, Writing – original draft. CS: Data curation, Writing – review & editing, Formal analysis. HZ: Data curation, Writing – review & editing. YC: Data curation, Writing – review & editing. MC: Data curation, Writing – review & editing. LW: Data curation, Validation, Writing – review & editing. WS: Data curation, Writing – review & editing. YT: Data curation, Writing – review & editing. GM: Data curation, Writing – review & editing. BH: Data curation, Writing – review & editing. SY: Data curation, Writing – review & editing. JZ: Data curation, Writing – review & editing. JW: Data curation, Writing – review & editing. YJL: Data curation, Writing – review & editing. YG: Methodology, Writing – review & editing. MS: Data curation, Writing – review & editing. YW: Data curation, Writing – review & editing. YYL: Data curation, Writing – review & editing. XS: Funding acquisition, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Project of Natural Science Foundation of China (82270019, 82070011), the Key Project of Jiangsu Commission of Health (K2019004), fundings for Clinical Trials from the Affiliated Drum Tower Hospital, Medical School of Nanjing University (2023-LCYJ-MS-18).

## Acknowledgments

We would like to express our gratitude to KingMed for providing the tNGS test and to all those who participated in this study.

## Conflict of interest

Author YG was employed by Hugobiotech Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1398190/full#supplementary-material

## References

Armstrong, A. E., Rossoff, J., Hollemon, D., Hong, D. K., Muller, W. J., and Chaudhury, S. (2019). Cell-free DNA next-generation sequencing successfully detects infectious pathogens in pediatric oncology and hematopoietic stem cell transplant patients at risk for invasive fungal disease. *Pediatr. Blood Cancer* 66, e27734. doi: 10.1002/pbc.27734

Burnham, P., Khush, K., and De Vlaminck, I. (2017). Myriad applications of circulating cell-free DNA in precision organ transplant monitoring. *Ann. Am. Thorac. Soc.* 14, S237–S241. doi: 10.1513/AnnalsATS.201608-634MG

Chen, X., Ding, S., Lei, C., Qin, J., Guo, T., Yang, D., et al. (2020). Blood and bronchoalveolar lavage fluid metagenomic next-generation sequencing in pneumonia. *Can. J. Infect. Dis. Med. Microbiol.* 2020, 6839103. doi: 10.1155/2020/6839103

Dai, Z., Cai, M., Yao, Y., Zhu, J., Lin, L., Fang, L., et al. (2021). Comparing the diagnostic value of bronchoalveolar lavage fluid galactomannan, serum galactomannanan, and serum 1,3-beta-d-glucan in non-neutropenic respiratory disease patients with invasive pulmonary aspergillosis. *Med. (Baltimore)* 100, e25233. doi: 10.1097/MD.0000000000025233

Denning, D. W., Cadranel, J., Beigelman-Aubry, C., Ader, F., Chakrabarti, A., Blot, S., et al. (2016). Chronic pulmonary aspergillosis: rationale and clinical guidelines for diagnosis and management. *Eur. Respir. J.* 47, 45–68. doi: 10.1183/13993003.00583-2015

Donnelly, J. P., Chen, S. C., Kauffman, C. A., Steinbach, W. J., Baddley, J. W., Verweij, P. E., et al. (2020). Revision and update of the consensus definitions of invasive fungal disease from the european organization for research and treatment of cancer and the mycoses study group education and research consortium. *Clin. Infect. Dis.* 71, 1367–1376. doi: 10.1093/cid/ciz1008

Ebinger, A., Fischer, S., and Hoper, D. (2021). A theoretical and generalized approach for the assessment of the sample-specific limit of detection for clinical metagenomics. *Comput. Struct. Biotechnol. J.* 19, 732–742. doi: 10.1016/j.csbj.2020.12.040

El-Baba, F., Gao, Y., and Soubani, A. O. (2020). Pulmonary aspergillosis: what the generalist needs to know. *Am. J. Med.* 133, 668–674. doi: 10.1016/j.amjmed.2020.02.025

Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryke, D., et al. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* 27, 115–124. doi: 10.1038/s41591-020-1105-z

Guinea, J., Torres-Narbona, M., Gijon, P., Munoz, P., Pozo, F., Pelaez, T., et al. (2010). Pulmonary aspergillosis in patients with chronic obstructive pulmonary disease: incidence, risk factors, and outcome. *Clin. Microbiol. Infect.* 16, 870–877. doi: 10.1111/j.1469-0691.2009.03015.x

He, P., Wang, J., Ke, R., Zhang, W., Ning, P., Zhang, D., et al. (2022). Comparison of metagenomic next-generation sequencing using cell-free DNA and whole-cell DNA for the diagnoses of pulmonary infections. *Front. Cell Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.1042945

He, Q., Li, H., Rui, Y., Liu, L., He, B., Shi, Y., et al. (2018). Pentraxin 3 gene polymorphisms and pulmonary aspergillosis in chronic obstructive pulmonary disease patients. *Clin. Infect. Dis.* 66, 261–267. doi: 10.1093/cid/cix749

Ji, X. C., Zhou, L. F., Li, C. Y., Shi, Y. J., Wu, M. L., Zhang, Y., et al. (2020). Reduction of human DNA contamination in clinical cerebrospinal fluid specimens improves the sensitivity of metagenomic next-generation sequencing. *J. Mol. Neurosci.* 70, 659–666. doi: 10.1007/s12031-019-01472-z

Latgé, J. P., and Chamilos, G. (2019). Aspergillus fumigatus and aspergillosis in 2019. *Clin. Microbiol. Rev.* 33, e00140-18. doi: 10.1128/CMR.00140-18

Liang, M., Fan, Y., Zhang, D., Yang, L., Wang, X., Wang, S., et al. (2022). Metagenomic next-generation sequencing for accurate diagnosis and management of lower respiratory tract infections. *Int. J. Infect. Dis.* 122, 921–929. doi: 10.1016/j.ijid.2022.07.060

Liu, L., Yuan, M., Shi, Y., and Su, X. (2021). Clinical Performance of BAL Metagenomic Next-Generation Sequence and Serum (1,3)-β-D-Glucan for Differential Diagnosis of Pneumocystis jirovecii Pneumonia and Pneumocystis jirovecii Colonisation. *Front. Cell Infect. Microbiol.* 11. doi: 10.3389/fcimb.2021.784236

Lu, Y., Liu, L., Li, H., Chen, B., Gu, Y., Wang, L., et al. (2023). The clinical value of Aspergillus-specific IgG antibody test in the diagnosis of nonneutropenic invasive pulmonary aspergillosis. *Clin. Microbiol. Infect.* 29 (6), 797.e1–797.e7. doi: 10.1016/j.cmi.2023.02.002

Miao, Q., Ma, Y., Wang, Q., Pan, J., Zhang, Y., Jin, W., et al. (2018). Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin. Infect. Dis.* 67, S231–s240. doi: 10.1093/cid/ciy693

Patterson, T. F., Thompson, G. R., Denning, D. W., Fishman, J. A., Hadley, S., Herbrecht, R., et al. (2016). Practice guidelines for the diagnosis and management of aspergillosis: 2016 update by the infectious diseases society of America. *Clin. Infect. Dis.* 63, e1–e60. doi: 10.1093/cid/ciw326

Sehgal, I. S., Dhooria, S., Choudhary, H., Aggarwal, A. N., Garg, M., Chakrabarti, A., et al. (2019). Utility of serum and bronchoalveolar lavage fluid galactomannan in diagnosis of chronic pulmonary aspergillosis. *J. Clin. Microbiol.* 57, e01821-18. doi: 10.1128/JCM.01821-18

Shi, Y., Peng, J. M., Hu, X. Y., Yang, Q. W., and Wang, Y. (2023). Metagenomic next-generation sequencing for detecting Aspergillosis pneumonia in immunocompromised patients: a retrospective study. *Front. Cell Infect. Microbiol.* 13. doi: 10.3389/fcimb.2023.1209724

Taccone, F. S., Van den Abeele, A. M., Bulpa, P., Misset, B., Meersseman, W., Cardoso, T., et al. (2015). Epidemiology of invasive aspergillosis in critically ill patients: clinical presentation, underlying conditions, and outcomes. *Crit. Care* 19, 7. doi: 10.1186/s13054-014-0722-7

Ullmann, A. J., Aguado, J. M., Arikan-Akdagli, S., Denning, D. W., Groll, A. H., Lagrou, K., et al. (2018). Diagnosis and management of Aspergillus diseases: executive summary of the 2017 ESCMID-ECMM-ERS guideline. *Clin. Microbiol. Infect.* 24 Suppl 1, e1–e38. doi: 10.1016/j.cmi.2018.01.002

Wang, C., You, Z., Fu, J., Chen, S., Bai, D., Zhao, H., et al. (2022). Application of metagenomic next-generation sequencing in the diagnosis of pulmonary invasive fungal disease. *Front. Cell Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.949505

Wei, P., Wu, L., Li, Y., Shi, J., Luo, Y., Wu, W., et al. (2022). Metagenomic next-generation sequencing for the detection of pathogenic microorganisms in patients with pulmonary infection. *Am. J. Transl. Res.* 14, 6382–6388.

Wilson, M. R., Naccache, S. N., Samayoa, E., Biagtan, M., Bashir, H., Yu, G., et al. (2014). Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N. Engl. J. Med.* 370, 2408–2417. doi: 10.1056/NEJMoa1401268

Yu, L., Zhang, Y., Zhou, J., Zhang, Y., Qi, X., Bai, K., et al. (2022). Metagenomic next-generation sequencing of cell-free and whole-cell DNA in diagnosing central nervous system infections. *Front. Cell Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.951703

Zhan, W., Liu, Q., Yang, C., Zhao, Z., Yang, L., Wang, Y., et al. (2023). Evaluation of metagenomic next-generation sequencing diagnosis for invasive pulmonary aspergillosis in immunocompromised and immunocompetent patients. *Mycoses* 66, 331–337. doi: 10.1111/myc.13557

Zhang, Y., Cui, P., Zhang, H. C., Wu, H. L., Ye, M. Z., Zhu, Y. M., et al. (2020). Clinical application and evaluation of metagenomic next-generation sequencing in suspected adult central nervous system infection. *J. Transl. Med.* 18, 199. doi: 10.1186/s12967-020-02360-6

Zhou, W., Li, H., Zhang, Y., Huang, M., He, Q., Li, P., et al. (2017). Diagnostic value of galactomannan antigen test in serum and bronchoalveolar lavage fluid samples from patients with nonneutropenic invasive pulmonary aspergillosis. *J. Clin. Microbiol.* 55, 2153–2161. doi: 10.1128/JCM.00345-17

Zuo, Y. H., Wu, Y. X., Hu, W. P., Chen, Y., Li, Y. P., Song, Z. J., et al. (2023). The clinical impact of metagenomic next-generation sequencing (mNGS) test in hospitalized patients with suspected sepsis: A multicenter prospective study. *Diagnostics (Basel)* 13, 323. doi: 10.3390/diagnostics13020323

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 06 August 2024
DOI 10.3389/fcimb.2024.1439472



OPEN ACCESS

EDITED BY
Qing Wei,
Genskey Co. Ltd, China

REVIEWED BY
Xiao Tang,
Capital Medical University, China
Zhen Yang,
Nankai University, China
Xu Liu,
University of Oxford, United Kingdom

*CORRESPONDENCE
Yanfang Jiang
yanfangjiang@hotmail.com;
yanfangjiang@jlu.edu.cn

RECEIVED 28 May 2024
ACCEPTED 09 July 2024
PUBLISHED 06 August 2024

CITATION
Liu Y, Wu W, Xiao Y, Zou H, Hao S and Jiang Y (2024) Application of metagenomic next-generation sequencing and targeted metagenomic next-generation sequencing in diagnosing pulmonary infections in immunocompetent and immunocompromised patients.
*Front. Cell. Infect. Microbiol.* 14:1439472.
doi: 10.3389/fcimb.2024.1439472



# Application of metagenomic next-generation sequencing and targeted metagenomic next-generation sequencing in diagnosing pulmonary infections in immunocompetent and immunocompromised patients

Yong Liu[1], Wencai Wu[2], Yunping Xiao[1], Hongyan Zou[1], Sijia Hao[1] and Yanfang Jiang[1]*

[1]Genetic Diagnosis Center, The First Hospital of Jilin University, Changchun, China, [2]Key Laboratory of Biomedical Information Engineering of Ministry of Education, Biomedical Informatics & Genomics Center, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

**Background:** Metagenomic next-generation sequencing (mNGS) technology has been widely used to diagnose various infections. Based on the most common pathogen profiles, targeted mNGS (tNGS) using multiplex PCR has been developed to detect pathogens with predesigned primers in the panel, significantly improving sensitivity and reducing economic burden on patients. However, there are few studies on summarizing pathogen profiles of pulmonary infections in immunocompetent and immunocompromised patients in Jilin Province of China on large scale.

**Methods:** From January 2021 to December 2023, bronchoalveolar lavage fluid (BALF) or sputum samples from 546 immunocompetent and immunocompromised patients with suspected community-acquired pneumonia were collected. Pathogen profiles in those patients on whom mNGS was performed were summarized. Additionally, we also evaluated the performance of tNGS in diagnosing pulmonary infections.

**Results:** Combined with results of mNGS and culture, we found that the most common bacterial pathogens were *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii* in both immunocompromised and immunocompetent patients with high detection rates of *Staphylococcus aureus* and *Enterococcus faecium*, respectively. For fungal pathogens, *Pneumocystis jirovecii* was commonly detected in patients, while fungal infections in immunocompetent patients were mainly caused by *Candida albicans*. Most of viral infections in patients were caused by Human betaherpesvirus 5 and Human gammaherpesvirus 4. It is worth noting that, compared with immunocompetent patients (34.9%, 76/218), more mixed infections were found in immunocompromised patients (37.8%, 14/37). Additionally, taking final comprehensive clinical diagnoses as reference standard, total coincidence rate of BALF tNGS (81.4%, 48/59) was much higher than that of BALF mNGS (40.0%, 112/280).

**Conclusions:** Our findings supplemented and classified the pathogen profiles of pulmonary infections in immunocompetent and immunocompromised patients in Jilin Province of China. Most importantly, our findings can accelerate the development and design of tNGS specifically used for regional pulmonary infections.

KEYWORDS

**pulmonary infection, bronchoalveolar lavage fluid, TNGS, MNGs, immunocompetent and immunocompromised**

# Introduction

Although some progress has been made in the treatment of pulmonary infections, pulmonary infection is still an important cause of mortality worldwide (Ravimohan et al., 2018; D'Anna et al., 2021; Schneider et al., 2021). Therefore, early identification of pathogens is crucial for rapid clinical diagnosis and initial treatment. Although there are various detection methods, it is still a challenge to quickly and accurately diagnose pulmonary infections (Luyt et al., 2020; Turcios, 2020). Conventional methods, including microbial culture (Handel et al., 2021), microscopic smear (Yue et al., 2021), polymerase chain reaction (PCR) (Ramachandran and Wilson, 2020), nucleic acid hybridization, histopathology (Greninger and Naccache, 2019), and serological antibody detection (Simner et al., 2018), can only identify about 40% of pathogens, with long detection time and low positive detection rate, which cannot meet the current clinical needs (Qu et al., 2022; Shi et al., 2022). Besides, empirical therapy is often ineffective in treating immunocompromised patients with atypical pneumonia (Azoulay et al., 2019). Therefore, rapid and accurate identification of pathogens is very important for the clinical intervention of these patients.

At present, the metagenomic next-generation sequencing (mNGS) is a new and rapidly developing pathogen diagnosis technology (Han et al., 2019). mNGS has the advantages of short detection time and wide detection range (Gökdemir et al., 2022), which can accurately identify bacteria, fungi, viruses, parasites, and other pathogens (Indelli et al., 2021). Clinicians has extensively used mNGS to diagnose various infections, such as bloodstream infection, abdominal cavity infection, central nervous system infection, etc (Greninger and Naccache, 2019; Thakur, 2020; Li et al., 2022). mNGS has become a promising detection method for infectious diseases (Zhou et al., 2022).

Recently, based on the most common pathogen profiles, targeted mNGS (tNGS) using multiplex PCR (Xie et al., 2022), reducing economic burden on patients, has been developed to detect pathogens with predesigned primers in the panel (Li et al., 2021; Huang et al., 2023). However, the occurrence of infections in immunocompromised patients caused by rare (Zhan et al., 2021), regional (Ramirez et al., 2020), and emerging (El Zein et al., 2020; Fishman, 2023) pathogens limited the application of this tNGS in those patients. However, there are few studies on summarizing pathogen profiles of pulmonary infections in immunocompetent and immunocompromised patients in Jilin Province of China on large scale.

Given the advantage of tNGS and comprehensiveness of mNGS, it is necessary to summarize pathogen profiles of pulmonary infections in Jilin Province of China using mNGS to provide reference for designing and developing regional tNGS. Therefore, our study retrospectively enrolled immunocompetent and immunocompromised patients with suspected community-acquired pneumonia, evaluated the value of mNGS in the diagnosis of pulmonary infections by comparing the diagnostic performance of mNGS and conventional culture using sputum and bronchoalveolar lavage fluid (BALF) samples, and summarized pathogen profiles in both immunocompetent and immunocompromised patients. Additionally, we also summarized the coincidence rate of tNGS results against final comprehensive clinical diagnoses.

# Methods

## Ethics statement

This study was reviewed and approved by the Ethical Review Committee of the First hospital of Jilin University (approval no. 2024–612). All procedures followed were in strict compliance with the Ethical Review of Biomedical Research Involving Human Subjects (2016), the Declaration of Helsinki, and the International Ethical Guidelines for Biomedical Research Involving Human Subjects. The study was exempted from requiring informed consent by the Ethical Review Committee as it was a retrospective study and patient data were anonymized.

## Patients and sample collection

A total of 546 patients admitted to The First Hospital of Jilin University from January 2021 to December 2023 and diagnosed as suspected community-acquired pneumonia were retrospectively

enrolled (Figure 1). Specimens were subjected to conventional culture and mNGS/tNGS testing in parallel. All patients with suspected pulmonary infection had abnormal chest imaging results. The inclusion criteria were as follows: (1) Patients with symptoms such as fever, cough, sputum, dyspnea, and imaging abnormalities, such as lung shadows, space-occupying lesions, and other signs of pulmonary infection, (2) Patients on whom mNGS/tNGS was performed, and (3) Patients with complete clinical data. The exclusion criteria were as follows: (1) Patients on whom mNGS/tNGS was not performed and (2) Patients with incomplete clinical and laboratory data.

## mNGS pipeline

BALF and sputum samples were collected based on the standard clinical procedure. After collection, BALF or sputum samples were immediately inactivated at 65°C for 30 minutes. A 1.5 mL microcentrifuge tube with 0.5 mL sample and 1 g 0.5 mm glass beads was attached to a horizontal platform on a vortex mixer and agitated vigorously at 2800–3200 rpm for 30 min. Subsequently, 0.3 mL supernatant sample was separated into a new 1.5 mL microcentrifuge tube and DNA was extracted using the TIANamp Micro DNA Kit (DP316, TIANGEN BIOTECH) according to the manufacturer's recommendation. Qubit 4.0 (Thermo Fisher Scientific, MA, USA) was used to measure extracted DNA concentrations. QIAseq Ultralow Input Library Kit (QIAGEN, Hilden, Germany) was used to construct metagenomics libraries (Ji et al., 2020). Inspected and qualified library was sequenced on Nextseq 550 platform (Illumina, San Diego, USA).

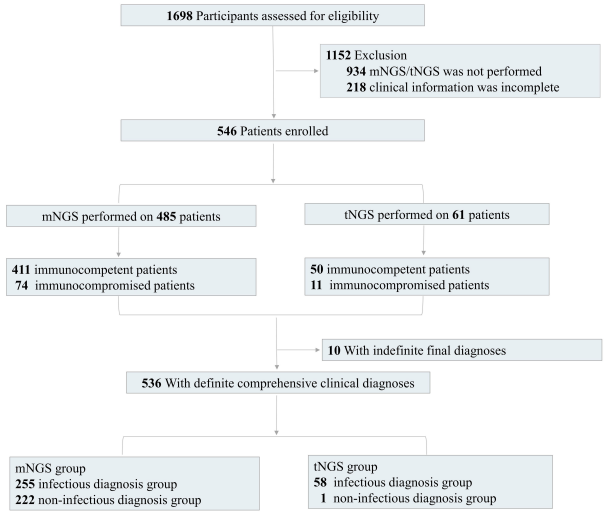



FIGURE 1
Flow diagram. We retrospectively assessed the eligibility of patients. According to the inclusion and exclusion criteria, we finally enrolled 546 patients with detection results of mNGS/tNGS and culture and complete clinical information, including 485 patients with mNGS results and 61 patients with tNGS results. We also divided the enrolled patients into immunocompetent and immunocompromised groups. There were 10 patients with indefinite final diagnoses.

## Commercial tNGS pipeline

Based on multiplex PCR and mNGS, in-house tNGS panel was designed to detect 273 pathogens, including 113 bacteria, 47 fungi, 101 viruses, and 12 parasites, causing infections in different systems according to the public pathogen databases and the published studies. After nucleic acid extraction, multiplex PCR with the designed primers was used to construct libraries. Library concentrations were quantified using Qubit 4.0 and Illumina NextSeq platform was used for high-throughput sequencing.

## Analysis of sequencing data

Raw sequencing data were exposed to quality control, including the removal of low-quality reads (Q score cutoff, 20). High-quality sequencing data were generated using in-house software, followed by computational subtraction of human host sequences mapped to the human reference genome (hg38) using Burrows-Wheeler Alignment (Li and Durbin, 2009). The remaining data were blasted against in-house classification reference databases, which were constructed according to the published microbial genome databases, including reference sequence database at National Center for Biotechnology Information. The constructed databases contain 25,863 pathogens, including 12,142 bacteria, 2,680 fungi, 10,061 viruses (including DNA and RNA viruses), 654 parasites, 206 mycobacteria, and 120 mycoplasma/chlamydia.

## Interpretation of mNGS/tNGS results and diagnostic assessment

As control, negative and positive controls were also set with the same procedure and bioinformatics analysis. Strictly map reads number (SMRN) and genomic coverage were analyzed. SMRN represents the number of sequences that are strictly aligned with the microorganism at species level.

During the interpretation process, positive mNGS/tNGS results were defined as follows:

1) Bacteria, fungi, parasites, mycoplasmas, and chlamydiae: When the microorganism was not detected in the negative control ('No template' control, NTC) and genome coverage of detected sequences belonged to this microorganism ranked top10 among the microbes in the same genus or when its ratio of $SMRN_{sample}$ to $SMRN_{NTC}$ was > 10 if the $SMRN_{NTC}\neq 0$. Besides, the $SMRN_{sample}$ of bacteria, fungi, mycoplasmas, and chlamydiae should be ≥ 3, while the $SMRN_{sample}$ of parasites should be ≥ 100. $SMRN_{sample}$ of *Mycobacterium tuberculosis* should be ≥ 1.

2) Viruses: When the virus was not detected in NTC or when $SMRN_{sample}/SMRN_{NTC}$ was > 5 if the $SMRN_{NTC}\neq 0$. Besides, $SMRN_{sample}$ of viruses should be ≥ 1.

Subsequently, positive mNGS/tNGS results were further defined according to whether the detected microbes by mNGS/tNGS were the most commonly reported pathogens or the infections caused by the microbes were in accordance with clinical features of patients.

For diagnostic assessment, 2–3 clinical adjudicators independently made the final comprehensive clinical diagnoses and defined the causative pathogens, according to clinical characteristics, laboratory examinations, response of the patients to treatment, mNGS/tNGS results, and clinical experiences. Based on final comprehensive clinical diagnoses, we divided the enrolled patients into infectious diagnosis group, non-infectious diagnosis group, and indefinite clinical diagnosis group, which were used as reference standard to evaluate the performance of mNGS/tNGS.

The definition of infectious diagnoses was based on 1) At least one of culture result, clinical characteristics, or clinical experiences suggested pulmonary infections, or 2) For the patients without any laboratory examinations or with negative laboratory examinations, pneumonia was relieved after treatment according to the mNGS/tNGS results. The definition of non-infectious diagnoses was based on 1) No pathogens were detected by both culture and mNGS/tNGS, and 2) The inflammation was relieved after the use of glucocorticoid or immunosuppressant. Indefinite clinical diagnosis group included patients whose clinical characteristics or laboratory examinations were not adequate for diagnoses and patients lost during follow-up duration.

Subsequently, causative pathogens can be defined. The non-infectious diagnosis group was used to evaluate the specificity of mNGS, tNGS and culture. In addition, we can accurately divide the detection results into true-positive, false-negative, false-positive, and true-negative results. Guided by mNGS/tNGS results, we also adjusted the therapeutic regimens, which was approved by the Ethical Review Committee.

## Statistical analysis

Continuous variables were expressed as mean ± standard deviation (SD) and categorical variables were presented as numbers (percentage). We tested for differences in continuous variables using t test and categorical variables with chi-square test as appropriate. The $2 \times 2$ contingency tables were established to determine sensitivity, specificity, and total coincidence rate (TCR), and the results are presented with 95% confidence intervals (CIs). Data analyses were performed using SPSS 22.0 software and a two-tailed value of $P$ of $< 0.05$ was considered to represent a significant difference.

## Data availability

Sequencing data were deposited to the National Genomics Data Center under accession numbers CRA016476 and CRA017510. The authors declare that the main data supporting the findings are available within this article. The other data generated and analyzed for this study are available from the corresponding author upon reasonable request.

# Results

## Baseline characteristics of enrolled patients

A total of 546 patients [325 males (59.5%) and 221 females (40.5%)] with suspected community-acquired pulmonary infection were included in this study (Table 1). Most of the patients were over 40 years old (72.0%), while the patients aged from 40 to 70 years old accounted for 46.7% of the enrolled patients. Among them, 134 (24.5%) patients had underlying diseases, including diabetes ($n$ =42, 7.7%), lung cancer ($n$ =22, 4.0%), previous history of tuberculosis ($n$ =18, 3.3%), bronchiectasis ($n$ =16, 2.9%), chronic obstructive pulmonary disease ($n$ =14, 2.6%), extrapulmonary malignancies ($n$ =11, 2.0%), bronchial asthma ($n$ =9, 1.6%), and connective tissue disease ($n$ =2, 0.4%).

Additionally, 284 BALF and 201 sputum samples for mNGS test were respectively collected from 485 patients, including 411 immunocompetent patients and 74 immunocompromised patients. The tNGS test was performed on BALF samples from the other 61 patients, including 50 immunocompetent patients and 11 immunocompromised patients (Figure 1). The immunodeficiency

TABLE 1 Baseline characteristics of 546 patients.

| | Number of cases | Percentage (%) |
|---|---|---|
| **Sex** | | |
| Male | 325 | 59.5 |
| Female | 221 | 40.5 |
| **Age (years)** | | |
| ≤ 40 | 153 | 28.0 |
| >40, ≤70 | 255 | 46.7 |
| > 70 | 138 | 25.3 |
| **Underlying illness** | | |
| Bronchiectasis | 16 | 2.9 |
| Chronic obstructive pulmonary disease | 14 | 2.6 |
| Previous history of tuberculosis | 18 | 3.3 |
| Bronchial asthma | 9 | 1.6 |
| Lung cancer | 22 | 4.0 |
| Diabetes | 42 | 7.7 |
| Connective tissue disease | 2 | 0.4 |
| Extrapulmonary malignancies | 11 | 2.0 |
| **Sample type** | | |
| BALF | 345 (284 for mNGS, 61 for tNGS) | 63.2 |
| Sputum | 201 (all for mNGS) | 36.8 |

BALF, bronchoalveolar lavage fluid.

10.3389/fcimb.2024.1439472

was mainly caused by the above underlying diseases, including diabetes, lung cancer, extrapulmonary malignancies, connective tissue disease, or the other diseases with long-term immunosuppressive treatment.

## Comparison between mNGS and culture

mNGS and culture were simultaneously performed on BALF and sputum samples and the diagnostic value of mNGS and culture using different kinds of samples was summarized (Figure 2). For BALF and sputum samples, the positive rate of mNGS can reach up to 88.0%, which was much higher than that of culture (50.2%) (Figure 2IA). For BALF samples, the positive rate of mNGS (87.3%) was much higher than that of culture (44.8%) (Figure 2IIA). For sputum samples, a similar trend (89.1% vs 58.3%) was found (Figure 2IIIA).

For BALF and sputum samples, sensitivity [50.8% (95% CI 44.7–56.9%)] of mNGS was much higher than that of culture [22.5% (95% CI 17.3–27.8%)], taking final comprehensive clinical diagnoses as reference standard (Figure 2IB). The specificity of culture can reach up to 56.1%, while the specificity of mNGS was only 18.0% (95% CI 13.0–23.1%). Besides, there was no significant difference in TCR between mNGS [35.6% (95% CI 31.3–39.9%)] and culture [38.2% (95% CI 33.8–42.7%)]. Similar trends were found in the comparison between BALF mNGS and BALF culture (Figure 2IIB) or sputum mNGS and sputum culture (Figure 2IIIB). In addition, the sensitivity [53.7% (95% CI 46.0–61.4%)] and TCR [40.0% (95% CI 34.3–45.7%)] of BALF mNGS were higher than those of sputum mNGS [45.7% (95% CI 35.7–55.8%) and 29.3% (95% CI 23.0–35.6%)]. The above results indicate that compared with sputum sample, BALF sample was preferred sample for mNGS in detecting pathogens of pulmonary infections.

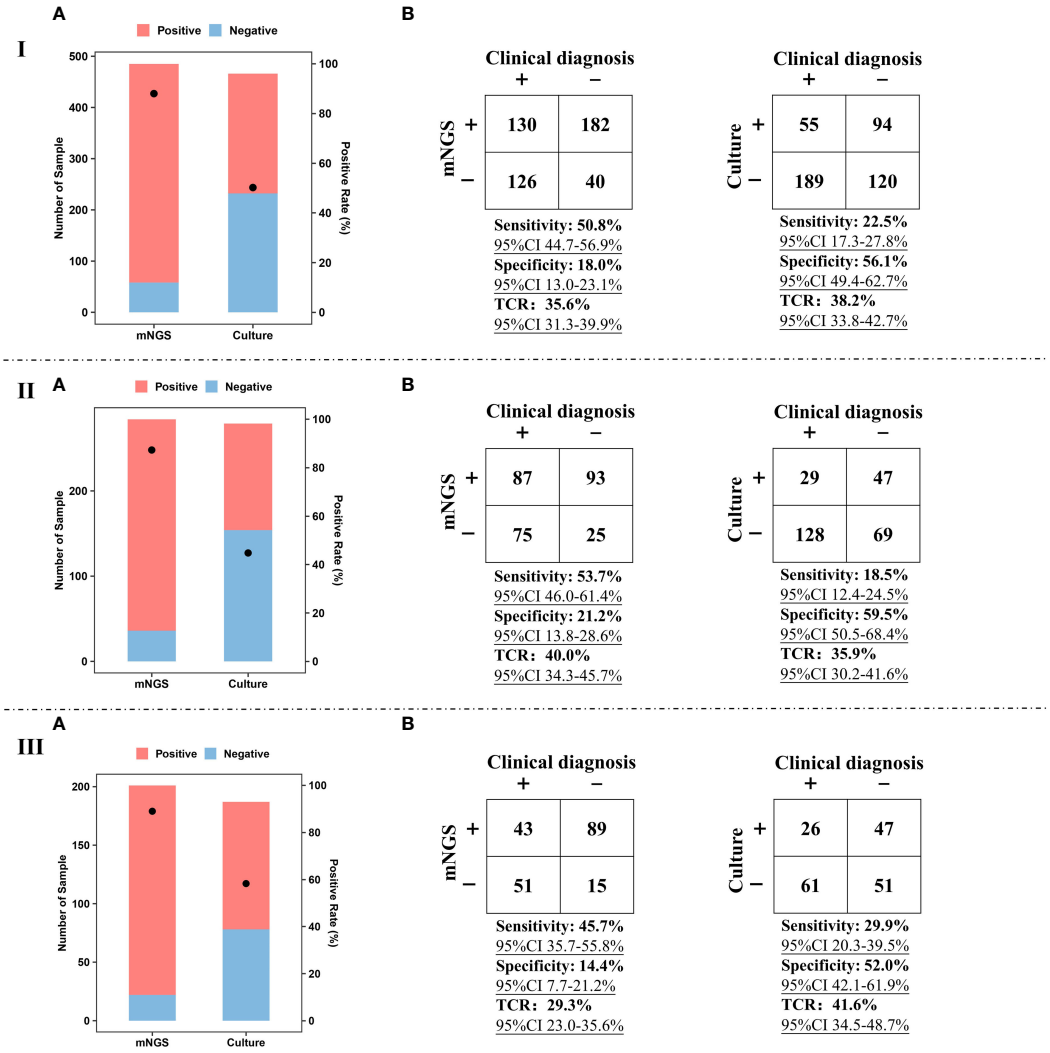



FIGURE 2
Performance of mNGS and culture. I, Positive rate **(A)**, sensitivity, specificity, and TCR **(B)** of mNGS and culture using BALF and sputum samples. II, Positive rate **(A)**, sensitivity, specificity, and TCR **(B)** of mNGS and culture using BALF samples. III, Positive rate **(A)**, sensitivity, specificity, and TCR **(B)** of mNGS and culture using sputum samples. The 'TCR' is total coincidence rate, and 'CI' represents confidence interval.

10.3389/fcimb.2024.1439472

## mNGS in immunocompetent and immunocompromised patients with pulmonary infections

We divided the enrolled patients on whom mNGS was performed into immunocompetent ($n$ = 411) and immunocompromised ($n$ = 74) groups, and the diagnostic positive rates of mNGS and conventional culture in the two groups were calculated (Figure 3). It was found that the performance of mNGS was significantly higher than that of culture in both immunocompetent and immunocompromised groups. However, there was no significant difference in the positive rate of mNGS between immunocompromised (90.5%) (Figure 3IA) and immunocompetent (87.6%) (Figure 3IIA) groups ($P$>0.05). A similar trend was found in the positive rate of culture (47.1% and 50.8%) ($P$>0.05).

Among the infected patients in immunocompromised group, single infection accounted for 62.1% (23/37). For single infection, the most common pathogens were bacteria (35.1%, 13/37), followed by fungi (16.2%, 6/37) and viruses (10.8%, 4/37) (Figure 3IB). In addition to single infection, our results show that the number of patients with mixed infection accounted for 37.8% (14/37) of immunocompromised patients, including bacterial-fungal co-infection ($n$ = 2), bacterial-viral co-infection ($n$ = 4), fungal-viral co-infection ($n$ = 4), and bacterial-fungal-viral co-infection ($n$ = 2) (Figure 3IB).

The pathogen profiles of immunocompromised patients were summarized (Figure 3IC). We found that *Pseudomonas aeruginosa* ($n$ = 6), *Klebsiella pneumoniae* ($n$ = 4), *Acinetobacter baumannii* ($n$ = 4), and *Staphylococcus aureus* ($n$ = 4) were the most common bacterial pathogens in immunocompromised patients with pulmonary infection. The most common fungal pathogens were *Pneumocystis jirovecii* ($n$ = 8) and *Candida parapsilosis* ($n$ = 3). Besides, viral infection was mainly caused by Human betaherpesvirus 5 (CMV) ($n$ = 11), Human gammaherpesvirus 4 (EBV) ($n$ = 3), and human parvovirus B19 ($n$ = 2).

For the infected patients in immunocompetent group, single infection accounted for 65.1% (142/218) (Figure 3IIB). For single infection, the most common pathogens were bacteria (45.4%, 99/218), followed by fungi (9.2%, 20/218), and viruses (6%, 13/218)

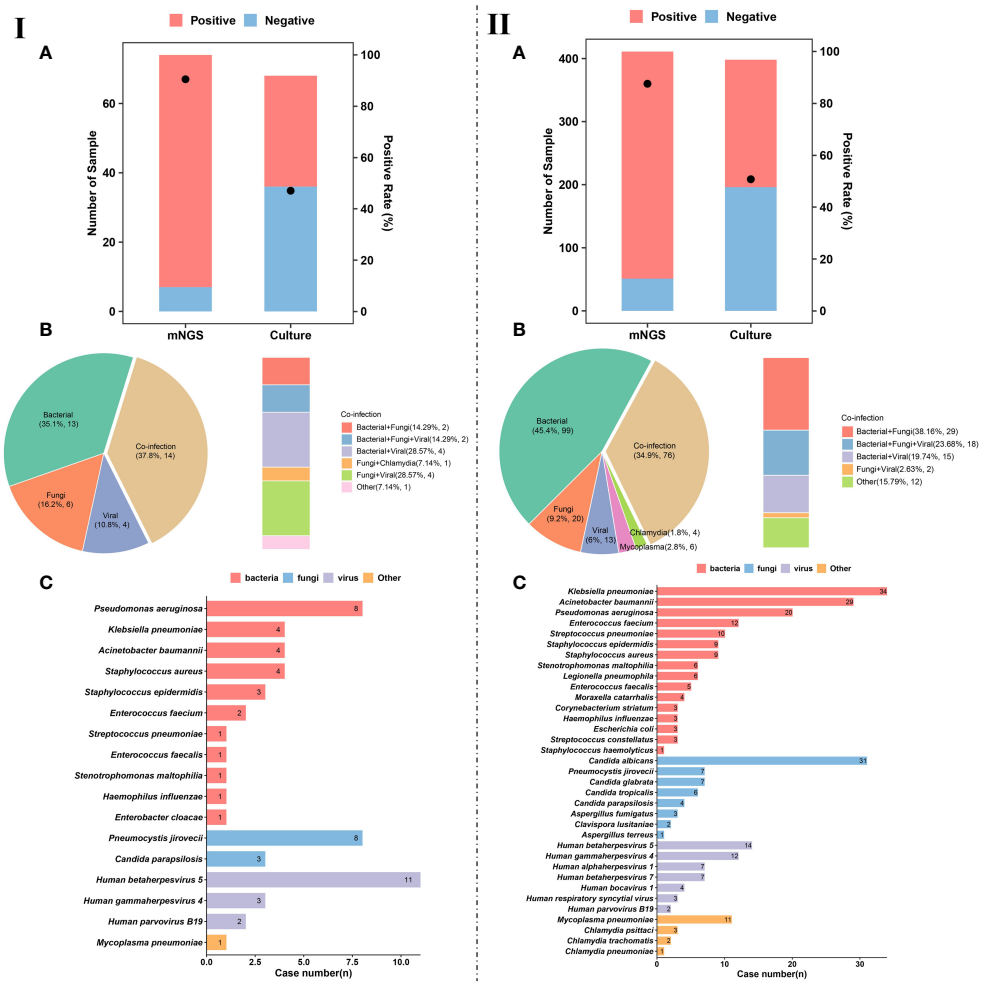



FIGURE 3
Pathogen detection and characteristics of pulmonary infection. **(IA)**, Number of patients and positive rate of mNGS and culture in immunocompromised patients. **(IIA)**, Number of patients and positive rate of mNGS and culture in immunocompetent patients. **(IB)**, Different kinds of infections ($n$ = 37) in immunocompromised patients. **(IIB)**, Different kinds of infections ($n$ = 218) in immunocompetent patients. **(IC)**, Pathogen profiles in immunocompromised patients. **(IIC)**, Pathogen profiles in immunocompetent patients.

(Figure 3IIB). In addition to single infection, our results show that mixed infection occurred in 34.9% (76/218) of immunocompetent patients, such as bacterial-fungal co-infection ($n$ = 29), bacterial-fungal-viral co-infection ($n$ = 18), bacterial-viral co-infection ($n$ = 15), and fungal-viral co-infection ($n$ = 12) (Figure 3IIB). It is worth noting that more mixed infections are found in immunocompromised patients than in immunocompetent patients.

The pathogen profiles of were also summarized (Figure 3IIC). We found that the most common bacterial pathogens were *K. pneumoniae* ($n$ = 34), *A. baumannii* ($n$ = 29), *P. aeruginosa* ($n$ = 20), and *Enterococcus faecium* ($n$ = 12). Fungal infection in immunocompetent patients with pulmonary infection was mainly caused by *Candida albicans* ($n$ = 31), *P. jirovecii* ($n$ = 7), and *Candida glabrata* ($n$ = 7). The most common viral pathogens were CMV ($n$ = 14), EBV ($n$ = 12), Human alphaherpesvirus 1 ($n$ = 7). At the same time, we also detected *Mycoplasma pneumoniae*, *Chlamydia psittaci*, *Chlamydia trachomatis*, and Chlamydia pneumoniae from 11, 3, 2, and 1 patients, respectively.

## Adjustment of antibiotics and prognosis

mNGS plays an important role in providing reference for clinical therapy. The effects of mNGS on adjustment of antibiotics in both immunocompromised and immunocompetent patients were summarized (Figure 4). mNGS exhibited positive impact on most of patients, including 41.1% of immunocompetent patients ($n$ = 167) and 31.1% of immunocompromised patients ($n$ = 23) receiving de-escalation treatment and 31.8% of immunocompetent patients ($n$ = 129) and 41.9% of immunocompromised patients ($n$ = 31) receiving escalation treatment (Figures 4IA, IIA). In addition, the empirical treatment on the 27.1% of immunocompetent patients ($n$ = 110) and 27% of immunocompromised patients ($n$ = 20) was not changed.

Although prognosis is influenced by a variety of complex factors, we have summarized the outcomes among patients receiving escalation, de-escalation, and maintain treatments (Figure 4). Escalation treatment did not result in improved prognosis in all of patients. Only 68.2% of immunocompetent patients and 45.2% of immunocompromised patients receiving escalation treatment had improved outcome, while deaths accounted for 13.2% and 22.6% of cases, respectively (Figures 4IB, IIB). In addition, de-escalation treatment had no impact on most of patients, including 57.5% of immunocompetent patients and 73.9% of immunocompromised patients, and no patients died (Figures 4IC, IIC). The highest proportion of patients showing improved prognosis was observed in immunocompetent individuals receiving maintain treatment, while 10.0% of immunocompromised patients died (Figures 4ID, IID).

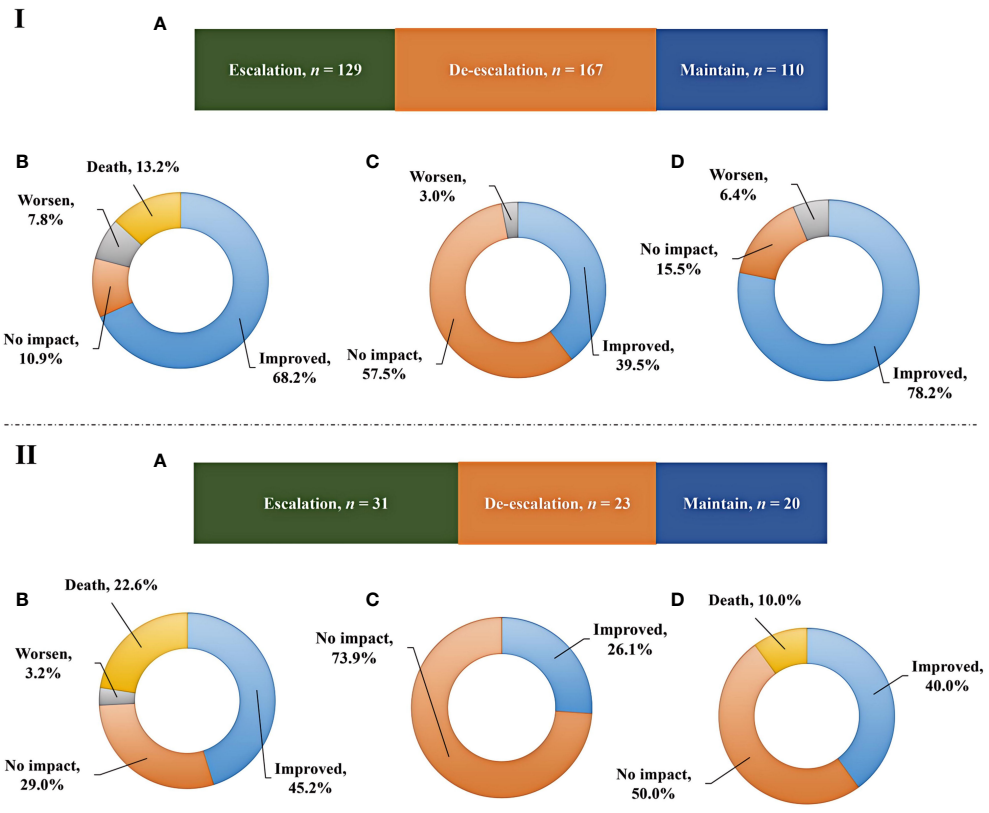



FIGURE 4
Adjustment of antibiotics and prognosis. **(IA, IIA)**, Impact of mNGS on adjustments of antibiotics in immunocompetent and immunocompromised patients, respectively. **(IB, IC, ID)**, Outcomes among immunocompetent patients receiving escalation, de-escalation, and maintain treatments, respectively. **(IIB, IIC, IID)**, Outcomes among immunocompromised patients receiving escalation, de-escalation, and maintain treatments, respectively.

## Application of commercial tNGS in clinics

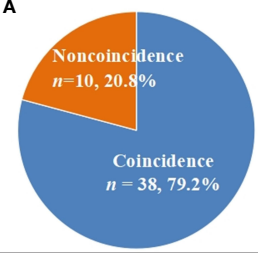
Given high cost of mNGS and its positive impact on patients, we further explored the role of commercial tNGS in diagnosing pulmonary infections in clinics. It was found that the coincidence rates of tNGS against final comprehensive clinical diagnoses in immunocompetent and immunocompromised patients can reach up to 79.2% (38/48) (Figure 5A) and 90.9% (10/11) (Figure 5B), respectively, which were much higher than that of BALF mNGS (Figure 2IIB). Further analysis of noncoincidence between tNGS and final comprehensive clinical diagnoses (Figure 5C) showed that some causative pathogens were partially detected or not detected by tNGS. Although commercial tNGS can provided much more accurate clinical reference for clinics, further improvement, such as optimizing the primer pool or designing a regional-specific tNGS, can be further implemented to enhance the suitability of tNGS for regional pulmonary infections, based on our findings of this study.

# Discussion

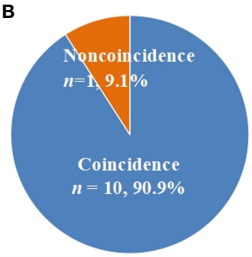
This is a retrospective study on evaluating the diagnostic performance of mNGS and tNGS in detecting pulmonary infections in immunocompetent and immunocompromised patients. The results showed that the overall positive rate of mNGS in detecting pathogens of pulmonary infections was significantly higher than that of culture ($P<0.05$) in both immunocompetent and immunocompromised patients. This result is consistent with previous study that the proportion of clinically relevant pathogens found by mNGS was significantly higher than that found by conventional methods (Joensen et al., 2017). Compared with the current conventional microbial methods, mNGS has more advantages in microbial detection (Joensen et al., 2017). The above findings, together with our results, confirm that mNGS can improve the detection rate of pathogens, which is critical to the accurate and rapid diagnosis of pulmonary infection and to guiding the treatment and prevention of patients.

In addition, this study compared the positive detection rates of mNGS and culture not only between BALF and sputum samples but also between immunocompetent and immunocompromised patients. We found that mNGS detection was superior to conventional culture in detecting bacteria, fungi, mycoplasma, and viruses, improving the clinical guidance on the use of antibiotics. This is consistent with the research of Langelier et al. that the detection rate of pathogens in HSCT patients with acute respiratory diseases detected by BALF mNGS was higher than that

A




B




C

| Patients | Group | Causative pathogens according to final clinical diagnosis | tNGS result | Noncoincidence reason for tNGS results |
|---|---|---|---|---|
| Case 1 | Immunocompetent | *K. pneumoniae* | *C. albicans* | Not detected |
| Case 2 | Immunocompetent | SARS-CoV-2 | *P. aeruginosa*<br>Human bocavirus | Partially detected |
| Case 3 | Immunocompetent | *P. aeruginosa* | *Haemophilus influenzae*<br>*Tropheryma whipplei*<br>*P. jirovecii* | Not detected |
| Case 4 | Immunocompetent | *C. albicans* | *C. tropicalis*<br>*Acinetobacter calcoaceticus* | Not detected |
| Case 5 | Immunocompetent | *Stenotrophomonas maltophilia*<br>*Serratia marcescens* | *Elizabethkingia meningoseptica* | Not detected |
| Case 6 | Immunocompetent | *P. jirovecii*<br>*P. aeruginosa* | *P. jirovecii* | Partially detected |
| Case 7 | Immunocompromised | *A. baumannii* | *Clavispora lusitaniae* | Not detected |
| Case 8 | Immunocompetent | *Stenotrophomonas maltophilia* | *Acinetobacter pittii*<br>*C. albicans*<br>SARS-CoV-2 | Not detected |
| Case 9 | Immunocompetent | *C. albicans*<br>*K. pneumoniae* | *C. albicans*<br>Human coronavirus OC43 | Partially detected |
| Case 10 | Immunocompetent | *Clostridioides difficile* | None | Not detected |
| Case 11 | Immunocompetent | CMV | *K. pneumoniae*<br>*Rhinovirus* C<br>Human coronavirus 229E<br>*Mycoplasma hominis* | Not detected |

FIGURE 5
Performance of tNGS. **(A)**, Coincidence rate of tNGS against final comprehensive clinical diagnoses in immunocompetent patients. **(B)**, Coincidence rate of tNGS against final comprehensive clinical diagnoses in immunocompromised patients. **(C)**, Noncoincidence reason for tNGS results.

detected by conventional methods (Langelier et al., 2018). In addition, BALF and sputum mNGS had higher sensitivity (53.7% vs. 18.5% and 45.7% vs. 29.9%, $p<0.001$), but lower specificity (21.2% vs. 59.5% and 14.4% vs. 52.0%, $p<0.001$) than culture, which was consistent with the result of Wang et al. that compared with conventional tests, mNGS has a higher sensitivity in detecting pathogens of lung infection (Wang et al., 2019). Huang et al. also proposed that conventional pathogen detection methods have limitation in accurately and comprehensively detecting microorganisms (Huang et al., 2020).

Increasing evidences suggest that mNGS is an important means for clinical diagnosis of infectious diseases (He et al., 2022; Yu et al., 2022; Qian et al., 2023; Zhang et al., 2023). The comparison between the immunocompetent and immunocompromised groups showed that although the positive rates of mNGS detection were higher than those of the conventional detection, there was no significant difference in the positive rates between the two groups. Besides, compared with patients with normal immune function, more mixed infections were found in patients with immunodeficiency. In conclusion, mNGS is superior to conventional pathogen detection in etiological diagnosis of pulmonary infection, especially for the immunocompromised patients.

As far as we know, mNGS is often used in clinical practice for etiological diagnosis of pulmonary infection and reasonable use of antibiotics. Li et al. showed that mNGS has the advantage in diagnosing mixed pulmonary infection of immunocompromised patients, which is crucial for the accurate diagnosis (Li et al., 2020). In our study, mNGS exhibited positive impact on most of patients, including 73.0% of immunocompromised patients and 72.9% of immunocompetent patients. The abuse and unreasonable use of antibiotics will make more opportunistic lung infections and become a thorny clinical problem for immunocompromised patients (Bouglé et al., 2022; Griffith and Daley, 2022; Smith et al., 2022). Overuse of antibiotics in immunocompromised patients will lead to drug resistance, while reasonable use of antibiotics can reduce the waste of medical resources (Ramirez et al., 2020; Martin-Loeches et al., 2022). mNGS can help clinicians evaluate empiric antibacterial treatment more comprehensively and effectively adjust therapeutic regimens of the immunocompromised patients with pulmonary infections.

Theoretically speaking, tNGS integrating the high sensitivity of PCR and the high throughput of mNGS can successfully enrich pathogens with predesigned primers in the panel (Li et al., 2021; Huang et al., 2023). However, some most common pathogens were detected in some cases but not detected or partially detected in other cases using the same commercial tNGS pipeline in our study. Actually, multiplex PCR scaling to large panels for broad range of most common pathogens causes the nonlinear increase of primer dimer species that reduces tNGS mapping rates (Xie et al., 2022) and subsequently decreases its performance. Additionally, due to the epidemiology of pathogens characterized by geographical specificity (Ramirez et al., 2020), rarity (Zhan et al., 2021), and novelty (El Zein et al., 2020; Fishman, 2023), commercial tNGS cannot fully meet the requirements of regional pathogen detection. Accordingly, based on our pathogen profiles summarized on a large scale, designing and developing of regional tNGS can be conducted according to the published studies (Xie et al., 2022; Xia et al., 2023).

## Limitation

At the same time, it is worth mentioning that this study also has some limitations. First of all, only sputum and BALF samples were used to evaluate the performance of mNGS or tNGS, and future study should include other samples, such as blood sample. Comparison of mNGS using various body fluid samples to analyze the lung infection may provide more valuable information. Secondly, since there were few negative cases with non-infectious diseases, there might be slight bias in calculating specificity. Thirdly, mNGS of both DNA and RNA should be simultaneously performed to detect comprehensive causative pathogens.

# Conclusion

Our research unravels that mNGS has the advantages of being faster and more sensitive than conventional pathogen culture method in patients with pulmonary infection. In addition, mNGS can identify more mixed infections in both immunocompetent and immunocompromised patients. Additionally, BALF tNGS exhibited better performance than BALF mNGS. Both mNGS and tNGS can identify the pathogen of pulmonary infection as early as possible, help clinicians adjust the treatment of antibiotics in time, and greatly improve the diagnosis of suspected pulmonary infection. Given high cost of mNGS and high sensitivity of tNGS, future development of pathogen molecular detection may focus on designing and developing of regional tNGS.

# Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://ngdc.cncb.ac.cn/search/all?&q=CRA016476, https://ngdc.cncb.ac.cn/search/all?q=CRA017510, CRA016476 and CRA017510.

# Ethics statement

The studies involving humans were approved by Ethical Review Committee of the First hospital of Jilin University. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from A total of 546 patients admitted to The First Hospital of Jilin University from January 2021 to December 2023 diagnosed as suspected pulmonary infection were retrospectively enrolled, and we only summarized the detection results of those patients. All of the

detection using human samples had been finished and we did not do any experiments using the human samples in this study. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

## Author contributions

YL: Conceptualization, Formal Analysis, Writing – original draft, Writing – review & editing. WW: Formal Analysis, Visualization, Writing – original draft. YX: Methodology, Writing – original draft. HZ: Methodology, Writing – original draft. SH: Data curation, Methodology, Writing – original draft. YJ: Conceptualization, Formal Analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China [grant number 22174137], Jilin Province Science and Technology Development [grant number 2023C013], Jilin Province Science and Technology Agency [grant numbers JJKH20211210KJ, JJKH20211164KJ, YDZJ202101ZYTS038, JLSWSRCZX2020–009, 20200901025SF, and 20200403084SF], and Beijing Medical Award Foundation [grant number YXJL-2021–1097-0645].

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or

financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1439472/full#supplementary-material

## References

Azoulay, E., Mokart, D., Kouatchet, A., Demoule, A., and Lemiale, V. (2019). Acute respiratory failure in immunocompromised adults. *Lancet Respir. Med.* 7, 173–186. doi: 10.1016/S2213-2600(18)30345-X

Bouglé, A., Tuffet, S., Federici, L., Leone, M., Monsel, A., Dessalle, T., et al. (2022). Comparison of 8 versus 15 days of antibiotic therapy for Pseudomonas aeruginosa ventilator-associated pneumonia in adults: a randomized, controlled, open-label trial. *Intensive Care Med.* 48, 841–849. doi: 10.1007/s00134-022-06690-5

D'Anna, S. E., Maniscalco, M., Cappello, F., Carone, M., Motta, A., Balbi, B., et al. (2021). Bacterial and viral infections and related inflammatory responses in chronic obstructive pulmonary disease. *Ann. Med.* 53, 135–150. doi: 10.1080/07853890.2020.1831050

El Zein, S., Hindy, J.-R., and Kanj, S. S. (2020). Invasive saprochaete infections: an emerging threat to immunocompromised patients. *Pathogens* 9, 922. doi: 10.3390/pathogens9110922

Fishman, J. A. (2023). Next-generation sequencing for identifying unknown pathogens in sentinel immunocompromised hosts. *Emerg. Infect. Dis.* 29, 431–432. doi: 10.3201/eid2902.221829

Gökdemir, F., İşeri, ÖD., Sharma, A., Achar, P. N., and Eyidoğan, F. (2022). Metagenomics next generation sequencing (mNGS): An exciting tool for early and accurate diagnostic of fungal pathogens in plants. *J. fungi (Basel Switzerland)* 8, 1195. doi: 10.3390/jof8111195

Greninger, A. L., and Naccache, S. N. (2019). Metagenomics to assist in the diagnosis of bloodstream infection. *J. Appl. Lab. Med.* 3, 643–653. doi: 10.1373/jalm.2018.026120

Griffith, D. E., and Daley, C. L. (2022). Treatment of mycobacterium abscessus pulmonary disease. *Chest* 161, 64–75. doi: 10.1016/j.chest.2021.07.035

Han, D., Li, Z., Li, R., Tan, P., Zhang, R., and Li, J. (2019). mNGS in clinical microbiology laboratories: on the road to maturity. *Crit. Rev. Microbiol.* 45, 668–685. doi: 10.1080/1040841X.2019.1681933

Handel, A. S., Muller, W. J., and Planet, P. J. (2021). Metagenomic next-generation sequencing (mNGS): SARS-CoV-2 as an example of the technology's potential pediatric infectious disease applications. *J. Pediatr. Infect. Dis. Soc.* 10, S69–s70. doi: 10.1093/jpids/piab108

He, P., Wang, J., Ke, R., Zhang, W., Ning, P., Zhang, D., et al. (2022). Comparison of metagenomic next-generation sequencing using cell-free DNA and whole-cell DNA for the diagnoses of pulmonary infections. *Front. Cell Infect. Microbiol.* 12, 1042945. doi: 10.3389/fcimb.2022.1042945

Huang, C., Huang, Y., Wang, Z., Lin, Y., Li, Y., Chen, Y., et al. (2023). Multiplex PCR-based next generation sequencing as a novel, targeted and accurate molecular approach for periprosthetic joint infection diagnosis. *Front. Microbiol.* 14, 1181348. doi: 10.3389/fmicb.2023.1181348

Huang, J., Jiang, E., Yang, D., Wei, J., Zhao, M., Feng, J., et al. (2020). Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions. *Infection Drug resistance* 13, 567–576. doi: 10.2147/IDR

Indelli, P. F., Ghirardelli, S., Violante, B., and Amanatullah, D. F. (2021). Next generation sequencing for pathogen detection in periprosthetic joint infections. *EFORT Open Rev.* 6, 236–244. doi: 10.1302/2058-5241.6.200099

Ji, X.-C., Zhou, L.-F., Li, C.-Y., Shi, Y.-J., Wu, M.-L., Zhang, Y., et al. (2020). Reduction of human DNA contamination in clinical cerebrospinal fluid specimens improves the sensitivity of metagenomic next-generation sequencing. *J. Mol. Neurosci.* 70, 659–666. doi: 10.1007/s12031-019-01472-z

Joensen, K. G., Engsbro, A. L. Ø., Lukjancenko, O., Kaas, R. S., Lund, O., Westh, H., et al. (2017). Evaluating next-generation sequencing for direct clinical diagnostics in diarrhoeal disease. *Eur. J. Clin. Microbiol. Infect. Dis.* 36, 1325–1338. doi: 10.1007/s10096-017-2947-2

Langelier, C., Zinter, M. S., Kalantar, K., Yanik, G. A., Christenson, S., O'Donovan, B., et al. (2018). Metagenomic sequencing detects respiratory pathogens in hematopoietic cellular transplant patients. *Am. J. Respir. Crit. Care Med.* 197, 524–528. doi: 10.1164/rccm.201706-1097LE

Li, B., Xu, L., Guo, Q., Chen, J., Zhang, Y., Huang, W., et al. (2021). GenSeizer: a multiplex PCR-based targeted gene sequencing platform for rapid and accurate identification of major mycobacterium species. *J. Clin. Microbiol.* 59, e00584–20. doi: 10.1128/JCM.00584-20

Li, D., Gai, W., Zhang, J., Cheng, W., Cui, N., and Wang, H. (2022). Metagenomic next-generation sequencing for the microbiological diagnosis of abdominal sepsis patients. *Front. Microbiol.* 13, 816631. doi: 10.3389/fmicb.2022.816631

Li, G., Huang, J., Li, Y., and Feng, J. (2020). The value of combined radial endobronchial ultrasound-guided transbronchial lung biopsy and metagenomic next-generation sequencing for peripheral pulmonary infectious lesions. *Can. Respir. J.* 2020, 2367505. doi: 10.1155/2020/2367505

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Luyt, C. E., Bouadma, L., Morris, A. C., Dhanani, J. A., Kollef, M., Lipman, J., et al. (2020). Pulmonary infections complicating ARDS. *Intensive Care Med.* 46, 2168–2183. doi: 10.1007/s00134-020-06292-z

Martin-Loeches, I., Garduno, A., Povoa, P., and Nseir, S. (2022). Choosing antibiotic therapy for severe community-acquired pneumonia. *Curr. Opin. Infect. Dis.* 35, 133–139. doi: 10.1097/QCO.0000000000000819

Qian, Z., Xia, H., Zhou, J., Wang, R., Zhu, D., Chen, L., et al. (2023). Performance of metagenomic next-generation sequencing of cell-free DNA from vitreous and aqueous humor for diagnoses of intraocular infections. *J. Infect. Dis.* 229, 252–261. doi: 10.1093/infdis/jiad363

Qu, J., Zhang, J., Chen, Y., Huang, Y., Xie, Y., Zhou, M., et al. (2022). Aetiology of severe community acquired pneumonia in adults identified by combined detection methods: a multi-centre prospective study in China. *Emerging Microbes infections* 11, 556–566. doi: 10.1080/22221751.2022.2035194

Ramachandran, P. S., and Wilson, M. R. (2020). Metagenomics for neurological infections - expanding our imagination. *Nat. Rev. Neurol.* 16, 547–556. doi: 10.1038/s41582-020-0374-y

Ramirez, J. A., Musher, D. M., Evans, S. E., Dela Cruz, C., Crothers, K. A., Hage, C. A., et al. (2020). Treatment of community-acquired pneumonia in immunocompromised adults: A consensus statement regarding initial strategies. *Chest* 158, 1896–1911. doi: 10.1016/j.chest.2020.05.598

Ravimohan, S., Kornfeld, H., Weissman, D., and Bisson, G. P. (2018). Tuberculosis and lung damage: from epidemiology to pathophysiology. *Eur. Respir.* 27, 170077. doi: 10.1183/16000617.0077-2017

Schneider, J. L., Rowe, J. H., Garcia-de-Alba, C., Kim, C. F., Sharpe, A. H., and Haigis, M. C. (2021). The aging lung: Physiology, disease, and immunity. *Cell* 184, 1990–2019. doi: 10.1016/j.cell.2021.03.005

Shi, Y., Wu, J., Liu, T., Yue, L., Liu, Y., Gu, Y., et al. (2022). Analysis of metagenomic next-generation sequencing results of 25 pus samples. *Infection Drug Resistance* 15, 6515–6524. doi: 10.2147/IDR.S385925

Simner, P. J., Miller, S., and Carroll, K. C. (2018). Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin. Infect. Dis.* 66, 778–788. doi: 10.1093/cid/cix881

Smith, S., Rowbotham, N. J., and Charbek, E. (2022). Inhaled antibiotics for pulmonary exacerbations in cystic fibrosis. *Cochrane Database systematic Rev.* 8, Cd008319. doi: 10.1002/14651858.CD008319.pub3

Thakur, K. T. (2020). Application of pathogen discovery/metagenomic sequencing in CNS HIV. *Curr. HIV/AIDS Rep.* 17, 507–513. doi: 10.1007/s11904-020-00514-1

Turcios, N. L. (2020). Cystic fibrosis lung disease: An overview. *Respir. Care* 65, 233–251. doi: 10.4187/respcare.06697

Wang, J., Han, Y., and Feng, J. (2019). Metagenomic next-generation sequencing for mixed pulmonary infection diagnosis. *BMC Pulm Med.* 19, 252. doi: 10.1186/s12890-019-1022-4

Xia, H., Zhang, Z., Luo, C., Wei, K., Li, X., Mu, X., et al. (2023). MultiPrime: A reliable and efficient tool for targeted next-generation sequencing. *iMeta* 2, e143. doi: 10.1002/imt2.143

Xie, N. G., Wang, M. X., Song, P., Mao, S., Wang, Y., Yang, Y., et al. (2022). Designing highly multiplex PCR primer sets with Simulated Annealing Design using Dimer Likelihood Estimation (SADDLE). *Nat. Commun.* 13, 1881. doi: 10.1038/s41467-022-29500-4

Yu, L., Zhang, Y., Zhou, J., Zhang, Y., Qi, X., Bai, K., et al. (2022). Metagenomic next-generation sequencing of cell-free and whole-cell DNA in diagnosing central nervous system infections. *Front. Cell Infect. Microbiol.* 12, 951703. doi: 10.3389/fcimb.2022.951703

Yue, R., Wu, X., Li, T., Chang, L., Huang, X., and Pan, L. (2021). Early Detection of Legionella pneumophila and Aspergillus by mNGS in a Critically Ill Patient With Legionella Pneumonia After Extracorporeal Membrane Oxygenation Treatment: Case Report and Literature Review. *Front. Med.* 8, 686512. doi: 10.3389/fmed.2021.686512

Zhan, Y., Xu, T., He, F., Guan, W.-J., Li, Z., Li, S., et al. (2021). Clinical evaluation of a metagenomics-based assay for pneumonia management. *Front. Microbiol.* 12, 751073. doi: 10.3389/fmicb.2021.751073

Zhang, C., Li, Z., Wang, M., Zhou, J., Yu, W., Liu, H., et al. (2023). High specificity of metagenomic next-generation sequencing using protected bronchial brushing sample in diagnosing pneumonia in children. *Front. Cell Infect. Microbiol.* 13, 1165432. doi: 10.3389/fcimb.2023.1165432

Zhou, X., Peng, S., Song, T., Tie, D., Tao, X., Jiang, L., et al. (2022). Neurosyphilis with ocular involvement and normal magnetic resonance imaging results affirmed by metagenomic next-generation sequencing. *Front. Cell. Infect. Microbiol.* 12, 985373. doi: 10.3389/fcimb.2022.985373

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 19 August 2024
DOI 10.3389/fcimb.2024.1442062



OPEN ACCESS

EDITED BY
Qing Wei,
Genskey Co. Ltd, China

REVIEWED BY
Xinyu Gu,
Henan University of Science and Technology, China
Yanan Zhao,
China Medical University, China

*CORRESPONDENCE
Yigang Tong
tongyigang@gmail.com
Lixin Xie
xielx301@126.com

†These authors have contributed equally to this work

RECEIVED 01 June 2024
ACCEPTED 01 August 2024
PUBLISHED 19 August 2024

CITATION
Shi Y, Zhang W, Li L, Wu W, Li M, Xiao K, Wang K, Sheng Z, Xie F, Wang X, Shi X, Tong Y and Xie L (2024) Evaluation of phage-based decontamination in respiratory intensive care unit environments using ddPCR and 16S rRNA targeted sequencing techniques.
*Front. Cell. Infect. Microbiol.* 14:1442062.
doi: 10.3389/fcimb.2024.1442062



# Evaluation of phage-based decontamination in respiratory intensive care unit environments using ddPCR and 16S rRNA targeted sequencing techniques

Yinghan Shi[1†], Weihua Zhang[1†], Lina Li[2†], Wencai Wu[3], Mengzhe Li[4], Kun Xiao[1], Kaifei Wang[1], Zhaojun Sheng[1], Fei Xie[1], Xiuli Wang[1], Xin Shi[1], Yigang Tong[4*] and Lixin Xie[1*]

[1]College of Pulmonary & Critical Care Medicine, Chinese PLA General Hospital, Beijing, China, [2]The First Medical Centre, Chinese PLA General Hospital, Beijing, China, [3]Key Laboratory of Biomedical Information Engineering of Ministry of Education, Biomedical Informatics & Genomics Center, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China, [4]College of Life Science and Technology, Beijing University of Chemical Technology, Beijing, China

**Background:** *Klebsiella pneumoniae* is a major cause of hospital-acquired infections (HAIs), primarily spread through environmental contamination in hospitals. The effectiveness of current chemical disinfectants is waning due to emerging resistance, which poses environmental hazards and fosters new resistance in pathogens. Developing environmentally friendly and effective disinfectants against multidrug-resistant organisms is increasingly important.

**Methods:** This study developed a bacteriophage cocktail targeting two common carbapenem-resistant *Klebsiella pneumoniae* (CRKP) strains, ST11 KL47 and ST11 KL64. The cocktail was used as an adjunctive disinfectant in a hospital's respiratory intensive care unit (RICU) via ultrasonic nebulization. Digital PCR was used to quantify CRKP levels post-intervention. The microbial community composition was analyzed via 16S rRNA sequencing to assess the intervention's impact on overall diversity.

**Results:** The phage cocktail significantly reduced CRKP levels within the first 24 hours post-treatment. While a slight increase in pathogen levels was observed after 24 hours, they remained significantly lower than those treated with conventional disinfectants. 16S rRNA sequencing showed a decrease in the target pathogens' relative abundance, while overall species diversity remained stable, confirming that phages selectively target CRKP without disrupting ecological balance.

**Discussion:** The findings highlight the efficacy and safety of phage-based biocleaners as a sustainable alternative to conventional disinfectants. Phages selectively reduce multidrug-resistant pathogens while preserving microbial diversity, making them a promising tool for infection control.

KEYWORDS

**bacteriophages, hospital-acquired infections, drug-resistant bacterial, targeted metagenomics in pathogen, 16S rRNA**

# Introduction

*K. pneumoniae* is a significant cause of Hospital-acquired infections (HAIs), which are predominantly transmitted through contamination of the hospital environment or medical equipment (Odoyo et al., 2023). The risk factors for these infections include environmental contamination, reuse or improper disinfection of medical devices, inadequate hand hygiene practices by healthcare workers, and poor ventilation systems (Peters et al., 2022). Research indicates that while current disinfection measures play a crucial role in controlling the spread of these pathogens, their effectiveness is diminishing due to some pathogens developing resistance to existing disinfectants. Furthermore, frequent use of chemical disinfectants can cause equipment corrosion, harm the environment and health of personnel, and may encourage the development of new resistances in pathogens (Boyce, 2016; Kampf, 2018). Given these challenges, the development of new disinfectants, especially those that are both environmentally friendly and highly effective against multidrug-resistant organisms, has become increasingly urgent.

Bacteriophages are viruses that specifically target bacteria, acting as natural bacterial predators, and offer a unique mechanism of action. Bacteriophage therapy is experiencing a renaissance in response to the escalating global antimicrobial resistance crisis (Salmond and Fineran, 2015; Gordillo Altamirano and Barr, 2019). The historical significance of bacteriophages in combating bacterial infections, dating back almost a century, has resurfaced as a promising solution to address antibiotic-resistant pathogens. With applications ranging from combating antibiotic resistance to food safety and environmental health (Moye et al., 2018; Tie et al., 2018; Nick et al., 2022).

Studies have shown the effectiveness of using atomized bacteriophages to reduce bacterial loads on hard surfaces. Studies by Hussain et al. have demonstrated the efficacy of atomized bacteriophages in reducing *Acinetobacter baumannii* levels on hard surfaces (Ho et al., 2016). Furthermore, the work by Maria D'Accolti et al. explored the combined use of phages and a probiotic-based sanitation system to efficiently remove hospital-acquired infection pathogens from various hard surfaces, showcasing a rapid reduction of argeted pathogens. These studies collectively support the use of atomized bacteriophages as a promising strategy for decontaminating hard surfaces and reducing bacterial levels effectively (D'Accolti et al., 2018).

While the efficacy of bacteriophages as biological disinfectants has been established, their impact on CRKP in clinical settings has yet to be thoroughly investigated. Additionally, further research is needed to explore the impact of bacteriophage intervention on environmental microbial communities to ensure the safety and efficacy of this approach in real-world applications. In this study, we developed a phage cocktail targeting two prevalent *K. pneumoniae* strains, ST11 KL47 and ST11 KL64, found in clinical settings. We utilized these bacteriophages as an adjunctive disinfectant to augment environmental cleaning practices. The efficacy of the phage cocktail aerosol in clearing CRKP from the environment was assessed. The primary objective was to evaluate the impact of the bacteriophage treatment on the clearance of environmentally persistent drug-resistant bacteria and to analyze its effects on the overall microbial community structure. This approach aimed to establish a scientifically sound basis for the broader application of phage therapy in infection control practices, particularly against antibiotic-resistant pathogens.

# Materials and methods

## Setting and environmental sampling

The study was conducted in the respiratory intensive care unit of a hospital in Beijing, China. In addition, this study was approved by the institutional review board of the PLA Hospital (309202305011312). To assess the baseline microbial community composition and the molecular characteristics of *K.pneumoniae* on environmental surfaces in the Respiratory Intensive Care Unit (RICU) prior to any phage or chemical disinfectant intervention, bedside environmental samples were collected from patients diagnosed with CRKP infection(n=9). For *K.pneumoniae* infection diagnosis, Metagenomic next-generation sequencing (mNGS) pipeline was performed on six out of these patients, in brief, mNGS and hybridization capture-based targeted mNGS were used to detectpathogens. Approximately 1 mL of sample was centrifuged at 12,000 g for 5 minutes to separate pathogens and human cells. The precipitate was treated with Benzonase and Tween 20, followed by DNA extraction using the QIAamp UCP Pathogen Mini Kit. DNA concentrations were measured with a Qubit 4.0. Libraries prepared from the DNA were sequenced on the NextSeq 550 platform (Illumina). For targeted NGS, libraries were enriched with microbial probes and sequenced data filtered to remove unsuitable reads. Species-level microbial identification was performed using an NCBI database. While the remaining three were identified via microbiological culture of respiratory samples.

This initial sampling was conducted before the application of chlorinated disinfectants or bacteriophage-based sanitation methods. Surface swabs were taken from areas frequently contacted by staff and patients, including floors, bed linens, bed frames and bedside tables, with a total of 36 environmental hard surface samples, each covering an area of 100 square centimeters. These swabs were preserved in phosphate-buffered saline (PBS) and transported to the laboratory in a chilled condition and processed within 2 hours of collection for subsequent DNA extraction and microbial culture.

## CRKP isolates and identification

Each swab will be streaked onto MacConkey agar plates. *K. pneumoniae* was identified by standard methods and confirmed using the "BD Phoenix Automated Microbiology System". Antimicrobial drug susceptibility testing was performed using the BD system and the minimal inhibitory concentration (MIC) of imipenem and meropenemresults were interpreted using the Clinical and Laboratory Standards Institute 2016. The resistant breakpoint of CRKP to imipenem and meropenem was defined as a MIC > 4 μg/mL.

## Source and selection of active bacteriophage

In this study, bacteriophages were isolated from sewage using the double-layer agar plate method. Briefly, untreated wastewater samples were centrifuged and then filtered through a 0.22μm filter (Millipore, USA) to remove bacteria and other particles. A volume of 500 μL of the filtrate was co-cultured with 500 μL of *K. pneumoniae* (OD600 = 0.8) in 5 mL of LB medium, and incubated overnight at 37°C with shaking. The mixed culture was then centrifuged at 12,000 x g for 2 minutes and the supernatant containing the phages was collected by filtration through a 0.22μm filter. Then, CRKP isolated from the ward environment samples was added to soft LB agar (0.5%, haibo, qingdao) and then poured onto a regular LB agarplate. Subsequently, 2 μl aliquots of different phage solutions (approximately $10^8$ plaques forming units (PFUs)) were spotted on the bacterialplate. Finally, the plates were incubated at 37 °C for 24 h. Generation of a clearance zone surrounding the phage inoculation spots indicates that the CRKP host was susceptible to the inoculated phage.

## Genomic identification of *K. pneumoniae*

The *K. pneumoniae* strains isolated from environmental samples were cultured in LB liquid medium to the logarithmic growth stage. The genomic DNA of *K. pneumoniae* isolates was extracted by Bacterial Genomic DNA Extraction kit (Beijing Solarbio Science & Technology Co., Ltd.) and whole genome sequencing of 16 K. *pneumoniae* isolates obtained from environmental samples was performed using Illumina HiSeq-150. The raw sequencing reads of *K. pneumoniae* was assembled using SPAdes v3.13.0. The MLST, resistance, and virulence of *K. pneumoniae* was predicted using the integrated tool Pathogenwatch(https://pathogen.watch/).

## Environmental decontamination by phage-containing aerosols

In this study, we evaluated the efficacy of bacteriophage disinfectants compared to chemical disinfectants for decontaminating environmental surfaces contaminated with *K. pneumoniae*, in a hospital setting. Two wards, housing patients infected with *K. pneumoniae*, were elected for environmental cleaning.

Initial environmental sampling (T1) was performed prior to the application of any disinfectant to establish a baseline pathogen load. Then use a chemical disinfectant, and a second sampling (T2) was taken six hours later to evaluate the immediate efficacy of the chemical agents. Subsequently, apart from chemical disinfectants, a bacteriophage-based disinfectant was additionally used. The phages solution was aerosolized using an ultrasonic nebulizer, with the phage concentration in the ultrasonic humidifier set at $10^8$ PFU (plaque-forming units) per milliliter, and approximately 500 mL utilized per room. Further samples were collected at 24 (T3), 48 (T4), and 72 (T5) hours post-application of the bacteriophage disinfectant, enabling the assessment of its long-term disinfection performance.

The hospital daily decontamination schedule was not changed during the bacteriophage decontamination process. For environmental decontamination, the phage stock solution was diluted in normal saline, and the phage aerosols were then generated with an ultrasonic humidifier (Rimei Electronic Technology Co. LTD) for 20 min, to ensure homogenization distribution of the phage aerosols throughout the ward.Briefly, we attached a disposable sterile sampling bag to the outlet of the ultrasonic nebulizer to collect the aerosolized phage cocktail. After initiating the machine and allowing the nebulization process to complete, we used a pipette to aspirate the condensed phage cocktail droplets from the sterile sampling bag. The collected droplets were then mixed thoroughly, serially diluted, and the phage titer was determined using the double-layer agar method.

## DNA extraction, PCR amplification

The genomic DNA was extracted by SDS method, and the DNA concentration and purity were detected by agarose gel electrophoresis. According to the concentration results, the DNA was diluted to l ng/μL with sterile water. Specific primers were used for PCR amplification using diluted genomic DNA as a template. The amplification process consisted of predenaturation at 98°C for 1 min, 30 cycles (denaturation at 98°C for 10 s, annealing at 50°C for 30 s, extension at 72°C for 30 s), and finally extension at 72°C for 5 min. Then the mixed PCR products were purified. Samples were divided into room, position, time, and sample types.

## Droplet digital PCR analysis and amplicon sequencing

The ddPCR assays were conducted to quantify *K. pneumoniae* DNA in the samples, using specific primers and probes designed for *K. pneumoniae. The* ddPCR assay was performed in a reaction volume of 20μL using a commercial ddPCR kit (Xinyi, Beijing, China). Twenty microliters of PCR mixture,50μL Droplet generation oil and 5μL sealant were mixed, and droplet generation was performed using the test instrument MicroDrop -100A. The droplet emulsion was thermally cycled in the following conditions: predenaturing at 95°C for 10 min, 45 cycles of PCR at 95°C for 30 s, and at 60°C for 1 min, and the test instrument was cooled at 12°C for 5 min, and then, the reaction was finished. The ddPCR system partitioned the DNA samples into approximately 20,000 nanoliter-sized droplets, with PCR amplification occurring in each droplet independently. The PCR reaction plate was placed in the MicroDrop-100B biochip reader for detection. Analyze data results using QuantDrop data analysis software.

To assess the broader impact on the microbial community, 16S rRNA gene sequencing was performed on the same samples. The V3-V4 region of the 16S rRNA gene was amplified using universal primers. The TruSeq DNA PCR-Free Sample preparation Kit was used to construct the library from the purified amplified fragments. After qualified, the library was sequenced using Illumina NovaSeq sequencing platform.

## 16S rRNA reads sequencing data processing

The original data obtained by sequencing was spliced and filtered to remove interference data and obtain effective data that can be used for subsequent analysis. Then, the valid data were grouped into Operational Taxonomic Units (OTUs) with 97% consistency, and the OTUs sequence was compared with the SSUrRNA database for annotation at kingdom, phylum, class, order, family, and genus levels. The species composition of each sample is counted. QIIME version 2 was used to estimate alpha and beta diversity (Bolyen et al., 2019). OTU abundances were used to calculate the alpha diversity metrics, including OTU richness (unique OTUs), ChaoI richness estimation, and Shannon's diversity indices. For overall comparison of significant differences among bacterial communities (i.e., beta diversity), principal coordinates analysis (PCoA) was performed.

# Results

## CRKP isolates and their phage susceptibility

*K. pneumoniae* was cultured in 16 out of 36 environmental samples taken before the disinfection procedure (Table 1). The study findings indicated a universal presence of *K. pneumoniae* on the floors near all patient bed units. Pillowcases also showed a significant presence of the pathogen, with five out of nine (n=5) testing positive. In contrast, bed rails and bedside tables exhibited lower detection rates. These isolates were identified as CRKP using the BD Phoenix Automated Microbiology System. The whole genome sequencing results indicate that the *K. pneumoniae* isolates belong to the ST11 KL47 (n=6) and ST11 KL64 (n=10) sequence types (Figure 1).

TABLE 1 Microbial culture results for K. pneumoniae in samples collected from various sites adjacent to bedside units in hospital wards housing nine infected patients.

| | Unit1 | Unit2 | Unit3 | Unit4 | Unit5 | Unit6 | Unit7 | Unit8 | Unit9 |
|---|---|---|---|---|---|---|---|---|---|
| P1 | + | + | + | + | + | + | + | + | + |
| P2 | – | + | – | – | – | – | – | – | – |
| P3 | + | – | – | – | – | – | – | – | – |
| P4 | + | + | + | + | – | – | + | – | – |

The table is structured with the horizontal axis representing the nine bedside units, and the vertical axis indicating different locations within each room: P1 for the floor, P2 for the bedside table, P3 for the bed rail, and P4 for the bedsheet. '-' indicates no colonies were cultured, while '+' indicates presence of colonies.

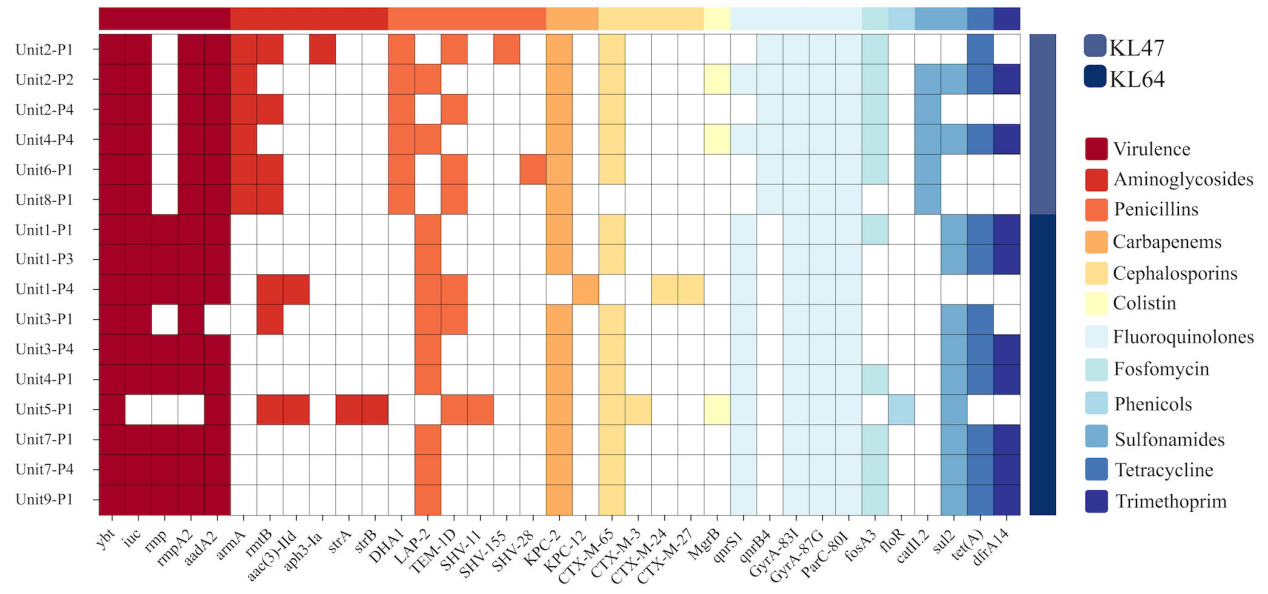



FIGURE 1
Genotypic characterisation and analysis of 16 K.pneumoniae isolates. This figure displays the distribution of prevalent sequence types ST11 KL64 and ST11 KL47, and account of virulence factors and antibiotic resistance genes. Virulence genes include ybt, iuc, rmpA and rmpA2, which are linked to the mucoid phenotype. Resistance genes span multiple antibiotic classes: Aminoglycosides (aadA2, armA), Penicillins(TEM-1D, SHV-11, SHV-15, SHV-28), Carbapenems(KPC-2, KPC-12), Cephalosporins (CTX-M-65, CTX-M-3, CTX-M-24, CTX-M-27), and Fluoroquinolone(GyrA and ParC).

10.3389/fcimb.2024.1442062

From each of these 16 culture dishes, one bacterial strain was selected for phage typing. The isolated strains were susceptible to *K. pneumoniae* type ST11 KL47 or ST11 KL64 as host infection (Table 2).The phages with the best lytic performance were composed into a phage cocktail, consisting of one phage against *K. pneumoniae* type KL47 and two phages against *K. pneumoniae* type KL64 (Supplementary Figure 1).

The phage stock solution with a titer of $10^8$PFU/mL was introduced into the ultrasonic atomizer. After undergoing ultrasonic atomization, the titer of the recovered phage dropped to $10^7$PFU/mL. This one-log decrease in titer suggests that the phage was adversely affected by the physical processes, such as heat generation or vibrations, associated with the ultrasonic atomization (Supplementary Figure 2).

## *K. pneumoniae* quantification

Although only 16 samples were culturably positive, we detected *K. pneumoniae* at different molecular levels in all 36 samples by ddpcr.The quantitative results of *K. pneumoniae* in each sample are shown in the figure (Figure 2).The findings revealed that bed linens exhibited the highest total copy numbers of the pathogen ($3.03\times10^4$ copies/uL), indicating significant reservoirs of the organism in these locations, the floors also demonstrated a considerable presence ($1.69\times10^4$ copies/uL), while lower copy numbers were observed on bed frames ($1.33\times10^4$ copies/uL) and bedside tables ($1.34\times10^4$ copies/uL) (Supplementary Table 1).These findings highlight that floors and bed linen are frequent areas of contact between health care workers and patients in the healthcare environment, and these areas can be an important route for pathogen transmission, especially bed linens, which are in direct contact with the face and respiratory tract of patients, and can easily become a source of cross-infection.

## 16S rRNA sequencing analysis: microbial community composition and diversity

Microbial community composition and diversity analysis in environmental samples from the RICU prior to disinfection were conducted using 16S rRNA sequencing. The α-diversity analysis across different bedside units shows similar microbial community compositions ($P > 0.05$), reflecting the stability of microbial communities within the ward environment (Figure 3A). In the hard surface environment samples from unsterilized bedsheets, we counted the top 15 species. The results show that, in addition to *Klebsiella*, common nosocomial pathogens such as *Acinetobacter* (17%) and *Elizabethkingia* (8%) were observed to hold dominant positions across all facilities surveyed. In contrast, *Staphylococcus* accounted for a comparatively smaller proportion of the microbial community (2.6%). Furthermore, the presence of *Akkermansia* (5.5%)—a genus associated with environmental symbiosis or opportunistic infections—as well as *Corynebacterium* (2.5%) and *Bacteroides* (2.9%), indicates that although *Klebsiella* is widespread, it does not dominate the microbial ecosystem (Figure 3B).

TABLE 2 Phage sensitivity matrix for K. pneumoniae colonies cultured from various locations within different rooms, assessed using the double-layer agar spot test.

| Bacteria Phage | Unit1-P1 | Unit1-P3 | Unit1-P4 | Unit2-P1 | Unit2-P2 | Unit2-P4 | Unit3-P1 | Unit3-P4 | Unit4-P1 | Unit4-P4 | Unit5-P1 | Unit6-P1 | Unit7-P1 | Unit7-P4 | Unit8-P1 | Unit9-P1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PCCM_KpP42 | – | – | – | + | + | + | – | – | – | + | – | + | – | – | + | – |
| PCCM_KpP216 | – | – | – | + | + | + | – | – | – | + | – | + | – | – | + | – |
| PCCM_KpP416 | – | – | – | + | + | + | – | – | – | + | – | + | – | – | + | – |
| PCCM_KpP24 | + | + | + | – | – | – | + | + | + | – | + | – | + | + | – | + |
| PCCM_KpP22 | + | + | + | – | – | – | + | + | + | – | + | – | + | + | – | + |
| PCCM_KpP112 | + | + | + | – | – | – | + | + | + | – | + | – | + | + | – | + |
| PCCM_KpP1119 | + | + | + | – | – | – | + | + | + | – | + | – | + | + | – | + |
| PCCM_KpP27X | + | + | + | – | – | – | + | + | + | – | + | – | + | + | – | + |
| PCCM_KpP1172 | + | + | + | – | – | – | + | + | + | – | + | – | + | + | – | + |

The table is structured with the horizontal axis listing the K. pneumoniae colonies identified by room and specific location, while the vertical axis denotes the phages used for typing, where '-' indicates phage insensitivity and '+' denotes phage sensitivity.

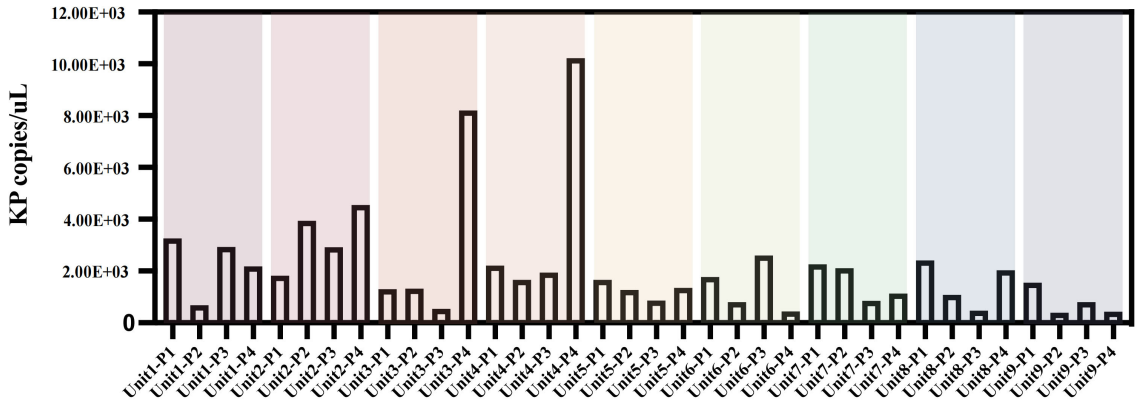



FIGURE 2
Copies of K.pneumoniae pathogens in bedside environmental samples from 9 patients without disinfection. Unit represents different bedside elements: P1 for the floor, P2 for the bedside table, P3 for the bed rail, and P4 for the bedsheet.

## Effect of bacteriophage aerosol on environmental microorganisms

In this investigation, ddPCR was employed to quantitatively analyse the genomic copies of *K. pneumoniae* at five distinct time intervals, to compare the immediate and sustained impacts of chemical and bacteriophage disinfectants (Figure 4). There were no significant differences in the absolute abundance of *K. pneumoniae* before and 6 h after chemical disinfection, which were respectively $8.97\times10^3$ and $9.13\times10^3$ copies/μL in Unit 1, and $1.31\times10^4$ and $1.48\times10^4$ μL in Unit 2.In contrast, a significant reduction in *K. pneumoniae* load was observed in samples treated with additional bacteriophage disinfectant for 24 hours compared to those treated with chemical disinfectants, and the absolute abundance of *K. pneumoniae* decreased to $3.49\times10^3$ and $4.02\times10^3$ copies/μL in Unit 1 and 2, respectively.Further assessments at 48 and 72 hours post-bacteriophage application showed that the majority of samples maintained low pathogen levels, with only slight increases.The absolute abundance of *K. pneumoniae* in unit 1 and Unit 2 was $6.07\times10^3$ and $1.12\times10^4$ copies/μL after 48 hours of application of phage cocktail, and after 72 hours in unit 1 and Unit 2 was $5.55\times10^3$ and $6.7\times10^3$ copies/μL. Notably, a sharp increase in pathogen copy numbers was observed 48 hours post-bacteriophage application on the floor near the second bed unit, potentially linked to high-risk procedures such as suctioning, which had been conducted in the vicinity.

To assess the impact of phage treatment on the broader microbial community, 16S rRNA sequencing was performed on the same environmental sample. Our analysis of α and β diversity indices pre- and post-bacteriophage intervention on hard surface samples in a healthcare setting revealed a remarkable stability in microbial community structure (Figures 5A, B). This observation

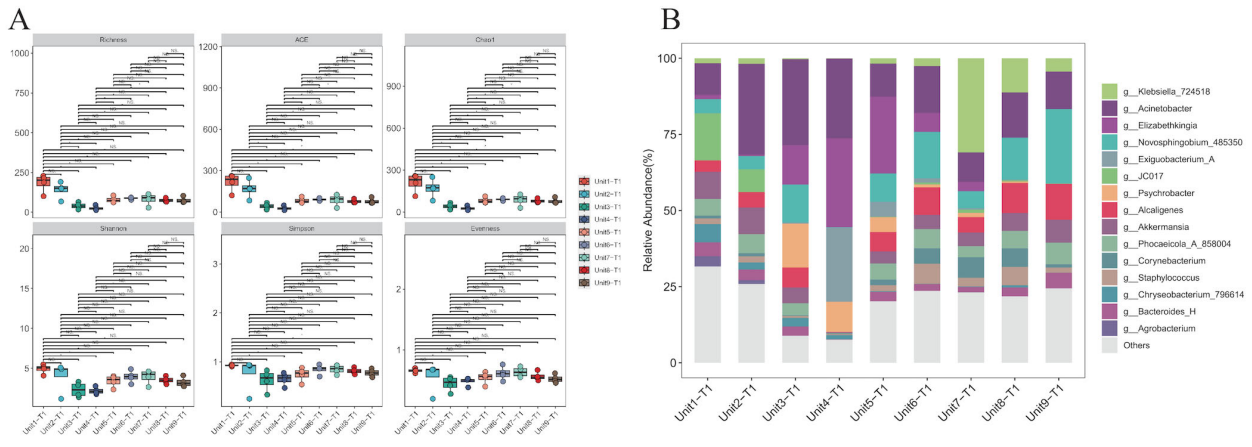



FIGURE 3
Species diversity and species composition of environmental hard surface samples without disinfection. **(A)** Analysis of alpha diversity of samples on hard surfaces near the bedside of K.pneumoniae infected patients without prior disinfection. **(B)** Composition of microbial genera on hard surfaces adjacent to the bedsides of patients infected with K. pneumoniae without prior disinfection. Common hospital pathogens include Klebsiella, Acinetobacter, Staphylococcus, and Elizabethkingia. Environmental symbiotic bacteria or opportunistic bacteria include Akkermansia, Corynebacterium, and Bacteroides.

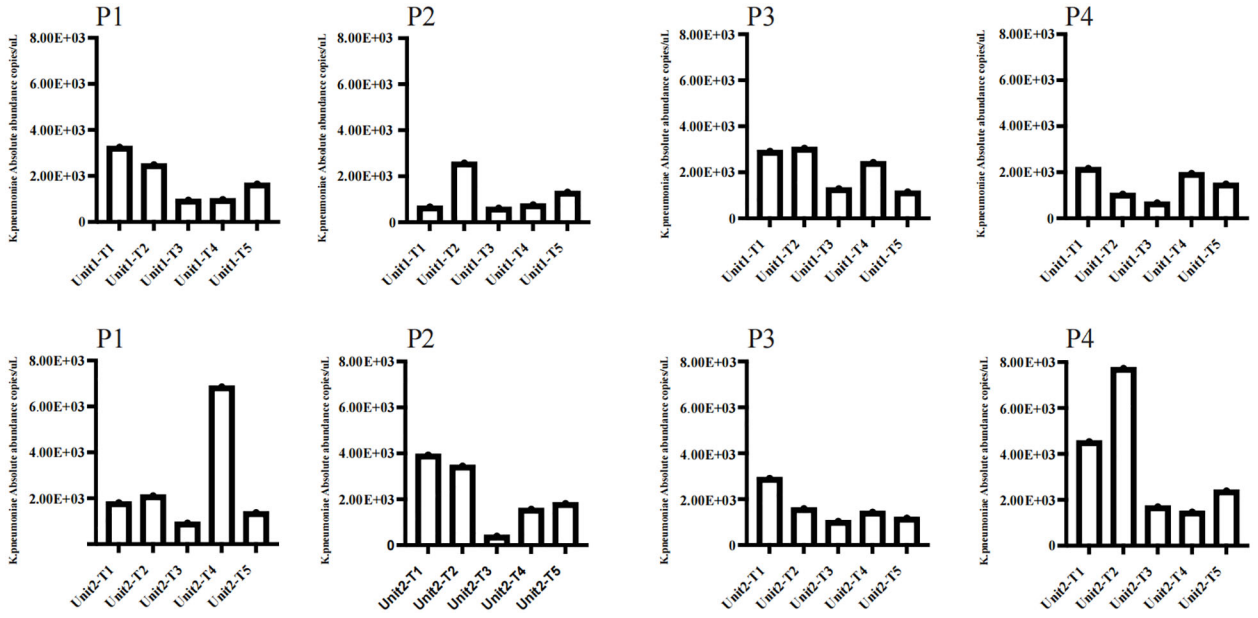



FIGURE 4
Employed ddPCR to quantitatively assess the genomic copy numbers of *K. pneumoniae* in two rooms disinfected with phage cocktails at five distinct time intervals. The sampling points included P1 for the floor, P2 for the bedside table, P3 for the bed rail, and P4 for the bedsheet. The time intervals were as follows: (T1) Before using any disinfectant; (T2) 6 hours after disinfection with chemical disinfectants; (T3) 24 hours after disinfection with phage cocktails; (T4) 48 hours after disinfection with phage cocktails; (T5) 72 hours after disinfection with phage cocktails.

underscores the host-specific action of bacteriophages, which selectively target specific pathogens without disturbing the overall microbial equilibrium (P >0.05).

Initially, *Klebsiella* constituted 6.5% and 7.3% of the microbial community. Post-application of chemical disinfectants, the relative abundance of *Klebsiella* increased marginally to 6.8% in one sample, but alarmingly to 73% in another, suggesting potential disparities in disinfectant application or microbial resistance. Conversely, bacteriophage cocktail treatments demonstrated substantial efficacy. After 24 hours, the prevalence of *Klebsiella* turn into 16% and 15%. This reduction persisted at 48 hours with further declines to 1.2% and 12%, albeit a slight resurgence was observed at 72 hours (3.3% and 14%), possibly due to phage titre decrease or bacterial adaptation. These findings underscore the potential of phage-based disinfectants in significantly reducing *Klebsiella* load, highlighting their advantages over traditional chemical disinfectants, especially given the variable efficacy and potential for resistance development observed with the latter (Figures 5C).

## Discussion

Hospital environmental contamination is a key factor in healthcare-associated infections (HAIs) (Cruz-López et al., 2023; Freier et al., 2023). Traditionally, pathogen detection in environmental samples relies on microbial culture and colony counting (Song et al., 2018). This method is limited in rapidly and accurately identifying specific pathogens. In our study, using both microbial culture and ddPCR to detect *K. pneumoniae* on untreated ward surfaces, we found only 16 positive samples via culture, whereas ddPCR detected *K. pneumoniae* DNA in all samples. This highlights ddPCR's sensitivity in detecting low-abundance or non-viable bacteria.

Regular monitoring of *K. pneumoniae* in hospital environments is crucial for preventing HAIs. In China, the ST11 CRKP clone, which has undergone virulence evolution, is prevalent in clinical settings (Wang et al., 2024). Our study detected the ST11KL64 and ST11KL47 strains of CRKP on hard surfaces in a respiratory intensive care unit, indicating their potential role in transmitting HAIs due to their survival capabilities. As antibiotic resistance worsens, traditional chemical disinfectants face increasing challenges in controlling HAIs. These chemicals can disrupt microbial communities, harm beneficial microbes, and promote resistant microbial populations (Nordholt et al., 2021; Romero-Fierro et al., 2021). Recently, phage therapy, a targeted biological control strategy, has gained renewed interest for treating infections caused by resistant bacteria.The use of bacteriophages as biological disinfectants in hospital environments has shown promising feasibility, and the safety of using bacteriophages as biological disinfectants in hospital environments is well-documented and supported by multiple aspects (Cisek et al., 2017). Firstly, bacteriophages are naturally occurring entities found widely in the environment and within the human body, where they coexist without causing significant adverse effects (Lin et al., 2017). Their high specificity means they target only particular bacterial strains, thus sparing human cells and beneficial microbiota from harm. This minimizes collateral damage to the host's microbiome, which is an important consideration in maintaining overall health. Moreover, bacteriophages biodegrade after completing their lifecycle, leaving no long-term residues in the environment. This characteristic reduces the ecological impact of their use in disinfection processes.

10.3389/fcimb.2024.1442062

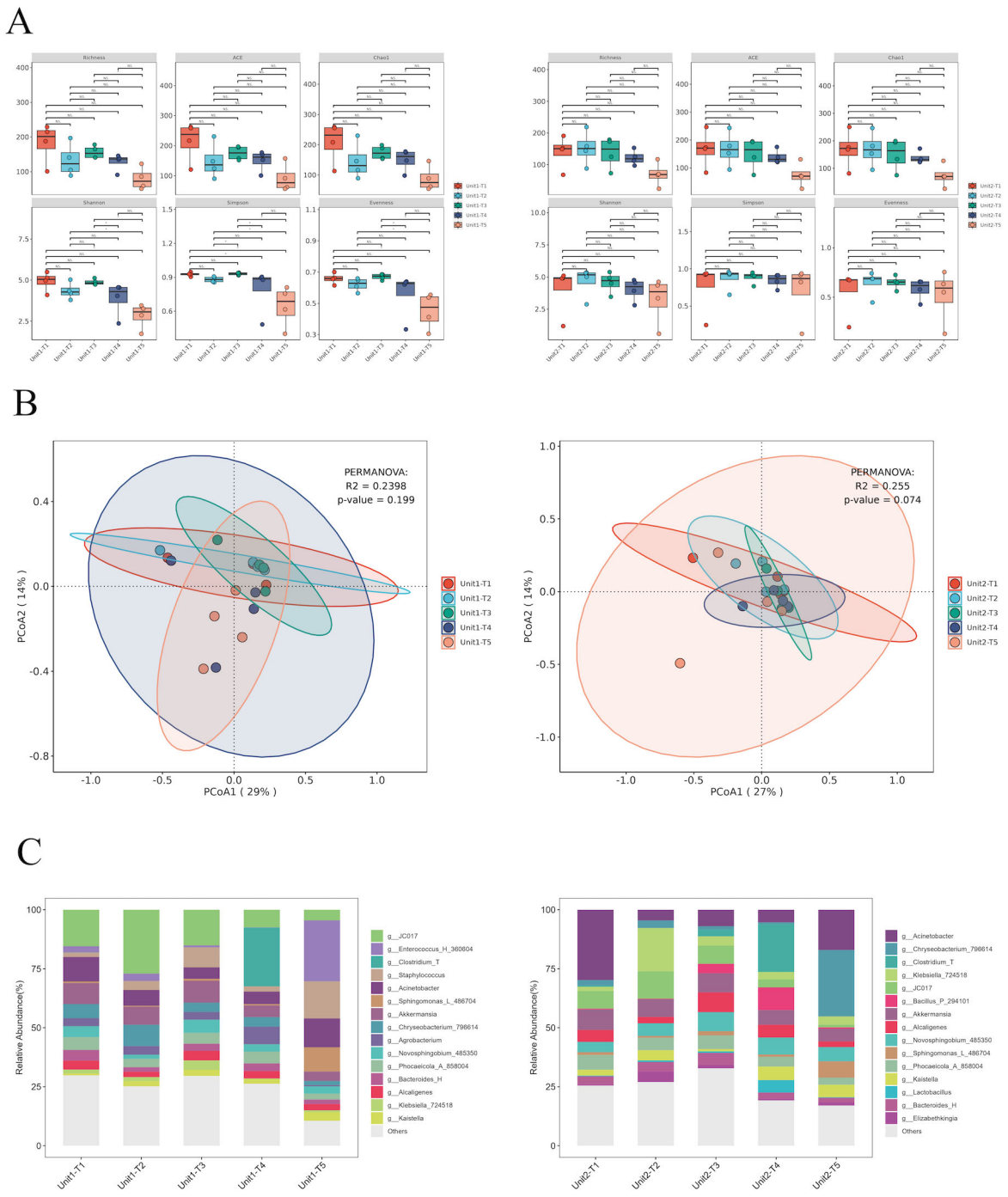



FIGURE 5
Species diversity and species composition of environmental hard surface samples under different treatments of Unit-1 and Unit-2: **(A)** Analysis of species α diversity in environmental hard surface samples before and after phage cocktail intervention. **(B)** Analysis of species β diversity in environmental hard surface samples before and after phage cocktail intervention. **(C)** Temporal changes in the composition of microbial genera on hard surfaces within two hospital rooms treated with phage cocktail disinfectants. T1: Before using any disinfectant, T2: 6 hours after disinfection with chemical disinfectants, T3: 24 hours after disinfection with phage cocktails, T4: 48 hours after disinfection with phage cocktails, T5: 72 hours after disinfection with phage cocktails.

This study investigates the potential of using phage cocktails to clean clinical environments and analyzes changes in environmental microbial communities before and after intervention via 16S rRNA amplicon sequencing. Compared to traditional chemical disinfectants, phages, as biological agents, offer distinct advantages in the cleaning process: they specifically target and eliminate bacteria, reducing harm to beneficial microbes and thus maintaining microbial balance. They can also keep target pathogens at low levels over extended periods. Additionally, employing phages reduces reliance on antibiotics and decreases the risk of developing antibiotic

resistance. These characteristics make phage cleaning methods especially important and promising in modern medical settings.

This study has certain limitations. Firstly, the short duration of the study may not fully capture the long-term efficacy and potential rebound of pathogens. While we employed the latest ddPCR technology to enhance the sensitivity of pathogen detection, logistical and resource constraints prevented us from implementing extended monitoring in this study. This limitation may have led to an incomplete observation of pathogen dynamics over a longer period. Additionally, although we demonstrated the short-term efficacy of bacteriophages, the number of pathogens began to rebound within 24 hours post-treatment. This suggests that the high specificity of bacteriophages may only target specific host bacteria. Over time, some bacteria may develop resistance to the phages through mutations in phage receptor sites. These resistant strains can survive the initial phage attack and begin to repopulate, causing a rebound in bacterial numbers. Future research needs to explore how to optimize the composition of phage cocktails to cover a broader range of pathogen species and improve their persistence in complex hospital settings.

Finally, while bacteriophages show potential as environmental disinfectants, a number of factors need to be considered before practical application, including bacteriophage preparation, stability, cost-effectiveness, and regulatory issues. Future studies should aim to address these practical issues while evaluating the universality and efficiency of phages in different types of hospital Settings. By optimizing phage preparations and applications, a new hospital disinfection strategy that is both environmentally friendly and efficient is expected to be realized.

## Data availability statement

The data presented in the study are deposited in the online repository. The names of the 16S rRNA repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, BioProject ID: PRJNA1090370. The names of the Targeted pathogen detection repository/repositories and accession number(s) can be found below https://ngdc.cncb.ac.cn/sso/login, BioProject ID: PRJNA1122928.

## Author contributions

YS: Data curation, Methodology, Writing – original draft, Conceptualization, Formal Analysis, Funding acquisition, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. WZ: Data curation, Writing – original draft, Writing – review & editing. LL: Investigation, Writing – original draft, Writing – review & editing. WW: Formal Analysis, Methodology, Writing – review & editing. ML: Supervision, Writing – review & editing. KX: Supervision, Writing – review & editing. KW: Supervision, Writing – review & editing. ZS: Resources, Writing – review & editing. FX: Resources, Writing – review & editing. XW: Resources, Writing – review & editing. XS: Investigation, Writing – review & editing. YT: Writing – review & editing, Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft. LX: Writing – review & editing, Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from the National Natural Science Foundation of China (Grant No.82341119) and Beijing Municipal Natural Science Foundation (No. 7222181).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2024.1442062/full#supplementary-material

## References

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Boyce, J. M. (2016). Modern technologies for improving cleaning and disinfection of environmental surfaces in hospitals. *Antimicrob. Resist. Infect. Control* 5, 10. doi: 10.1186/s13756-016-0111-x

Cisek, A. A., Dąbrowska, I., Gregorczyk, K. P., and Wyżewski, Z. (2017). Phage therapy in bacterial infections treatment: One hundred years after the discovery of bacteriophages. *Curr. Microbiol.* 74, 277–283. doi: 10.1007/s00284-016-1166-x

Cruz-López, F., Martínez-Meléndez, A., and Garza-González, E. (2023). How does hospital microbiota contribute to healthcare-associated infections? *Microorganisms* 11. doi: 10.3390/microorganisms11010192

D'Accolti, M., Soffritti, I., Piffanelli, M., Bisi, M., Mazzacane, S., and Caselli, E. (2018). Efficient removal of hospital pathogens from hard surfaces by a combined use of bacteriophages and probiotics: potential as sanitizing agents. *Infect. Drug Resist.* 11, 1015–1026. doi: 10.2147/idr.S170071

Freier, L., Zacharias, N., Gemein, S., Gebel, J., Engelhart, S., Exner, M., et al. (2023). Environmental contamination and persistence of clostridioides difficile in hospital wastewater systems. *Appl. Environ. Microbiol.* 89, e0001423. doi: 10.1128/aem.00014-23

Gordillo Altamirano, F. L., and Barr, J. J. (2019). Phage therapy in the postantibiotic era. *Clin. Microbiol. Rev.* 32. doi: 10.1128/cmr.00066-18

Ho, Y. H., Tseng, C. C., Wang, L. S., Chen, Y. T., Ho, G. J., Lin, T. Y., et al. (2016). Application of Bacteriophage-containing Aerosol against Nosocomial Transmission of Carbapenem-Resistant Acinetobacter baumannii in an Intensive Care Unit. *PLoS One* 11, e0168380. doi: 10.1371/journal.pone.0168380

Kampf, G. (2018). Biocidal agents used for disinfection can enhance antibiotic resistance in gram-negative species. *Antibiotics (Basel)* 7. doi: 10.3390/antibiotics7040110

Lin, D. M., Koskella, B., and Lin, H. C. (2017). Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World J. Gastrointest Pharmacol. Ther.* 8, 162–173. doi: 10.4292/wjgpt.v8.i3.162

Moye, Z. D., Woolston, J., and Sulakvelidze, A. (2018). Bacteriophage applications for food production and processing. *Viruses* 10. doi: 10.3390/v10040205

Nick, J. A., Dedrick, R. M., Gray, A. L., Vladar, E. K., Smith, B. E., Freeman, K. G., et al. (2022). Host and pathogen response to bacteriophage engineered against Mycobacterium abscessus lung infection. *Cell* 185, 1860–1874.e1812. doi: 10.1016/j.cell.2022.04.024

Nordholt, N., Kanaris, O., Schmidt, S. B. I., and Schreiber, F. (2021). Persistence against benzalkonium chloride promotes rapid evolution of tolerance during periodic disinfection. *Nat. Commun.* 12, 6792. doi: 10.1038/s41467-021-27019-8

Odoyo, E., Matano, D., Tiria, F., Georges, M., Kyanya, C., Wahome, S., et al. (2023). Environmental contamination across multiple hospital departments with multidrug-resistant bacteria pose an elevated risk of healthcare-associated infections in Kenyan hospitals. *Antimicrob. Resist. Infect. Control* 12, 22. doi: 10.1186/s13756-023-01227-x

Peters, A., Schmid, M. N., Parneix, P., Lebowitz, D., de Kraker, M., Sauser, J., et al. (2022). Impact of environmental hygiene interventions on healthcare-associated infections and patient colonization: a systematic review. *Antimicrob. Resist. Infect. Control* 11, 38. doi: 10.1186/s13756-022-01075-1

Romero-Fierro, D., Bustamante-Torres, M., Hidalgo-Bonilla, S., and Bucio, E. (2021). "Microbial degradation of disinfectants," in *Recent Advances in Microbial Degradation*. Eds. Inamuddin,, M. I. Ahamed and R. Prasad (Springer Singapore, Singapore), 91–130.

Salmond, G. P., and Fineran, P. C. (2015). A century of the phage: past, present and future. *Nat. Rev. Microbiol.* 13, 777–786. doi: 10.1038/nrmicro3564

Song, D., Liu, H., Dong, Q., Bian, Z., Wu, H., and Lei, Y. (2018). Digital, rapid, accurate, and label-free enumeration of viable microorganisms enabled by custom-built on-glass-slide culturing device and microscopic scanning. *Sensors (Basel)* 18. doi: 10.3390/s18113700

Tie, K., Yuan, Y., Yan, S., Yu, X., Zhang, Q., Xu, H., et al. (2018). Isolation and identification of Salmonella pullorum bacteriophage YSP2 and its use as a therapy for chicken diarrhea. *Virus Genes* 54, 446–456. doi: 10.1007/s11262-018-1549-0

Wang, Q., Wang, R., Wang, S., Zhang, A., Duan, Q., Sun, S., et al. (2024). Expansion and transmission dynamics of high risk carbapenem-resistant Klebsiella pneumoniae subclones in China: An epidemiological, spatial, genomic analysis. *Drug Resist. Update* 74, 101083. doi: 10.1016/j.drup.2024.101083

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 19 August 2024
DOI 10.3389/fcimb.2024.1438982



OPEN ACCESS

EDITED BY
Jiemin Zhou,
Vision Medicals Co, Ltd, China

REVIEWED BY
Ping Allen Wu,
Wenzhou Medical University, China
Jing Li,
Huzhou University, China
Chunfu Wang,
Air Force Medical University, China

*CORRESPONDENCE
Na Du
du_na@jlu.edu.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 27 May 2024
ACCEPTED 18 July 2024
PUBLISHED 19 August 2024

CITATION
Sun L, Zhang K, Liu Y, Che L, Zhang P, Wang B and Du N (2024) Metagenomic next-generation sequencing targeted and metagenomic next-generation sequencing for pulmonary infection in HIV-infected and non-HIV-infected individuals.
*Front. Cell. Infect. Microbiol.* 14:1438982.
doi: 10.3389/fcimb.2024.1438982



# Metagenomic next-generation sequencing targeted and metagenomic next-generation sequencing for pulmonary infection in HIV-infected and non-HIV-infected individuals

Luyao Sun[1†], Kaiyu Zhang[1†], Yong Liu[2], Lihe Che[1], Peng Zhang[1], Bin Wang[1] and Na Du[1*]

[1]Department of Infectious Diseases, The First Hospital of Jilin University, Changchun, China, [2]Genetic Diagnosis Center, The First Hospital of Jilin University, Changchun, China

**Background:** When individuals infected with human immunodeficiency virus (HIV) experience pulmonary infections, they often exhibit severe symptoms and face a grim prognosis. Consequently, early, rapid, and accurate pathogen diagnosis is vital for informing effective treatment strategies. This study aimed to use metagenomic next-generation sequencing (mNGS) and targeted mNGS (tNGS) to elucidate the characteristics of pulmonary infections in HIV and non-HIV individuals.

**Methods:** This study enrolled 90 patients with pulmonary infection at the Department of Infectious Diseases of The First Hospital of Jilin University from June 2022 to May 2023, and they were divided into HIV (n=46) and non-HIV (n=44) infection groups. Their bronchoalveolar lavage fluid (BALF) was collected for mNGS analysis to evaluate the differences in pulmonary infection pathogens, and tNGS detection was performed on BALF samples from 15 HIV-infected patients.

**Results:** A total of 37 pathogens were identified in this study, including 21 bacteria, 5 fungi, 5 viruses, 5 mycobacteria, and 1 mycoplasma. The sensitivity of mNGS was 78.9% (71/90), which is significantly higher than that of conventional methods (CTM) (39/90, P=1.5E-8). The combination of mNGS with CTM can greatly enhance the sensitivity of pathogen detection. The prevalence of *Pneumocystis jirovecii* (82.6% *vs.* 9.1%), cytomegalovirus (CMV) (58.7% *vs.* 0%), and Epstein-Barr virus (EBV) (17.4% *vs.* 2.3%) was significantly higher in the HIV infection group than in the non-HIV infection group ($P<0.05$). Although no statistically significant difference was observed, the detection rate of *Mycobacteria* was higher in HIV-infected patients (17.4%) than in the non-HIV group (6.8%). Furthermore, the tNGS results of BALF from 15 HIV-infected patients were not entirely consistent with the mNGS results., and the concordance rate of tNGS for the detection of main pathogens reached 86.7% (13/15).

**Conclusion:** Next-generation sequencing (NGS) can accurately detect pathogens in the BALF of patients with pulmonary infection. The sensitivity of tNGS is comparable to that of mNGS. Therefore, this technique should be promoted in the clinic for better patient outcomes.

KEYWORDS

**bronchoalveolar lavage fluid, metagenomic next-generation sequencing (mNGS), targeted metagenomic next-generation sequencing (tNGS), HIV infection, pathogens**

# Introduction

Pulmonary infection is one of the most common causes of fatality among humans due to its rapid onset, quick progression, and multiple comorbidity (Kalil et al., 2016). The World Health Organization estimates that more than 20 million people die of pneumonia annually (McAllister et al., 2019). It is caused by many pathogens, including bacteria, viruses, and fungi, and is more likely to occur in immunocompromised patients, especially human immunodeficiency virus (HIV)-infected individuals (Cribbs et al., 2020). Such cases may present severe manifestation and worse outcomes, thus requiring early, rapid, and accurate detection of pathogens to guide an effective treatment strategy and improve the prognosis. However, conventional methods (CTM), such as culture, sputum smear, and molecular assays, have low diagnostic efficiency of only 30-40% (Afsar et al., 2018; Phetsuksiri et al., 2020). So, clinicians frequently prescribe empirical anti-infective treatment based on clinical manifestations and imaging data. Nevertheless, there are significant differences in the types and numbers of pathogens between HIV-infected and non-HIV-infected people (Tan et al., 2023).

Metagenomic next-generation sequencing (mNGS) has recently been applied for diagnosing gene mutations as well as many infectious diseases like pulmonary infection (Pourahmadiyan et al., 2022; Kullar et al., 2023). This technique has a significantly higher detection speed over CTM and can perform high-throughput sequencing analysis of DNA or RNA using specimens (Schlaberg et al., 2017). By comparing the sequence data with the reference genome or a particular database, clinicians can identify the variation and quantity of various microorganisms as well as drug-resistance genes (Ávila-Ríos et al., 2020). Many studies have reported the high sensitivity of mNGS for detecting pathogens (Ávila-Ríos et al., 2020; Zhang et al., 2022; Kullar et al., 2023; Wu et al., 2023). Nevertheless, the obtained specimens can affect the accuracy of mNGS. Sputum and throat swabs are not ideal specimens for detection tests because they are more likely to be contaminated, and they carry fewer germs than bronchoalveolar lavage fluid (BALF) (Wu et al., 2023). Despite BALF being more difficult to obtain than sputum and throat swabs from patients with pulmonary infection, it is a more reliable and acceptable sample for mNGS analysis (Zhang et al., 2022). In clinical applications of mNGS, following the initial mNGS detection, the targeted metagenomic next-generation sequencing (tNGS) method has gradually emerged. mNGS is relatively expensive and may be affected by human genomic interference. tNGS, on the other hand, combines multiplex PCR amplification with high-throughput sequencing, focusing on detecting known pathogenic microorganisms and drug resistance genes in the samples. tNGS reduces detection costs and increases the sensitivity of pathogen detection, which can simultaneously detect DNA and RNA, shorten detection time, and is characterized by high cost-effectiveness and customizability.

Currently, there are few researches on the characteristics of pathogens among HIV-infected patients complicated with pneumonia. This study retrospectively enrolled HIV-infected and non-HIV-infected patients, evaluated the diagnostic value of mNGS and tNGS in the diagnosis of pulmonary infections, and summarized the pathogen profiles in such cases. Additionally, the performance of CTM and mNGS in diagnosing pulmonary infections was also summarized.

# Methods

## Study subjects

This study recruited 90 patients diagnosed with lower respiratory tract infection at the Department of Infectious Diseases of The First Hospital of Jilin University from June 2022 to May 2023. There were 46 patients who tested positive for HIV infection and 44 patients who tested negative for HIV infection. All patients underwent HIV antigen-antibody testing. For patients with positive HIV screening antibodies, a confirmatory Western blot (WB) assay was performed for confirmation. The detailed demographic information, clinical data, and sequencing results of the patients were obtained from the electronic medical records. The research protocol was approved by the ethics committee of our hospital (No. 2021-022-01). Each individual provided written informed consent.

## BALF extraction and mNGS analysis

Before extracting the BALF, the routine preoperative preparation was performed, with 2% lidocaine (15-20 mL) being used for nebulization anesthesia. The BF-260 electronic bronchoscope (Olympus, Japan) was applied to extract the BALF with 100-150 mL of sterile physiological saline at 37°C. Then the specimen was placed in a bottle coated with silicone.

And then, 1 mL of BALF sample was centrifuged at 12,000 rpm for 5 min to collect the pathogens and human cells. Next, 50 μL of precipitate underwent depletion of host nucleic acids using 1 U of benzonase and 0.5% Tween 20 at 37°C for 5 min. Subsequently, total nucleic acids were extracted using a QIAamp UCP pathogen minikit (Qiagen, Germany). To generate the library, 30 μL of the nucleic acid eluate was used for DNA fragmentation, end repair, adapter ligation, and polymerase chain reaction (PCR) amplification using a Nextera DNA Flex kit (Illumina, San Diego, CA, USA). A Qubit dsDNA HS assay kit was used to measure the library concentration. An Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) was applied to assess the library quality with a high-sensitivity DNA kit. Besides, peripheral blood mononuclear cell samples from healthy donors were prepared using the same protocol, and sterile deionized water was extracted alongside the specimens to serve as a non-template control (NTC) (Deng et al., 2022; Tao et al., 2022). The pooled libraries were sequenced on a Nextseq 550 sequencing system (Illumina, San Diego, CA, USA) with a single-end sequencing kit.

## tNGS pipeline

Based on multiplex PCR and mNGS, in-house tNGS panel was designed to detect 273 pathogens causing infections in different systems according to the public pathogen databases and the published studies. The 273 pathogens included 113 bacteria, 47 fungi, 101 viruses, and 12 parasites. After DNA extraction, multiplex PCR with the designed primers was used to construct libraries. Library concentrations were quantified using Qubit 4.0. Illumina NextSeq platform was used for high-throughput sequencing.

## Bioinformatics analysis

Trimmomatic software was used to remove the low-quality reads, adapter contamination, duplicate reads, as well as reads shorter than 35 bp. Low-complexity reads were also removed. Human sequence data were identified and excluded by mapping to a human reference genome (hg38) with the Burrows–Wheeler Aligner (BWA) software. The remaining sequencing information was aligned with the most recent databases for bacteria, viruses, fungi, and protozoa provided by the National Center for Biotechnology Information. Reads that met the criteria for being considered unique were those with alignment lengths greater than 80%, sequence identities over 90%, and suboptimal to optimal alignment score ratios less than 0.8.

## Criteria for a positive mNGS/tNGS result

The specifically mapped read number (SMRN) of each microbial taxonomy was normalized to the SMRN per 20 million total sequencing reads to give the standardized SMRN (SDSMRN). The criteria for reporting the positive mNGS/tNGS results were as follows (Qin et al., 2021): (1) SDSMRN≥3 for bacteria (mycobacteria excluded); (2) SDSMRN≥3 for fungi/DNA viruses; (3) SDSMRN≥1 for RNA viruses; (4) SDSMRN≥100 for parasites; (5) SDSMRN≥3 for *Mycoplasma/Chlamydia* spp.; (6) SDSMRN≥1 (or SDSMRNG≥1 at the genus level) for *Mycobacterium tuberculosis* complex; (7) SDSMRN ≥3 for *Nocardia* spp. as a positive result.

## Statistical analysis

SPSS version 22.0 software (IBM Corp., Armonk, NY, USA) was utilized for data analysis. The independent t-test was performed to compare the differences in counting data between the two groups. The chi-squared test was used to evaluate the categorical variables. A *P*-value below 0.05 was regarded as a significant difference.

## Data availability

Sequencing data were deposited to the National Genomics Data Center under accession number CRA016496. The authors declare that the main data supporting the findings are available within this article. The other data generated and analyzed for this study are available from the corresponding author upon reasonable request.

# Results

## Clinical characteristics

In this study of 90 patients, 46 were identified as HIV-positive. Within this HIV-infected cohort, a significantly higher proportion were male compared to females, marking a clear contrast with the non-HIV group. Furthermore, the ages of patients in the HIV group were significantly lower than those in the non-HIV group. Across all these patients with pulmonary infections, the predominant symptoms were fever, cough, and expectoration, with no significant differences between the two groups. Patients infected with HIV have significantly lower counts of CD4 and the CD4/CD8 ratio compared to patients without HIV. In addition, the occurrence of underlying diseases was significantly greater in the non-HIV group compared to the HIV-infected group, with hypertension being the most common among these conditions (Table 1).

TABLE 1 Basic characteristics.

| | Total (n=90) | HIV (n=46) | Non-HIV (n=44) | *P*-value |
|---|---|---|---|---|
| **Sex** | | | | |
| Male | 67(74.4%) | 44(95.7) | 23(52.3%) | 1.6E-6 |
| Female | 23(25.6%) | 2(4.3%) | 21(47.7%) | |
| **Age** (Median[IQR]) | 52(40-63.5) | 43.5(33.25-53) | 59(51.75-69) | 9.3E-7 |
| **Clinical characteristics** | | | | |
| Fever | 82(91.1%) | 41(89.1%) | 41(93.2%) | 0.7 |
| Cough | 69(76.7%) | 35(76.1%) | 34(77.3%) | 1 |
| Expectoration | 53(58.9%) | 23(50%) | 30(68.2%) | 0.1 |
| Diarrhea | 16(17.8%) | 11(23.9%) | 5(11.4%) | 0.2 |
| **Immunity index** | | | | |
| CD4 [Median(IQR)] | 128.5 (13.725-484.75) | 13.75 (6.1-35.25) | 511.5 (315-759.5) | 3.3E-12 |
| CD4/CD8 [Median(IQR)] | 0.225 (0.0525-1.93) | 0.055 (0.02-0.1) | 2.02 (0.84-3.205) | 2.2E-13 |
| **Comorbidities** | | | | |
| Hypertension | 16(17.8%) | 2(4.3%) | 14(31.8%) | 7.2E-4 |
| Coronary heart disease | 2(2.2%) | 0 | 2(4.5%) | 0.2 |
| Diabetes | 7(7.7%) | 0 | 7(15.9%) | 5.1E-3 |
| Tumor | 7(7.7%) | 0 | 7(15.9%) | 5.1E-3 |
| Interstitial lung disease | 1(1.1%) | 0 | 1(2.3%) | 0.5 |

## Pathogen profiles of HIV group and non-HIV group

Two physicians independently determined the causative pathogens for each patient based on the results of mNGS and CTM, as well as the patient's treatment strategies, and prognosis. In cases of disagreement, a third physician provided the final determination. According to the final clinical diagnosis results, a total of 37 pathogens were observed in this study, including 21 bacteria, 5 fungi, 5 viruses, 5 mycobacteria, and 1 mycoplasma. We found a notably higher incidence of fungal and viral infections among HIV-infected patients compared to those without HIV infection (Table 2).

The pathogen spectrum in patients from both the HIV and non-HIV groups was shown in Figure 1. The results indicated that the primary reason for the significantly higher rate of fungal and viral infections in HIV patients was attributed to *Pneumocystis jirovecii*, cytomegalovirus (CMV), and Epstein-Barr virus (EBV). The prevalence of *P. jirovecii* in HIV patients was as high as 82.6% (38/46), which was much higher than in the non-HIV group. For the rarely pathogenic agents CMV and EBV, their potential for opportunistic infection should not be underestimated due to the unique characteristics of HIV patients. The occurrence rates of other pathogens do not differ between the two groups, but it is noteworthy that the frequency of *Mycobacteria* in HIV patients is also higher than that in non-HIV patients, although the difference is not statistically significant.

## Comparison of mNGS with CTM

We further compared the mNGS detection results with CTM (Figure 2A). The sensitivity of mNGS reached 78.9% (71/90), which is markedly superior to that of CTM (39/90, *P*=1.5E-8). The superiority of mNGS over CTM was demonstrated in both the HIV and non-HIV groups (Figure 2B). The negative results of mNGS included 10 patients with no microorganisms detected, and 9 cases where the detected pathogen was excluded as the causative agents. All of these 19 patients were not infected with HIV, and their CTM also showed negative results.

The microorganisms identified by mNGS and CTM was depicted in Figure 3. In HIV patients, CTM missed many cases of *P. jirovecii* (n=16) and CMV (n=13), which were all detected by

TABLE 2 Pathogen types in the BALF from the HIV and non-HIV infection groups.

| Pathogen type | HIV infection group | Non-HIV infection group | *P* |
|---|---|---|---|
| Fungi | 91.3% (42/46) | 20.5% (9/44) | 3.7E-12 |
| Viruses | 67.4% (31/46) | 6.8% (3/44) | 1.1E-9 |
| Bacteria | 34.8% (16/46) | 47.7% (21/44) | 0.3 |

BALF, bronchoalveolar lavage fluid.

mNGS. Similarly, in non-HIV patients, mNGS detected bacterial and fungal infections that CTM failed to identify. mNGS only missed 3 pathogenic agents detected by CTM, including one case of *Klebsiella pneumoniae*, one case of *Acinetobacter baumannii*, and one case of SARS-CoV-2. In this study, two patients died due to multi-organ failure, including one HIV-positive individual and one non-HIV-positive individual. The vast majority of the remaining patients attained improved prognoses through treatment strategies informed by pathogen detection results (Figure 2C). Thus, it can be seen that mNGS detection can greatly reduce the rate of missed diagnoses due to the insufficient sensitivity of CTM.

## The performance of tNGS

BALF samples from 15 HIV patients underwent concurrent tNGS detection to assess its performance. Among them, tNGS identified the primary causative pathogens in 13 patients, achieving a consistency rate of 86.7% (13/15) compared to mNGS. tNGS identified some additional pathogens, mainly viruses, compared to mNGS (Figure 4A). However, the pathogenicity of most of these viruses was not considered. Additionally, tNGS successfully identified all causative pathogens in 3 cases. In other cases, CMV was the main pathogen that tNGS did not detect. (Figure 4B). Hence, the diagnostic potential of tNGS for pulmonary infections in HIV patients is promising.

# Discussion

It is crucial to identify pathogens and perform accurate treatment regimens on patients with pulmonary infection. The conventional detection methods include sputum smear, culture, and other molecular tests, but they have low diagnostic rates, which may delay the timing of treatment (Caetano Mota et al., 2012). In this study, we performed mNGS analysis on the BALF from HIV-infected and non-HIV-infected patients with pulmonary infection and found significant differences in the pathogen profiles between the two study groups. tNGS detection was simultaneously performed on BALF from 15 HIV-infected patients to observe the clinical utility of tNGS in HIV-infected patients.

In this study, patients in the non-HIV group were found to have a higher prevalence of bacterial infections, while HIV patients were more commonly infected with fungi and viruses. Particularly notable was the significantly higher proportion of *P. jirovecii*, CMV, EBV, and tuberculosis in HIV-infected patients compared to those without HIV infection. This indicated that they had multi-pathogenic pneumonia and were at a higher risk of developing

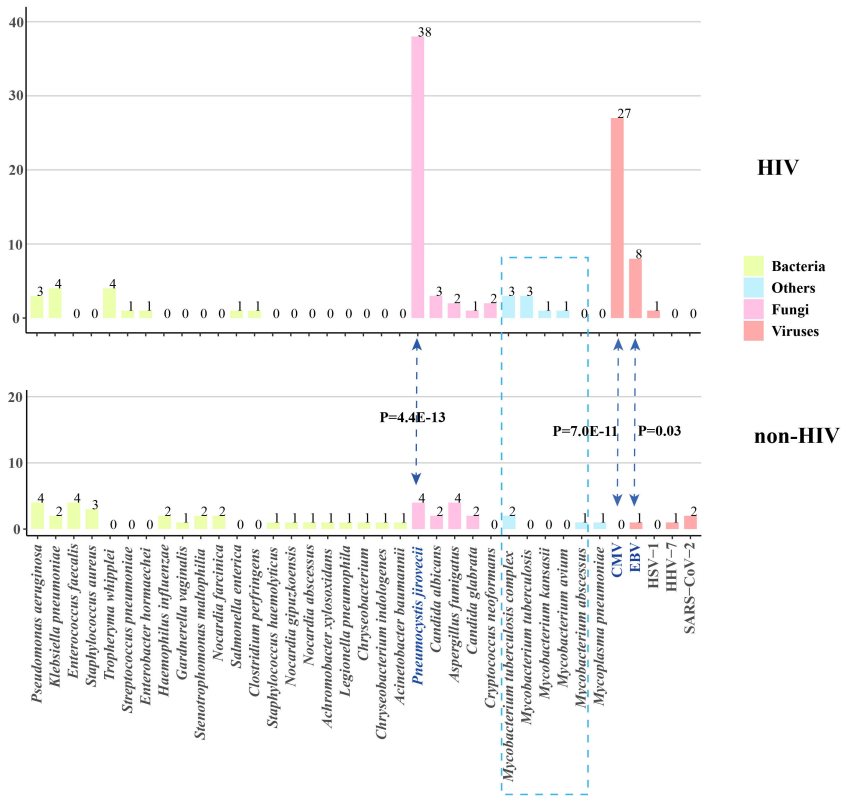



FIGURE 1
Pathogen profiles of patients with and without HIV infection. CMV, cytomegalovirus; EBV, Epstein-Barr virus; HSV, herpes simplex virus; HHV, human herpes virus; HIV, human immunodeficiency virus.

opportunistic infections. Opportunistic infections frequently occur in individuals with acquired immune deficiency (Shi et al., 2022), and the current study showed that the HIV-positive patients were infected with more than one type of pathogens, such as *P. jirovecii*, *M. tuberculosis*, and nontuberculous mycobacteria, as well as viruses, suggesting that they were susceptible to developing a mixed infection. Previous research has revealed that HIV-infected individuals with a lower $CD4^+$ T-cell count and $CD4^+/CD8^+$ ratio are prone to pneumonia (Xie et al., 2023). Therefore, monitoring $CD4^+$ lymphocytes with mNGS analysis is a favorable detection method for HIV-infected patients with suspected lung infection. Moreover, mNGS can identify the underlying pathogenic germs as well as vulnerable individuals like the HIV-infected population.

This study revealed that the constituent ratio of *P. jirovecii* ranked first among the HIV-infected patients and was higher than that among the non-HIV-infected patients. This finding is consistent with the previous studies demonstrating that this pathogen is more likely to occur in immunosuppressed hosts (Souza Lopes et al., 2016). *P. jirovecii* can aggravate lung exudation and increase the mortality of infected patients, especially those with an impaired immune function (Cheng et al., 2023). Therefore, early detection of such life-threatening bacteria can help clinicians determine the right treatment strategy to prevent infected patients from developing severe pneumonia.

Different hospitals have distinct detection capabilities and methods, which can greatly affect the detection rate of pathogens. In China, smear staining remains the primary detection method for *P. jirovecii* (Li et al., 2020). However, there is no unified operation, quality-control process, or diagnostic standard, which brings great diagnostic difficulties to clinicians. Over the years, culture has been used as the gold standard for the diagnosis of *M. tuberculosis* and nontuberculous mycobacteria, but the bacterial culture requires harsh conditions for operators and laboratories (Jain et al., 2017). In addition, its procedure takes a long time, which cannot be fed back to clinicians in time. Moreover, the culture method shows low sensitivity. With the advancement of mNGS, the detection rates of *M. tuberculosis* and nontuberculous mycobacteria have been greatly increased, especially for drug-resistant strains, indicating that a large proportion of patients can receive effective, timely, and accurate treatment. Serum markers and immunologic assays also showed low specificity in the detection of *Cryptococcus*, various viruses, and other pathogens (Jarvis et al., 2014). Additionally, non-HIV-infected people with community-acquired or nosocomial pneumonia had a very low chance of being detected with true

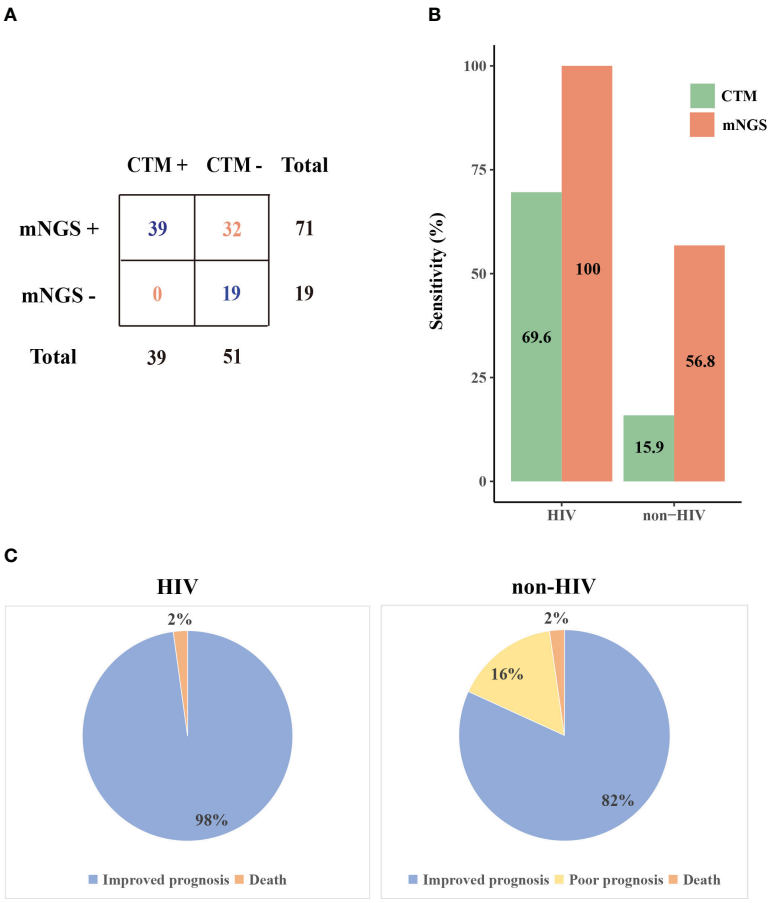



FIGURE 2
The performance of mNGS and CTM in patients with different prognoses. **(A)** Comparison of mNGS and CTM detection **(B)** The sensitivity of mNGS and CTM in HIV and non-HIV group **(C)** Prognosis of patients in HIV and non-HIV group. CTM, conventional methods.

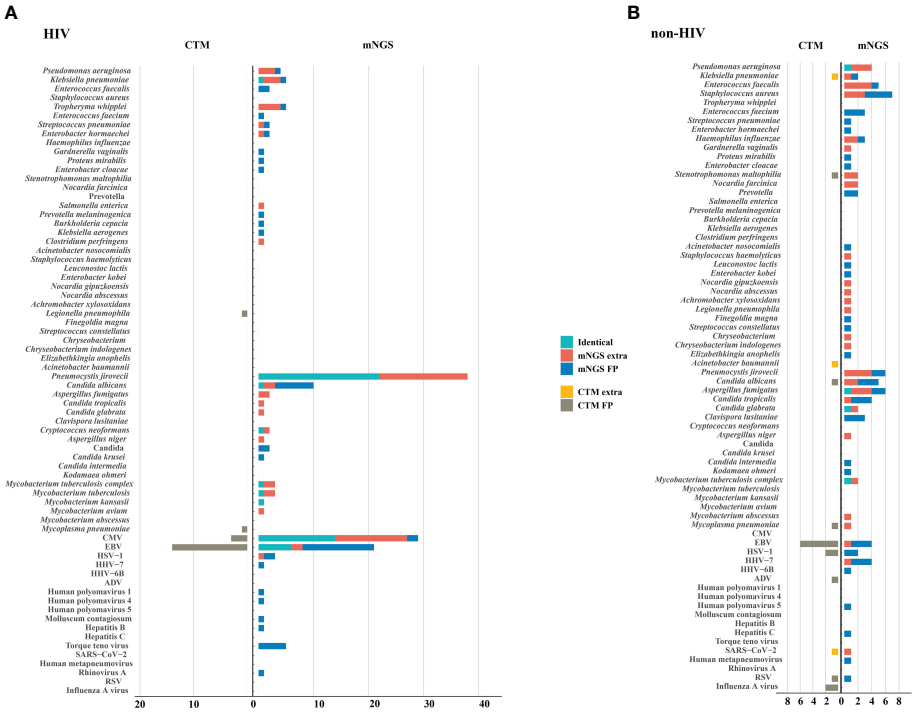



FIGURE 3
Potential pathogens identified by mNGS and CTM. **(A)** HIV group; **(B)** non HIV group. The pathogens consisted of identical detection (the green) of mNGS to CTM, extra detection (the red) of mNGS, false positive detection (the blue) of mNGS, extra detection (the yellow) of CTM, false positive detection (the grey) of CTM. The false positive detection of mNGS represents those microorganisms were detected only by mNGS and diagnosed as non-causitive pathogens. The false positive detection of CTM represents those microorganisms were detected only by CTM and diagnosed as non-causitive pathogens. CMV, cytomegalovirus; EBV, Epstein-Barr virus; HSV, herpes simplex virus; HHV, human herpes virus; HIV, human immunodeficiency virus; ADV, adenovirus; RSV, respiratory syncytial virus.

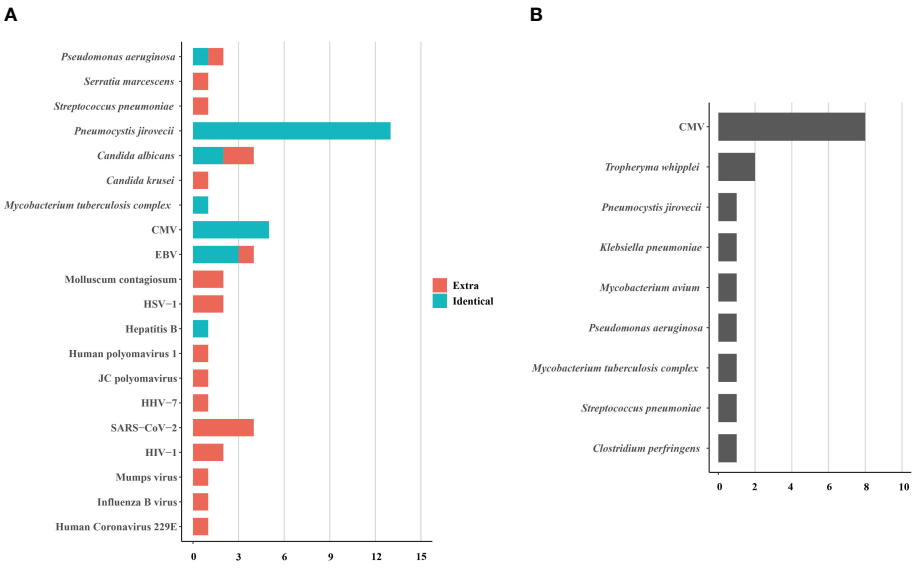



FIGURE 4
The pathogens identified **(A)** and not identified **(B)** by tNGS in 15 HIV-infected patients. The pathogens consisted of identical detection (the green) of tNGS to mNGS and extra detection (the red) of tNGS. CMV, cytomegalovirus; EBV, Epstein-Barr virus; HSV, herpes simplex virus; HHV, human herpes virus; HIV, human immunodeficiency virus.

pathogenic pathogens when only CTM were used, and the accuracy rate of sputum smear and culture methods was less than 50% (Caetano Mota et al., 2012).

Many studies have shown that mNGS is an irreplaceable method for identifying pathogens in complex infections. Compared with CTM, mNGS shows higher sensitivity because it can simultaneously detect multiple pathogens and mutated genes as well as requires a small amount of sample, which can help achieve a precise anti-infection treatment (Morganti et al., 2019; Pourahmadiyan et al., 2022). Although the adoption rate of mNGS remains low in resource-poor areas of the world, it is expected to be extensively used for the detection of complex infectious diseases with the continuous progress of mNGS technology and the significant decline in its cost.

Compared to the high economic burden of mNGS, tNGS is more economically feasible for clinical application. tNGS is a method that combines gene-targeted PCR amplification and high-throughput sequencing, and its application value for pathogen detection in clinical settings has been evaluated in several literature studies (Mertes et al., 2011; Gaston et al., 2022). tNGS can rapidly identify pathogens within 24-48 hours and complete the detection of drug-resistant genes and virulence genes (Onda et al., 2018). However, its applicability to patient populations is relatively limited. If the patient presents with fever of unknown origin, unclear etiology, and uncertain infectious pathogens, clinicians may require a broader range of testing to assist in diagnosis, especially for some rare and specific infectious pathogens. However, for some populations, such as HIV-infected individuals, opportunistic infectious pathogens are predominant. The tNGS detection range can almost meet the needs of the majority of patients, and it is more economical, making it more widely applicable in clinical practice.

### Limitation

This study had some limitations. First, it was a small-sample, single-center research study, which might limit the accuracy of the study. Further research should be performed with more samples from distinct areas. Second, there is no unified detection process for mNGS, and we did it according to the published literature, which might cause false-positive or false-negative results. So, the data interpretation was carried out by three experienced laboratory personnel to lower the result bias as much as possible. Finally, due to the limited budget, not all HIV patients underwent tNGS testing, and its performance still requires further research.

## Conclusions

It is important to perform mNGS or tNGS in immunosuppressed patients combined with pulmonary infection since these methods can quickly and precisely reveal pathogens to help determine an effective anti-infective regimen. Therefore, both mNGS and tNGS should be promoted in the clinic for better patient outcomes.

## Data availability statement

The data presented in the study are deposited in the China National Center for Bioinformation - National Genomics Data Center - Genome Sequence Archive, submission numbers CRA016496 and CRA017709.

## Ethics statement

The studies involving humans were approved by the ethics committee of The First Hospital of Jilin University (No. 2021-022-01). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

LS: Data curation, Investigation, Writing – original draft. KZ: Data curation, Writing – review & editing, Conceptualization, Visualization. YL: Writing – review & editing, Data curation, Software. LC: Formal analysis, Project administration, Resources, Writing – review & editing. PZ: Methodology, Investigation, Validation, Writing – review & editing. BW: Writing – review & editing, Data curation, Investigation. ND: Conceptualization, Funding acquisition, Writing – review & editing, Supervision, Visualization.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the Jilin Provincial Department of Science and Technology, Norman Bethune Special Project(20210101434JC).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

# References

Afsar, I., Gunes, M., Er, H., and Gamze Sener, A. (2018). Comparison of culture, microscopic smear and molecular methods in diagnosis of tuberculosis. *Rev. Esp Quimioter* 31, 435–438.

Caetano Mota, P., Carvalho, A., Valente, I., Braga, R., and Duarte, R. (2012). Predictors of delayed sputum smear and culture conversion among a Portuguese population with pulmonary tuberculosis. *Rev. Port Pneumol* 18, 72–79. doi: 10.1016/j.rppneu.2011.12.005

Cheng, Q. W., Shen, H. L., Dong, Z. H., Zhang, Q. Q., Wang, Y. F., Yan, J., et al. (2023). Pneumocystis jirovecii diagnosed by next-generation sequencing of bronchoscopic alveolar lavage fluid: A case report and review of literature. *World J. Clin. cases* 11, 866–873. doi: 10.12998/wjcc.v11.i4.866

Cribbs, S. K., Crothers, K., and Morris, A. (2020). Pathogenesis of HIV-related lung disease: immunity, infection, and inflammation. *Physiol. Rev.* 100, 603–632. doi: 10.1152/physrev.00039.2018

Deng, W., Xu, H., Wu, Y., and Li, J. (2022). Diagnostic value of bronchoalveolar lavage fluid metagenomic next-generation sequencing in pediatric pneumonia. *Front. Cell Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.950531

Gaston, D. C., Miller, H. B., Fissel, J. A., Jacobs, E., Gough, E., Wu, J., et al. (2022). Evaluation of metagenomic and targeted next-generation sequencing workflows for detection of respiratory pathogens from bronchoalveolar lavage fluid specimens. *J. Clin. Microbiol.* 60, e0052622. doi: 10.1128/jcm.00526-22

Jain, D., Ghosh, S., Teixeira, L., and Mukhopadhyay, S. (2017). Pathology of pulmonary tuberculosis and non-tuberculous mycobacterial lung disease: Facts, misconceptions, and practical tips for pathologists. *Semin. Diagn. Pathol.* 34, 518–529. doi: 10.1053/j.semdp.2017.06.003

Jarvis, J. N., Bicanic, T., Loyse, A., Meintjes, G., Hogan, L., Roberts, C. H., et al. (2014). Very low levels of 25-hydroxyvitamin D are not associated with immunologic changes or clinical outcome in South African patients with HIV-associated cryptococcal meningitis. *Clin. Infect. Dis.* 59, 493–500. doi: 10.1093/cid/ciu349

Kalil, A. C., Metersky, M. L., Klompas, M., Muscedere, J., Sweeney, D. A., Palmer, L. B., et al. (2016). Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 clinical practice guidelines by the infectious diseases society of america and the American thoracic society. *Clin. Infect. Dis.* 63, e61–e111. doi: 10.1093/cid/ciw353

Kullar, R., Chisari, E., Snyder, J., Cooper, C., Parvizi, J., and Sniffen, J. (2023). Next-generation sequencing supports targeted antibiotic treatment for culture negative orthopedic infections. *Clin. Infect. Dis.* 76, 359–364. doi: 10.1093/cid/ciac733

Li, T., Shi, J., Xu, F., and Xu, X. (2020). Clinical characteristics of pneumocystis pneumonia after parental renal transplantation. *Infect. Drug Resist.* 13, 81–88. doi: 10.2147/IDR

McAllister, D. A., Liu, L., Shi, T., Chu, Y., Reed, C., Burrows, J., et al. (2019). Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *Lancet Glob Health* 7, e47–e57. doi: 10.1016/S2214-109X(18)30408-X

Mertes, F., Elsharawy, A., Sauer, S., Van Helvoort, J. M., van der Zaag, P. J., Franke, A., et al. (2011). Targeted enrichment of genomic DNA regions for next-generation sequencing. *Brief Funct. Genomics* 10, 374–386. doi: 10.1093/bfgp/elr033

Morganti, S., Tarantino, P., Ferraro, E., D'amico, P., Viale, G., Trapani, D., et al. (2019). Complexity of genome sequencing and reporting: Next generation sequencing (NGS) technologies and implementation of precision medicine in real life. *Crit. Rev. Oncol. Hematol.* 133, 171–182. doi: 10.1016/j.critrevonc.2018.11.008

Onda, Y., Takahagi, K., Shimizu, M., Inoue, K., and Mochida, K. (2018). Multiplex PCR targeted amplicon sequencing (MTA-seq): simple, flexible, and versatile SNP genotyping by highly multiplexed PCR amplicon sequencing. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00201

Phetsuksiri, B., Rudeeaneksin, J., Srisungngam, S., Bunchoo, S., Klayut, W., Nakajima, C., et al. (2020). Comparison of loop-mediated isothermal amplification, microscopy, culture, and PCR for diagnosis of pulmonary tuberculosis. *Jpn J. Infect. Dis.* 73, 272–277. doi: 10.7883/yoken.JJID.2019.335

Pourahmadiyan, A., Alipour, P., Golchin, N., and Tabatabaiefar, M. A. (2022). Next-generation sequencing reveals a novel pathogenic variant in the ATM gene. *Int. J. Neurosci.* 132, 558–562. doi: 10.1080/00207454.2020.1826944

Qin, H., Peng, J., Liu, L., Wu, J., Pan, L., Huang, X., et al. (2021). A retrospective paired comparison between untargeted next generation sequencing and conventional microbiology tests with wisely chosen metagenomic sequencing positive criteria. *Front. Med. (Lausanne)* 8. doi: 10.3389/fmed.2021.686247

Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W., and Weinstock, G. (2017). Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch. Pathol. Lab. Med.* 141, 776–786. doi: 10.5858/arpa.2016-0539-RA

Shi, Y., Su, J., Chen, R., Wei, W., Yuan, Z., Chen, X., et al. (2022). The role of innate immunity in natural elite controllers of HIV-1 infection. *Front. Immunol.* 13. doi: 10.3389/fimmu.2022.780922

Souza Lopes, A. C., Rodrigues, J. F., Cabral, A. B., Da Silva, M. E., Leal, N. C., Da Silveira, V. M., et al. (2016). Occurrence and analysis of irp2 virulence gene in isolates of Klebsiella pneumoniae and Enterobacter spp. from microbiota and hospital and community-acquired infections. *Microb. Pathog.* 96, 15–19. doi: 10.1016/j.micpath.2016.04.018

Tan, Y., Chen, Z., Zeng, Z., Wu, S., Liu, J., Zou, S., et al. (2023). Microbiomes detected by bronchoalveolar lavage fluid metagenomic next-generation sequencing among HIV-infected and uninfected patients with pulmonary infection. *Microbiol. Spectr.* 11, e0000523. doi: 10.1128/spectrum.00005-23

Tao, Y., Yan, H., Liu, Y., Zhang, F., Luo, L., Zhou, Y., et al. (2022). Diagnostic performance of metagenomic next-generation sequencing in pediatric patients: A retrospective study in a large children's medical center. *Clin. Chem.* 68, 1031–1041. doi: 10.1093/clinchem/hvac067

vila-Ríos, S., Parkin, N., Swanstrom, R., Paredes, R., Shafer, R., Ji, H., et al. (2020). Next-generation sequencing for HIV drug resistance testing: laboratory, clinical, and implementation considerations. *Viruses* 12, 617. doi: 10.3390/v12060617

Wu, W., Han, X., Zhao, H., Sun, H., and Sun, Q. (2023). Application value of next-generation sequencing of bronchial alveolar lavage fluid in emergency patients with infection. *Cell Mol. Biol. (Noisy-le-grand)* 69, 45–49. doi: 10.14715/cmb/2023.69.8.7

Xie, Y., Dai, B., Zhou, X., Liu, H., Wu, W., Yu, F., et al. (2023). Diagnostic value of metagenomic next-generation sequencing for multi-pathogenic pneumonia in HIV-infected patients. *Infect. Drug Resist.* 16, 607–618. doi: 10.2147/IDR.S394265

Zhang, D., Yang, X., Wang, J., Xu, J., and Wang, M. (2022). Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in patients with lower respiratory tract infections. *J. Int. Med. Res.* 50, 3000605221089795. doi: 10.1177/03000605221089795

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Review
PUBLISHED 15 November 2024
DOI 10.3389/fcimb.2024.1458316



OPEN ACCESS

EDITED BY
Jiemin Zhou,
Vision Medicals Co, Ltd., China

REVIEWED BY
Benjamin M. Liu,
George Washington University, United States
Halis Akalin,
Bursa Uludağ University, Türkiye

*CORRESPONDENCE
Xin Zhang
zhangxin19731@hotmail.com

†The authors have contributed equally to this work

RECEIVED 02 July 2024
ACCEPTED 31 October 2024
PUBLISHED 15 November 2024

CITATION
Zhao Y, Zhang W and Zhang X (2024) Application of metagenomic next-generation sequencing in the diagnosis of infectious diseases. *Front. Cell. Infect. Microbiol.* 14:1458316. doi: 10.3389/fcimb.2024.1458316



# Application of metagenomic next-generation sequencing in the diagnosis of infectious diseases

Yu Zhao[1†], Wenhui Zhang[2†] and Xin Zhang[1*]

[1]Department of Urology Surgery, Beijing Chao-Yang Hospital Affiliated to Capital Medical University, Beijing, China, [2]Department of Hepatobiliary Surgery, Beijing Chao-Yang Hospital Affiliated to Capital Medical University, Beijing, China

Metagenomic next-generation sequencing (mNGS) is a transformative approach in the diagnosis of infectious diseases, utilizing unbiased high-throughput sequencing to directly detect and characterize microbial genomes from clinical samples. This review comprehensively outlines the fundamental principles, sequencing workflow, and platforms utilized in mNGS technology. The methodological backbone involves shotgun sequencing of total nucleic acids extracted from diverse sample types, enabling simultaneous detection of bacteria, viruses, fungi, and parasites without prior knowledge of the infectious agent. Key advantages of mNGS include its capability to identify rare, novel, or unculturable pathogens, providing a more comprehensive view of microbial communities compared to traditional culture-based methods. Despite these strengths, challenges such as data analysis complexity, high cost, and the need for optimized sample preparation protocols remain significant hurdles. The application of mNGS across various systemic infections highlights its clinical utility. Case studies discussed in this review illustrate its efficacy in diagnosing respiratory tract infections, bloodstream infections, central nervous system infections, gastrointestinal infections, and others. By rapidly identifying pathogens and their genomic characteristics, mNGS facilitates timely and targeted therapeutic interventions, thereby improving patient outcomes and infection control measures. Looking ahead, the future of mNGS in infectious disease diagnostics appears promising. Advances in bioinformatics tools and sequencing technologies are anticipated to streamline data analysis, enhance sensitivity and specificity, and reduce turnaround times. Integration with clinical decision support systems promises to further optimize mNGS utilization in routine clinical practice. In conclusion, mNGS represents a paradigm shift in the field of infectious disease diagnostics, offering unparalleled insights into microbial diversity and pathogenesis. While challenges persist, ongoing technological advancements hold immense potential to consolidate mNGS as a pivotal tool in the armamentarium of modern medicine, empowering clinicians with precise, rapid, and comprehensive pathogen detection capabilities.

KEYWORDS

**metagenomics next-generation sequencing (mNGS), infectious disease, pathogens, microorganisms, diagnosis**

# 1 Introduction

Infectious diseases, a collective term for diseases caused by pathogenic microorganisms, remain a major threat to global public health, and traditional pathogenetic diagnostic methods no longer meet the needs of clinical diagnosis and treatment (Smiley Evans et al., 2020). Rapid identification of pathogens from infected body fluid compartments is essential, as empirical antimicrobial therapy is often suboptimal, leading to increased morbidity and mortality (Singal et al., 2014; Glimåker et al., 2015; Lucas et al., 2016; Costales and Butler-Wu, 2018). In patients with severe infections, early detection of the causative microorganism is essential for early clinical interventions to be instituted and appropriate antimicrobials to be administered (Kumar et al., 2009; Liesenfeld et al., 2014; Barlam et al., 2016; Messacar et al., 2017). However, timely and accurate diagnosis remains extremely challenging for many patients. Many common pathogens are difficult or impossible to culture *in vitro*, deep infections often require invasive biopsies of infected tissues for diagnosis, and the use of broad-spectrum antibiotics prior to pathogen identification often confounds the specific diagnosis, leading to more effective and less toxic antimicrobial therapy (Fenollar and Raoult, 2007; Fishman, 2007; Tomblyn et al., 2009; Mancini et al., 2010; Paul et al., 2010; Kumar, 2014; Barlam et al., 2016). Previous studies have shown that a significant proportion of unknown pathogens are present in severe pneumonia, bacteremia, eye infections and central nervous system (CNS) infections (Li et al., 2018; Blauwkamp et al., 2019; Wilson et al., 2019). Metagenomic next-generation sequencing (mNGS) is useful when conventional microbiological tests fail to identify infection in suspected cases. It is capable of simultaneously detecting virtually all known pathogens from clinical samples (Simner et al., 2018; Chiu and Miller, 2019; Gu et al., 2019). Compared with traditional pathogenic diagnostic methods (culture, mass spectrometry, immune-associated antigen-antibody detection and nucleic acid detection technology, etc.), mNGS is a non-targeted, broad-spectrum pathogenicity screening technology, which has been developed rapidly in recent years and has been widely used in the precise diagnosis of infectious disease pathogenic microorganisms, especially in the diagnosis of infections caused by critical, difficult, rare, and new-emerging pathogens. Therefore, this paper introduces the basic principles and sequencing platform of mNGS, evaluates its strengths and weaknesses, summarizes its applications in various organ system infections, and finally looks forward to the future development.

# 2 Overview of mNGS

mNGS is a next-generation macro-genome-based sequencing technology that enables rapid sequencing of nucleic acids in samples (human and pathogenic microorganisms) and compare them to human genome sequences and pathogenic microbial genome sequences to learn the species and proportions of microorganisms in the sample. It is a technique to obtain nucleic acid sequences from samples (human and pathogenic microorganisms) by rapid sequencing on a second-generation sequencing platform, and compare them with human genome sequence libraries and pathogenic microorganism genome sequence libraries to know the types and proportions of microorganisms in the samples (Yi et al., 2024). It provided an ideal approach for genomic analysis of all microorganisms in a sample, not just those suitable for culture (Wooley et al., 2010).

## 2.1 Technical principle

The process of high-throughput sequencing of pathogens consists of two main parts (Gu et al., 2019): the wet lab part (laboratory testing) and the dry lab part (bioinformatic analysis). The wet lab part includes sample collection, nucleic acid extraction, library construction and high-throughput sequencing. The dry experimental part includes quality control of data, removal of human sequences, sequence comparison of microbial species sequences, and analysis of drug resistance or virulence genes (Figure 1).

mNGS is a NGS assay allowing for comprehensive detection of all genes in all organisms in a given sample (Liu et al., 2024). It can be used for bacterial, fungal, parasitic, and various viral infections and is primarily a sequencing comparison process for nucleic acids extracted from infected samples. Because of the different processes for targeting deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) in nucleic acids, an assessment should be made as to which method of testing to use before finalizing the test. DNA testing is recommended when infection by pathogens whose nucleic acids are DNA, such as bacteria, fungi, DNA viruses, parasites, etc., is suspected; RNA testing is recommended if RNA viral infection is suspected; and co-testing of DNA and RNA is recommended if it is not possible to determine which type of viral infection is involved. In addition, the diagnosis of infectious diseases requires that specific samples must first be collected from the site of primary infection before the samples can be preprocessed (Gu et al., 2019). For example, bronchoalveolar lavage fluid (BALF) and sputum are typically recommended for lung infections, while cerebrospinal fluid (CSF) is recommended for CNS infections. While library construction, sequencing, and bioinformatics analysis are the same for different samples, pretreatment and nucleic acid extraction vary depending on the sample source.

## 2.2 Sequencing platforms

The most commonly used mNGS sequencing platform is the Illumina platform, which is based on the core principle of sequencing by synthesis (SBS), which consists of four main steps: DNA library construction, BALF fluid Flowcell adsorption, bridge polymerase chain reaction (PCR) amplification, and SBS (Bentley et al., 2008). Illumina's sequencing principle of adding only one deoxy-ribonucleoside triphosphate (dNTP) at a time makes it possible to solve the problem of inaccurate sequencing due to the polymerization of identical bases (e.g., when the DNA strand contains repetitive sequences such as AAAAAAA). The

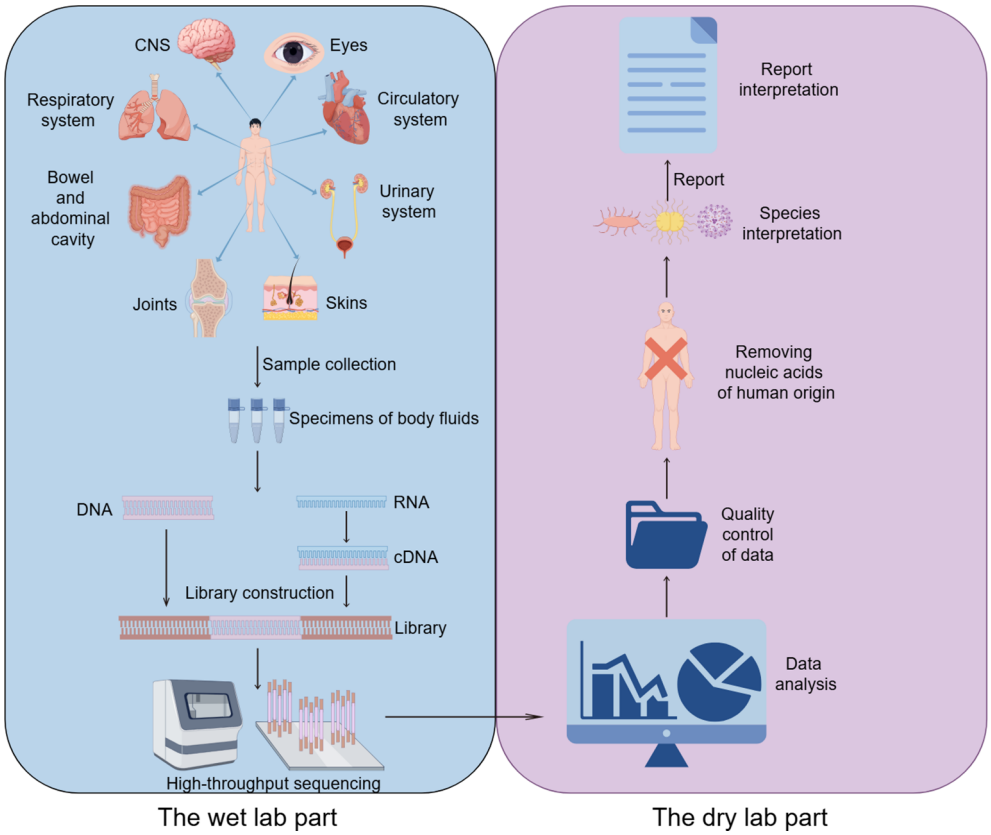



FIGURE 1
mNGS workflow in clinical application. CNS, Central nervous system.

sequencing principle of adding only one dNTP at a time makes Illumina a good solution to the problem of inaccurate sequencing due to the polymerization of the same base (e.g., when the DNA strand contains repetitive sequences such as AAAAAA, most sequencing platforms are prone to errors of over-reading or under-reading one base). Currently, Illumina sequencing has an error rate as low as 0.1% (e.g., HiSeq series), with base substitutions being the main source of error.

The Thermo Fisher Ion Torrent next-generation sequencing (NGS) platform is based on the principle of hydrogen ion semiconductor sequencing for non-destructive high-throughput sequencing of nucleic acid fragments. Using natural bases without any artificial modification during the synthetic extension of the nucleic acid chain, the ATGC base biosignal of the nucleic acid fragment to be tested is quickly and accurately converted into digital information by semiconductor technology. Without the need for complex, expensive and environmentally demanding optical detection and scanning imaging systems, and without the use of artificially modified bases, the Ion Torrent platform is more cost-effective, smaller, and faster than other sequencing technologies, completing the sequencing of a single 200 bp sequence in 2 to 2.5 hours.

In 2016, Beijing genomics institute (BGI) announced the BGISEQ-500 sequencing platform, which has a general NGS workflow and stepwise sequencing program similar to that of the Illumina series; however, the two templates are distinctly different (Goodwin et al., 2016). The follow-on DNA nanospheres technology in the BGISEQ-500 platform, which is specifically used for library preparation, is different from the library construction protocol used in the Illumina series (Drmanac et al., 2010). the BGISEQ-500 utilizes both single-end (SE) and paired-end (PE) modes, comparable to the latest Illumina model, the HiSeq4000. the BGISEQ-500 has published relatively high throughput data, and may be suitable for high throughput transcriptome studies.

Nanopore sequencing permits the inclusion of bacteria and fungi with marker genes of different sizes in the same sequencing library by detecting the electrical signals of DNA/RNA as it passes through nanopore proteins (Han et al., 2024). Although nanopore metagenome sequencing based on real-time analytical pathways reduces the detection time to less than 6 hours, it still faces challenges such as insufficient sensitivity, high sequencing errors, and elevated detection costs.

In 2021, the Association of Biomolecular Resource Facilities (ARBF) led an ABRF NGS Phase II study published in Nature Biotechnology, based on multiple sequencing platforms from Illumina, Pacific Biosciences, Thermo Fisher Scientific, Oxford Nanopore Technologies, and Genapsys, The team sequenced the same human genome family, three individual strains, and a mixture of ten bacterial metagenomes in multiple laboratories based on multiple sequencing platforms from Illumina, Pacific Biosciences, Thermo Fisher Scientific, BGI, Oxford Nanopore Technologies, and

Genapsys. The data from each platform were compared in a comprehensive and systematic way to analyze the performance differences and sequencing quality of each sequencing platform. The data show that among the short-read-long sequencing platforms, Illumina's HiSeq 4000 and HiSeq X10 platforms provide the most consistent and highest genome coverage, while BGI's BGISEQ-500 and MGISEQ-2000 platforms provide the lowest sequencing error rate. Among the long read-length sequencing platforms, the PacBio CCS has the highest reference-based mapping rate and the lowest non-mapping rate. Both the PacBio CCS and Oxford Nanopore's PromethION, MinION platforms show the best sequence localization performance in both repeat sequence-rich regions and across homopolymer assays. The NovaSeq 6000 uses the 2×250 bp read chemistry is the most powerful instrument for capturing known insertion and deletion (INDEL) events.

# 3 Advantages and limitations of mNGS in clinical applications

## 3.1 Advantages of mNGS in clinical applications

mNGS is increasingly recognized for its groundbreaking capabilities in the field of infectious disease diagnostics. One of the principal advantages is its comprehensive and unbiased approach; it does not require prior hypotheses about which pathogens might be present. This allows for the simultaneous detection and identification of a wide array of pathogens—including bacteria, viruses, fungi, and parasites——from a single sample (Rodino and Simner, 2024). Firstly, it can identify nearly any pathogen present in a sample without needing specific probes or primers for each one. This is particularly beneficial for detecting rare pathogens, those presenting atypically, or those for which no targeted diagnostics exist. Secondly, it is especially useful for diagnosing infections in immunocompromised patients, where the range of possible infecting organisms is broader and often includes lower common pathogens. Thirdly, it can detect genes responsible for resistance to antimicrobials, providing crucial information for guiding treatment decisions. Besides, Ithas the advantage of timeliness compared to culture methods, which typically take 2 - 3 days to obtain results, and even more than a week for fussy bacteria, such as mycobacterium tuberculosis (MTB) (Mu et al., 2021). The average turnaround time for conventional mNGS is 48 hours (Han et al., 2019; Gu et al., 2021). One study reported a turnaround time of only 6 hours for the detection of pathogenic microorganisms using the mNGS technology on a nanopore platform (Gu et al., 2021; Mu et al., 2021). Finally, It is also highly effective in situations where patients have already been treated with antibiotics, which can inhibit pathogen growth in cultures and lead to negative results despite ongoing infection (Zhang et al., 2019). In summary, mNGS can provide comprehensive and rapid results that can guide clinicians to more precise and effective treatments.

## 3.2 Limitations of mNGS in clinical applications

First, mNGS is unable to determine whether the sequences it detects are from live or dead pathogens, so it still does not solve the perennial problem of identifying colonizing and pathogenic pathogens. Thus, the detection of DNA only indicates what organisms are present, not whether they are biologically active, and even blood specimens is still unable to differentiate pathogenic bacteria from transient bacteremia and from microbial nucleic acid fragments contained in leukocytes. Perhaps the detection of RNA would help in this regard, as the presence of RNA could indicate that the organism is transcriptionally active (Liu et al., 2024). As a result, mNGS cannot serve as first line diagnostic assay due to its low sensitivity (Liu, 2024; Liu et al., 2024).

Second, it is limited to roughly determining the species of pathogenic microorganism and estimating the approximate proportion of microorganisms (quantifying pathogen reads as a percentage of the total number of sequences reads). If the pathogenic microorganism is a particularly small proportion of the genus, it is highly unlikely to produce a negative result. Therefore, a negative result may simply reflect a sample with a high non-microbial nucleic acid component (denominator) and/or a low microbial nucleic acid component (molecule), rather than a lack of pathogens (Liu et al., 2024).

Third, some low levels of intracellular bacteria, e.g., MTB, Legionella, Brucella, and fungi with thick cell walls will be detected at lower rates (the latter require special treatment to disrupt the cell wall and expose the DNA). Therefore, for all types of pathogens, nucleic acid recoveries may not be equal under the same DNA extraction technique. Therefore, different extraction methods should be used to detect specimens of different target microorganisms.

Fourth, there are no standardized procedures and standards to avoid contamination of nucleic acids in the steps from specimen collection to processing and the environment, so that the results of different laboratories tend to be similar.

Fifth, the relatively short reading sequence (300 bp) of mNGS makes it difficult to obtain the full-length sequence of drug resistance genes, and it is not possible to correlate the drug resistance genes with the corresponding microbial species. The length of three-generation sequencing can reach 1200bp, which can potentially cover the full length of drug-resistant genes, and is a good help for the determination of drug-resistant genes.

Sixth, the positive controls should cover the range of microorganisms likely to be encountered in blood samples, such as: enveloped and non-enveloped viruses, RNA and DNA genomes, Gram-positive and negative bacteria, mycobacteria and parasites. However, we do not know in advance what pathogens are present in the specimen to be tested. Negative controls are equally problematic because the water and sample matrix do not adequately reflect the background present in normal healthy blood (other) specimens.

Last but not least, interpreting the results generated by the sequencing lab is one of the biggest headaches in mNGS clinical practice today. Some institutions have implemented precision

medicine teams-composed of microbiology, computational biology, infectious disease, and other clinicians-to discuss results and provide interpretation of results prior to reporting. This may be the best option at this time. Furthermore, the expensive pricelimits its widespread clinical use.

## 3.3 Comparison between mNGS and targeted NGS

tNGS is a targeted high-throughput sequencing technology for specific genes or genomes. tNGS, unlike mNGS, performs high-throughput sequencing of only specific gene sequences, thereby increasing detection sensitivity while eliminating interference from host nucleic acids (Yi et al., 2024). tNGS is primarily designed for the detection of dozens to hundreds of known pathogenic microorganisms and their drug-resistant genes in a sample. Depending on the enrichment strategy, there are two main technical routes for tNGS enrichment: one is PCR amplicon enrichment, i.e., enrichment of small viral genomes by PCR amplification of viral genomes with hundreds to thousands of base pairs using primers complementary to known nucleotide sequences before NGS sequencing; Another type of enrichment is hybridization-targeted probe enrichment, i.e., small RNA/DNA probes that are usually first designed to be complementary to the pathogen reference sequence (Pham et al., 2023; Chen et al., 2024; Yi et al., 2024). Unlike methods based on specific PCR amplicons, probe-targeted enrichment allows the entire genome to be covered by overlapping probes that are used in a hybridization reaction to capture complementary DNA sequences that bind to their sequences (Chen et al., 2024). Thus, tNGS combines the advantages of PCR and NGS.

A study conducted in 2003 on the molecular diagnosis of infective endocarditis by PCR amplification and direct sequencing of valvular tissue DNA can be considered a prototype of tNGS. Its results showed significant concordance between tNGS results and histopathologic evaluation, with concordance rates as high as 93.1% (27/29) for positive samples and 92.9% (13/14) for negative samples (Gauduchon et al., 2003). Several subsequent studies have validated this finding (Marín et al., 2007; Vondracek et al., 2011; Maneg et al., 2016). However, relying solely on tNGS for clinical testing is not advisable because of its occasionally limited predictive power for negative samples (Maneg et al., 2016). In the challenge posed by the COVID-19 pandemic, tNGS has been used for infectious disease surveillance and genotyping (Cheng et al., 2023; Ramos et al., 2023). In addition, Chao et al. reported the use of tNGS for pathogen identification in patients with acute lower respiratory tract infections. The positive rate of tNGS was as high as 95.6% based on the gold standard sputum culture (Chao et al., 2020). Recent studies have also reported the use of tNGS in the identification of rare pathogens, including Legionella pneumophila, Chlamydia psittaci, Whipple's bacillus, Aspergillus fumigatus, and Cryptococcus neoformans (Du and Chen, 2023; Li et al., 2023; Ren et al., 2023; Zhang et al., 2023). To date, the value of tNGS has been demonstrated for clinical applications in the areas of bloodstream infections, central nervous system infections, and tuberculosis (Cabibbe et al., 2020; Deng et al., 2020; Mensah et al., 2020; Jouet et al., 2021; Mesfin et al., 2021; Chen et al., 2022; Kunasol et al., 2022; Sibandze et al., 2022; Yang et al., 2022; Jiang et al., 2023). tNGS cannot be run on its own, but is used in conjunction with conventional assays. This approach may contribute to an effective and accurate clinical diagnosis. By combining the ubiquity of conventional testing with the high specificity and sensitivity of tNGS, clinicians are better able to make a more accurate diagnosis. This integrated diagnostic strategy may improve patient prognosis through timely and appropriate therapeutic interventions. In conclusion, tNGS may have a role in the diagnosis of infectious diseases. By addressing the limitations of current assays, tNGS could provide a more refined, accurate and comprehensive approach to pathogen detection (Chen et al., 2024). In summary, tNGS may bridge the diagnostic gap between traditional assays and mNGS.

There are significant differences between mNGS and tNGS (Table 1). Firstly, tNGS sequences only specific regions or specific genes, usually for known genes, pathogens, or specific genomic loci, i.e., this method requires the target sequence to be set before the experiment begins. In contrast, mNGS sequencing is wide-ranging and can sequence all DNA/RNA fragments in a sample without bias. This means it can recognize all genomic information in the

TABLE 1 The comparison between mNGS and tNGS.

| | mNGS | tNGS |
|---|---|---|
| **Methodology** | 1 Direct extraction of DNA/RNA<br>2 Without predefining specific pathogens<br>3 High-throughput sequencing | 1 Targeted enrichment and ultra-multiplex PCR<br>2 With predefining specific pathogens<br>3 High-throughput sequencing |
| **Sequencing Scope** | Wide | Narrow |
| **Sample Size** | Small | Small |
| **Data Volume** | Large | Small |
| **Analysis Complexity** | Sophisticated | Simpler |
| **Sensitivity** | Normal | High |
| **Specificity** | Normal | High |
| **Cost and Time** | High and long | Low and short |
| **Flexibility** | Flexible | Inflexible |
| **Clinical Application Time** | Short | Short |
| **Application** | Unknown pathogens:<br>1 Analysis of complex microbial communities<br>2 Detection of infectious diseases<br>3 Screening for drug resistance genes | Known pathogens:<br>1 Genetic mutation<br>2 Genetic testing for genetic diseases<br>3 Cancer genetic testing |
| **Targeting pathogens** | Bacteria, viruses, mycoplasma, etc. | Bacteria, viruses, fungi, mycoplasma, etc. |

mNGS, Metagenomic next-generation sequencing; tNGS, Targeted metagenomic next-generation sequencing; DNA, Deoxyribonucleic acid; RNA, Ribonucleic acid; PCR, Polymerase chain reaction.

sample, including pathogens, human genome, microbiota, etc. Therefore, tNGS possesses greater sensitivity and specificity but poor flexibility. Secondly, tNGS is usually used for the detection of known targets, such as gene mutation, genetic disease related gene detection, cancer gene detection, etc. It is very effective for rapid and precise detection of known pathogens, and is suitable for diagnosing specific diseases or for personalized medicine. However, mNGS is suitable for identification of unknown pathogens, analysis of complex microbial communities, detection of infectious diseases, and screening of drug resistance genes. Due to its extensive sequencing capabilities, mNGS can be used clinically to discover novel pathogens or complex sources of infection. Thirdly, the results of tNGS are more direct and precise due to clear targeting, relatively small amount of data, and simple analysis process. Its data processing is faster and suitable for rapid response in clinical diagnosis. On the other hand, mNGS is more complicated to analyze due to the large amount of data generated and the inclusion of a large amount of irrelevant information (e.g., host genes, environmental strays, etc.), which requires powerful bioinformatics tools to filter and interpret the data, and the process of data processing and analysis is time-consuming and the results may have uncertainties. Finally, tNGS is less costly and faster to analyze, making it suitable for diagnostics or research with specific targets. In contrast, mNGS, due to its high coverage, is more costly and relatively time-consuming, and is suitable for complex, unresolved infection cases or studies that require extensive exploration. In summary, tNGS is a targeted, low-cost sequencing method for rapid and accurate detection of known targets, while mNGS is a broad, unbiased sequencing method for identification of unknown pathogens and analysis of complex environments.

In addition to tNGS, there are a variety of microbiological testing methods available, each with its own advantages and disadvantages, as shown in Table 2. In conclusion, mNGS is not currently a replacement for current conventional microbiological testing methods, but should be viewed as a complement to these traditional methods.

TABLE 2 Comparison of testing methods for diagnosing infectious diseases.

| Diagnostic test | Advantages | Disadvantages |
|---|---|---|
| **Serology** | 1 Low cost<br>2 Suitable for acute infections | 1 False-negatives<br>2 False-positives |
| **Culture** | 1 Able to accommodate large sample volumes<br>2 Low cost<br>3 Wide range of applications | 1 Sensitivity is limited by antibiotics and antifungal drugs<br>2 Sensitivity limited by picky microorganisms<br>3 Limited use in viral assays<br>4 Long time to produce results, especially in antacid and fungal cultures |
| **Microscopy and staining (eg, Gram stain, auramine–rhodamine, calcofluor-white)** | 1 Rapid<br>2 Low cost | 1 Low sensitivity<br>2 Low specificity |
| **Matrix-assisted laser desorption/ionization time of-flight mass spectrometry** | 1 High specificity<br>2 Rapid after culture | 1 Requires culture-positive isolate |
| **Direct PCR** | 1 Simple<br>2 Rapid<br>3 Low cost<br>4 Potential for quantitative PCR | 1 Dependent on assumptions<br>2 Primers are not always effective<br>3 Limited to a small part of the genome |
| **Multiplex PCR** | 1 Fast<br>2 Detect a wide range of microorganisms | 1 Low specificity<br>2 False positives |
| **Targeted universal multiplex PCR (eg, 16S, ITS) for Sanger sequencing** | 1 Ability to distinguish multiple species within a pathogen type | 1 Primers are not always effective<br>2 Limited to a small part of the genome |
| **Targeted universal multiplex PCR (eg, 16S, ITS) for NGS** | 1Ability to distinguish multiple species within a pathogen type<br>2 Multiplexing capability<br>3 Potential for quantitation | 1 Primers are not always effective<br>2 Expensive<br>3 time consuming<br>4 Often requires more than one amplification<br>5 Limited to a small part of the genome |
| **Amplicon sequencing** | 1 Bacteria, fungi<br>2 Low expenditure<br>3 Low biomass required, no host contamination<br>4 Low volume of data generated, easy to analyze<br>5 Reliable database | 1 Not applicable to viruses<br>2 Low species resolution<br>3 Functional genes not available<br>4 Low biomass may lead to false negatives<br>5 Results of community diversity analyses varied across variable regions |

*(Continued)*

TABLE 2 Continued

| Diagnostic test | Advantages | Disadvantages |
|---|---|---|
| **Targeted NGS** | 1 Suitable for initial screening of hospitalized patients<br>2 Unaffected by human genome and background flora<br>3 High ability to detect engulfed pathogens<br>4 High ability to detect drug resistance and virulence<br>5 Ability to add new targets according to clinical needs | 1 Short clinical application time<br>2 Inability to recognize new pathogens<br>3 Incomplete database |
| **mNGS** | 1 Identify viruses, fungi, archaea and protozoa<br>2 Timeliness<br>3 Without needing specific probes or primers for each one<br>4 It is especially useful for diagnosing infections in immunocompromised patients<br>5 It can detect genes responsible for resistance to antimicrobials<br>6 It applies when you are already receiving antibiotics | 1 High DNA quality requirement<br>2 Host contamination<br>3 Not easy to assembly and complex analysis process<br>4 False positive results<br>5 Expensive |
| **2bRAD-M** | 1 High technical reproducibility<br>2 High species resolution<br>3 It can be used for low biomass, heavily degraded samples, high host contamination samples<br>4 Simultaneous detection of bacteria, fungi and archaea<br>5 Host snp analysis, human genetic analysis<br>6 Microbial diversity analysis and host SNP analysis can be combined for GWAS analysis | 1 Expensive<br>2 It cannot detect small and short genes, such as viruses<br>3 It can do GWAS analysis, but the number of snp is low<br>Cannot recognize new pathogens |
| **MobiMicrobe** | 1 Reliable genomes<br>2 High quality genome assembly<br>3 Precise genomic analysis at the strain level to discover new uncultured strains<br>4 Mining inter-strain relationships to analyze horizontal gene transfer<br>5 Single-cell level of host-phage binding | 1 Low genome coverage of Gram-negative bacteria |

PCR, Polymerase chain reaction; ITS, Internal transcribed spacer sequencing; NGS, Next-generation sequencing; mNGS, Metagenomic next-generation sequencing; DNA, Deoxyribonucleic acid; SNP, Single nucleotide polymorphism; GWAS, Genome-wide association study.

# 4 Application of mNGS in infections in different organ systems

## 4.1 Bloodstream Infections

The composition of causative organisms varies from sepsis to sepsis; in recent years there has been an increase in the number of cases of gram-negative, anaerobic and fungal sepsis, but gram-positive organisms remain the most common (Zhou et al., 2016). There is also a subset of culture-negative sepsis patients for whom the causative organism remains undetermined. In patients with severe sepsis, failure to diagnose the pathogen in a timely manner can lead to receiving inappropriate and mismatched antimicrobials, which in turn can lead to high mortality rates (Kumar et al., 2009; Gupta et al., 2016). Traditional diagnostic methodologies for septic pathogens encompass the cultivation and isolation of microorganisms, serological detection of pathogen-specific antibodies, antigen identification, and molecular characterization through nucleic acid analysis, predominantly via PCR. Whereas conventional molecular techniques often employ specific primers or probes targeting a restricted array of pathogens, mNGS enable comprehensive characterization of all DNA or RNA within a sample. This approach facilitates a holistic analysis of the entire microbiome and the human host's genome or transcriptome in clinical specimens (Wensel et al., 2022).

Multiple studies and case reports indicate that genomic DNA or RNA fragments from pathogens involved in infections—whether circulating or non-circulating—can be detected as cell free DNA (cfDNA) or cell-free RNA (cfRNA) in purified plasma (De Vlaminck et al., 2015; Long et al., 2016; Gosiewski et al., 2017; Pan et al., 2017). These findings demonstrate the potential of mNGS for rapid and accurate identification of the pathogens responsible for sepsis (Abril et al., 2016; Hong et al., 2018). Moreover, it can provide detailed information on the abundance of pathogens and their genetic relationships. This technology, therefore, offers significant advantages in diagnosing and understanding the dynamics of infections associated with sepsis (Hong et al., 2018). One study showed that 76% of patients with positive routine blood cultures tested positive for cfDNA mNGS, and only 4% of cfDNA mNGS did not match routine bacterial cultures, and pathogens were accurately determined by cfDNA mNGS combined with analysis of the patient's clinical presentation in 32.8% of patients with routine blood culture-negative suspected bacteremia (Zhang et al., 2022). This suggests that mNGS can diagnose pathogen infections more accurately than blood cultures. Another study showed that the diagnostic sensitivity was significantly higher than that of blood cultures, providing additional useful information for the development of patient treatment plans (Long et al., 2016). In summary, the advantages in the diagnosis and therapeutic guidance of bloodstream infections are undeniable, as it is effective in reducing the time required for pathogen identification regardless of the microbial type and is less affected by antibiotic administration (Abril et al., 2016). In addition, this method is highly desired for patients infected with rare fungi, mycobacteria and parasites (Miao et al., 2018). In addition, mNGS detects viral

infections or mixed infections and guides physicians in the correct and targeted use of antibiotics for septic patients (Hu et al., 2018; Wilson et al., 2018; Xing et al., 2019).

## 4.2 CNS infections

A variety of pathogenic microorganisms can infect the central nervous system, often manifesting as meningitis, encephalitis and abscesses, which may be life-threatening. However, routine microbiological testing is often insufficient to detect all neuroinvasive pathogens, especially rare ones. In addition, obtaining relevant samples for detection of pathogenic pathogens requires invasive procedures such as lumbar puncture or brain biopsy, which are limited by the availability and volume of CSF or brain tissue. As a result, the etiology of CNS infections is often unspecified, which occurs in up to 50 per cent of encephalitis (Glaser et al., 2003; Granerod et al., 2010). Numerous studies have reported the use of mNGS in CSF and brain tissue to detect viruses, bacteria, fungi, and parasites (Xing et al., 2020). In addition, it has proven valuable in diagnosing subacute or chronic meningitis (Wilson et al., 2018). Elevated CSF leukocyte and protein levels, as well as a decreased percentage of glucose in the CSF may be associated with an increase in mNGS detection of CNS infection (Zhang et al., 2020b). A systematic review recommended NGS as a first-line diagnostic test for chronic and recurrent infections and a second-line technique for cases of acute encephalitis (Brown et al., 2018). The detection rate is higher for diagnosing the CNS than traditional pathogen diagnostic methods (Xing et al., 2020; Zhang et al., 2020b). In addition, it may help to rule out active infection in patients with suspected autoimmune encephalitis, providing favourable information for clinicians' judgement and reducing concerns about missed microbial infections (Wilson et al., 2019; Xing et al., 2020). mNGS results are less affected by the use of antibiotics prior to the collection of CNS samples, and therefore the technique has advantages over other methods for CNS in which antibiotics have been administered (Zhang et al., 2019; Zhang et al., 2020b). Studies have shown that it has high sensitivity, specificity and positive predictive value (PPV) for the diagnosis of CSF tuberculous meningitis (Wang et al., 2019). Indeed, the sensitivity of mNGS was significantly higher than that of culture alone, and the combination of mNGS with conventional methods significantly increased the detection rate. mNGS also has value in identifying complex and rare pathogens present in culture-negative and unconfirmed cases. mNGS has been shown to be useful in detecting the presence of microorganisms such as Listeria monocytogenes, the species that cause brucellosis, *Naegleria fowleri*, the parasites that cause neurocysticercosis, and *Vibrio traumaticus* (Mongkolrattanothai et al., 2017; Fan et al., 2018b; Fan et al., 2018a; Liu et al., 2018; Wang et al., 2018; He et al., 2019). In CNS toxoplasmosis, it can help in the diagnosis when Toxoplasma IgG is negative, CSF PCR is negative, imaging is atypical, or there is a lack of response to anti-Toxoplasma treatment (Hu et al., 2018). In addition, it can be used to dynamically monitor disease progression by analysing semi-quantitative values (Zhang et al., 2020b).

However, mNGS has some shortcomings in diagnosing CNS. While it frequently detects DNA viruses, particularly herpesviruses, its ability to enhance the diagnosis of viral encephalitis and meningitis has not shown significant improvement. One possible reason for this limitation is the underrepresentation of RNA detection methods in current mNGS protocols (Guan et al., 2016; Tyler, 2018; Xia et al., 2019; Fang et al., 2020). Since RNA-based mNGS has not been widely implemented, this restricts the detection of RNA viruses, which are often significant causative agents of these conditions. The lack of comprehensive viral RNA detection may thus hinder the overall effectiveness in diagnosing these serious infections (Xing et al., 2020). Besides, the detection rate of pathogenic microorganisms by mNGS showed a decreasing trend with the prolongation of treatment time (Ai et al., 2018).

## 4.3 Respiratory infections

Upper respiratory tract infections along with lower respiratory tract infections are one of the common diseases and lead to significant mortality (Milucky et al., 2020). Upper respiratory tract infections along with lower respiratory tract infections are one of the common diseases and lead to significant mortality. Undoubtedly, identification and characterization of pathogens is crucial for precise treatment of patients and improved prognosis. However, in the clinical setting, pathogens are often not rapidly identified, which leads physicians to use antibiotics only empirically, which in turn leads to frequent and inappropriate use of antibiotics, and which limits the sensitivity and reliability of culture-based surveillance. mNGS can detect and characterize a wide range of pathogens with relative rapidity and precision, which can contribute to the timely and accurate treatment of lung infections, especially for critically ill patients and patients with mixed infections (Li et al., 2018; van Rijn et al., 2019; Wang et al., 2019; Xie et al., 2019; Huang et al., 2020). In addition, compared with traditional sputum culture or BALF culture, mNGS was able to identify MTB, *nontuberculosis mycobacteria* (NTM), Nocardia, and various Actinomycetes (Miao et al., 2018). For the diagnosis of invasive fungi in the lungs, it is also helpful (Li et al., 2018; Wang et al., 2019). The sensitivity in the diagnosis of mixed lung infections and severe unresponsive pneumonia was evident (Li et al., 2018; Wang et al., 2020). In the analysis of immunocompromised patients, mNGS was even detected with 100% accuracy (Li et al., 2020). Interestingly, it has been reported to be more specific than BALF in transbronchial lung biopsy (TBLB) tissues; however, the BALF assay has been shown to have higher sensitivity in the diagnosis of peripheral lung infectious lesions (Liu et al., 2019). A study showed that for bacteria and fungi, the positive detection rate of mNGS was significantly higher than that of the culture method (91.94% vs 51.61%, $P < 0.001$), especially for polymicrobial infections (70.97% vs 12.90%, $P < 0.001$). Compared with the culture method, the diagnostic sensitivity of mNGS was 100%, the specificity was 16.67%, and the PPV and negative predictive value (NPV) were 56.14% and 100%, respectively (Wang et al., 2024). In addition, it more often detects NTM than MTB, Aspergillus or Cryptococcus in BALF (Miao et al., 2018). For human

immunodeficiency virus-infected (HIV-infected) patients with suspected lung infections, it can quickly and accurately identify the pathogens that cause lung infections (Hou et al., 2024). Patients with corona virus disease 2019 (COVID-19) may be at increased risk of developing fungal infections as well as concurrent bacterial or viral infections, and mNGS could be a powerful tool for identifying these infections (Huang et al., 2023). Another study showed that BALF mNGS is a valuable tool for differentiating between colonization and infection of *Aspergillu* (Jiang et al., 2024).

## 4.4 Digestive system infection

Although the study of the gut microbiome is a very popular and relevant topic, the use of mNGS technology to diagnose related infectious diseases such as diarrhea has been rarely reported. A case report reporting a definitive diagnosis of *Enterocytozoon bieneusi microsporidiosis* using mNGS suggests that in patients with recurrent unexplained diarrhea with wasting associated with hematological malignancies we should consider the possibility of infection by atypical pathogens. mNGS can help to rule out malignancy and diagnose infection (Zhou et al., 2022). A case report of a definitive diagnosis of *Encephalitozoon hellem* (E. hellem) infection by mNGS in a 9-year-old boy following hematopoietic stem-cell transplantation (HSCT), which led to his timely treatment with albendazole and eventual recovery from reduced immunosuppressive therapy (Shang et al., 2024). A case of Disseminated *Talaromyces marneffei* infection after renal transplantation was also diagnosed and effectively treated by mNGS (Xu et al., 2023). The use of mNGS to assist in the diagnosis and treatment of severe Cryptosporidium infection was reported (Liu et al., 2023; Shan et al., 2023). A patient with diarrhoea was identified as having the causative pathogens MTB complex and *Leptospira* spp. by mNGS of CSF, urine, plasma and sputum clinical samples (Shi et al., 2022). In addition, some cases of *Disseminated histoplasmosis* infection, *Chlamydia psittaci* infection, *Bunyaviridae virus* infection and T. *marneffei* infection diagnosed by mNGS have been reported (Chen et al., 2020; Wang et al., 2022; Zhan et al., 2022; Zheng et al., 2023; Wang et al., 2024). Even, one case reported misdiagnosed tuberculosis being corrected as *Nocardia farcinica* infection by mNGS (Pan et al., 2021). These suggest that mNGS can facilitate diagnosis and timely therapeutic decisions.

## 4.5 Urinary tract infection

UTI is one of the most common infections, affecting 150 million people worldwide each year. The most common cause of UTIs is pathogens (mainly E. coli) in the feces that rise up the urinary tract (Flores-Mireles et al., 2015). Urine culture is the gold standard for the diagnosis of UTI. However, cultures need a long time with low detection rates and limited diagnostic accuracy (Schmiemann et al., 2010). Although PCR methods allow for the rapid detection of pathogens, including non-culturable microorganisms, directly from clinical samples as compared to urine cultures, PCR methods are limited to amplification of predetermined target species and do not meet the diagnostic needs for microorganisms that cannot be predetermined in advance (Smith and Osborn, 2009). The ability of mNGS to detect microorganisms that cause UTIs quickly, accurately, and without prior predetermination allows for the identification of rare, complicated urinary tract infections and is virtually unaffected by prior antibiotic exposure (Miao et al., 2018). A study showed that based on the gold standard of routine culture, mNGS had a sensitivity of 81.4%, a specificity of 92.3%, a PPV of 96.6%, a NPV of 64.9%, and an overall accuracy of 84.4%; while when evaluated based on a composite standard, the sensitivity and the specificity increased to 89.9% and 100%, respectively, and the PPV was 100% and accuracy increased to 92.4% (Wang et al., 2023). Another study shows mNGS-based targeted antibiotic therapy significantly improves urinalysis and urinary symptoms in patients (Jia et al., 2023). Besides, the role of it in the pathogen diagnosis of urinary tract infections in patients after cutaneous ureterostomy, recurrent urinary tract infections in renal transplant recipients, and scrub typhus has been reported (Liu et al., 2021; Duan et al., 2022; Huang et al., 2023). Consequently, mNGS is a technique that offers significant advantages over culture, especially in the case of mixed infections and urinary tract infections that are difficult to diagnose and treat. It helps improve pathogen detection, guides change in treatment strategies, and is a useful complement to urine culture (Jia et al., 2023).

## 4.6 Bone and joint infection

BJI is very serious infection, especially periprosthetic joint infection (PJI), and may even be life-threatening (Malizos, 2017). PJI occurs most often after arthroplasty and multiple revision surgeries. The key to the treatment of PJI is the definitive diagnosis of the causative pathogen. Microbiological cultures are still the primary method for diagnosing PJI, however negative pathogen cultures are a great challenge for clinicians to make treatment decisions (Parvizi et al., 2014; Ahmed and Haddad, 2019). PCR technology has been used to enhance the accuracy of pathogen diagnosis, but it is unable to identify pathogens beyond those that are pre-designed for the disease (Ryu et al., 2014; Hischebeth et al., 2017; Villa et al., 2017). PCR of 16S ribosomal RNA genes has the potential to detect the majority of bacteria but it is not able to identify fungi or multiple microbial infections, nor is it able to distinguish contaminating bacteria from true infecting pathogens (Huang et al., 2018). However, mNGS can provide a comprehensive microbial profile without targeted pre-amplification (Huang et al., 2020). mNGS was reported to be able to identify known pathogens in 94.8% of culture-positive PJI cases and new potential pathogens in 43.9% of culture-negative infections (Thoendel et al., 2018). In one study, in addition to the collection of routine microbiological culture samples, some samples were collected for intraoperative culture to optimize the culture method based on the preoperative mNGS results. Preoperative aspiration of synovial fluid detected by mNGS provides more etiologic information than preoperative cultures, which can guide the optimization of intraoperative cultures and improve the

sensitivity of intraoperative cultures (Fang et al., 2021). Another study evaluated the diagnostic value of mNGS using three types of specimens, periprosthetic tissue, synovial fluid, and prosthetic ultrasound-treated fluid, and concluded that it can be used as an accurate diagnostic tool for the detection of pathogens in patients with PJI in all 3 specimens, and that due to its excellence in identifying pathogens, mNGS in artificial ultrasound fluid offers the greatest value and may partially replace traditional tests such as bacterial cultures in these patients (He et al., 2021).

## 4.7 Intra-abdominal infection

Intra-abdominal infection is one of the most common postoperative complications of abdominal surgery. The incidence of postoperative intra-abdominal infection (PIAI) is about 3-10% (Sánchez-Velázquez et al., 2018). Intra-abdominal infections are one of the most common postoperative complications of abdominal surgery. The incidence of postoperative PIAI is about 3-10%. Etiologic evidence remains key to the diagnosis of PIAI to date. However, the diagnosis of causative microorganisms leaves much to be desired (Zhu et al., 2023). First, traditional microbial culture methods are inadequate for isolation of picky microorganisms and anaerobes; second, certain microorganisms are often masked by rapidly proliferating microorganisms, making them difficult to identify; and third, empirical antibiotic treatment prior to sample collection may compromise the sensitivity of culture methods. Finally, traditional methods often take several days to produce results, which may delay appropriate treatment and increase the risk of antibiotic misuse. It was shown that the median sample-to-answer turnaround time for mNGS was significantly lower compared to culture-based methods (<24 h vs. 59.5-111 h). I was detected in a much wider range of assays than culture-based methods (Zhu et al., 2023).

For patients with abdominal sepsis, plasma mNGS can provide early, noninvasive and rapid microbiologic diagnosis. It facilitates rapid detection of pathogenic bacteria compared to traditional peritoneal drainage (PD) smear, culture, and blood culture methods. Paired plasma and PD fluid mNGS improves the microbiologic diagnosis of acute intra-abdominal infections (IAI) compared to microbiologic testing (CMT). The combination of plasma and PD mNGS predicts poor prognosis. It optimizes the use of empiric antibiotics (Li et al., 2022). Patients with spontaneous bacterial peritonitis (SBP) usually receive only empiric antibiotic therapy because pathogens can only be identified in a small number of patients using conventional culture techniques. Compared with conventional culture methods, mNGS improves the detection rate of ascites pathogens (including bacteria, viruses, and fungi), and has significant advantages in diagnosing rare pathogens and pathogens that are difficult to culture; moreover, it may be an effective method to improve the diagnosis of ascites infections in patients with cirrhosis, to guide early antibiotic therapy, and to reduce complications associated with abdominal infections, and relevant cases have been reported (Li et al., 2022; Lei et al., 2023; Shi et al., 2024). Infectious pancreatic necrosis (IPN)-associated pathogens can be identified by plasma mNGS, which has more valuable diagnostic properties and a shorter turnaround time, and may be useful for appropriate treatment (Hong et al., 2022; Lin et al., 2022). mNGS assay improves the detection of pathogenic microorganisms in PD-associated peritonitis, greatly reduces the time to detection, and has good concordance with microbiologic cultures (Zhang et al., 2023). mNGS is advantageous in diagnosing pathogens that are difficult to culture (Ye et al., 2022). mNGS is recommended for patients with PD-associated peritonitis who have received prior antibiotic therapy (Nie et al., 2023).

## 4.8 Skin and soft tissue infection

The diversity of pathogens in skin and soft tissue infections (SSTIs) and the tendency of clinicians to choose broad-spectrum antibiotics increases the prevalence of drug-resistant pathogens, so pathogen identification is important in the rational use of antibiotics (Esposito et al., 2018). Traditional culture techniques are time-consuming and less accurate, and even in many cases the microbial etiology fails to be clearly diagnosed. mNGS has emerged as a technology capable of more accurate diagnosis of pathogens. it has been shown in one study to detect twice as many pathogens compared to traditional culture (67.7% vs. 32.3%), with a particularly high detection rate for anaerobic bacteria, viruses, MTB, and NTM, and was superior in detecting viruses and rare pathogens, even recognizing multiple pathogens in a single specimen (Wang et al., 2020). In addition, positive results significantly contribute to targeted antibiotic therapy and may improve prognosis in the presence of negative culture results. It overcomes the limitations of current diagnostic tests (as 16S rRNA PCR) by allowing universal pathogen detection and new organism detection without *a priori* knowledge of a specific organism. However, in immunocompromised patients, it is less sensitive, possibly due to lower bacterial load (Parize et al., 2017). In conclusion, mNGS can reduce the time to pathogenic diagnosis of SSTI cases and make it possible to give targeted antibiotic therapy in the early stages of infectious diseases.

## 4.9 Intraocular infections

Intraocular infections are often caused by bacteria, fungi, viruses, and parasites, and in some cases can lead to visual impairment or even loss of vision due to poor diagnosis and treatment (Durand, 2017; Dave et al., 2019; Li et al., 2020; Zhang et al., 2020a; Fan et al., 2021). Although intraocular infections have a lower mortality rate compared to other infections, they are still an important cause of blindness (Durand, 2017; Ma et al., 2019). There are a number of cases of intraocular infections that are treated promptly, but the pathogen responsible for the intraocular infection is not diagnosed in time and only antibiotics can be used empirically, eventually leading to worsening of the infection and necrosis of the eye. Instead, only empirical antibiotics can be used, which ultimately leads to worsening of the infection and necrosis of the eyeball, and ultimately the patient's eyeball is removed (Tsai and Tseng, 2001; Lu et al., 2016). Therefore, timely and accurate

diagnosis and treatment have a crucial impact on the prognosis of intraocular infections. The presence of immune privilege in the eye and the blood-ocular barrier makes it difficult to detect intraocular infection pathogens from the blood, so the ideal samples for diagnosing intraocular infection pathogens are vitreous humor (VH) and aqueous humor (AH) (Doan et al., 2016; Deshmukh et al., 2019; Maitray et al., 2019; Li et al., 2020; Fan et al., 2021). The sensitivity of the culture method is only about 40%, and PCR fails to identify pathogens that are not set up for the test (Taravati et al., 2013; Doan et al., 2016; Borroni et al., 2019; Borroni et al., 2022). One study showed that the sensitivity of the mNGS using VH samples had a sensitivity of 92.2% and an overall compliance rate of 81.3%, whereas mNGS using AH samples had a sensitivity of 85.4% and an overall compliance rate of 75.4% (Qian et al., 2024). This indicates that mNGS demonstrates high sensitivity and high overall compliance in the diagnosis of intraocular infections.

## 5 Discussion

mNGS has not been in clinical use for a long time, but with the rapid development of diagnostic molecular microbiology in recent years, it has attracted extensive attention from laboratories and clinics due to its advantages of short time-consumption and wide detection range. Compared with traditional clinical microbiology detection methods, it not only effectively improves the detection rate of pathogenic microorganisms, but also makes up for the shortcomings of traditional detection methods, especially in the precise diagnosis and treatment of certain difficult and serious infectious diseases. So far, although it has been applied in several systemic infections and achieved remarkable results, it has not been widely used in clinical practice because of some shortcomings. Several findings demonstrate the potential benefits of adding it to routine diagnostic workflows for the detection and discovery of rare or novel pathogens, identifying key determinants of clinical benefit, particularly in immunocompromised patients and individuals with brain biopsies or fecal samples (Regnault et al., 2022; Fourgeaud et al., 2023; Pérot et al., 2023; Riller et al., 2023; Fourgeaud et al., 2024).

We think there is a long way to go before mNGS becomes a routine test in the clinic. First, the detection of DNA can only indicate which pathogens are present, not whether they are surviving pathogens. it is also unable to differentiate between pathogenic organisms and one-time bacteremia, so the inclusion of the detection of RNA in mNGS is very necessary. Second, false-negative results are easily produced when the proportion of pathogenic microorganisms in the genus microbacterium is low, so the amounts of precise microorganisms rather than a general determination of species is also in urgent need of improvement. Furthermore, the accuracy can be improved by stipulating uniform procedures and standards to avoid contamination of nucleic acids during specimen collection, transportation, and standardized testing. Finally, the future direction of mNGS is to build a professional team of analysts and interpreters, and to ensure that such a team is trained to a professional and uniform standard, so as to ensure the objectivity and reliability of the result.

Although a great deal of work needs to be improved, we believe that in the next few years, with the development of sequencing technology, the cost and turnaround time of mNGS will continue to decrease, the experimental process will be easy to manipulate, and the whole process will be automated, and it will become an even more mature method of detecting pathogens and play a historic role in clinical diagnosis and treatment.

## Author contributions

YZ: Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. WZ: Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. XZ: Funding acquisition, Supervision, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## References

Abril, M. K., Barnett, A. S., Wegermann, K., Fountain, E., Strand, A., Heyman, B. M., et al. (2016). Diagnosis of capnocytophaga canimorsus sepsis by whole-genome next-generation sequencing. *Open Forum Infect. Dis.* 3, ofw144. doi: 10.1093/ofid/ofw144

Ahmed, S. S., and Haddad, F. S. (2019). Prosthetic joint infection. *Bone Joint Res.* 8, 570–572. doi: 10.1302/2046-3758.812.BJR-2019-0340

Ai, J. W., Weng, S. S., Cheng, Q., Cui, P., Li, Y. J., Wu, H. L., et al. (2018). Human endophthalmitis caused by pseudorabies virus infection, China, 2017. *Emerg. Infect. Dis.* 24, 1087–1090. doi: 10.3201/eid2406.171612

Barlam, T. F., Cosgrove, S. E., Abbo, L. M., MacDougall, C., Schuetz, A. N., Septimus, E. J., et al. (2016). Implementing an antibiotic stewardship program: guidelines by the

infectious diseases society of America and the society for healthcare epidemiology of America. *Clin. Infect. Dis.* 62, e51–e77. doi: 10.1093/cid/ciw118

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature.* 456, 53–59. doi: 10.1038/nature07517

Blauwkamp, T. A., Thair, S., Rosen, M. J., Blair, L., Lindner, M. S., Vilfan, I. D., et al. (2019). Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat. Microbiol.* 4, 663–674. doi: 10.1038/s41564-018-0349-6

Borroni, D., Paytuví-Gallart, A., Sanseverino, W., Gómez-Huertas, C., Bonci, P., Romano, V., et al. (2022). Exploring the healthy eye microbiota niche in a multicenter study. *Int. J. Mol. Sci.* 23. doi: 10.3390/ijms231810229

Borroni, D., Romano, V., Kaye, S. B., Somerville, T., Napoli, L., Fasolo, A., et al. (2019). Metagenomics in ophthalmology: current findings and future prospectives. *BMJ Open Ophthalmol.* 4, e000248. doi: 10.1136/bmjophth-2018-000248

Brown, J. R., Bharucha, T., and Breuer, J. (2018). Encephalitis diagnosis using metagenomics: application of next generation sequencing for undiagnosed cases. *J. Infect.* 76, 225–240. doi: 10.1016/j.jinf.2017.12.014

Cabibbe, A. M., Spitaleri, A., Battaglia, S., Colman, R. E., Suresh, A., Uplekar, S., et al. (2020). Application of targeted next-generation sequencing assay on a portable sequencing platform for culture-free detection of drug-resistant tuberculosis from clinical samples. *J. Clin. Microbiol.* 58. doi: 10.1128/JCM.00632-20

Chao, L., Li, J., Zhang, Y., Pu, H., and Yan, X. (2020). Application of next generation sequencing-based rapid detection platform for microbiological diagnosis and drug resistance prediction in acute lower respiratory infection. *Ann. Transl. Med.* 8, 1644. doi: 10.21037/atm-20-7081

Chen, D., Chang, C., Chen, M., Zhang, Y., Zhao, X., Zhang, T., et al. (2020). Unusual disseminated Talaromyces marneffei infection mimicking lymphoma in a non-immunosuppressed patient in East China: a case report and review of the literature. *BMC Infect. Dis.* 20, 800. doi: 10.1186/s12879-020-05526-1

Chen, W., Wu, Y., and Zhang, Y. (2022). Next-generation sequencing technology combined with multiplex polymerase chain reaction as a powerful detection and semiquantitative method for herpes simplex virus type 1 in adult encephalitis: A case report. *Front. Med. (Lausanne).* 9, 905350. doi: 10.3389/fmed.2022.905350

Chen, Q., Yi, J., Liu, Y., Yang, C., Sun, Y., Du, J., et al. (2024). Clinical diagnostic value of targeted next–generation sequencing for infectious diseases (Review). *Mol. Med. Rep.* 30. doi: 10.3892/mmr.2024.13277

Cheng, L. L., Li, S. Y., and Zhong, N. S. (2023). New characteristics of COVID-19 caused by the Omicron variant in Guangzhou. *Zhonghua Jie He He Hu Xi Za Zhi.* 46, 441–443. doi: 10.3760/cma.j.cn112147-20230311-00125

Chiu, C. Y., and Miller, S. A. (2019). Clinical metagenomics. *Nat. Rev. Genet.* 20, 341–355. doi: 10.1038/s41576-019-0113-7

Costales, C., and Butler-Wu, S. M. (2018). A real pain: diagnostic quandaries and septic arthritis. *J. Clin. Microbiol.* 56. doi: 10.1128/JCM.01358-17

Dave, T. V., Dave, V. P., Sharma, S., Karolia, R., Joseph, J., Pathengay, A., et al. (2019). Infectious endophthalmitis leading to evisceration: spectrum of bacterial and fungal pathogens and antibacterial susceptibility profile. *J. Ophthalmic Inflammation Infect.* 9, 9. doi: 10.1186/s12348-019-0174-y

Deng, X., Achari, A., Federman, S., Yu, G., Somasekar, S., Bártolo, I., et al. (2020). Author Correction: Metagenomic sequencing with spiked primer enrichment for viral diagnostics and genomic surveillance. *Nat. Microbiol.* 5, 525. doi: 10.1038/s41564-020-0671-7

Deshmukh, D., Joseph, J., Chakrabarti, M., Sharma, S., Jayasudha, R., Sama, K. C., et al. (2019). New insights into culture negative endophthalmitis by unbiased next generation sequencing. *Sci. Rep.* 9, 844. doi: 10.1038/s41598-018-37502-w

De Vlaminck, I., Martin, L., Kertesz, M., Patel, K., Kowarsky, M., Strehl, C., et al. (2015). Noninvasive monitoring of infection and rejection after lung transplantation. *Proc. Natl. Acad. Sci. U S A.* 112, 13336–13341. doi: 10.1073/pnas.1517494112

Doan, T., Wilson, M. R., Crawford, E. D., Chow, E. D., Khan, L. M., Knopp, K. A., et al. (2016). Illuminating uveitis: metagenomic deep sequencing identifies common and rare pathogens. *Genome Med.* 8, 90. doi: 10.1186/s13073-016-0344-6

Drmanac, R., Sparks, A. B., Callow, M. J., Halpern, A. L., Burns, N. L., Kermani, B. G., et al. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science.* 327, 78–81. doi: 10.1126/science.1181498

Du, Z. M., and Chen, P. (2023). Co-infection of Chlamydia psittaci and Tropheryma whipplei: A case report. *World J. Clin. Cases.* 11, 7144–7149. doi: 10.12998/wjcc.v11.i29.7144

Duan, W., Yang, Y., Zhao, J., Yan, T., and Tian, X. (2022). Application of metagenomic next-generation sequencing in the diagnosis and treatment of recurrent urinary tract infection in kidney transplant recipients. *Front. Public Health* 10, 901549. doi: 10.3389/fpubh.2022.901549

Durand, M. L. (2017). Bacterial and fungal endophthalmitis. *Clin. Microbiol. Rev.* 30, 597–613. doi: 10.1128/CMR.00113-16

Esposito, S., Noviello, S., De Caro, F., and Boccia, G. (2018). New insights into classification, epidemiology and microbiology of SSTIs, including diabetic foot infections. *Infez Med.* 26, 3–14.

Fan, W., Han, H., Chen, Y., Zhang, X., Gao, Y., Li, S., et al. (2021). Antimicrobial nanomedicine for ocular bacterial and fungal infection. *Drug Delivery Transl. Res.* 11, 1352–1375. doi: 10.1007/s13346-021-00966-x

Fan, S., Qiao, X., Liu, L., Wu, H., Zhou, J., Sun, R., et al. (2018a). Next-generation sequencing of cerebrospinal fluid for the diagnosis of neurocysticercosis. *Front. Neurol.* 9, 471. doi: 10.3389/fneur.2018.00471

Fan, S., Ren, H., Wei, Y., Mao, C., Ma, Z., Zhang, L., et al. (2018b). Next-generation sequencing of the cerebrospinal fluid in the diagnosis of neurobrucellosis. *Int. J. Infect. Dis.* 67, 20–24. doi: 10.1016/j.ijid.2017.11.028

Fang, X., Cai, Y., Mei, J., Huang, Z., Zhang, C., Yang, B., et al. (2021). Optimizing culture methods according to preoperative mNGS results can improve joint infection diagnosis. *Bone Joint J.* 103-b, 39–45. doi: 10.1302/0301-620X.103B1.BJJ-2020-0771.R2

Fang, M., Weng, X., Chen, L., Chen, Y., Chi, Y., Chen, W., et al. (2020). Fulminant central nervous system varicella-zoster virus infection unexpectedly diagnosed by metagenomic next-generation sequencing in an HIV-infected patient: a case report. *BMC Infect. Dis.* 20, 159. doi: 10.1186/s12879-020-4872-8

Fenollar, F., and Raoult, D. (2007). Molecular diagnosis of bloodstream infections caused by non-cultivable bacteria. *Int. J. Antimicrob. Agents.* 30 Suppl 1, S7–15. doi: 10.1016/j.ijantimicag.2007.06.024

Fishman, J. A. (2007). Infection in solid-organ transplant recipients. *N Engl. J. Med.* 357, 2601–2614. doi: 10.1056/NEJMra064928

Flores-Mireles, A. L., Walker, J. N., Caparon, M., and Hultgren, S. J. (2015). Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat. Rev. Microbiol.* 13, 269–284. doi: 10.1038/nrmicro3432

Fourgeaud, J., Regnault, B., Faury, H., Da Rocha, N., Jamet, A., Stirnemann, J., et al. (2023). Fetal Zika virus infection diagnosed by metagenomic next-generation sequencing of amniotic fluid. *Ultrasound Obstet Gynecol.* 61, 116–117. doi: 10.1002/uog.26074

Fourgeaud, J., Regnault, B., Ok, V., Da Rocha, N., Sitterlé, É, Mekouar, M., et al. (2024). Performance of clinical metagenomics in France: a prospective observational study. *Lancet Microbe* 5, e52–e61. doi: 10.1016/S2666-5247(23)00244-6

Gauduchon, V., Chalabreysse, L., Etienne, J., Célard, M., Benito, Y., Lepidi, H., et al. (2003). Molecular diagnosis of infective endocarditis by PCR amplification and direct sequencing of DNA from valve tissue. *J. Clin. Microbiol.* 41, 763–766. doi: 10.1128/JCM.41.2.763-766.2003

Glaser, C. A., Gilliam, S., Schnurr, D., Forghani, B., Honarmand, S., Khetsuriani, N., et al. (2003). In search of encephalitis etiologies: diagnostic challenges in the California Encephalitis Project, 1998-2000. *Clin. Infect. Dis.* 36, 731–742. doi: 10.1086/cid.2003.36.issue-6

Glimåker, M., Johansson, B., Grindborg, Ö, Bottai, M., Lindquist, L., and Sjölin, J. (2015). Adult bacterial meningitis: earlier treatment and improved outcome following guideline revision promoting prompt lumbar puncture. *Clin. Infect. Dis.* 60, 1162–1169. doi: 10.1093/cid/civ011

Goodwin, S., McPherson, J. D., and McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* 17, 333–351. doi: 10.1038/nrg.2016.49

Gosiewski, T., Ludwig-Galezowska, A. H., Huminska, K., Sroka-Oleksiak, A., Radkowski, P., Salamon, D., et al. (2017). Comprehensive detection and identification of bacterial DNA in the blood of patients with sepsis and healthy volunteers using next-generation sequencing method - the observation of DNAemia. *Eur. J. Clin. Microbiol. Infect. Dis.* 36, 329–336. doi: 10.1007/s10096-016-2805-7

Granerod, J., Tam, C. C., Crowcroft, N. S., Davies, N. W., Borchert, M., and Thomas, S. L. (2010). Challenge of the unknown. A systematic review of acute encephalitis in non-outbreak situations. *Neurology.* 75, 924–932. doi: 10.1212/WNL.0b013e3181f11d65

Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryke, D., et al. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* 27, 115–124. doi: 10.1038/s41591-020-1105-z

Gu, W., Miller, S., and Chiu, C. Y. (2019). Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol.* 14, 319–338. doi: 10.1146/annurev-pathmechdis-012418-012751

Guan, H., Shen, A., Lv, X., Yang, X., Ren, H., Zhao, Y., et al. (2016). Detection of virus in CSF from the cases with meningoencephalitis by next-generation sequencing. *J. Neurovirol.* 22, 240–245. doi: 10.1007/s13365-015-0390-7

Gupta, S., Sakhuja, A., Kumar, G., McGrath, E., Nanchal, R. S., and Kashani, K. B. (2016). Culture-negative severe sepsis: nationwide trends and outcomes. *Chest.* 150, 1251–1259. doi: 10.1016/j.chest.2016.08.1460

Han, D., Li, Z., Li, R., Tan, P., Zhang, R., and Li, J. (2019). mNGS in clinical microbiology laboratories: on the road to maturity. *Crit. Rev. Microbiol.* 45, 668–685. doi: 10.1080/1040841X.2019.1681933

Han, D., Yu, F., Zhang, D., Hu, J., Zhang, X., Xiang, D., et al. (2024). Molecular rapid diagnostic testing for bloodstream infections: Nanopore targeted sequencing with pathogen-specific primers. *J. Infect.* 88, 106166. doi: 10.1016/j.jinf.2024.106166

He, R., Wang, Q., Wang, J., Tang, J., Shen, H., and Zhang, X. (2021). Better choice of the type of specimen used for untargeted metagenomic sequencing in the diagnosis of periprosthetic joint infections. *Bone Joint J.* 103-b, 923–930. doi: 10.1302/0301-620X.103B5.BJJ-2020-0745.R1

He, R., Zheng, W., Long, J., Huang, Y., Liu, C., Wang, Q., et al. (2019). Vibrio vulnificus meningoencephalitis in a patient with thalassemia and a splenectomy. *J. Neurovirol.* 25, 127–132. doi: 10.1007/s13365-018-0675-8

Hischebeth, G. T. R., Gravius, S., Buhr, J. K., Molitor, E., Wimmer, M. D., Hoerauf, A., et al. (2017). Novel diagnostics in revision arthroplasty: implant sonication and multiplex polymerase chain reaction. *J. Vis. Exp.* 130. doi: 10.3791/55147

Hong, D. K., Blauwkamp, T. A., Kertesz, M., Bercovici, S., Truong, C., and Banaei, N. (2018). Liquid biopsy for infectious diseases: sequencing of cell-free plasma to detect pathogen DNA in patients with invasive fungal disease. *Diagn. Microbiol. Infect. Dis.* 92, 210–213. doi: 10.1016/j.diagmicrobio.2018.06.009

Hong, D., Wang, P., Zhang, J., Li, K., Ye, B., Li, G., et al. (2022). Plasma metagenomic next-generation sequencing of microbial cell-free DNA detects pathogens in patients with suspected infected pancreatic necrosis. *BMC Infect. Dis.* 22, 675. doi: 10.1186/s12879-022-07662-2

Hou, M., Wang, Y., Yuan, H., Zhang, Y., Luo, X., Xin, N., et al. (2024). The diagnostic value of metagenomics next-generation sequencing in HIV-infected patients with suspected pulmonary infections. *Front. Cell Infect. Microbiol.* 14, 1395239. doi: 10.3389/fcimb.2024.1395239

Hu, Z., Weng, X., Xu, C., Lin, Y., Cheng, C., Wei, H., et al. (2018). Metagenomic next-generation sequencing as a diagnostic tool for toxoplasmic encephalitis. *Ann. Clin. Microbiol. Antimicrob.* 17, 45. doi: 10.1186/s12941-018-0298-1

Huang, C., Chang, S., Ma, R., Shang, Y., Li, Y., Wang, Y., et al. (2023). COVID-19 in pulmonary critically ill patients: metagenomic identification of fungi and characterization of pathogenic microorganisms. *Front. Cell Infect. Microbiol.* 13, 1220012. doi: 10.3389/fcimb.2023.1220012

Huang, J., Jiang, E., Yang, D., Wei, J., Zhao, M., Feng, J., et al. (2020). Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions. *Infect. Drug Resist.* 13, 567–576. doi: 10.2147/IDR.S235182

Huang, Z., Li, W., Lee, G. C., Fang, X., Xing, L., Yang, B., et al. (2020). Metagenomic next-generation sequencing of synovial fluid demonstrates high accuracy in prosthetic joint infection diagnostics: mNGS for diagnosing PJI. *Bone Joint Res.* 9, 440–449. doi: 10.1302/2046-3758.97.BJR-2019-0325.R2

Huang, Z., Wu, Q., Fang, X., Li, W., Zhang, C., Zeng, H., et al. (2018). Comparison of culture and broad-range polymerase chain reaction methods for diagnosing periprosthetic joint infection: analysis of joint fluid, periprosthetic tissue, and sonicated fluid. *Int. Orthop.* 42, 2035–2040. doi: 10.1007/s00264-018-3827-9

Huang, R., Yuan, Q., Gao, J., Liu, Y., Jin, X., Tang, L., et al. (2023). Application of metagenomic next-generation sequencing in the diagnosis of urinary tract infection in patients undergoing cutaneous ureterostomy. *Front. Cell Infect. Microbiol.* 13, 991011. doi: 10.3389/fcimb.2023.991011

Jia, K., Huang, S., Shen, C., Li, H., Zhang, Z., Wang, L., et al. (2023). Enhancing urinary tract infection diagnosis for negative culture patients with metagenomic next-generation sequencing (mNGS). *Front. Cell Infect. Microbiol.* 13, 1119020. doi: 10.3389/fcimb.2023.1119020

Jiang, Z., Gai, W., Zhang, X., Zheng, Y., Jin, X., Han, Z., et al. (2024). Clinical performance of metagenomic next-generation sequencing for diagnosis of pulmonary Aspergillus infection and colonization. *Front. Cell Infect. Microbiol.* 14, 1345706. doi: 10.3389/fcimb.2024.1345706

Jiang, J., Lv, M., Yang, K., Zhao, G., and Fu, Y. (2023). A case report of diagnosis and dynamic monitoring of Listeria monocytogenes meningitis with NGS. *Open Life Sci.* 18, 20220738. doi: 10.1515/biol-2022-0738

Jouet, A., Gaudin, C., Badalato, N., Allix-Béguec, C., Duthoy, S., Ferré, A., et al. (2021). Deep amplicon sequencing for culture-free prediction of susceptibility or resistance to 13 anti-tuberculous drugs. *Eur. Respir. J.* 57. doi: 10.1183/13993003.02338-2020

Kumar, A. (2014). An alternate pathophysiologic paradigm of sepsis and septic shock: implications for optimizing antimicrobial therapy. *Virulence.* 5, 80–97. doi: 10.4161/viru.26913

Kumar, A., Ellis, P., Arabi, Y., Roberts, D., Light, B., Parrillo, J. E., et al. (2009). Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. *Chest.* 136, 1237–1248. doi: 10.1378/chest.09-0087

Kunasol, C., Dondorp, A. M., Batty, E. M., Nakhonsri, V., Sinjanakhom, P., Day, N. P. J., et al. (2022). Comparative analysis of targeted next-generation sequencing for Plasmodium falciparum drug resistance markers. *Sci. Rep.* 12, 5563. doi: 10.1038/s41598-022-09474-5

Lei, Y., Guo, Q., Liu, J., Huang, H., and Han, P. (2023). Staphylococcus cohnii infection diagnosed by metagenomic next generation sequencing in a patient on hemodialysis with cirrhotic ascites: a case report. *Front. Cell Infect. Microbiol.* 13, 1240283. doi: 10.3389/fcimb.2023.1240283

Li, Z., Breitwieser, F. P., Lu, J., Jun, A. S., Asnaghi, L., Salzberg, S. L., et al. (2018). Identifying corneal infections in formalin-fixed specimens using next generation sequencing. *Invest. Ophthalmol. Vis. Sci.* 59, 280–288. doi: 10.1167/iovs.17-21617

Li, Y., Deng, X., Hu, F., Wang, J., Liu, Y., Huang, H., et al. (2018). Metagenomic analysis identified co-infection with human rhinovirus C and bocavirus 1 in an adult suffering from severe pneumonia. *J. Infect.* 76, 311–313. doi: 10.1016/j.jinf.2017.10.012

Li, D., Gai, W., Zhang, J., Cheng, W., Cui, N., and Wang, H. (2022). Metagenomic next-generation sequencing for the microbiological diagnosis of abdominal sepsis patients. *Front. Microbiol.* 13, 816631. doi: 10.3389/fmicb.2022.816631

Li, H., Gao, H., Meng, H., Wang, Q., Li, S., Chen, H., et al. (2018). Detection of pulmonary infectious pathogens from lung biopsy tissues by metagenomic next-generation sequencing. *Front. Cell Infect. Microbiol.* 8, 205. doi: 10.3389/fcimb.2018.00205

Li, B., He, Q., Rui, Y., Chen, Y., Jalan, R., and Chen, J. (2022). Rapid detection for infected ascites in cirrhosis using metagenome next-generation sequencing: A case series. *Liver Int.* 42, 173–179. doi: 10.1111/liv.15083

Li, Y., Sun, B., Tang, X., Liu, Y. L., He, H. Y., Li, X. Y., et al. (2020). Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in critically ill patients. *Eur. J. Clin. Microbiol. Infect. Dis.* 39, 369–374. doi: 10.1007/s10096-019-03734-5

Li, S., Tong, J., Li, H., Mao, C., Shen, W., Lei, Y., et al. (2023). L. pneumophila Infection Diagnosed by tNGS in a Lady with Lymphadenopathy. *Infect. Drug Resist.* 16, 4435–4442. doi: 10.2147/IDR.S417495

Li, J. J., Yi, S., and Wei, L. (2020). Ocular microbiota and intraocular inflammation. *Front. Immunol.* 11, 609765. doi: 10.3389/fimmu.2020.609765

Liesenfeld, O., Lehman, L., Hunfeld, K. P., and Kost, G. (2014). Molecular diagnosis of sepsis: New aspects and recent developments. *Eur. J. Microbiol. Immunol. (Bp).* 4, 1–25. doi: 10.1556/EuJMI.4.2014.1.1

Lin, C., Bonsu, A., Li, J., Ning, C., Chen, L., Zhu, S., et al. (2022). Application of metagenomic next-generation sequencing for suspected infected pancreatic necrosis. *Pancreatology.* 22, 864–870. doi: 10.1016/j.pan.2022.07.006

Liu, B. M. (2024). Epidemiological and clinical overview of the 2024 Oropouche virus disease outbreaks, an emerging/re-emerging neurotropic arboviral disease and global public health threat. *J. Med. Virol.* 96, e29897. doi: 10.1002/jmv.29897

Liu, N., Kan, J., Cao, W., Cao, J., Jiang, E., Zhou, Y., et al. (2019). Metagenomic next-generation sequencing diagnosis of peripheral pulmonary infectious lesions through virtual navigation, radial EBUS, ultrathin bronchoscopy, and ROSE. *J. Int. Med. Res.* 47, 4878–4885. doi: 10.1177/0300060519866953

Liu, M. F., Liu, Y., Xu, D. R., Wan, L. G., and Zhao, R. (2021). mNGS helped diagnose scrub typhus presenting as a urinary tract infection with high D-dimer levels: a case report. *BMC Infect. Dis.* 21, 1219. doi: 10.1186/s12879-021-06889-9

Liu, B. M., Mulkey, S. B., Campos, J. M., and DeBiasi, R. L. (2024). Laboratory diagnosis of CNS infections in children due to emerging and re-emerging neurotropic viruses. *Pediatr. Res.* 95, 543–550. doi: 10.1038/s41390-023-02930-6

Liu, X., Wang, J., Liu, J., Li, X., Guan, Y., Qian, S., et al. (2023). Cryptosporidiosis diagnosed using metagenomic next-generation sequencing in a healthy child admitted to pediatric intensive care unit: a case report. *Front. Cell Infect. Microbiol.* 13, 1269963. doi: 10.3389/fcimb.2023.1269963

Liu, P., Weng, X., Zhou, J., Xu, X., He, F., Du, Y., et al. (2018). Next generation sequencing based pathogen analysis in a patient with neurocysticercosis: a case report. *BMC Infect. Dis.* 18, 113. doi: 10.1186/s12879-018-3015-y

Long, Y., Zhang, Y., Gong, Y., Sun, R., Su, L., Lin, X., et al. (2016). Diagnosis of sepsis with cell-free DNA by next-generation sequencing technology in ICU patients. *Arch. Med. Res.* 47, 365–371. doi: 10.1016/j.arcmed.2016.08.004

Lu, X., Ng, D. S., Zheng, K., Peng, K., Jin, C., Xia, H., et al. (2016). Risk factors for endophthalmitis requiring evisceration or enucleation. *Sci. Rep.* 6, 28100. doi: 10.1038/srep28100

Lucas, M. J., Brouwer, M. C., and van de Beek, D. (2016). Neurological sequelae of bacterial meningitis. *J. Infect.* 73, 18–27. doi: 10.1016/j.jinf.2016.04.009

Ma, L., Jakobiec, F. A., and Dryja, T. P. (2019). A review of next-generation sequencing (NGS): applications to the diagnosis of ocular infectious diseases. *Semin. Ophthalmol.* 34, 223–231. doi: 10.1080/08820538.2019.1620800

Maitray, A., Rishi, E., Rishi, P., Gopal, L., Bhende, P., Ray, R., et al. (2019). Endogenous endophthalmitis in children and adolescents: Case series and literature review. *Indian J. Ophthalmol.* 67, 795–800. doi: 10.4103/ijo.IJO_710_18

Malizos, K. N. (2017). Global forum: the burden of bone and joint infections: A growing demand for more resources. *J. Bone Joint Surg. Am.* 99, e20. doi: 10.2106/JBJS.16.00240

Mancini, N., Carletti, S., Ghidoli, N., Cichero, P., Burioni, R., and Clementi, M. (2010). The era of molecular and other non-culture-based methods in diagnosis of sepsis. *Clin. Microbiol. Rev.* 23, 235–251. doi: 10.1128/CMR.00043-09

Maneg, D., Sponsel, J., Müller, I., Lohr, B., Penders, J., Madlener, K., et al. (2016). Advantages and limitations of direct PCR amplification of bacterial 16S-rDNA from resected heart tissue or swabs followed by direct sequencing for diagnosing infective endocarditis: A retrospective analysis in the routine clinical setting. *BioMed. Res. Int.* 2016, 7923874. doi: 10.1155/2016/7923874

Marín, M., Muñoz, P., Sánchez, M., Del Rosal, M., Alcalá, L., Rodríguez-Créixems, M., et al. (2007). Molecular diagnosis of infective endocarditis by real-time broad-range polymerase chain reaction (PCR) and sequencing directly from heart valve tissue. *Med. (Baltimore).* 86, 195–202. doi: 10.1097/MD.0b013e31811f44ec

Mensah, B. A., Aydemir, O., Myers-Hansen, J. L., Opoku, M., Hathaway, N. J., Marsh, P. W., et al. (2020). Antimalarial drug resistance profiling of plasmodium falciparum infections in Ghana using molecular inversion probes and next-generation sequencing. *Antimicrob. Agents Chemother.* 64. doi: 10.1128/AAC.01423-19

Mesfin, A. B., Araia, Z. Z., Beyene, H. N., Mebrahtu, A. H., Suud, N. N., Berhane, Y. M., et al. (2021). First molecular-based anti-TB drug resistance survey in Eritrea. *Int. J. Tuberc Lung Dis.* 25, 43–51. doi: 10.5588/ijtld.20.0558

Messacar, K., Parker, S. K., Todd, J. K., and Dominguez, S. R. (2017). Implementation of rapid molecular infectious disease diagnostics: the role of diagnostic and antimicrobial stewardship. *J. Clin. Microbiol.* 55, 715–723. doi: 10.1128/JCM.02264-16

Miao, Q., Ma, Y., Wang, Q., Pan, J., Zhang, Y., Jin, W., et al. (2018). Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin. Infect. Dis.* 67, S231–Ss40. doi: 10.1093/cid/ciy693

Milucky, J., Pondo, T., Gregory, C. J., Iuliano, D., Chaves, S. S., McCracken, J., et al. (2020). The epidemiology and estimated etiology of pathogens detected from the upper

respiratory tract of adults with severe acute respiratory infections in multiple countries, 2014-2015. *PLoS One* 15, e0240309. doi: 10.1371/journal.pone.0240309

Mongkolrattanothai, K., Naccache, S. N., Bender, J. M., Samayoa, E., Pham, E., Yu, G., et al. (2017). Neurobrucellosis: unexpected answer from metagenomic next-generation sequencing. *J. Pediatr. Infect. Dis. Soc* 6, 393–398. doi: 10.1093/jpids/piw066

Mu, S., Hu, L., Zhang, Y., Liu, Y., Cui, X., Zou, X., et al. (2021). Prospective evaluation of a rapid clinical metagenomics test for bacterial pneumonia. *Front. Cell Infect. Microbiol.* 11, 684965. doi: 10.3389/fcimb.2021.684965

Nie, S., Zhang, Q., Chen, R., Lin, L., Li, Z., Sun, Y., et al. (2023). Rapid detection of pathogens of peritoneal dialysis-related peritonitis, especially in patients who have taken antibiotics, using metagenomic next-generation sequencing: a pilot study. *Ren Fail.* 45, 2284229. doi: 10.1080/0886022X.2023.2284229

Pan, W., Ngo, T. T. M., Camunas-Soler, J., Song, C. X., Kowarsky, M., Blumenfeld, Y. J., et al. (2017). Simultaneously Monitoring Immune Response and Microbial Infections during Pregnancy through Plasma cfRNA Sequencing. *Clin. Chem.* 63, 1695–1704. doi: 10.1373/clinchem.2017.273888

Pan, L., Pan, X. H., Xu, J. K., Huang, X. Q., Qiu, J. K., Wang, C. H., et al. (2021). Misdiagnosed tuberculosis being corrected as Nocardia farcinica infection by metagenomic sequencing: a case report. *BMC Infect. Dis.* 21, 754. doi: 10.1186/s12879-021-06436-6

Parize, P., Muth, E., Richaud, C., Gratigny, M., Pilmis, B., Lamamy, A., et al. (2017). Untargeted next-generation sequencing-based first-line diagnosis of infection in immunocompromised adults: a multicentre, blinded, prospective study. *Clin. Microbiol. Infect.* 23, 574.e1–574.e6. doi: 10.1016/j.cmi.2017.02.006

Parvizi, J., Erkocak, O. F., and Della Valle, C. J. (2014). Culture-negative periprosthetic joint infection. *J. Bone Joint Surg. Am.* 96, 430–436. doi: 10.2106/JBJS.L.01793

Paul, M., Shani, V., Muchtar, E., Kariv, G., Robenshtok, E., and Leibovici, L. (2010). Systematic review and meta-analysis of the efficacy of appropriate empiric antibiotic therapy for sepsis. *Antimicrob. Agents Chemother.* 54, 4851–4863. doi: 10.1128/AAC.00627-10

Pérot, P., Fourgeaud, J., Rouzaud, C., Regnault, B., Da Rocha, N., Fontaine, H., et al. (2023). Circovirus hepatitis infection in heart-lung transplant patient, France. *Emerg. Infect. Dis.* 29, 286–293. doi: 10.3201/eid2902.221468

Pham, J., Su, L. D., Hanson, K. E., and Hogan, C. A. (2023). Sequence-based diagnostics and precision medicine in bacterial and viral infections: from bench to bedside. *Curr. Opin. Infect. Dis.* 36, 228–234. doi: 10.1097/QCO.0000000000000936

Qian, Z., Xia, H., Zhou, J., Wang, R., Zhu, D., Chen, L., et al. (2024). Performance of metagenomic next-generation sequencing of cell-free DNA from vitreous and aqueous humor for diagnoses of intraocular infections. *J. Infect. Dis.* 229, 252–261. doi: 10.1093/infdis/jiad363

Ramos, N., Panzera, Y., Frabasile, S., Tomás, G., Calleros, L., Marandino, A., et al. (2023). A multiplex-NGS approach to identifying respiratory RNA viruses during the COVID-19 pandemic. *Arch. Virol.* 168, 87. doi: 10.1007/s00705-023-05717-6

Regnault, B., Evrard, B., Plu, I., Dacheux, L., Troadec, E., Cozette, P., et al. (2022). First case of lethal encephalitis in western europe due to european bat lyssavirus type 1. *Clin. Infect. Dis.* 74, 461–466. doi: 10.1093/cid/ciab443

Ren, H. Q., Zhao, Q., Jiang, J., Yang, W., Fu, A. S., and Ge, Y. L. (2023). Acute heart failure due to pulmonary aspergillus fumigatus and cryptococcus neoformans infection associated with COVID-19. *Clin. Lab.* 69. doi: 10.7754/Clin.Lab.2023.230407

Riller, Q., Fourgeaud, J., Bruneau, J., De Ravin, S. S., Smith, G., Fusaro, M., et al. (2023). Late-onset enteric virus infection associated with hepatitis (EVAH) in transplanted SCID patients. *J. Allergy Clin. Immunol.* 151, 1634–1645. doi: 10.1016/j.jaci.2022.12.822

Rodino, K. G., and Simner, P. J. (2024). Status check: next-generation sequencing for infectious-disease diagnostics. *J. Clin. Invest.* 134. doi: 10.1172/JCI178003

Ryu, S. Y., Greenwood-Quaintance, K. E., Hanssen, A. D., Mandrekar, J. N., and Patel, R. (2014). Low sensitivity of periprosthetic tissue PCR for prosthetic knee infection diagnosis. *Diagn. Microbiol. Infect. Dis.* 79, 448–453. doi: 10.1016/j.diagmicrobio.2014.03.021

Sánchez-Velázquez, P., Pera, M., Jiménez-Toscano, M., Mayol, X., Rogés, X., Lorente, L., et al. (2018). Postoperative intra-abdominal infection is an independent prognostic factor of disease-free survival and disease-specific survival in patients with stage II colon cancer. *Clin. Transl. Oncol.* 20, 1321–1328. doi: 10.1007/s12094-018-1866-8

Schmiemann, G., Kniehl, E., Gebhardt, K., Matejczyk, M. M., and Hummers-Pradier, E. (2010). The diagnosis of urinary tract infection: a systematic review. *Dtsch Arztebl Int.* 107, 361–367. doi: 10.3238/arztebl.2010.0361

Shan, H., Wei, C., Zhang, J., He, M., and Zhang, Z. (2023). Case report: severe diarrhea caused by cryptosporidium diagnosed by metagenome next-generation sequencing in blood. *Infect. Drug Resist.* 16, 5777–5782. doi: 10.2147/IDR.S422799

Shang, Y., Ren, Y., Liu, L., Chen, X., Liu, F., Li, X., et al. (2024). Encephalitozoon hellem infection after haploidentical allogeneic hematopoietic stem cell transplantation in children: a case report. *Front. Immunol.* 15, 1396260. doi: 10.3389/fimmu.2024.1396260

Shi, P., Liu, J., Liang, A., Zhu, W., Fu, J., Wu, X., et al. (2024). Application of metagenomic next-generation sequencing in optimizing the diagnosis of ascitic infection in patients with liver cirrhosis. *BMC Infect. Dis.* 24, 503. doi: 10.1186/s12879-024-09396-9

Shi, J., Wu, W., Wu, K., Ni, C., He, G., Zheng, S., et al. (2022). The diagnosis of leptospirosis complicated by pulmonary tuberculosis complemented by metagenomic next-generation sequencing: A case report. *Front. Cell Infect. Microbiol.* 12, 922996. doi: 10.3389/fcimb.2022.922996

Sibandze, D. B., Kay, A., Dreyer, V., Sikhondze, W., Dlamini, Q., DiNardo, A., et al. (2022). Rapid molecular diagnostics of tuberculosis resistance by targeted stool sequencing. *Genome Med.* 14, 52. doi: 10.1186/s13073-022-01054-6

Simner, P. J., Miller, S., and Carroll, K. C. (2018). Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin. Infect. Dis.* 66, 778–788. doi: 10.1093/cid/cix881

Singal, A. K., Salameh, H., and Kamath, P. S. (2014). Prevalence and in-hospital mortality trends of infections among patients with cirrhosis: a nationwide study of hospitalised patients in the United States. *Aliment Pharmacol. Ther.* 40, 105–112. doi: 10.1111/apt.2014.40.issue-1

Smiley Evans, T., Shi, Z., Boots, M., Liu, W., Olival, K. J., Xiao, X., et al. (2020). Synergistic China-US ecological research is essential for global emerging infectious disease preparedness. *Ecohealth.* 17, 160–173. doi: 10.1007/s10393-020-01471-2

Smith, C. J., and Osborn, A. M. (2009). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol. Ecol.* 67, 6–20. doi: 10.1111/j.1574-6941.2008.00629.x

Taravati, P., Lam, D., and Van Gelder, R. N. (2013). Role of molecular diagnostics in ocular microbiology. *Curr. Ophthalmol. Rep.* 1. doi: 10.1007/s40135-013-0025-1

Thoendel, M. J., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., et al. (2018). Identification of prosthetic joint infection pathogens using a shotgun metagenomics approach. *Clin. Infect. Dis.* 67, 1333–1338. doi: 10.1093/cid/ciy303

Tomblyn, M., Chiller, T., Einsele, H., Gress, R., Sepkowitz, K., Storek, J., et al. (2009). Guidelines for preventing infectious complications among hematopoietic cell transplantation recipients: a global perspective. *Biol. Blood Marrow Transplant.* 15, 1143–1238. doi: 10.1016/j.bbmt.2009.06.019

Tsai, Y. Y., and Tseng, S. H. (2001). Risk factors in endophthalmitis leading to evisceration or enucleation. *Ophthalmic Surg. Lasers.* 32, 208–212. doi: 10.3928/1542-8877-20010501-06

Tyler, K. L. (2018). Acute viral encephalitis. *N Engl. J. Med.* 379, 557–566. doi: 10.1056/NEJMra1708714

van Rijn, A. L., van Boheemen, S., Sidorov, I., Carbo, E. C., Pappas, N., Mei, H., et al. (2019). The respiratory virome and exacerbations in patients with chronic obstructive pulmonary disease. *PLoS One* 14, e0223952. doi: 10.1371/journal.pone.0223952

Villa, F., Toscano, M., De Vecchi, E., Bortolin, M., and Drago, L. (2017). Reliability of a multiplex PCR system for diagnosis of early and late prosthetic joint infections before and after broth enrichment. *Int. J. Med. Microbiol.* 307, 363–370. doi: 10.1016/j.ijmm.2017.07.005

Vondracek, M., Sartipy, U., Aufwerber, E., Julander, I., Lindblom, D., and Westling, K. (2011). 16S rDNA sequencing of valve tissue improves microbiological diagnosis in surgically treated patients with infective endocarditis. *J. Infect.* 62, 472–478. doi: 10.1016/j.jinf.2011.04.010

Wang, S., Chen, Y., Wang, D., Wu, Y., Zhao, D., Zhang, J., et al. (2019). The feasibility of metagenomic next-generation sequencing to identify pathogens causing tuberculous meningitis in cerebrospinal fluid. *Front. Microbiol.* 10, 1993. doi: 10.3389/fmicb.2019.01993

Wang, Y., Chen, T., Zhang, S., Zhang, L., Li, Q., Lv, Q., et al. (2023). Clinical evaluation of metagenomic next-generation sequencing in unbiased pathogen diagnosis of urinary tract infection. *J. Transl. Med.* 21, 762. doi: 10.1186/s12967-023-04562-0

Wang, J., Han, Y., and Feng, J. (2019). Metagenomic next-generation sequencing for mixed pulmonary infection diagnosis. *BMC Pulm Med.* 19, 252. doi: 10.1186/s12890-019-1022-4

Wang, J., Jia, P., Zhang, D., Zhao, Y., Sui, X., Jin, Z., et al. (2024). Diagnosis of a familial psittacosis outbreak with clinical analysis and metagenomic next-generation sequencing under COVID-19: A case series. *Infect. Drug Resist.* 17, 1099–1105. doi: 10.2147/IDR.S440400

Wang, Q., Li, J., Ji, J., Yang, L., Chen, L., Zhou, R., et al. (2018). A case of Naegleria fowleri related primary amoebic meningoencephalitis in China diagnosed by next-generation sequencing. *BMC Infect. Dis.* 18, 349. doi: 10.1186/s12879-018-3261-z

Wang, H., Lu, Z., Bao, Y., Yang, Y., de Groot, R., Dai, W., et al. (2020). Clinical diagnostic application of metagenomic next-generation sequencing in children with severe nonresponding pneumonia. *PLoS One* 15, e0232610. doi: 10.1371/journal.pone.0232610

Wang, Q., Miao, Q., Pan, J., Jin, W., Ma, Y., Zhang, Y., et al. (2020). The clinical value of metagenomic next-generation sequencing in the microbiological diagnosis of skin and soft tissue infections. *Int. J. Infect. Dis.* 100, 414–420. doi: 10.1016/j.ijid.2020.09.007

Wang, C., Yin, X., Ma, W., Zhao, L., Wu, X., Ma, N., et al. (2024). Clinical application of bronchoalveolar lavage fluid metagenomics next-generation sequencing in cancer patients with severe pneumonia. *Respir. Res.* 25, 68. doi: 10.1186/s12931-023-02654-5

Wang, N., Zhao, C., Tang, C., and Wang, L. (2022). Case report and literature review: disseminated histoplasmosis infection diagnosed by metagenomic next-generation sequencing. *Infect. Drug Resist.* 15, 4507–4514. doi: 10.2147/IDR.S371740

Wensel, C. R., Pluznick, J. L., Salzberg, S. L., and Sears, C. L. (2022). Next-generation sequencing: insights to advance clinical investigations of the microbiome. *J. Clin. Invest.* 132. doi: 10.1172/JCI154944

Wilson, M. R., O'Donovan, B. D., Gelfand, J. M., Sample, H. A., Chow, F. C., Betjemann, J. P., et al. (2018). Chronic meningitis investigated via metagenomic next-generation sequencing. *JAMA Neurol.* 75, 947–955. doi: 10.1001/jamaneurol.2018.0463

Wilson, M. R., Sample, H. A., Zorn, K. C., Arevalo, S., Yu, G., Neuhaus, J., et al. (2019). Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N Engl. J. Med.* 380, 2327–2340. doi: 10.1056/NEJMoa1803396

Wooley, J. C., Godzik, A., and Friedberg, I. (2010). A primer on metagenomics. *PLoS Comput. Biol.* 6, e1000667. doi: 10.1371/journal.pcbi.1000667

Xia, H., Guan, Y., Zaongo, S. D., Xia, H., Wang, Z., Yan, Z., et al. (2019). Progressive multifocal leukoencephalopathy diagnosed by metagenomic next-generation sequencing of cerebrospinal fluid in an HIV patient. *Front. Neurol.* 10, 1202. doi: 10.3389/fneur.2019.01202

Xie, Y., Du, J., Jin, W., Teng, X., Cheng, R., Huang, P., et al. (2019). Next generation sequencing for diagnosis of severe pneumonia: China, 2010-2018. *J. Infect.* 78, 158–169. doi: 10.1016/j.jinf.2018.09.004

Xing, X. W., Zhang, J. T., Ma, Y. B., He, M. W., Yao, G. E., Wang, W., et al. (2020). Metagenomic next-generation sequencing for diagnosis of infectious encephalitis and meningitis: A large, prospective case series of 213 patients. *Front. Cell Infect. Microbiol.* 10, 88. doi: 10.3389/fcimb.2020.00088

Xing, X. W., Zhang, J. T., Ma, Y. B., Zheng, N., Yang, F., and Yu, S. Y. (2019). Apparent performance of metagenomic next-generation sequencing in the diagnosis of cryptococcal meningitis: a descriptive study. *J. Med. Microbiol.* 68, 1204–1210. doi: 10.1099/jmm.0.000994

Xu, L., Chen, X., Yang, X., Jiang, H., Wang, J., Chen, S., et al. (2023). Disseminated Talaromyces marneffei infection after renal transplantation: A case report and literature review. *Front. Cell Infect. Microbiol.* 13, 1115268. doi: 10.3389/fcimb.2023.1115268

Yang, H. H., He, X. J., Nie, J. M., Guan, S. S., Chen, Y. K., and Liu, M. (2022). Central nervous system aspergillosis misdiagnosed as Toxoplasma gondii encephalitis in a patient with AIDS: a case report. *AIDS Res. Ther.* 19, 40. doi: 10.1186/s12981-022-00468-x

Ye, P., Xie, C., Wu, C., Yu, C., Chen, Y., Liang, Z., et al. (2022). The application of metagenomic next-generation sequencing for detection of pathogens from dialysis effluent in peritoneal dialysis-associated peritonitis. *Perit Dial Int.* 42, 585–590. doi: 10.1177/08968608221117315

Yi, X., Lu, H., Liu, X., He, J., Li, B., Wang, Z., et al. (2024). Unravelling the enigma of the human microbiome: Evolution and selection of sequencing technologies. *Microb. Biotechnol.* 17, e14364. doi: 10.1111/1751-7915.14364

Zhan, L., Huang, K., Xia, W., Chen, J., Wang, L., Lu, J., et al. (2022). The diagnosis of severe fever with thrombocytopenia syndrome using metagenomic next-generation sequencing: case report and literature review. *Infect. Drug Resist.* 15, 83–89. doi: 10.2147/IDR.S345991

Zhang, Y., Amin, S., Lung, K. I., Seabury, S., Rao, N., and Toy, B. C. (2020a). Incidence, prevalence, and risk factors of infectious uveitis and scleritis in the United States: A claims-based analysis. *PLoS One* 15, e0237995. doi: 10.1371/journal.pone.0237995

Zhang, C., Cheng, H., Zhao, Y., Chen, J., Li, M., Yu, Z., et al. (2022). Evaluation of cell-free DNA-based next-generation sequencing for identifying pathogens in bacteremia patients. *Pol. J. Microbiol.* 71, 499–507. doi: 10.33073/pjm-2022-043

Zhang, Y., Cui, P., Zhang, H. C., Wu, H. L., Ye, M. Z., Zhu, Y. M., et al. (2020b). Clinical application and evaluation of metagenomic next-generation sequencing in suspected adult central nervous system infection. *J. Transl. Med.* 18, 199. doi: 10.1186/s12967-020-02360-6

Zhang, X. X., Guo, L. Y., Liu, L. L., Shen, A., Feng, W. Y., Huang, W. H., et al. (2019). The diagnostic value of metagenomic next-generation sequencing for identifying Streptococcus pneumoniae in paediatric bacterial meningitis. *BMC Infect. Dis.* 19, 495. doi: 10.1186/s12879-019-4132-y

Zhang, Y., Jiang, X., Ye, W., and Sun, J. (2023). Clinical features and outcome of eight patients with Chlamydia psittaci pneumonia diagnosed by targeted next generation sequencing. *Clin. Respir. J.* 17, 915–930. doi: 10.1111/crj.13681

Zhang, Q. Y., Jin, B., Feng, Y., Qian, K., Wang, H., Wan, C., et al. (2023). Etiological diagnostic value of metagenomic next-generation sequencing in peritoneal dialysis-related peritonitis. *Zhonghua Gan Zang Bing Za Zhi.* 39, 8–12. doi: 10.3760/cma.j.cn441217-20220729-00748

Zhang, W., Wu, T., Guo, M., Chang, T., Yang, L., Tan, Y., et al. (2019). Characterization of a new bunyavirus and its derived small RNAs in the brown citrus aphid, Aphis citricidus. *Virus Genes.* 55, 557–561. doi: 10.1007/s11262-019-01667-x

Zheng, X., Wu, C., Jiang, B., Qin, G., and Zeng, M. (2023). Clinical analysis of severe Chlamydia psittaci pneumonia: Case series study. *Open Life Sci.* 18, 20220698. doi: 10.1515/biol-2022-0698

Zhou, L., Guan, Z., Chen, C., Zhu, Q., Qiu, S., Liu, Y., et al. (2022). The successful treatment of Enterocytozoon bieneusi Microsporidiosis with nitazoxanide in a patient with B-ALL: A Case Report. *Front. Cell Infect. Microbiol.* 12, 1072463. doi: 10.3389/fcimb.2022.1072463

Zhou, Y., Wylie, K. M., El Feghaly, R. E., Mihindukulasuriya, K. A., Elward, A., Haslam, D. B., et al. (2016). Metagenomic approach for identification of the pathogens associated with diarrhea in stool specimens. *J. Clin. Microbiol.* 54, 368–375. doi: 10.1128/JCM.01965-15

Zhu, R., Hong, X., Zhang, D., Xiao, Y., Xu, Q., Wu, B., et al. (2023). Application of metagenomic sequencing of drainage fluid in rapid and accurate diagnosis of postoperative intra-abdominal infection: a diagnostic study. *Int. J. Surg.* 109, 2624–2630. doi: 10.1097/JS9.0000000000000500

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 24 December 2024
DOI 10.3389/fcimb.2024.1509630



OPEN ACCESS

EDITED BY
Qing Wei,
Genskey Co. Ltd, China

REVIEWED BY
Monica Viveros-Rogel,
National Institute of Medical Sciences and Nutrition Salvador Zubirán, Mexico
Tatina Todorova Todorova,
Medical University of Varna, Bulgaria

*CORRESPONDENCE
Jingzhe Han
420612049@qq.com

RECEIVED 11 October 2024
ACCEPTED 29 November 2024
PUBLISHED 24 December 2024

CITATION
Tang J, Wang K, Xu H and Han J (2024) Metagenomic next-generation sequencing of cerebrospinal fluid: a diagnostic approach for varicella zoster virus-related encephalitis. *Front. Cell. Infect. Microbiol.* 14:1509630. doi: 10.3389/fcimb.2024.1509630



# Metagenomic next-generation sequencing of cerebrospinal fluid: a diagnostic approach for varicella zoster virus-related encephalitis

Jin Tang[1], Kaimeng Wang[2], Haoming Xu[3] and Jingzhe Han[4]*

[1]Department of General Practice, The Second Affiliated Hospital of Xi 'an Jiaotong University, Xi'an, China, [2]Department of Neurology, Hebei General Hospital, Shijiazhuang, China, [3]Department of Geriatric Respiratory, Hebei General Hospital, Shijiazhuang, China, [4]Department of Neurology, Harrison International Peace Hospital, Hengshui, China

**Purpose:** Varicella zoster virus-related encephalitis (VZV-RE) is a rare and often misdiagnosed condition caused by an infection with the VZV. It leads to meningitis or encephalitis, with patients frequently experiencing poor prognosis. In this study, we used metagenomic next-generation sequencing (mNGS) to rapidly and accurately detect and identify the VZV pathogen directly from cerebrospinal fluid (CSF) samples, aiming to achieve a definitive diagnosis for encephalitis patients.

**Methods:** In this retrospective study, we analyzed the clinical characteristics and laboratory evaluations of 28 patients at the Harrison International Peace Hospital in Hebei, China, between 2018 and 2024. These patients were diagnosed with neurological disorders using mNGS techniques applied to CSF.

**Results:** In this cohort of 28 patients, 11 were females and 17 males, with a median age of 65 (IQR: 42.3-70). VZV-RE presented with a range of clinical manifestations, the most common being headaches (81.2%), fever>38°C (42.9%), and vomiting (42.9%). Less frequent symptoms include personality changes (10.7%), speech impairments (21.4%), cranial nerve involvement (21.4%), altered consciousness (17.9%) and convulsions (3.6%). Herpes zoster rash was observed in 35.7% of the cases. Neurological examination revealed nuchal rigidity in only 5 patients. CSF analysis indicated mild pressure and protein levels increase, with all patients having negative bacterial cultures. Abnormal electroencephalogram (EEG) findings were noted in 10.7% (N=3), and encephalorrhagia on Magnetic Resonance Imaging (MRI) was observed in 3.6%. VZV-RE was confirmed through mNGS analysis of CSF within three days of admission. All patients received empiric treatment with acyclovir or valacyclovir, with 21.4% receiving hormonotherapy, and 7.14% receiving immunoglobulin therapy. At the three-month follow-up, 10.7% of the patients had persistent neurologic sequelae, and the mortality rate was 3.6%.

**Conclusions:** Performing mNGS on CSF offers a rapidly and precisely diagnostic method for identifying causative pathogens in patients with VZV central nervous system (CNS) infections, especially when traditional CNS examination results are negative. Furthermore, the cases reported highlight the positive therapeutic effect of ganciclovir in treating VZV infections.

KEYWORDS

**varicella zoster virus, encephalitis, metagenomic next-generation sequencing, cerebrospinal fluid, diagnose**

# 1 Introduction

Varicella zoster virus (VZV), also known as human herpesvirus type 3, is a highly contagious, neurotropic, double-stranded DNA virus, with humans as its only natural host (Bhattacharya et al., 2024). Following the initial infection, which may range from asymptomatic to causing varicella or chickenpox, VZV remains dormant in the sensory ganglia of cranial and spinal nerves (Lewandowski et al., 2024). Reactivation of VZV can lead to various clinical syndromes, most commonly herpes zoster, characterized by a dermatomal vesicular rash and neuropathic pain known as shingles. Additionally, VZV reactivation can trigger neurological complications such as meningitis, encephalitis, cerebellitis, and cranial nerve palsies (e.g., Ramsay-Hunt syndrome). Moreover, it can lead to vasculopathy, myelopathy, and retinal necrosis, underscoring the virus's potential to affect multiple nervous system regions (Kennedy, 2023). Reactivation can occur in individuals with compromised immunity, such as the elderly or immunosuppressed, potentially leading to secondary central nervous system (CNS) infection.

Encephalitis has a high incidence and mortality rate. VZV is one of the most frequent etiologic agent of encephalitis, with or without associated herpes zoster rash (Mirouse et al., 2022). However, quick and accurate pathogen identification is the key for prompt clinical management (Chen et al., 2020). Cerebrospinal fluid (CSF) viral nucleic acid detection, including Polymerase Chain Reaction (PCR) and metagenomic Next-Generation Sequencing (mNGS), is currently the primary method for the etiological diagnosis of viral meningoencephalitis (Ramachandran and Wilson, 2020). However, PCR has limitations in clinical practice, with the process of virus isolation being labor-intensive, and there is a risk of missing detections due to the specificity of primers (Han et al., 2023). Broad-spectrum and unbiased metagenomic next-generation sequencing (mNGS) is commonly used for the nucleic acid detection of neurotropic viruses (Piantadosi et al., 2021). Since its introduction in 2014, mNGS has revolutionized diagnosing CNS infections, offering key advantages such as non-specific primer use, rapid detection, and exceptionally high sensitivity, particularly identifying previously unknown or infrequently encountered pathogens. Its application in analyzing CSF has marked a transformative shift in infectious disease diagnostics. The ability of mNGS to swiftly and accurately identify pathogens has dramatically improved the accuracy and efficiency of clinical diagnostics (Chen et al., 2024).

This study presents a summary of the characteristics of VZV-RE and proposes a mNGS detection method capable of rapidly and accurately identifying pathogens directly from CSF samples. This approach has been meticulously designed and optimized to address the challenges of diagnosing meningoencephalitis of indeterminate etiology, representing a significant advancement in clinical virology.

# 2 Materials and methods

## 2.1 Study design

This retrospective investigation was conducted at Harrison International Peace Hospital in Hebei, China, from 2018 to 2024, involving a total of 28 patients. The study population included patients who tested positive for VZV in their CSF samples using mNGS. Exclusion criteria: (1) Patients with concomitant central nervous system tumors; (2) Patients with encephalitis due to multiple pathogen infections; (3) Patients with autoimmune encephalitis; (4) Patients with incomplete clinical data. Subsequent follow-up evaluations were scheduled for the enrolled participants at one and three months post-hospital discharge.

The study was approved and conducted under the supervision of the Harrison International Peace Hospital Ethics Committee, Hengshui, Hebei (approval number: 2023109).

## 2.2 Diagnostic criteria of CNS infection

Encephalitis is diagnosed in patients with an unexplained altered mental state lasting more than 24 hours, accompanied by fever over 38°C within 72 hours, new-onset seizures, focal neurological signs, CSF pleocytosis, neuroimaging changes suggesting encephalitis, and EEG abnormalities consistent with the condition (Werner et al., 2016).

## 2.3 Data collection

Collect the following data for all enrolled patients: demographic information (age, gender), past medical history, history of alcohol use and immunosuppressive factors. Record clinical manifestations, physical examination findings, time from onset to lumbar puncture, laboratory test results (including blood and cerebrospinal fluid analysis). Obtain neuroimaging via computed tomography (CT) or magnetic resonance imaging (MRI), moreover electroencephalogram (EEG) results if performed. Analyze and assess the results of metagenomic testing in cerebrospinal fluid. Additionally, track treatment regimens(empiric antibiotics, antiviral medications, corticosteroids, immunoglobulins), complications, clinical outcomes (length of hospital stay, mortality), and follow-up results (sequelae).

## 2.4 mNGS of cerebrospinal fluid

All these patients underwent mNGS concurrently during their first lumbar puncture examination. The mNGS of CSF was meticulously carried out through the following streamlined process: (1) Specimen Collection: A volume of 1-2 mL of CSF was collected from patients via lumbar puncture, aliquoted into test tubes, and immediately stored at –80°C for 30 minutes prior to being utilized for mNGS. (2) Sample Extraction and Quality Control: Genomic DNA was extracted from the CSF samples using a micro-sample genomic DNA extraction kit (DP316, TIANGEN BIOTECH, Beijing, China). The DNA was then fragmented into 200-300 base pair (bp) fragments using a DNA cutting ultrasonic disruptor (Bioruptor Pico, Diagenode, Belgium). Following the quality control assessment of fragment sizes with a 2100 Bioanalyzer, the concentration of the DNA library was determined by quantitative PCR. (3) Library Construction: DNA libraries were constructed through a series of processes including end-repair, poly(A)-tailing, adapter ligation, and PCR amplification. Roller amplification technology was employed to amplify single-stranded circular DNA by 2–3 fold, resulting in the formation of DNA nanospheres. (4) Sequencing: The DNA nanospheres were loaded onto sequencing chips and sequenced using the BGISEQ-50 sequencing platform at the Institute of Medical Laboratory, Beijing Golden Key Gene Technology Co., Ltd., Beijing, China.

After the sequencing data underwent analysis and quality control, reads characterized by low quality, low complexity, or a length of less than 35 base pairs were discarded. The resulting high-quality sequencing data were aligned against the BWA human genome database. Subsequently, the remaining gene fragment data were compared against a microbial gene database to identify potential pathogenic microorganisms, including bacteria, viruses, and fungi. The primary indicators for analysis were the type of virus detected and the number of viral copy sequences. A positive result was defined as the detection of more than one viral copy sequence in the cerebrospinal fluid. All species included in the curated pathogen reference databases were collected from books, such as the Manual of Clinical Microbiology, Diagnosis and Illustration of Clinical Microbiology, and NCBI RefSeq genome database(ftp://ftp.ncbi.nlm.nih.gov/genomes/). Strictly only one typical high-quality representative strain whose genome sequence was downloaded from the NCBI RefSeq genome database or NCBI GenBank genome database was selected for each species. Currently, our curated database contains 12895 bacterial genomes or scaffolds, 11120 whole genome sequences of viral taxa, 1582 whole genome sequences of fungal taxa, 312 whole genome sequences of parasites, 184 mycoplasma and 177 mycobacterium.

## 2.5 Statistical methods

All data analyses were performed using SPSS 27.0 (IBM, Armonk, NY, USA). Continuous data following normal distribution were expressed as means ± standard deviations (Mean ± SD), while non-normally distributed continuous data were represented as medians with interquartile ranges (IQRs). Categorical variables were shown as counts and percentages for each category.

# 3 Results

## 3.1 The characteristics of the patients

The clinical characteristics of these 28 patients (11 females and 17 males) with VZV-RE are summarized in Table 1. The patients had underlying conditions such as diabetes, hypertension, and coronary heart disease, totaling eight individuals, all of whom were immunocompetent. Two patients had systemic lupus erythematosus and rheumatoid arthritis, both of whom were on long-term treatment regimens involving corticosteroids and immunosuppressive drugs, who were identified as immunocompromised patients. There were 15 individuals ≥ 65 years of age with a median age of 65 (IQR: 42.3-70 years). VZV-RE presented with a range of clinical manifestations. A herpes zoster rash was observed in 35.7% (N = 10), with 3 individuals presenting with the rash prior to the onset of symptoms. The most common neurological symptom was headache (81.2%, N = 22), subsequently vomiting and fever > 38°C (42.9%, N = 12). Additionally, some less common symptoms, such as personality changes (10.7%, N=3) and convulsions (3.6%, N = 1) were also observed. For other detailed information, refer to Table 1.

## 3.2 Diagnosis of VZV-RE through mNGS

We used CSF mNGS to identify the DNA sequence of VZV. The patients' lumbar puncture pressure values, laboratory evaluation, clinical treatments and outcome are presented in Table 2. The median CSF pressure was 150 mmH2O (IQR: 111.3-260mmH2O), indicating a range of pressure values in the cohort. The CSF WBC count was $109.5 \times 10^6$ cells/L (IQR: 31.3-315.3 cells/L). The median CSF protein level reflecting significant variability. The CSF glucose mainly within the normal range, though some values were near the lower limit. All patients exhibited negative bacterial cultures, and no Cryptococcus, Mycobacterium tuberculosis, or other specific pathogens were identified upon examination. Abnormal electroencephalogram (EEG) results were noted in 10.7% (N = 3) of

TABLE 1 Clinical and epidemiological characteristics of patients with neurological symptoms and VZV-mNGS-positive CSF (N=28).

| Characteristic | N (%) |
|---|---|
| Female sex | 11 (39.1) |
| Age (years), median (IQR) | 65 (42.3, 70) |
| Underlying diseases | 10 (35.7) |
| Immunosuppressive agents | 1 (3.6) |
| Time from neurological symptoms onset, days, means ± standard | 5.1 ± 3.5 |
| Time from onset to lumbar puncture, days, means ± standard | 6.5 ± 3.8 |
| Time from onset to medication, days, means ± standard | 5.6 ± 3.6 |
| Fever>38°C | 12 (42.9) |
| Headache | 22 (81.2) |
| Herpes zoster rash | 10 (35.7) |
| Vomiting | 12 (42.9) |
| Personality changes | 3 (10.7) |
| Speech impairment | 6 (21.4) |
| Disturbance of consciousness | 5 (17.9) |
| Cranial nerve involvement | 6 (21.4) |
| Convulsions | 1 (3.6) |
| Meningeal irritation | 5 (17.9) |

IQR, interquartile range.

TABLE 2 Laboratory evaluation and clinical outcome.

| Laboratory features | N (%) |
|---|---|
| CSF pressure (mmH2O), median (IQR) | 150 (111.3, 260) |
| CSF WBC ($\times 10^6$ cells/L), median (IQR) | 109.5 (31.3, 315.3) |
| CSF protein (0.15-0.45mg/dl), median (IQR) | 0.78 (0.61, 1.431) |
| CSF glucose (2.5-4.5mg/dl), median (IQR), median (IQR) | 3.33 (2.78, 3.99) |
| CSF Chloride (120-130mmol/L), means ± standard | 121.15 ± 3.69 |
| Negative bacterial culture | 28 (100) |
| Abnormal electroencephalogram | 3 (10.7) |
| Encephalorrhagia | 1 (3.6) |
| Empiric treatment with acyclovir/valacyclovir | 28 (100) |
| Hormonotherapy | 6 (21.4) |
| Immunoglobulin | 2 (7.14) |
| Neurologic sequelae at 3 months* | 3 (10.7) |
| Death | 1 (3.6) |

IQR, Interquartile range; *Defined as non-resolution of neurologic signs and symptoms, or general deterioration; CSF, Cerebrospinal fluid; WBC, White blood cells.

the subjects, suggesting the presence of slow-wave activity or mild abnormalities. Only one patient presented with hemorrhage involving the temporal lobe and brainstem. All patients commonly received empiric treatment with acyclovir or valacyclovir. Three individuals were administered methylprednisolone, while one individual was treated with dexamethasone for anti-inflammatory and immunomodulatory effects. One of them combined received immunoglobulin therapy for 5 days. Another patient received immunoglobulin therapy alone for 3 days, due to the detection of Gamma-Aminobutyric Acid antibodies receptor(GABAR) antibodies. At the time of hospital discharge, one patient had died. The majority of the remaining patients demonstrated full recovery during the subsequent three-month follow-up. Nonetheless, two patients persisted with symptoms of headache and dizziness. These outcomes highlight the severity of the cases and the importance of vigilant treatment and follow-up.

mNGS identified a number of pathogens with indeterminate clinical significance. Apart from VZV, we detected Enterococcus faecalis, Citrobacter koseri, Staphylococcus epidermidis, Cutibacterium acnes, Micrococcus luteus, and Moraxella osloensis. These findings, in conjunction with the patients' clinical presentations and laboratory tests, are considered background contaminants without pathogenic significance. Additionally, some patients had concurrent infections with Epstein-Barr virus (EBV), Human Herpesvirus 6 (HHV-6), Hepatitis B Virus(HBV), and so on. However, a review of the clinical data for these subjects indicates that these viral findings are unlikely to explain their clinical syndromes.

## 4 Discussion

Following primary infection, VZV may reactivate, causing trigeminal neuralgia and painful dermatomal rashes characteristic of Herpes Zoster, most commonly affecting the abdominal and dorsal areas (Sun et al., 2024). Encephalitis and meningitis are the most severe complications of VZV infection, typically occurring in older individuals, those with compromised immune systems, and in association with factors such as gender, concurrent malignancies, and a history of solid organ transplantation (Yuan et al., 2023). The clinical manifestations caused by VZV infection are varied, with some patients presenting atypical symptoms. In our case series, the primary symptoms included atypical presentations such as headache, fever, and vomiting, along with mental disturbances, dysarthria, and cranial nerve damage. The absence of a rash or its delayed appearance months after the onset of symptoms leads to misdiagnosis, complicating accurate diagnosis (Dulin et al., 2024). Although reactivation of VZV is more commonly reported in the elderly (Deobhakta and Gilden, 2024), in our study, one patient was a 15-year-old immunocompromised individual with systemic lupus erythematosus (SLE) undergoing steroid therapy. Additionally, five cases (17.9%) involved patients under 50-year-old, none with immunosuppressive conditions or extraordinary medical history. These patients presented with VZV meningitis without a rash, suggesting that VZV reactivation can occur in immunocompetent young individuals, potentially spreading directly to the leptomeninges and causing aseptic meningitis. This observation

aligns with the findings by Lee et al (Lee et al., 2021). The only fatal case involved a 75-year-old female with a history of diabetes and coronary heart disease. Despite long-term treatment for these chronic conditions, her immune function was normal. She was admitted with fever and mental disturbances and was initially suspected of having diabetic ketoacidosis due to elevated ketone bodies. After corrective treatment, the patient remained comatose with worsening hypoxemia. Brain MRI revealed bilateral temporal lobe necrotic foci with hemorrhagic transformation. Routine CSF analysis suggested viral encephalitis and VZV-RE was confirmed through mNGS of the CSF. Despite a month of antiviral therapy in combination with corticosteroids and immunoglobulins, the patient remained unconscious and ultimately died from multiple organ dysfunction and respiratory failure. This case highlights that prognosis is closely linked to age, underlying conditions, and the presence of cerebral hemorrhagic complications. Patients with diabetes are predisposed to diabetes-related cerebrovascular issues and tend to have poorer vascular health. Moreover, VZV-RE in elderly patients is associated with an elevated risk of cerebral hemorrhage, which significantly worsens outcomes. The involvement of cortical hemorrhage in VZV-RE may indicate a poor prognosis. Early recognition and prompt antiviral therapy are crucial for improving patient outcomes. Previous studies have shown that VZV reactivation can spread axonal spread along cranial nerve ganglia, infecting arterial adventitia and leading to vasculitis. Cerebral hemorrhage contributes to a 60% mortality rate in such cases (Wu et al., 2022).

In this study, all patients underwent lumbar puncture, revealing an inflammatory response with elevated WBC count of CSF. This confirmed the diagnosis of infectious meningitis despite negative Gram stain, India ink staining, acid-fast staining as well as CSF culture results. Conventional diagnostic methods often present challenges in identifying the causative agents of encephalitis. Although PCR detection of VZV in CSF is the gold standard for the diagnosis of VZV-RE (Kriger et al., 2024). In contrast, mNGS technology allows for the direct detection of pathogenic microorganisms from brain tissue and CSF, significantly reducing diagnostic time, particularly for patients with relapsing encephalitis. Even with extensive clinical laboratory testing, approximately 60% of acute meningoencephalitis cases remain undiagnosed regarding etiology (Wang et al., 2022). Although we did not perform PCR validation, there is literature supporting that CSF VZV-specific PCR may not increase the sensitivity for diagnosing VZV CNS infections when mNGS is conducted on CSF samples. CSF mNGS is the most sensitive microbial test for diagnosing VZV central nervous system infections and can unexpectedly identify pathogens that traditional diagnostic tests fail to detect (Zhu et al., 2021). Recently, mNGS has been utilized as a powerful tool for detecting pathogens in CNS infectious, particularly in cases involving latent or chronic viral infections such as those caused by the Herpesviridae family. This technology, with its unique ability to reduce diagnostic time to within three days (Zhang et al., 2024), was applied to all patients in our study within three days of admission. This rapid turnaround time underscores the pivotal role of mNGS in identifying pathogenic CNS infections, thereby underlining the significance of the technology.

mNGS provides the ability to identify pathogens within prior knowledge of the target. Theoretically, with sufficiently long read lengths and multiple hits to the microbial genome, NGS can accurately pinpoint the causative pathogen. Our study detected many unique reads corresponding to neurotropic viral genomes in the patient's CSF. The advancement and application of mNGS have given medical laboratory technicians an unprecedented ability to identify the pathogens responsible for encephalitis, as mNGS detects pathogens in a target-independent manner. This technology has the potential to drastically reduce the time required for diagnosis (Wang et al., 2020). Although mNGS is a high-cost test that requires specialized equipment and laboratory infrastructure, which may limit the widespread adoption of this technology. Concurrently, the costs associated with this testing are decreasing annually due to continuous technological improvements. It is anticipated that in the near future, the costs will be within a reasonable range, facilitating its broader application in clinical settings.

We acknowledge several limitations in our study. The sample size was insufficient for epidemiological analysis and group comparisons. All subjects underwent mNGS testing only, without PCR confirmation. Further research on VZV-RE is necessary, with an emphasis on expanding the sample size and thoroughly analyzing the clinical and imaging characteristics to facilitate early diagnosis and treatment and ultimately improving patient outcomes.

## 5 Conclusion

The clinical manifestations of VZV-RE are varied, and VZV reactivation causing aseptic meningitis in immunocompetent adults, with or without herpes zoster, is more common than previously recognized. Prognosis tends to be worse in elderly patients, particularly those with cerebral hemorrhage. This study emphasizes the feasibility of using mNGS on CSF as a diagnostic tool for CNS infections. In theory, unbiased NGS can identify all potential pathogens in a single test, which is highly significant for providing rapid, accurate diagnosis and facilitating targeted antimicrobial treatment.

## Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/bioproject/1199460.

## Ethics statement

The studies involving humans were approved by the Harrison International Peace Hospital Ethics Committee, Hengshui, Hebei. (approval 2023109). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent

was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

## Author contributions

JT: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. KW: Data curation, Formal analysis, Writing – original draft. HX: Data curation, Formal analysis, Investigation, Writing – original draft. JH: Supervision, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Hebei Province Medical Science Research Project (grant number: 20240959).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note



## References

Bhattacharya, A., Jan, L., Burlak, O., Li, J., Upadhyay, G., Williams, K., et al. (2024). Potent and long-lasting humoral and cellular immunity against varicella zoster virus induced by mRNA-LNP vaccine. *NPJ Aging Health* 4, 1–8. doi: 10.1038/s41541-024-00865-5

Chen, W., Liu, G., Cui, L., Tian, F., Zhang, J., Zhao, J., et al. (2024). Evaluation of metagenomic and pathogen-targeted next-generation sequencing for diagnosis of meningitis and encephalitis in adults: A multicenter prospective observational cohort study in China. *J. Infect.* 88 (5), 106143. doi: 10.1016/j.jinf.2024.106143

Chen, L., Xu, Y., Liu, C., Huang, H., Zhong, X., Ma, C., et al. (2020). Clinical features of aseptic meningitis with varicella zoster virus infection diagnosed by next-generation sequencing: case reports. *BMC Infect. Dis.* 20 (1), 435. doi: 10.1186/s12879-020-05155-8

Deobhakta, A. A., and Gilden, D. H. (2024). Potent and long-lasting humoral and cellular immunity against varicella zoster virus induced by mRNA-LNP vaccine. *NPJ Aging Health* 4, 1–8. doi: 10.1038/s41541-024-00865-5

Dulin, M., Chevret, S., Salmona, M., Jacquier, H., Bercot, B., Molina, J.-M., et al. (2024). New insights into the therapeutic management of varicella zoster virus meningitis: A series of 123 polymerase chain reaction–confirmed cases. *Open Forum Infect. Dis.* 11, ofae340. doi: 10.1093/ofid/ofae340

Han, J., Si, Z., Wei, N., Cao, D., Ji, Y., Kang, Z., et al. (2023). Next-generation sequencing of cerebrospinal fluid for the diagnosis of VZV-associated rhombencephalitis. *J. Integr. Neurosci.* 22, 36. doi: 10.31083/j.jin2202036

Kennedy, P. G. E. (2023). The spectrum of neurological manifestations of Varicella-zoster virus reactivation. *Viruses* 15, 968. doi: 10.3390/v15080968

Kriger, O., Dovrat, S., Fratty, I. S., Leshem, E., Oikawa, M. T., Sofer, D., et al. (2024). Don't rash it! The clinical significance of positive Varicella zoster virus PCR in cerebrospinal fluid of patients with neurological symptoms. *J. Clin. Virol.* 171, 105648. doi: 10.1016/j.jcv.2024.105648

Lee, G. H., Kim, J., Kim, H. W., and Cho, J. W. (2021). Herpes simplex viruses (1 and 2) and varicella-zoster virus infections in an adult population with aseptic meningitis or encephalitis: a nine-year retrospective clinical study. *Medicine* 100, e27856. doi: 10.1097/MD.0000000000027856

Lewandowski, D., Toczylowski, K., Kowalska, M., Krasnodębska, M., Krupienko, I., Nartowicz, K., et al. (2024). Varicella-zoster disease of the central nervous system in immunocompetent children: case series and a scoping review. *Vaccines* 12, 183. doi: 10.3390/vaccines12020183

Mirouse, A., Sonneville, R., Razazi, K., Merceron, S., Argaud, L., Bigé, N., et al. (2022). Neurologic outcome of VZV encephalitis one year after ICU admission: a multicenter cohort study. *Ann. Intensive Care* 12, 32. doi: 10.1186/s13613-022-01002-y

Piantadosi, A., Mukerji, S. S., Ye, S., Leone, M. J., Freimark, L. M., Park, D., et al. (2021). Enhanced virus detection and metagenomic sequencing in patients with meningitis and encephalitis. *mBio* 12, e01143–e01121. doi: 10.1128/MBio.01143-21

Ramachandran, P. S., and Wilson, M. R. (2020). Metagenomics for neurological infections — expanding our imagination. *Nat. Rev. Neurol.* 16, 547–556. doi: 10.1038/s41582-020-0374-y

Sun, Z., Chen, E., Yi, D., and Xiao, S. (2024). Alveolar osteonecrosis and tooth exfoliation following herpes zoster infection: A case report and review of the literature. *Aust. Dent. J.* doi: 10.1111/adj.12748

Wang, Q., Wang, K., Zhang, Y., Lu, C., Yan, Y., Huang, X., et al. (2020). Neonatal Ureaplasma parvum meningitis: a case report and literature review. *Transl. Pediatr.* 9, 174–179. doi: 10.21037/tp2020.04.06

Wang, L. P., Yuan, Y., Lu, Q. B., Shi, L. S., Ren, X., Zhou, S. X., et al. (2022). Etiological and epidemiological features of acute meningitis or encephalitis in China: a nationwide active surveillance study. *Lancet Reg. Health West Pac.* 20, 100361. doi: 10.1016/j.lanwpc.2021.100361

Werner, R. N., Nikkels, A. F., Marinović, B., Schäfer, M., Czarnecka-Operacz, M., Agius, A. M., et al. (2016). European consensus-based (S2k) guideline on the management of herpes zoster - guided by the European Dermatology Forum (EDF) in cooperation with the European Academy of Dermatology and Venereology (EADV), part 1: diagnosis. *J. Eur. Acad. Dermatol. Venereol* 31, 9–19. doi: 10.1111/jdv.13995

Wu, H., Wang, R., Li, Y., Sun, X., Li, J., and Bi, X. (2022). Cerebrovascular complications after adult-onset varicella-zoster virus encephalitis in the central nervous system: A literature review. *Neuropsychiatr. Dis. Treat* 18, 449–462. doi: 10.2147/NDT.S343846

Yuan, Y., Zhang, Y., Wang, J., Liu, H., Zhang, H., and Yan, Y. (2023). Immune changes and their relationship with prognosis in patients with varicella-zoster virus encephalitis/encephalitis. *Am. J. Transl. Res* 15 (2), 1421–1429.

Zhang, M., Chen, L., Zhao, H., Qiao, T., Jiang, L., Wang, C., et al. (2024). Metagenomic next-generation sequencing for diagnosis of infectious encephalitis and meningitis: a retrospective study of 90 patients. *Neurol. Res.* 46 (2), 187–194. doi: 10.1080/01614412.2024.2175902

Zhu, Y., Xu, M., Ding, C., Peng, Z., Wang, W., Sun, B., et al. (2021). Metagenomic next-generation sequencing vs traditional microbiological tests for diagnosing varicella-zoster virus central nervous system infection. *Front. Public Health* 9. doi: 10.3389/fpubh.2021.738412

frontiers | Frontiers in Cellular and Infection Microbiology

TYPE Original Research
PUBLISHED 18 February 2025
DOI 10.3389/fcimb.2025.1497198



OPEN ACCESS

EDITED BY
Qing Wei,
Genskey Co. Ltd, China

REVIEWED BY
Zheng Jin Tu,
Cleveland Clinic, United States
Kaiyu Cui,
Hubei University of Chinese Medicine, China

*CORRESPONDENCE
Bing Wei
weibing1999@hotmail.com
Jie Ma
majie@medmail.com.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 16 September 2024
ACCEPTED 27 January 2025
PUBLISHED 18 February 2025

CITATION
Yang K, Zhao J, Wang T, Wang Z, Sun R, Gu D, Liu H, Wang W, Zhang C, Zhao C, Guo Y, Ma J and Wei B (2025) Clinical application of targeted next-generation sequencing in pneumonia diagnosis among cancer patients.
*Front. Cell. Infect. Microbiol.* 15:1497198.
doi: 10.3389/fcimb.2025.1497198



# Clinical application of targeted next-generation sequencing in pneumonia diagnosis among cancer patients

Ke Yang[1†], Jiuzhou Zhao[1†], Tingjie Wang[1], Zhizhong Wang[1], Rui Sun[1], Dejian Gu[2], Hao Liu[2], Weizhen Wang[1], Cuiyun Zhang[1], Chengzhi Zhao[1], Yongjun Guo[1], Jie Ma[1*] and Bing Wei[1*]

[1]Department of Molecular Pathology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China, [2]Medical Department, Geneplus-Beijing Co., Ltd., Beijing, China

**Background:** Cancer patients are highly susceptible to infections due to their immunocompromised state from both the malignancy and intensive treatments. Accurate and timely identification of causative pathogens is crucial for effective management and treatment. Targeted next-generation sequencing (tNGS) has become an important tool in clinical infectious disease diagnosis because of its broad microbial detection range and acceptable cost. However, there is currently a lack of systematic research to evaluate the diagnostic value of this method in cancer patients.

**Methods:** To evaluate the diagnostic value of tNGS for cancer patients with pneumonia, a retrospective analysis was conducted on 148 patients with suspected pneumonia who were treated at the Henan Cancer Hospital. The tNGS results were compared with conventional microbiological tests (CMT) and clinical diagnoses based on symptoms and imaging studies to assess the diagnostic performance of tNGS in cancer patients with pneumonia.

**Results:** Among these 148 patients, 130 were ultimately diagnosed with pneumonia. tNGS demonstrated significantly higher sensitivity (84.62% vs. 56.92%) and diagnostic accuracy (85.81% vs. 62.16%) compared to the CMT method. The tNGS method identified more pathogens than CMT method (87.50% vs 57.14%), regardless of whether they were bacteria, fungi, or viruses, primarily due to its broader pathogen detection range and higher sensitivity compared to the CMT method. tNGS had significantly higher diagnostic accuracy for *Pneumocystis jirovecii* and *Legionella pneumophila* than the CMT method,

but for most pathogens, tNGS showed higher sensitivity but with a correspondingly lower specificity compared to CMT.

**Conclusion:** tNGS demonstrates higher sensitivity and a broader pathogen detection spectrum compared to CMT, making it a valuable diagnostic tool for managing pneumonia in cancer patients.

KEYWORDS

**targeted next-generation sequencing, pathogen, cancer patients, CMT, pneumonia**

# 1 Introduction

Cancer patients are highly susceptible to infections due to their immunocompromised state, which results from both the malignancy and the intensive treatments they undergo. Accurate and timely identification of the causative pathogens is crucial for the effective management and treatment of these infections. Traditional diagnostic methods, such as blood cultures, often have significant limitations, including low sensitivity (Babady, 2016), limited scope of pathogen detection (Hill et al., 2024; Rodino and Simner, 2024) and prolonged turnaround times. These limitations can delay appropriate treatment and negatively impact patient outcomes (Deng et al., 2022; Gu et al., 2023).

Multiplex polymerase chain reaction(PCR) is a commonly used molecular detection technique in clinical settings. Its ultra-high sensitivity provides support for precise clinical diagnosis (Hou et al., 2020; Zhang et al., 2022). However, considering its limited target coverage, the clinical application is somewhat restricted. Next-generation sequencing (NGS) has emerged as a powerful tool in the detection of infectious agents. Metagenomic NGS(mNGS) has garnered significant clinical acclaim in recent years, offering the unprecedented ability to detect all nucleic acids in a sample without prior assumptions (Chiu and Miller, 2019). Numerous studies have demonstrated that mNGS far surpasses traditional diagnostic methods in terms of diagnostic value for infections (Liang et al., 2022; Serpa et al., 2022; Fourgeaud et al., 2024). However, the cost of over 500 USD per test has led to its recommendation primarily for the diagnosis of critically ill patients.

As the new favorite in clinical settings following metagenomic NGS (mNGS), the clinical performance of tNGS has been demonstrated through several studies. In the diagnosis of respiratory tract infections, the diagnostic performance of tNGS is comparable to that of mNGS, but the cost is less than half (Wei et al., 2024; Zhang et al., 2024). In patients with lower respiratory tract infections and pneumonia, tNGS has shown better sensitivity than traditional methods and has diagnostic value comparable to that of mNGS (Gaston et al., 2022; Li et al., 2022; Zhang et al., 2023; Zhang et al., 2024), but its specific application in the context of cancer patients warrants further investigation. This population presents unique challenges, such as a higher prevalence of opportunistic infections (Deng et al., 2023) and the presence of non-pathogenic microorganisms that may complicate the interpretation of sequencing results.

In this study, we aim to evaluate the diagnostic performance of tNGS in cancer patients with pneumonia. We compared tNGS results with those obtained from conventional microbiological tests, and a clinical diagnosis to assess its sensitivity, specificity, and accuracy. Furthermore, we have analyzed the pathogen spectrum detected by tNGS in this patient population. Through this comprehensive analysis, we hope to elucidate the potential of tNGS as a reliable and efficient diagnostic tool for managing infections in cancer patients.

# 2 Methods

## 2.1 Sample enrollment and microbiology testing

To evaluate the diagnostic performance of tNGS in pneumonia of cancer patients, a retrospective study was conducted, enrolling 237 cancer patients with suspected pneumonia who had undergone tNGS testing at the Henan Cancer Hospital from April 2022 to April 2024 (Figure 1). Clinical information and biochemical indicators were collected through the hospital's electronic medical record system, and ultimately, 148 patients with complete clinical information and diagnostic outcomes were included in the analysis. Patients were eligible for enrolment if they (1) were at least 18 years of age; (2) were undergoing cancer treatment; (3) had suspected pneumonia. Patients with suspected pneumonia met both of the following criteria: (1) at least one compatible symptom, such as new-onset fever, cough, or dyspnea; (2) new-onset radiological findings on chest images.

This study received approval from the Ethics Committees to ensure compliance with ethical guidelines. Patients' identification remained anonymous throughout the study, and informed consent was waived due to the retrospective and observational nature of the research. All samples were aliquoted into 1ml portions and stored at -80°C after clinical collection.

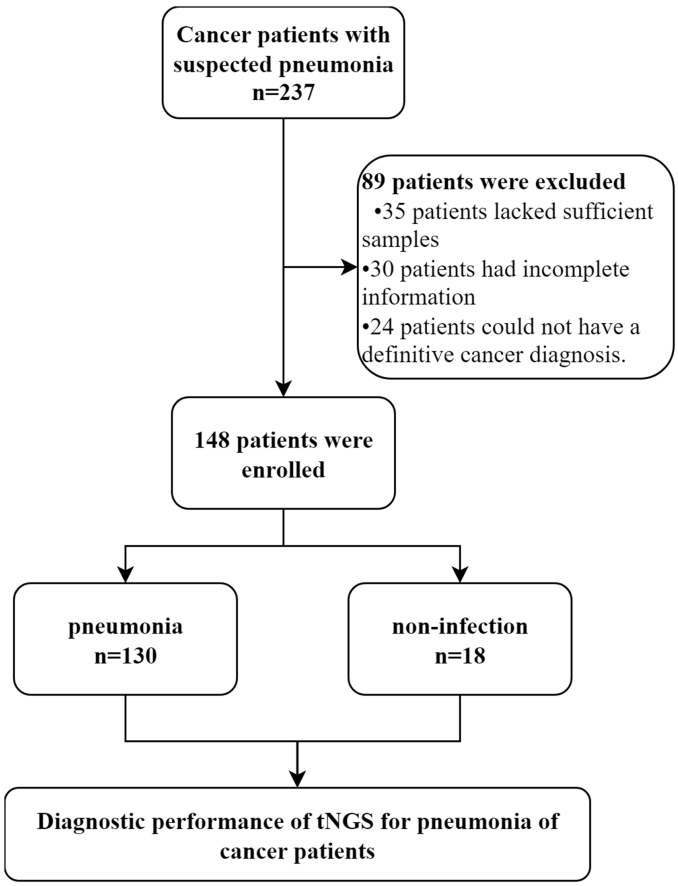



FIGURE 1
Flowchart of samples analyzed by tNGS. We conducted a retrospective analysis of eligible patients and ultimately included 148 participants. Patients were divided into pneumonia and non-infection groups based on their final clinical diagnoses to explore the diagnostic performance of tNGS.

## 2.2 Conventional microbiological tests

All CMTs were performed on every sample, except for viral PCRs, which were performed at the clinician's discretion. A total of 148 patients underwent BALF sampling, from which a portion was aliquoted for testing by tNGS. Additionally, this BALF was subjected to CMTs alongside other relevant samples, such as blood and urine. The CMTs encompassed a range of diagnostic methods including culture, antigen detection, multiplex PCR, and Xpert PCR assays.

## 2.3 tNGS workflow for BALF

tNGS was performed once the sample was obtained. As previous reported (Wei et al., 2024; Zhang et al., 2024), BALF samples with high viscosity were diluted 1:1 with 0.1 M dithiothreitol before nucleic acid extraction. A volume of 400μL of BALF, lysis buffer, protease K mixture, binding buffer, and 1.2 g glass bead were agitated vigorously at 4500 rpm for a total of 130 s by FastPrep-24™ 5G Instrument (MP Biomedical, CA, USA). Nucleic acid extraction was performed using the VAMNE Magnetic Pathogen DNA/RNA kit (Vazyme, Nanjing, China). Then, the RNA was reverse-transcribed into cDNA using Hieff NGS® ds-cDNA Synthesis Kit (Yeasen, Shanghai, China). Nucleic acids were quantified using a Qubit 3.0 fluorometer. A549 human cells (GenePlus, Suzhou, China) were used as negative controls (NTC) to detect contamination, and A549 human cells spiked with *Staphylococcus aureus* (BeNa Culture Collection, Beijing, China) were used as positive controls (PTC). Following extraction, cDNA synthesis and library preparation were performed with the HieffNGS®C37P4 One PotcDNA&gDNA Library Prep Kit (Yeasen, Shanghai, China). Then, the cDNA was taken through library enrichment with NadPrep® NanoBlockers (Nanodigmbio, Nanjing, China) reagents to generate the product for targeted sequencing. Eight uniquely barcoded libraries were pooled to hybridize and capture by specific biotinylated probes for 4 hours using Geneplus design probes. Products were quantified with a Qubit 3.0 instrument using DNA HS Assay Kit (Vazyme, Nanjing, Jiangsu, China). Products were stored at −20°C until Sequencing. After library construction, sequencing was performed on the Gene +seq 100 platform(GenePlus, Suzhou, China). The sequencing read length was set to 100 base pairs (bp), with a preset data of 5 million reads. Clean reads were obtained by removing sequencing adapters, low-quality reads, or reads below 35 bp using fastp (version 0.23.1). The remaining reads were aligned to the human reference (hg38)

using Burrow-Wheeler Aligner (version 0.7.12-r1039) and human reads were filtered. The filtered reads were compared with the self-built pathogenic microorganism database, and the retained results were annotated. Microbial reads within the target range were normalized to reads per million (RPM), and only microorganisms above a predefined threshold were initial reported in this study. The threshold was set at RPM ≥6 for common pathogens (excluding mycobacteria) and ≥0.5 for fungi and mycobacteria. A manual review is conducted. The common colonizing microorganisms within the respiratory tract target range was reported separately from other microorganisms in different sections to avoid interfering with clinical judgment (Li et al., 2023; Zhang et al., 2024).

## 2.4 Clinical diagnosis as the reference standard

Two physicians with extensive clinical experience, each having worked in the respiratory department for over 5 years, independently reviewed all inpatient medical information and microbiological test results of the patients, including tNGS. Initially, they determined whether the patients had infectious diseases. Subsequently, they identified the pathogenic microorganisms based on clinical manifestations, laboratory tests, imaging examinations, microbiological test results (including CMT and tNGS), and treatment responses. Discrepancies between the two reviewers were first resolved through in-depth discussions; if consensus could not be reached, another senior reviewer was consulted.

## 2.5 Statistical methods

For continuous variables, report as the median and interquartile range (IQR). Categorical variables are represented by frequency and percentage. Inter-group comparisons are made using the unpaired t-test or the Mann-Whitney U test. For comparisons between groups of categorical variables, the chi-square test is used. In evaluating diagnostic performance, sample consistency is judged by referring to previous studies (Blauwkamp et al., 2019); partial pathogen consistency is considered as consistent. In the calculation of diagnostic performance, sensitivity, specificity, and accuracy are computed using the standard proportion formula, and the 95% confidence intervals for these proportions are determined using the Wilson method. To compare the differences in diagnostic performance and analyze diagnostic accuracy between two groups, a t-test can be employed.

All figures were drawn using GraphPad Prism version 9.5.0 (GraphPad Software LLC., San Diego, CA, USA). All analyses were performed with SPSS version 26.0 (SPSS Inc., Chicago, Illinois, USA) (Blauwkamp et al., 2019). A p-value less than 0.05 was considered statistically significant.

# 3 Results

## 3.1 Patient characteristics

A total of 237 patients with suspected pneumonia were considered for inclusion. Of these, 89 patients were excluded, some

TABLE 1 Baseline characteristics of the 148 patients enrolled.

| Characteristic | Number (n = 148) | Pneumonia (n=130) | non-infection (n=18) | p value |
|---|---|---|---|---|
| **Median age, years** | 61(16-83) | 60(16-81) | 66(41-83) | 0.122 |
| **Gender, n(%)** | | | | 0.928 |
| Male | 89 (60.14) | 78(60.00) | 11(61.11) | |
| Female | 59 (39.86) | 52(40.00) | 7(38.89) | |
| Laboratory findings | | | | |
| WBC ($10^9$/L), median (IQR) | 7.41 (5.75, 10.56) | 7.49(5.73, 10.76) | 6.41(5.58, 9.25) | 0.609 |
| NEUT%, median (IQR) | 73.44 (63.21, 82.63) | 73.44(63.35, 83.45) | 67.75(61.20, 80.13) | 0.371 |
| LYM%, median (IQR) | 17.42 (8.97, 27.82) | 16.52(8.65, 27.65) | 19.95(10.55, 28.925) | 0.446 |
| CRP(nmol/L),median (IQR) | 87.88(49.31, 158.13) | 89.32(51.22, 153.33) | 86.13(48.34, 153.27) | 0.433 |
| PCT(ng/mL),median (IQR) | 1.41(0.48,5.32) | 1.57(0.46, 5.77) | 1.26(0.37, 5.21) | 0.189 |
| Cancer types, n(%) | | | | |
| lung cancer | 103 (69.59) | 98(75.38) | 15(83.33) | 0.238 |
| hematological malignancies | 27 (18.24) | 27(20.77) | 0(0) | 0.03 |
| gastrointestinal cancer | 13 (8.78) | 11(8.46) | 2(11.11) | 0.709 |
| other cancers | 5 (3.38) | 4(3.08) | 1(5.56) | 0.585 |

IQR, interquartile range; WBC, White blood cell; PCT, procaicitonin; CRP, C-reactive protein; LYM, Lymphocyte; NEUT, Neutrophil. Other cancer types included two patients with ovarian cancer, two with breast cancer, and one with bladder cancer.
Bold formatting is used to denote the primary groupings.

TABLE 2 Diagnostic performance of tNGS in 148 patients.

| | Methods | Sensitivity (95% CI) | Specificity (95%CI) | PPV(95% CI) | NPV(95%CI) | Accuracy (95%CI) | p-value |
|---|---|---|---|---|---|---|---|
| **All samples** | **CMT** | 56.92% (48.33%-65.12%) | 100.00% (82.41%-100.00%) | 100.00% (95.07%-100.00%) | 24.32% (15.98%-35.21%) | 62.16% (54.13%-69.57%) | <0.0001 |
| | **tNGS** | 84.62% (77.43%-89.81%) | 94.44% (74.24%-99.01%) | 99.10% (95.07%-99.84%) | 45.95% (31.04%-61.62%) | 85.81% (79.28%-90.53%) | |
| **Fungal pneumonia** | **CMT** | 52.63% (37.26%-67.52%) | 99.09% (95.03%-99.84%) | 95.24% (77.33%-99.15%) | 85.83% (78.71%-90.84%) | 87.16% (80.82%-91.63%) | 0.86 |
| | **tNGS** | 81.58% (66.58%-90.78%) | 88.18% (80.82%-92.96%) | 70.45% (55.78%-81.84%) | 93.27% (86.75%-96.70%) | 86.49% (80.05%-91.08%) | |
| **Viral pneumonia** | **CMT** | 76.09% (62.06%-86.09%) | 100.00% (96.37%-100.00%) | 100.00% (90.11%-100.00%) | 90.27% (83.41%-94.48%) | 92.57% (87.18%-95.80%) | 0.0005 |
| | **tNGS** | 89.13% (76.96%-95.27%) | 78.43% (69.50%-85.30%) | 65.08% (52.75%-75.67%) | 94.12% (86.96%-97.46%) | 81.76% (74.76%-87.15%) | |
| **Bacterial pneumonia** | **CMT** | 57.53% (46.10%-68.22%) | 89.33% (80.34%-94.50%) | 84.00% (71.49%-91.66%) | 68.37% (58.62%-76.73%) | 73.65% (66.02%-80.08%) | 0.17 |
| | **tNGS** | 84.93% (75.00%-91.37%) | 76.00% (65.22%-84.25%) | 77.50% (67.21%-85.27%) | 83.82% (73.31%-90.72%) | 80.41% (73.28%-86.00%) | |

tNGS, targeted Next-generation sequencing; CMT, conventional microbiological tests; PPV, positive predictive value; NPV, negative predictive value; NA, not available. The difference of diagnostic accuracy based on the t-test of tNGS and CMT was marked as p-value.
Bold formatting is applied to emphasize the subgroups and detection methods.

because they were ultimately diagnosed as non-cancer patients, and others due to insufficient data for further analysis. One hundred and forty-eight patients met the criteria and were included in the final analysis. This group comprised 60.14% male patients. The patients were primarily composed of three types of cancer patients, with lung cancer accounting for 69.59%, hematological malignancies accounting for 18.24%, and gastrointestinal cancer accounting for 8.78%. Additionally, there were 5 patients mainly with ovarian, bladder, and breast cancer (Table 1). When comparing pneumonia and non-infection patients, no significant differences in characteristics were observed between the two groups. However, the proportion of hematological malignancies was higher in the pneumonia patient group.

## 3.2 Clinical diagnostic performance of tNGS and CMT

In the clinical diagnosis, 130 patients were confirmed to have infections, while 18 patients were non-infectious. tNGS identified 110 infected patients and found no microorganisms in 17/18 non-infectious patients, ultimately showing a sensitivity of 84.62% and a specificity of 94.44%, with a diagnostic accuracy of 85.81% (Table 2). The sensitivity of CMT in diagnosing was 56.92%, with a specificity of 100%, and a diagnostic accuracy of 62.16%. The diagnostic accuracy of tNGS was significantly higher than that of the CMT method. This was mainly because tNGS had significantly higher sensitivity for fungal and bacterial infections compared to the CMT method ($p<0.0001$). However, due to the excessive sensitivity of tNGS, there was no significant difference in accuracy between tNGS and CMT for bacterial and fungal infections. In terms of viral infections, PCR, as the diagnostic standard, had a higher diagnostic accuracy than the tNGS method. Nevertheless, the broader coverage of viruses by tNGS provided better sensitivity. Therefore, tNGS was a more sensitive method than CMT overall.

## 3.3 Pathogens identification performance of tNGS

The study further explored the diagnostic accuracy of the targeted Next-Generation Sequencing (tNGS) method for various pathogens. Ultimately, 168 pathogens were identified in these patients. Of these, tNGS identified 147(87.50%), while CMT identified 96(57.14%). The tNGS method detected a significantly higher number of pathogens compared to CMT ($p<0.0001$), including bacteria, fungi, and viruses (Figure 2A). A further comparison of the pathogen distribution profiles revealed that tNGS had higher sensitivity for detecting *Pneumocystis jirovecii* among fungi (Figure 2B). For viruses, tNGS was able to identify additional RNA viruses not covered by multiplex PCR, such as *Human respiratory viruses*, *Human metapneumovirus*, and *adenovirus*. In terms of bacterial detection, tNGS covered difficult-to-culture or unexpected species such as *Brucella melitensis*, *non-tuberculous mycobacteria*, and *Legionella pneumophila*. However, tNGS was less sensitive in detecting *Aspergillus* spp compared to the CMT. Further evaluation of the diagnostic performance of pathogens with a frequency count exceeding five times (*Mycobacterium tuberculosis*:4) showed higher diagnostic accuracy for *Pneumocystis jirovecii*, and *Legionella pneumophila*, but lower accuracy for *Aspergillus* spp, *Streptococcus pneumoniae*, and *rhinovirus* compared to CMT(Table 3). Additionally, in the diagnosis of *Epstein-Barr virus* (EBV), tNGS was able to detect more EBV in samples, resulting in significantly lower diagnostic accuracy than PCR. We conducted a comparative

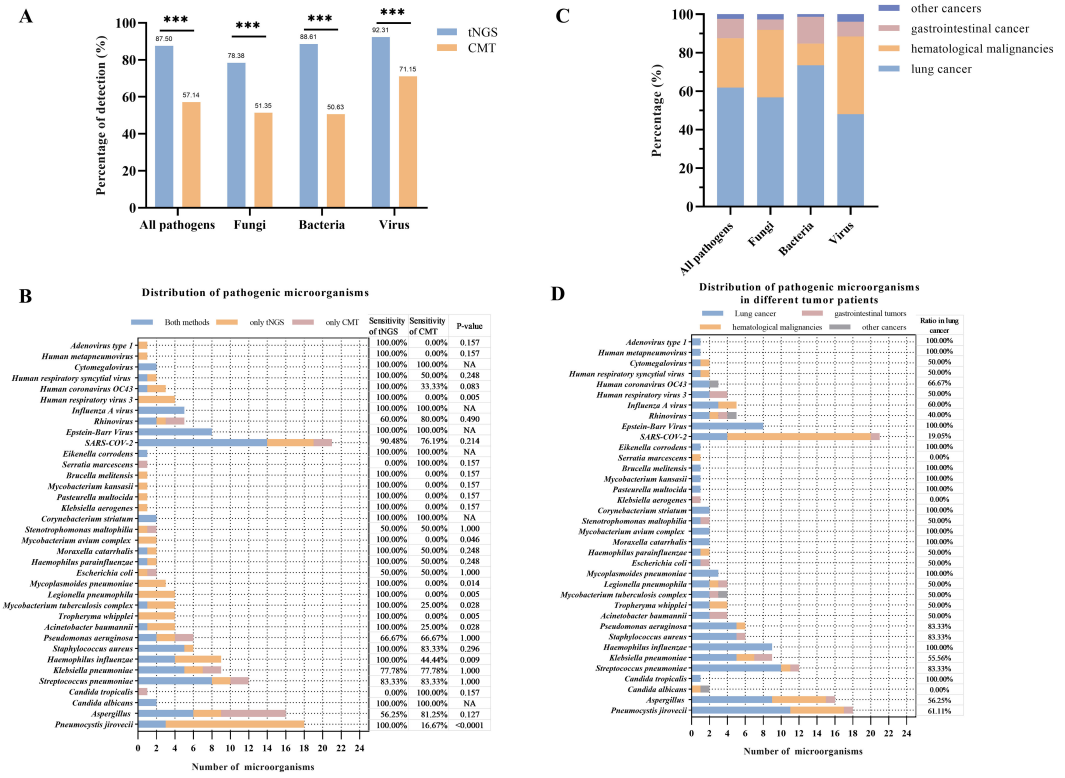



FIGURE 2
Comparison between tNGS and CMT in pathogens detection. **(A)** Percentage of pathogens detection in difference methods. **(B)** Distribution of pathogens and comparison between tNGS and CMT. The sensitivity of tNGS or CMT was calculated as follows: sensitivity = pathogens detected by tNGS or CMT/Total. The difference of sensitivity based on the t-test of tNGS and CMT was marked as p-value. **(C)** Percentage of pathogens founding in different cancer patients. **(D)** Distribution of pathogens in different cancer patients. **p<0.05, **p<0.01, ***p<0.001.*.

analysis of the RPM values between pathogens consistently detected by both methods and those detected exclusively by tNGS. The study findings indicate that the RPM values of *Pneumocystis jirovecii* detected exclusively by tNGS are comparatively lower, a phenomenon also observed in the analysis of viruses and bacteria (Supplementary Figure 1). This may be attributed to the higher sensitivity of the tNGS method.

## 3.4 Clinical diagnostic performance in different types of cancer

The performance of the tNGS method in the etiological diagnosis of pneumonia in different cancer patients was further explored. In patients with lung cancer, tNGS showed 80.68% of sensitivity and 93.33% of specificity (Table 4). In patients with gastrointestinal cancer and hematological malignancies, tNGS demonstrates a sensitivity exceeding 90% of tNGS. Whether in patients with lung cancer, gastrointestinal cancer, or hematological malignancies, the diagnostic accuracy of tNGS was higher than CMT. Compared with the overall patients, there was no significant difference in the diagnostic accuracy of tNGS among different cancer types ($p>0.05$). The distribution of pathogens among different cancer types was observed (Figure 2C). Most of the pathogens were identified in patients with lung cancer, accounting for about 60%, with 73.42% being bacterial pathogens. Patients with hematological malignancies accounted for 25% of identified pathogens, but only 11.39% were bacterial. In gastrointestinal cancer cases, the proportion of identified pathogens was 10.12%, with fungi being relatively rare at only 5.41%. *Pneumocystis jirovecii* and *Aspergillus* spp were more commonly found in patients with lung cancer and hematological malignancies, while common pathogens such as *Pseudomonas aeruginosa*, *Mycoplasma pneumoniae*, *Streptococcus pneumoniae*, and others were more prevalent in patients with lung cancer (Figure 2D). No differences were observed in the presence of *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and *Mycobacterium tuberculosis* in patients with different types of cancer. Interestingly, a higher number of *severe acute respiratory syndrome coronavirus 2* infections were found in patients with hematological malignancies, but no differences were observed in the presence of other viruses among patients with different types of cancer.

## 3.5 Clinical impact of tNGS: case report

A retrospective review of the treatment history of patients was conducted to further analyze the impact of tNGS on clinical treatment. It was found that tNGS results directly influenced antibiotic treatment decisions in many patients.

Case 1: A patient with lung cancer presented with symptoms of coughing and sputum production and was diagnosed with

TABLE 3 Diagnostic performance of tNGS in difference pathogens.

| Pathogens | tNGS | | | | | CMT | | | | | p-value |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Sensitivity (95% CI) | Specificity (95%CI) | PPV (95% CI) | NPV (95%CI) | Accuracy (95%CI) | Sensitivity (95% CI) | Specificity (95%CI) | PPV (95% CI) | NPV (95%CI) | Accuracy (95%CI) | |
| ***Pneumocystis jirovecii*** | 100.00% (82.41%-100.00%) | 96.92% (92.36%-98.80%) | 81.82% (61.48%-92.69%) | 100.00% (97.04%-100.00%) | 97.30% (93.26%-98.94%) | 16.67% (5.84%-39.22%) | 100.00% (97.13%-100.00%) | 100.00% (43.85%-100.00%) | 89.66% (83.63%-93.63%) | 89.86% (83.95%-93.76%) | 0.009 |
| ***Aspergillus* spp** | 56.25% (33.18%-76.90%) | 97.73% (93.53%-99.22%) | 75.00% (46.77%-91.11%) | 94.85% (89.76%-97.48%) | 93.24% (88.01%-96.29%) | 81.25% (56.99%-93.41%) | 100.00% (97.17%-100.00%) | 100.00% (77.19%-100.00%) | 97.78% (93.67%-99.24%) | 97.97% (94.21%-99.31%) | 0.047 |
| ***Streptococcus pneumoniae*** | 83.33% (55.20%-95.30%) | 78.68% (71.05%-84.72%) | 25.64% (14.57%-41.08%) | 98.17% (93.56%-99.50%) | 79.05% (71.81%-84.83%) | 83.33% (55.20%-95.30%) | 98.53% (94.80%-99.60%) | 83.33% (55.20%-95.30%) | 98.53% (94.80%-99.60%) | 97.30% (93.26%-98.94%) | <0.0001 |
| ***Klebsiella pneumoniae*** | 77.78% (45.26%-93.68%) | 97.12% (92.83%-98.88%) | 63.64% (35.38%-84.83%) | 98.54% (94.83%-99.60%) | 95.95% (91.44%-98.13%) | 77.78% (45.26%-93.68%) | 100.00% (97.31%-100.00%) | 100.00% (64.57%-100.00%) | 98.58% (94.98%-99.61%) | 98.65% (95.21%-99.63%) | 0.15 |
| ***Haemophilus influenzae*** | 100.00% (70.09%-100.00%) | 93.53% (88.15%-96.56%) | 50.00% (29.03%-70.97%) | 100.00% (97.13%-100.00%) | 93.92% (88.85%-96.77%) | 44.44% (18.88%-73.33%) | 100.00% (97.31%-100.00%) | 100.00% (51.01%-100.00%) | 96.53% (92.13%-98.51%) | 96.62% (92.34%-98.55%) | 0.27 |
| ***Staphylococcus aureus*** | 100.00% (60.97%-100.00%) | 100.00% (97.37%-100.00%) | 100.00% (60.97%-100.00%) | 100.00% (97.37%-100.00%) | 100.00% (97.47%-100.00%) | 83.33% (43.65%-96.99%) | 100.00% (97.37%-100.00%) | 100.00% (56.55%-100.00%) | 99.30% (96.15%-99.88%) | 99.32% (96.27%-99.88%) | 0.32 |
| ***Pseudomonas aeruginosa*** | 66.67% (30.00%-90.32%) | 100.00% (97.37%-100.00%) | 100.00% (51.01%-100.00%) | 98.61% (95.08%-99.62%) | 98.65% (95.21%-99.63%) | 66.67% (30.00%-90.32%) | 100.00% (97.37%-100.00%) | 100.00% (51.01%-100.00%) | 98.61% (95.08%-99.62%) | 98.65% (95.21%-99.63%) | 1.00 |
| ***Mycobacterium tuberculosis complex*** | 100.00% (51.01%-100.00%) | 97.92% (94.05%-99.29%) | 57.14% (25.05%-84.18%) | 100.00% (97.35%-100.00%) | 97.97% (94.21%-99.31%) | 25.00% (4.56%-69.94%) | 100.00% (97.40%-100.00%) | 100.00% (20.65%-100.00%) | 97.96% (94.17%-99.30%) | 97.97% (94.21%-99.31%) | 1.00 |
| ***Legionella pneumophila*** | 100.00% (51.01%-100.00%) | 100.00% (97.40%-100.00%) | 100.00% (51.01%-100.00%) | 100.00% (97.40%-100.00%) | 100.00% (97.47%-100.00%) | 0.00% (0.00%-48.99%) | 100.00% (97.40%-100.00%) | NA | 97.30% (93.26%-98.94%) | 97.30% (93.26%-98.94%) | 0.044 |
| **SARS-cov-2** | 90.48% (71.09%-97.35%) | 94.49% (89.06%-97.30%) | 73.08% (53.92%-86.30%) | 98.36% (94.22%-99.55%) | 93.92% (88.85%-96.77%) | 76.19% (54.91%-89.37%) | 100.00% (97.06%-100.00%) | 100.00% (80.64%-100.00%) | 96.21% (91.44%-98.37%) | 96.62% (92.34%-98.55%) | 0.27 |
| ***Epstein-Barr Virus*** | 100.00% (67.56%-100.00%) | 65.00% (56.79%-72.40%) | 14.04% (7.29%-25.32%) | 100.00% (95.95%-100.00%) | 66.89% (58.97%-73.96%) | 100.00% (67.56%-100.00%) | 100.00% (97.33%-100.00%) | 100.00% (67.56%-100.00%) | 100.00% (97.33%-100.00%) | 100.00% (97.47%-100.00%) | <0.0001 |
| ***Rhinovirus*** | 60.00% (23.07%-88.24%) | 95.10% (90.24%-97.61%) | 30.00% (10.78%-60.32%) | 98.55% (94.87%-99.60%) | 93.92% (88.85%-96.77%) | 80.00% (37.55%-96.38%) | 100.00% (97.38%-100.00%) | 100.00% (51.01%-100.00%) | 99.31% (96.17%-99.88%) | 99.32% (96.27%-99.88%) | 0.01 |

*(Continued)*





10.3389/fcimb.2025.1497198

10.3389/fcimb.2025.1497198

TABLE 3 Continued

| Pathogens | tNGS | | | | | CMT | | | | | p-value |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Sensitivity (95% CI) | Specificity (95%CI) | PPV (95% CI) | NPV (95%CI) | Accuracy (95%CI) | Sensitivity (95% CI) | Specificity (95%CI) | PPV (95% CI) | NPV (95%CI) | Accuracy (95%CI) | |
| ***Influenza A virus*** | 100.00% (56.55%-100.00%) | 97.90% (94.01%-99.28%) | 62.50% (30.57%-86.32%) | 100.00% (97.33%-100.00%) | 97.97% (94.21%-99.31%) | 100.00% (56.55%-100.00%) | 100.00% (97.38%-100.00%) | 100.00% (56.55%-100.00%) | 100.00% (97.38%-100.00%) | 100.00% (97.47%-100.00%) | 0.082 |

tNGS, targeted Next-generation sequencing; CMT, conventional microbiological tests; PPV, positive predictive value; NPV, negative predictive value; NA, not available. The difference of diagnostic accuracy based on the t-test of tNGS and CMT was marked as p-value. The names of the pathogens have been bolded.

pneumonia at our hospital. The patient was treated with anti-infective agents, but no improvement was observed. A tNGS test for respiratory pathogens was performed, revealing the presence of *Rhinovirus C* and *Pneumocystis jirovecii*. Given the patient's deteriorating pulmonary condition and the tNGS results, it was clinically considered that the infection was caused by both viral and fungal pathogens. Consequently, the patient was administered Levofloxacin and Piperacillin for anti-infective treatment, along with a combination of Sulfamethoxazole and Trimethoprim tablets for antifungal treatment, the patient showed signs of improvement.

Case 2: A patient with follicular lymphoma presented with fever and subsequently underwent a CT scan, which indicated symptoms of pulmonary infection. A tNGS test for respiratory pathogens detected *Pneumocystis jirovecii*. Symptomatic treatment with Sulfamethoxazole and Trimethoprim was administered. The patient has not experienced further fever and has shown symptomatic improvement following continued anti-infective treatment.

Case 3: A patient, post-debulking surgery for a malignant ovarian cancer, presented with symptoms of coughing and sputum production. Upon admission, biochemical tests and CT sacn indicated a pulmonary infection. tNGS test identified a specific pathogen: *Mycobacterium tuberculosis complex*. The patient was commenced on anti-tuberculosis therapy, which led to symptomatic relief. The patient was subsequently discharged following improvement.

These cases emphasize the important role of tNGS in guiding the diagnosis of infections in cancer patients and targeted antibiotic therapy, ultimately improving patient prognosis.

## 4 Discussion

Diagnosing pneumonia in cancer patients undergoing treatment is crucial, as it aids in managing patient infections and in decision-making regarding the timing of anti-cancer treatments. Targeted NGS is a common used for cancer-related gene detection, but its use for detecting infectious pathogens is a novel technique (Xia et al., 2023; Zhang et al., 2024). We conducted a retrospective analysis of the performance of this method in cancer patients. This study demonstrated that tNGS provides significantly higher sensitivity and a broader range of pathogen detection compared to CMT in cancer patients with pneumonia. Our findings highlight the potential of tNGS as a reliable diagnostic tool for managing infections in these patients.

As a new technology, understanding the clinical performance of tNGS is crucial. Previous studies have systematically evaluated the diagnostic value of tNGS in lower respiratory tract infections and pneumonia (Gaston et al., 2022; Zhang et al., 2024). However, research in immunocompromised patients has been relatively limited. Our study confirmed that tNGS is a more sensitive technique for diagnosing pneumonia in cancer patients compared to CMT methods (84.62% vs 56.92%, $p<0.0001$). tNGS also showed higher sensitivity in detecting bacteria, viruses, and fungi (Deng et al., 2023). These findings are consistent with previous studies, which indicated that the tNGS method maintains a high sensitivity in cancer patients with pneumonia (Deng et al., 2023; Mansoor et al., 2023). However, the increased sensitivity of tNGS may also

TABLE 4 The diagnostic performance of tNGS in patients with difference type of cancer.

| | Methods | Sensitivity (95% CI) | Specificity (95%CI) | PPV(95% CI) | NPV(95%CI) | Accuracy (95%CI) | p-value |
|---|---|---|---|---|---|---|---|
| **lung cancer** | **CMT** | 55.68% (45.28%-65.61%) | 100.00% (79.61%-100.00%) | 100.00% (92.73%-100.00%) | 27.78% (17.62%-40.89%) | 62.14% (52.49%-70.91%) | 0.0011 |
| | **tNGS** | 80.68% (71.22%-87.57%) | 93.33% (70.18%-98.81%) | 98.61% (92.54%-99.75%) | 45.16% (29.16%-62.23%) | 82.52% (74.06%-88.65%) | |
| **gastrointestinal cancer** | **CMT** | 36.36% (15.17%-64.62%) | 100.00% (34.24%-100.00%) | 100.00% (51.01%-100.00%) | 22.22% (6.32%-54.74%) | 46.15% (23.21%-70.86%) | 0.011 |
| | **tNGS** | 90.91% (62.26%-98.38%) | 100.00% (34.24%-100.00%) | 100.00% (72.25%-100.00%) | 66.67% (20.77%-93.85%) | 92.31% (66.69%-98.63%) | |
| **hematological malignancies** | **CMT** | 70.37% (51.52%-84.15%) | NA | 100.00% (83.18%-100.00%) | 0.00% (0.00%-32.44%) | 70.37% (51.52%-84.15%) | 0.011 |
| | **tNGS** | 96.30% (81.72%-99.34%) | NA | 100.00% (87.13%-100.00%) | 0.00% (0.00%-79.35%) | 96.30% (81.72%-99.34%) | |

tNGS, targeted Next-generation sequencing; CMT, conventional microbiological tests; PPV, positive predictive value; NPV, negative predictive value; NA, not available. The difference of diagnostic accuracy based on the t-test of tNGS and CMT was marked as p-value.
Bold formatting is applied to emphasize the subgroups and detection methods.

lead to the detection of non-pathogenic microorganisms, resulting in reduced specificity (Rodino and Simner, 2024). This emphasized the importance of carefully interpreting tNGS results, as the study suggests (Fourgeaud et al., 2024; Yin et al., 2024), by incorporating the clinical context into the interpretation of tNGS results to avoid overdiagnosing non-pathogenic microorganisms. In particular, the broader detection of *human herpesvirus* in viral analysis had also led to a decrease in the accuracy of viral infection diagnosis. Although this might be related to the higher reactivation of *herpesviruses* in immunocompromised populations and the neglected of herpesvirus diagnosis in clinical practice, the detection of herpesviruses had not played a significant role in antibiotic decision-making (Huang et al., 2023; Jiang et al., 2023). Therefore, the ability of tNGS to detect a broader range of pathogens makes it a sensitivity diagnostic tool for cancer patients who are at risk of opportunistic infections.

In terms of pathogen detection performance, tNGS identified a significantly greater number of pathogens, whether in bacteria, fungi, or viruses. This was mainly due to the ultrasensitivity of the NGS method (Supplementary Figure 1), which was reflected in the detection of most pathogens (Deng et al., 2020; Li et al., 2021; Peng et al., 2021; Sun et al., 2021). Additionally, tNGS was able to detect some difficult-to-culture or unexpected pathogens, such as *Pneumocystis jirovecii*, *Legionella pneumophila*, and *Brucella melitensis*, which fully demonstrated the advantage of the NGS method's broad coverage (Supplementary Table 1) (Zhang et al., 2024). However, it should be noted that tNGS was not as sensitive as the CMT method for *Aspergillus* spp, consistent with previous findings (Peng et al., 2021). Insufficient extraction of molds may be due to the simultaneous extraction of DNA and RNA required by tNGS. Moreover, CMT may have detected more using the galactomannan (GM) method, which might have resulted in insufficient detection when the *Aspergillus* spp load was low (Peng et al., 2021). Additionally, it was observed that the diagnostic accuracy of most pathogens was not much different from CMT, with better diagnostic accuracy for *Pneumocystis jirovecii* and *Legionella pneumophila* in tNGS. This may be due to the clinical application of PCR methods that covered these pathogens, making their detection limits comparable to tNGS (Murphy et al., 2020). In common pathogens such as *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*, the performance of both methods was similar, whis is also related to the widespread use of multiplex PCR (Zhang et al., 2022). In summary, tNGS was a more sensitive and broad-spectrum detection method compared to CMT, and except for *Aspergillus* spp, this method could provide more evidence to support clinical diagnosis.

Several studies have discussed the impact of mNGS on etiological diagnosis and antibiotic adjustment. Sun and colleagues reported that mNGS can guide antibiotic adjustments in critically ill pneumonia patients, with 87% of immunocompromised patients undergoing antibiotic adjustments based on mNGS results (Sun et al., 2021). Additionally, Wei et al.'s study mentioned that tNGS played a positive role in the pathogen diagnosis of 72.7% of patients and may have led to antibiotic treatment adjustments in 17% of patients (Wei et al., 2024). Our study also confirmed these findings. Given that the tNGS method costs only about half as much as mNGS and is even comparable to the price of multiplex PCR, the application of tNGS in primary or secondary testing may significantly enhance diagnostic accuracy and reduce the misuse of antibiotics. Our study also explored the diagnostic performance of tNGS in different cancer types, with no significant differences observed. This is similar to the impact of antibiotics on NGS detection performance, where nucleic acid levels are less influenced by patient characteristics (Azoulay et al., 2020). However, when comparing the distribution of pathogens among different cancer types, we found that fungi were detected more frequently in patients with lung cancer and hematological malignancies, which may be related to the higher proportion of these populations. Additionally, we observed that *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Staphylococcus aureus*, and *Pseudomonas aeruginosa* occurred more frequently in lung cancer patients compared to those with hematological malignancies or gastrointestinal cancer. These pathogens are more likely to colonize the respiratory tract and may be more prone to cause infections due to the immune impairment caused by lung cancer (Plummer et al., 2016; van Elsland and Neefjes, 2018; Hatta et al., 2021). Of course,

this may also be related to the uneven distribution of patients with different cancer types.

This study has several limitations. Firstly, due to the retrospective nature of this study and the relatively small sample size, potential biases may exist, such as the diagnostic value for pathogens with lower incidence rates, like Mycoplasma pneumoniae and the Mycobacterium tuberculosis complex, which require a larger population to clarify … Secondly, the PCR detection of DNA viruses, especially *herpesviruses*, was determined by clinical needs, which may introduce bias in the evaluation of the performance of DNA virus detection. Lastly, the distribution of patients with different types of cancer in our cohort was relatively concentrated, suggesting that the conclusions regarding the pathogen preferences of different cancer types need to be verified in a larger cohort.

## 5 Conclusion

In conclusion, tNGS shows great promise as a diagnostic tool for detecting infections in cancer patients. It offers higher sensitivity and broader pathogen detection capabilities compared to conventional methods. However, despite its limitations, tNGS can significantly enhance the management of infections in immunocompromised patients, potentially leading to better clinical outcomes. Future research should focus on optimizing the specificity of tNGS and integrating it into routine clinical practice to fully leverage its diagnostic potential.

## Data availability statement

The data presented in the study are deposited in the China National Center for Bioinformation - National Genomics Data Center repository, accession number PRJCA035966.

## Ethics statement

The studies involving humans were approved by Ethics Committee of Henan Cancer Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

KY: Data curation, Formal analysis, Writing – review & editing. JZ: Writing – original draft, Project administration. TW: Software, Writing – original draft. ZW: Visualization, Writing – original draft. RS: Investigation, Writing – original draft. DG: Data curation, Writing – original draft. HL: Visualization, Writing – original draft. WW: Investigation, Writing – review & editing. CYZ: Methodology, Writing – review & editing. CZZ: Supervision, Writing – original draft. YG: Project administration, Writing – original draft. JM: Conceptualization, Writing – original draft. BW: Conceptualization, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by Major Science and Technology Project of Henan Province (221100310100).

## Acknowledgments

We owe thanks to the patients in the study and their family members. We acknowledge the staffs of all departments for their assistance to this study.

## Conflict of interest

Authors DG and HL are employed by the company Genepus-Beijing.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2025.1497198/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Further analysis to additional detections by RPM of tNGS. The results were grouped based on whether the CMT was positive or negative (labeled as CMT+/tNGS+ or CMT-/tNGS+). *: p<0.05, **: p<0.01, ***: p<0.001.

# References

Azoulay, E., Russell, L., Van de Louw, A., Metaxa, V., Bauer, P., Povoa, P., et al. (2020). Diagnosis of severe respiratory infections in immunocompromised patients. *Intensive Care Med.* 46, 298–314. doi: 10.1007/s00134-019-05906-5

Babady, N. E. (2016). Laboratory diagnosis of infections in cancer patients: challenges and opportunities. *J. Clin. Microbiol.* 54, 2635–2646. doi: 10.1128/JCM.00604-16

Blauwkamp, T. A., Thair, S., Rosen, M. J., Blair, L., Lindner, M. S., Vilfan, I. D., et al. (2019). Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat. Microbiol.* 4, 663–674. doi: 10.1038/s41564-018-0349-6

Chiu, C. Y., and Miller, S. A. (2019). Clinical metagenomics. *Nat. Rev. Genet.* 20, 341–355. doi: 10.1038/s41576-019-0113-7

Deng, Q., Cao, Y., Wan, X., Wang, B., Sun, A., Wang, H., et al. (2022). Nanopore-based metagenomic sequencing for the rapid and precise detection of pathogens among immunocompromised cancer patients with suspected infections. *Front. Cell Infect. Microbiol.* 12, 943859. doi: 10.3389/fcimb.2022.943859

Deng, X., Achari, A., Federman, S., Yu, G., Somasekar, S., Bártolo, I., et al. (2020). Metagenomic sequencing with spiked primer enrichment for viral diagnostics and genomic surveillance. *Nat. Microbiol.* 5, 443–454. doi: 10.1038/s41564-019-0637-9

Deng, Z., Li, C., Wang, Y., Wu, F., Liang, C., Deng, W., et al. (2023). Targeted next-generation sequencing for pulmonary infection diagnosis in patients unsuitable for bronchoalveolar lavage. *Front. Med. (Lausanne)* 10, 1321515. doi: 10.3389/fmed.2023.1321515

Fourgeaud, J., Regnault, B., Ok, V., Da Rocha, N., Sitterlé, É, Mekouar, M., et al. (2024). Performance of clinical metagenomics in France: a prospective observational study. *Lancet Microbe* 5, e52–e61. doi: 10.1016/s2666-5247(23)00244-6

Gaston, D. C., Miller, H. B., Fissel, J. A., Jacobs, E., Gough, E., Wu, J., et al. (2022). Evaluation of metagenomic and targeted next-generation sequencing workflows for detection of respiratory pathogens from bronchoalveolar lavage fluid specimens. *J. Clin. Microbiol.* 60, e0052622. doi: 10.1128/jcm.00526-22

Gu, B., Zhuo, C., Xu, X., and El Bissati, K. (2023). Editorial: Molecular diagnostics for infectious diseases: Novel approaches, clinical applications and future challenges. *Front. Microbiol.* 14, 1153827. doi: 10.3389/fmicb.2023.1153827

Hatta, M. N. A., Mohamad Hanif, E. A., Chin, S. F., and Neoh, H. M. (2021). Pathogens and carcinogenesis: A review. *Biol. (Basel)* 10, 533–552. doi: 10.3390/biology10060533

Hill, J. A., Park, S. Y., Gajurel, K., and Taplitz, R. (2024). A systematic literature review to identify diagnostic gaps in managing immunocompromised patients with cancer and suspected infection. *Open Forum Infect. Dis.* 11, ofad616. doi: 10.1093/ofid/ofad616

Hou, D., Ju, M., Wang, Y., Zhang, D., Zhu, D., Zhong, M., et al. (2020). PCR coupled to electrospray ionization mass spectrometry for microbiological diagnosis and surveillance of ventilator-associated pneumonia. *Exp. Ther. Med.* 20, 3587–3594. doi: 10.3892/etm.2020.9103

Huang, L., Zhang, X., Pang, L., Sheng, P., Wang, Y., Yang, F., et al. (2023). Viral reactivation in the lungs of patients with severe pneumonia is associated with increased mortality, a multicenter, retrospective study. *J. Med. Virol.* 95, e28337. doi: 10.1002/jmv.28337

Jiang, X., Yan, J., Huang, H., Ai, L., Yu, X., Zhong, P., et al. (2023). Development of novel parameters for pathogen identification in clinical metagenomic next-generation sequencing. *Front. Genet.* 14, 126690. doi: 10.3389/fgene.2023.1266990

Li, N., Cai, Q., Miao, Q., Song, Z., Fang, Y., and Hu, B. (2021). High-throughput metagenomics for identification of pathogens in the clinical settings. *Small Methods* 5, 2000792. doi: 10.1002/smtd.202000792

Li, X., Liu, Y., Li, M., Bian, J., Song, D., and Liu, C. (2023). Epidemiological investigation of lower respiratory tract infections during influenza A (H1N1) pdm09 virus pandemic based on targeted next-generation sequencing. *Front. Cell. infection Microbiol.* 13, 1303456. doi: 10.3389/fcimb.2023.1303456

Li, S., Tong, J., Liu, Y., Shen, W., and Hu, P. (2022). Targeted next generation sequencing is comparable with metagenomic next generation sequencing in adults with pneumonia for pathogenic microorganism detection. *J. infection* 85, e127–e129. doi: 10.1016/j.jinf.2022.08.022

Liang, M., Fan, Y., Zhang, D., Yang, L., Wang, X., Wang, S., et al. (2022). Metagenomic next-generation sequencing for accurate diagnosis and management of lower respiratory tract infections. *Int. J. Infect. diseases: IJID* 122, 921–929. doi: 10.1016/j.ijid.2022.07.060

Mansoor, H., Hirani, N., Chavan, V., Das, M., Sharma, J., Bharati, M., et al. (2023). Clinical utility of target-based next-generation sequencing for drug-resistant TB. *Int. J. tuberculosis Lung Dis.* 27, 41–48. doi: 10.5588/ijtld.22.0138

Murphy, C. N., Fowler, R., Balada-Llasat, J. M., Carroll, A., Stone, H., Akerele, O., et al. (2020). Multicenter evaluation of the BioFire filmArray pneumonia/pneumonia plus panel for detection and quantification of agents of lower respiratory tract infection. *J. Clin. Microbiol.* 58, e00128–20. doi: 10.1128/jcm.00128-20

Peng, J. M., Du, B., Qin, H. Y., Wang, Q., and Shi, Y. (2021). Metagenomic next-generation sequencing for the diagnosis of suspected pneumonia in immunocompromised patients. *J. infection* 82, 22–27. doi: 10.1016/j.jinf.2021.01.029

Plummer, M., de Martel, C., Vignat, J., Ferlay, J., Bray, F., and Franceschi, S. (2016). Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet Glob Health* 4, e609–e616. doi: 10.1016/S2214-109X(16)30143-7

Rodino, K. G., and Simner, P. J. (2024). Status check: next-generation sequencing for infectious-disease diagnostics. *J. Clin. Invest.* 134, e178003. doi: 10.1172/JCI178003

Serpa, P. H., Deng, X., Abdelghany, M., Crawford, E., Malcolm, K., Caldera, S., et al. (2022). Metagenomic prediction of antimicrobial resistance in critically ill patients with lower respiratory tract infections. *Genome Med.* 14, 74. doi: 10.1186/s13073-022-01072-4

Sun, T., Wu, X., Cai, Y., Zhai, T., Huang, L., Zhang, Y., et al. (2021). Metagenomic next-generation sequencing for pathogenic diagnosis and antibiotic management of severe community-acquired pneumonia in immunocompromised adults. *Front. Cell. infection Microbiol.* 11. doi: 10.3389/fcimb.2021.661589

van Elsland, D., and Neefjes, J. (2018). Bacterial infections and cancer. *EMBO Rep.* 19, 46632. doi: 10.15252/embr.201846632

Wei, M., Mao, S., Li, S., Gu, K., Gu, D., Bai, S., et al. (2024). Comparing the diagnostic value of targeted with metagenomic next-generation sequencing in immunocompromised patients with lower respiratory tract infection. *Ann. Clin. Microbiol. antimicrobials* 23, 88. doi: 10.1186/s12941-024-00749-5

Xia, H., Zhang, Z., Luo, C., Wei, K., Li, X., Mu, X., et al. (2023). MultiPrime: A reliable and efficient tool for targeted next-generation sequencing. *iMeta* 2, e143. doi: 10.1002/imt2.143

Yin, Y., Zhu, P., Guo, Y., Li, Y., Chen, H., Liu, J., et al. (2024). Enhancing lower respiratory tract infection diagnosis: implementation and clinical assessment of multiplex PCR-based and hybrid capture-based targeted next-generation sequencing. *EBioMedicine* 107, 105307. doi: 10.1016/j.ebiom.2024.105307

Zhang, C., Chen, X., Wang, L., Song, J., Zhou, C., Wang, X., et al. (2022). Evaluation of a multiplex PCR kit for detection of 17 respiratory pathogens in hospitalized patients. *J. Thorac. Dis.* 14, 3386–3397. doi: 10.21037/jtd-22-544

Zhang, J., Dong, P., Liu, B., Xu, X., Su, Y., Chen, P., et al. (2023). Comparison of XBB and BA.5.2: differences in clinical characteristics and disease outcomes. *Archivos bronconeumologia* 59, 782–784. doi: 10.1016/j.arbres.2023.08.012

Zhang, P., Liu, B., Zhang, S., Chang, X., Zhang, L., Gu, D., et al. (2024). Clinical application of targeted next-generation sequencing in severe pneumonia: a retrospective review. *Crit. Care (London England)* 28, 225. doi: 10.1186/s13054-024-05009-8

# Frontiers in
# Cellular and Infection Microbiology

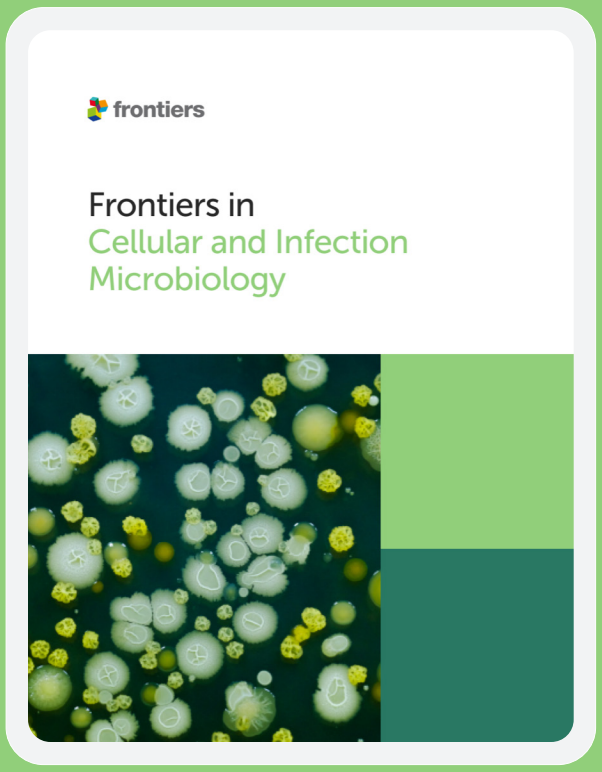
**Investigates how microorganisms interact with their hosts**

Explores bacteria, fungi, parasites, viruses, endosymbionts, prions and all microbial pathogens as well as the microbiota and its effect on health and disease in various hosts.

---

## Discover the latest Research Topics

See more →




**Frontiers**

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

**Contact us**

+41 (0)21 510 17 00
frontiersin.org/about/contact