

The Enhanced Importance of CD4 + T Cell Based Monitoring of Microbial Flora in Patients with HIV

Siyuan Yang

Beijing Institute of Infectious Diseases

Jiayi Liu

Beijing Ditan Hospital

Fengting Yu

Beijing Institute of Infectious Diseases

Yunxia Tang

Beijing Institute of Infectious Diseases

Xi Wang

Beijing Institute of Infectious Diseases

Linghang Wang

`linghang.wang@ccmu.edu.cn`

Beijing Ditan Hospital

Article

Keywords: Lung infection, Microbiome, HIV, 16S sequencing, ITS sequencing

Posted Date: July 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9465465/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on September 22nd, 2026. See the published version at <https://doi.org/10.1038/s41598-026-69120-2>.

Abstract

Individuals infected with human immunodeficiency virus (HIV) frequently display diminished CD4⁺ T Cell levels, heightening susceptibility to diverse infections and imposing substantial burdens alongside life-threatening risks. This study aimed to compare lung microbial communities between general infections and HIV-associated cases. Bronchoalveolar lavage fluid (BALF) samples from 221 patients with lung diseases or infection underwent 16S and ITS sequencing analysis. The cohort comprised with 18 HIV-negative, GM-negative individuals; 29 HIV-negative, GM-positive individuals; 104 HIV-positive, GM-negative individuals; and 70 HIV-positive, GM-positive individuals.

Through comparative analysis of microbial disparities across groups, we observed that GM test-positive individuals displayed pronounced fungal infections, which significantly amplified the detectable diversity of fungal species while exhibiting heterogeneous compositional profiles. Conversely, bacterial diversity demonstrated negligible overall variation. Notably, viral load exhibited an inverse correlation with CD4⁺ T Cell counts, and these immune cells further revealed a discernible association with both viral load dynamics and microbial community shifts. These insights underscore the critical necessity for real-time monitoring of CD4⁺ T Cell fluctuations and microbial equilibrium in HIV-infected patients.

1. Introduction

16S ribosomal RNA (16S rRNA) and internal transcribed spacer (ITS) sequencing have emerged as powerful analytical tools for microbial community characterization in bronchoalveolar lavage fluid (BALF) which confer substantial benefits in both research and clinical applications [1, 2]. Their high-throughput identification enables comprehensive profiling of bacterial and fungal populations, offering insights into the composition, diversity, and dynamics of the pulmonary microbiota [3, 4].

One of the primary advantages of 16S and ITS sequencing is their capability to detect low-abundance or fastidious microorganisms that might elude conventional culture-based methods [5]. This is especially relevant in BALF [6], where the microbial load is frequently low, and pathogens may be present in minute quantities, which is particularly crucial in immunocompromised patients, such as those with HIV or undergoing immunosuppressive therapy, where atypical or opportunistic infections are common and can have severe clinical implications.

Furthermore, 16S and ITS sequencing elucidate the role of the lung microbiome in modulating immune responses and influencing disease outcomes [7]. By identifying microbial communities associated with inflammation, immune suppression, or resistance to therapy, the therapeutic interventions, such as probiotics, prebiotics, or microbiome-targeted immunomodulators, could be accelerated.

In the context of HIV infection, where CD4⁺ T cell depletion predisposes patients to opportunistic infections, understanding the interaction between viral load, immune status, and microbial composition is vital for optimizing patient management and improving outcomes [8, 9]. In this study, we aimed to

investigate differences in lung microbial communities between individuals with general infections and those with HIV infections.

2. Materials and Methods

2.1. Patient Recruitment and Sample Collection

We recruited 221 HIV-infected patients admitted to Beijing Ditan Hospital, Beijing, China, between July and October 2019, based on strict inclusion and exclusion criteria. Clinical data were collected and organized for analysis. As part of the study, participants underwent bronchoscopy, and 10 ml of bronchoalveolar lavage fluid (BALF) was collected for microbiome analysis. Bronchoalveolar lavage was performed in the lung segment showing the most significant involvement on chest radiographs. The median time from admission to bronchoscopy was 1 day [interquartile range (IQR), 1–3 days]. The study was approved by the Ethics Committee of Beijing Ditan Hospital with the reference JDKLZ-2019-016-001, and all patients provided written informed consent prior to undergoing bronchoscopy. Additionally, all participants provided written informed consent for participation in the survey and biomarker testing. The research was carried out in line with the Declaration of Helsinki.

2.2. 16S/ITS Sequencing

DNA extraction and amplicon sequencing were performed at the Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University. Full-length 16S rDNA and internal transcribed spacer (ITS) rDNA regions were targeted for sequencing. DNA was extracted using the MagNA Pure LC 2.0 system and the MagNA Pure LC Total NA Isolation Kit (Roche, Mannheim, Baden-Württemberg, Germany) according to the manufacturer's instructions, and quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Eugene, Oregon, United States). Polymerase chain reaction (PCR) amplification of the target regions was performed using the following primers, which included Illumina adapter sequences and dual-index barcodes to label each sample: 341F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAATCC-3'). The PCR conditions were as follows: initial denaturation at 95°C for 3 minutes, followed by 25 cycles of denaturation at 95°C for 30 seconds, annealing at 65°C for 30 seconds, and extension at 72°C for 30 seconds, with a final extension step at 72°C for 5 minutes. PCR products were purified using AMPure XP beads (Catalog No. A63882; Beckman Coulter Inc., Fullerton, California, United States). Amplicon sequencing libraries were prepared following the 16S Metagenomic Sequencing Library Preparation protocol (Illumina, Inc., San Diego, CA, United States). Paired-end sequencing (250 bp × 2) was performed on the MiSeq platform (Illumina, Inc.) using the MiSeq v2 reagent kit (Illumina, Inc.).

2.3. 16S/ITS OTU Generation

Raw sequencing data were initially assessed for quality using FastQC (v0.11.9, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). For bioinformatics processing of MiSeq results, QIIME (v1.9.1) was employed [10]. Sequencing data were quality-filtered using the fastq_filter

function in USEARCH (v7.0.1001) to remove reads with an expected error threshold exceeding two [11]. An open-reference clustering strategy was applied to analyze the sequencing data. This strategy automatically annotates OTUs (Operational Taxonomic Units) with taxonomic information from a reference database. Additionally, a de novo OTU picking method was used to cluster sequences without relying on corresponding reference sequences in the database. The Greengenes database was utilized for taxonomic assignment of OTUs. OTUs were filtered as follows: (1) OTUs corresponding to common contaminants in BAL samples were removed using negative control samples as a reference; (2) OTUs with fewer than 10 reads in BAL samples were excluded; and (3) OTUs with read counts less than 1/5000 of the total reads across all samples were removed.

A strategy similar to the one described above was employed to generate ITS OTUs, with the following modifications and annotations: (1) Taxonomic classification was performed using the UNITE database [12]; (2) No phylogenetic tree was generated; (3) No negative template control (NTC) samples were included; and (4) OTUs with read counts less than 1/1000 of the total reads across all samples were removed.

2.4. Statistical methods

Diversity indices were calculated using QIIME and DADA2 [13]. Bray-Curtis distance was employed to determine principal coordinates, and PERMANOVA (Permutational Multivariate Analysis of Variance) was used to assess the accuracy of the Principal Coordinate Analysis (PCoA) results. Two sets of non parametric tests were conducted using Wilcoxon test, multiple non parametric tests using Kruskal Wallis test.

3. Results

3.1. Clinical and demographic information summary for enrolled individuals

Between July and October 2019, a total of 221 patients admitted to our hospital were enrolled in our study cohort based on CD4⁺ T Cell counts. All samples underwent 16S and ITS sequencing followed by comprehensive analysis (Fig. 1A). The clinical characteristics of the patients are summarized in Table 1. The median CD4⁺ T Cell count in the HIV-positive group was 39 cells/ μ l (IQR 9–94), and the median plasma HIV RNA concentration was 5.1 log₁₀ copies/ml (2.5–5.4). Among the HIV-positive patients, 11/74 (15%) had an HIV RNA concentration exceeding 500,000 copies/ml. Based on HIV and GM testing results, the samples were further categorized into four groups: 1. HIV-negative and GM testing-negative (HNGN); 2. HIV-negative and GM testing-positive (HNGP); 3. HIV-positive and GM testing-negative (HPGN); and 4. HIV-positive and GM testing-positive (HPGP).

Table 1
Summary of clinical features for all individuals

Features	HNGN	HNGP	HPGN	HPGP	statisitcs	P
Gender (Male/Total)	14/18	23/29	99/103	65/70	$\chi^2=0.60$	0.90
Age (median, Q1 ~ Q3)	52(41.5 ~ 61.8)	55(41 ~ 72)	35(29.5 ~ 44)	38(30 ~ 46)	H = 40.15	0.00
Treatment (ART/Total)	/	/	14/103	10/70	$\chi^2=0.02$	0.90
lymphocyte (median, Q1 ~ Q3)	1.33(1.07 ~ 2.19)	0.99(0.53 ~ 1.73)	0.9(0.57 ~ 1.4)	0.71(0.44 ~ 1.11)	H = 16.84	0.00
neutrophils (median, Q1 ~ Q3)	3.63(2.68 ~ 4.80)	6.45(4 ~ 9.98)	3.47(2.12 ~ 5.36)	2.78(1.58 ~ 3.8)	H = 27.49	0.00
CD4 (median, Q1 ~ Q3)	/	/	45(15.5 ~ 126)	63(19 ~ 122)	W = 3481	0.70
CD8 (median, Q1 ~ Q3)	/	/	559(298 ~ 816)	538(309 ~ 730)	W = 3799	0.55
HIV Load (median, Q1 ~ Q3)	/	/	207534(73068 ~ 588877)	205699(52374 ~ 642104)	W = 3142	0.68

Age distribution analysis revealed that HPGN individuals were significantly younger than HNGN individuals, and HNGP individuals were significantly younger than HPGP individuals. This observation may be attributed to differences in the average age of onset between HIV-infected individuals and those with other pulmonary conditions (Fig. 1B). The mean lymphocyte counts in HPGP and HPGN individuals was significantly lower than that in HNGN individuals, likely due to HIV infection (Fig. 1C). Although neutrophil levels showed statistically significant differences among the groups, the overall distribution patterns were relatively similar (Fig. 1D). Among HIV-positive individuals, no significant differences were observed in CD4⁺ T Cell counts, CD8⁺ T Cell counts, or viral load between GM testing-positive and GM testing-negative groups (Fig. 1E, Fig. 1F, Fig. 1G).

3.2. Diversity of microbe in HIV and non-HIV individuals

Compared to HNGN individuals, HNGP individuals exhibited an increase in both species' diversity indices and the number of observed species as detected by ITS sequencing. Similarly, compared to HPGN individuals, HPGP individuals also showed elevated species diversity indices and an increased number of observed species using ITS technology, confirming the presence of fungal infections (Fig. 2A). In contrast, bacterial diversity indices and the number of observed bacterial species, as detected by 16S sequencing, showed no significant differences overall (Fig. 2B). These findings suggest that GM testing-positive individuals harbor fungal infections, which contribute to an increased diversity of detectable

fungi, while bacterial diversity remains relatively consistent. This observation aligns with established clinical knowledge.

3.3. Landscape of microbial composition in HIV and non-HIV individuals

In ITS-based fungal profiling, regardless of HIV status, at the family level, GM testing-positive individuals exhibited a reduced relative abundance of the Pneumocystidaceae family, while the relative abundances of Aspergillaceae, Nectriaceae, and some undefined fungi were increased (Fig. 3A). At the genus level, *Pneumocystis* was less abundant, whereas *Aspergillus*, *Talaromyces*, and some undefined fungi showed increased relative abundances (Fig. 3B).

At the family level, among GM testing-positive individuals, those with HIV infection demonstrated an increased relative abundance of Pneumocystidaceae, while the relative abundances of Aspergillaceae, Nectriaceae, and some undefined fungi were decreased (Fig. 3A). At the genus level, *Pneumocystis* was more abundant, whereas *Aspergillus*, *Candida*, and some undefined fungi exhibited reduced relative abundances (Fig. 3B). These results highlight the predominant types of fungal infections associated with HIV infection and GM testing positivity.

In 16S-based bacterial profiling, at the family level, GM testing-positive individuals exhibited a reduced relative abundance of Prevotellaceae and Veillonellaceae, while the relative abundances of Micrococcaceae, Staphylococcaceae, Pasteurellaceae, and Neisseriaceae were increased (Fig. 3C). At the genus level, *Capnocytophaga*, *Prevotella*, *Fusobacterium*, and *Filifactor* showed a decreasing trend, whereas *Neisseria*, *Peptostreptococcus*, and *Shewanella* demonstrated an increasing trend (Fig. 3D).

3.4. Correlation of microbial composition and immune status in HIV and non-HIV individuals

To further investigate the relationship between patients' immune status, viral load, and infections, we defined CD4⁺ T Cell levels < 200 as "poor immune status" and > 200 as "good immune status". We found that 85% of HIV patients had CD4⁺ T Cell levels below 200 (Fig. 4A), which is likely due to the fact that HIV patients typically seek medical treatment only after experiencing infections. Further analysis of the relationship between viral load and CD4⁺ T Cell levels revealed a negative correlation between the viral load (log-transformed) and CD4⁺ T Cell levels (Fig. 4B). Statistical analysis showed no significant difference in the viral load between GM testing-positive and negative groups, regardless of whether patients were in the "poor" or "good" CD4⁺ T Cell groups (Fig. 4C). However, overall, the viral load was significantly higher in the "poor" group compared to the "good" group, both in GM testing-negative (Fig. 4D) and positive groups (Fig. 4E). In GM testing-positive patients, '*Pneumocystis*' levels decreased, while '*Candida*' levels increased (Fig. 4F). In patients with poorer immune status, '*Pneumocystis*' levels decreased, while '*Candida*' and '*Talaromyces*' levels increased (Fig. 4G), a trend that was particularly

pronounced in GM testing-positive patients (Fig. 4H, 4I). These results suggest that as the infection progresses, endogenous fungi may become one of the important pathogens that cause lung infections.

In summary, HIV-positive patients seeking medical attention frequently present with polymicrobial infections and exhibit lower CD4⁺ T Cell levels. CD4⁺ T Cell levels show a certain correlation with changes in microbial flora. These findings underscore the importance of real-time monitoring of CD4⁺ T Cell levels and microbial flora in HIV patients.

4. Discussion

The 16S rRNA gene, highly conserved among bacteria yet featuring variable regions, allows for accurate identification and classification of bacterial taxa at the genus or species level [14]. Similarly, ITS sequencing targets the non-coding regions between ribosomal RNA genes in fungi, providing high resolution for differentiating fungal species, including those challenging to culture or identify via traditional methods [15]. Both of them establish a robust platform for exploring the intricate interplay between microbial communities and respiratory health.

Clinically, 16S and ITS sequencing hold great promise for enhancing diagnostic accuracy and guiding personalized treatment strategies [16, 17]. For example, in patients suspected of having pneumonia or other respiratory infections, 16S/ITS could help assist in differentiating between bacterial, fungal, and polymicrobial infections, facilitating targeted antimicrobial therapy which is crucial in cases where empirical treatment has failed or when atypical pathogens are suspected. Moreover, the ability to monitor changes in microbial composition over time offer valuable insights into disease progression, treatment efficacy, and the risk of recurrent infections.

In this study, we utilized the sequencing technology for BALF, and found the majority (85%) of HIV patients exhibited low CD4⁺ T Cell levels (< 200) which was due to the HIV patients suffering infections when they came to the hospital. No surprisingly the viral load was negatively correlated with CD4⁺ T Cell levels which reflected the poor immune status for HIV individuals. Generally, in GM testing-positive patients, 'Pneumocystis' levels decreased, while 'Candida' levels increased. In patients with poorer immune status (CD4⁺ T Cell levels < 200), 'Pneumocystis' levels decreased, while 'Candida' and 'Talaromyces' levels increased, a trend that was particularly pronounced in GM testing-positive patients. Pneumocystis led to Pneumocystis pneumonia (PCP), which was one of the most common opportunistic infections in HIV patients [18]. In HIV patients, when the CD4⁺ T cell count was lower than 200/ μ l, the risk of Pneumocystis infection increases significantly [19, 20]. Candida was one of the normal flora of the human body and was usually present in parts such as the skin, oral cavity, gastrointestinal tract, and vagina [21]. In HIV patients, due to impaired immune function, the body's resistance to Candida decreases. Candida, which was originally in a symbiotic state, began to multiply in large numbers and invade the surrounding tissues, and finally lead to endogenous infection [22, 23]. These results suggest that as the infection progresses, endogenous fungi may become one of the important pathogens for patient infections.

5. Conclusions

In conclusion, 16S rRNA gene and ITS amplicon sequencing emerge as indispensable tools for multidimensional analysis of the bronchoalveolar lavage fluid (BALF) microbiome, enabling precise discrimination of bacterial and fungal phylogenetic signatures while revealing metabolic pathway dynamics through metagenomic inference. The convergence of third-generation sequencing platforms with machine learning-driven analytical frameworks is fundamentally redefining diagnostic precision in respiratory medicine, particularly for differentiating colonization from active infection. These technological advancements synergize with emerging concepts of immune-microbiome crosstalk, where longitudinal monitoring of CD4⁺ T lymphocyte kinetics and microbial community stability in HIV-associated pulmonary complications provides critical biomarkers for predicting opportunistic infection risks, guiding antimicrobial stewardship and optimizing immunoreconstitution therapies. Such integrative approaches establish a new paradigm in precision pulmonology, transforming passive disease management into proactive ecosystem surveillance capable of intercepting dysbiotic cascades before clinical manifestation, thereby improving therapeutic efficacy and reducing respiratory morbidity in patients with HIV.

Declarations

Author Contributions: Conceptualization, S.Y. and L.W.; methodology, S.Y. and J.L.; validation, S.Y. and J.L.; formal analysis, S.Y. and J.L.; investigation, S.Y., F.Y. and Y.T.; writing—original draft preparation, S.Y.; writing—review and editing, X.W. and L.W.; visualization, S.Y. ; supervision, X.W. ; project administration, X.W. and L.W.; funding acquisition, S.Y., X.W. and L.W.. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Key Research and Development Program of China (2023YFC2306003 to X.W.), National Natural Science Foundation of China (No. 32270635 to X.W.), “R&D Program of Beijing Municipal Education Commission (KM202410025003 to S.Y.), The National Key R&D Program of China (2020YFA0707600 to L.W.).

Institutional Review Board Statement: The animal study protocol was approved by the Ethics Committee of Beijing Ditan Hospital, with the reference JDKLZ-2019-016-001 (2019.04.03). The research was carried out in line with the Declaration of Helsinki.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original datasets presented in this study are included in the article. The data that support the findings of this study are openly available in the China National Center for Bioinformation at <https://ngdc.cncb.ac.cn/omix/releaseList>, accession number OMIX016691.

Conflict of Interest: The authors declare no conflict of interest.

Acknowledgement

None

References

1. Wang, Y. et al. Respiratory microbiota imbalance in children with *Mycoplasma pneumoniae* pneumonia. *Emerg. Microbes Infect.* **12**, 2202272 (2023).
2. Sun, Y. et al. Characterization of Lung and Oral Microbiomes in Lung Cancer Patients Using Culturomics and 16S rRNA Gene Sequencing. *Microbiol. Spectr.* **11**, e0031423 (2023).
3. Gao, B. et al. An Introduction to Next Generation Sequencing Bioinformatic Analysis in Gut Microbiome Studies. *Biomolecules* **11**, 530 (2021).
4. Jervis Bardy, J. & Psaltis, A. J. Next Generation Sequencing and the Microbiome of Chronic Rhinosinusitis: A Primer for Clinicians and Review of Current Research, Its Limitations, and Future Directions. *Ann. Otol Rhinol Laryngol.* **125**, 613–621 (2016).
5. Fukuda, K., Ogawa, M., Taniguchi, H. & Saito, M. Molecular Approaches to Studying Microbial Communities: Targeting the 16S Ribosomal RNA Gene. *J. UOEH.* **38**, 223–232 (2016).
6. Yamasaki, K. et al. Possible role of anaerobes in the pathogenesis of nontuberculous mycobacterial infection. *Respirology* **20**, 758–765 (2015).
7. Schneeberger, P. et al. Microbiota analysis optimization for human bronchoalveolar lavage fluid. *Microbiome* **7**, 141 (2019).
8. Sattentau, Q. J. & Moore, J. P. The role of CD4 in HIV binding and entry. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **342**, 59–66 (1993).
9. Ciuffi, A., Mohammadi, P., Golumbeanu, M., di Iulio, J. & Telenti, A. Bioinformatics and HIV latency. *Curr. HIV/AIDS Rep.* **12**, 97–106 (2015).
10. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
11. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
12. Nilsson, R. H. et al. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* ;47:D259-259D264. (2019).
13. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods.* **13**, 581–583 (2016).
14. Clarridge, J. E. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin. Microbiol. Rev.* ;**17** (2004). :840 – 62, table of contents.
15. Bellemain, E. et al. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol.* **10**, 189 (2010).

16. Church, D. L. et al. Performance and Application of 16S rRNA Gene Cycle Sequencing for Routine Identification of Bacteria in the Clinical Microbiology Laboratory. *Clin. Microbiol. Rev.* **33**, e00053–e00019 (2020).
17. Alsharksi, A. N., Sirekbasan, S., Gürkök-Tan, T. & Mustapha, A. From Tradition to Innovation: Diverse Molecular Techniques in the Fight Against Infectious Diseases. *Diagnostics (Basel)*. **14**, 2876 (2024).
18. McDonald, E. G. et al. Pneumocystis jirovecii pneumonia in people living with HIV: a review. *Clin. Microbiol. Rev.* **37**, e0010122 (2024).
19. Wang, J. Y. et al. Risk of Pneumocystis jirovecii Pneumonia in Patients With HIV in Taiwan: Evidence from a Cross-sectional Study. *Vivo* **39**, 1054–1066 (2025).
20. Kann, G. et al. Risk factors for IRIS in HIV-associated Pneumocystis-pneumonia following ART initiation. *J. Infect.* **83**, 347–353 (2021).
21. Talapko, J. et al. Candida albicans-The Virulence Factors and Clinical Manifestations of Infection. *J. Fungi (Basel)*. **7**, 79 (2021).
22. Maurya, V. et al. Oropharyngeal candidiasis and Candida colonization in HIV positive patients in northern India. *J. Infect. Dev. Ctries.* **7**, 608–613 (2013).
23. Awoyeni, A. et al. Isolation and evaluation of Candida species and their association with CD4 + T cells counts in HIV patients with diarrhoea. *Afr. Health Sci.* **17**, 322–329 (2017).

Figures

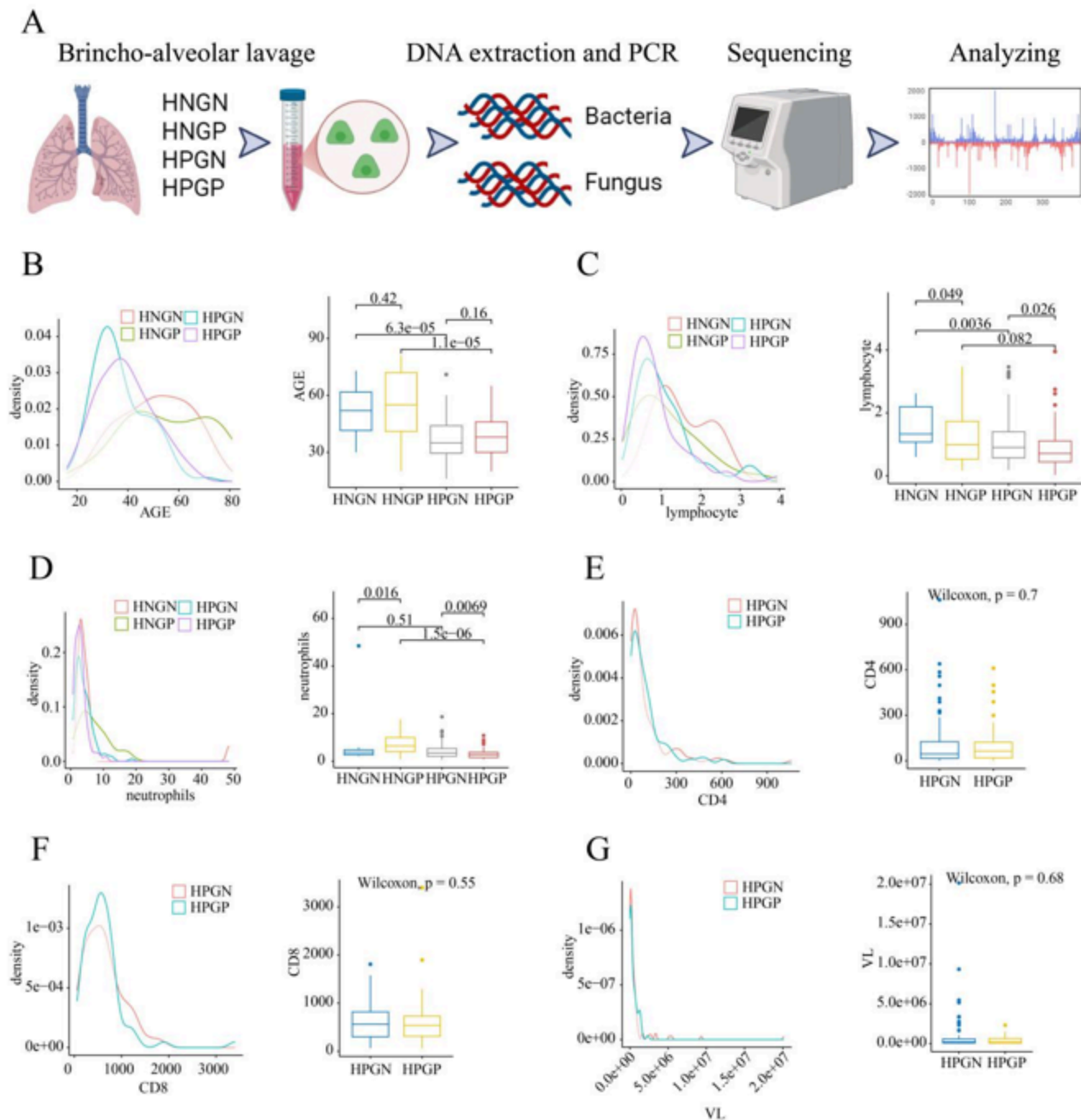


Figure 1

Experimental design process and demographic information.

A. Schematic diagram of experimental design process

B. Density map and box plot of age for enrolled individuals among HNGN, HNGP, HPGN, and HPGP.

C. Density map and box plot of lymphocyte levels for enrolled individuals among HNGN, HNGP, HPGN, and HPGP.

D. Density map and box plot of neutrophils levels for enrolled individuals among HNGN, HNGP, HPGN, and HPGP.

E. Density map and box plot of CD4⁺ T Cell levels for enrolled individuals among HPGN, and HPGP.

F. Density map and box plot of CD8⁺ T Cell levels for enrolled individuals among HPGN, and HPGP.

G. Density map and box plot of virus load for enrolled individuals among HPGN, and HPGP.

HNGN: HIV-negative and GM testing negative, HNGP: HIV-negative and GM testing positive, HPGN: HIV-positive and GM testing negative, HPGP: HIV-positive and GM testing positive.

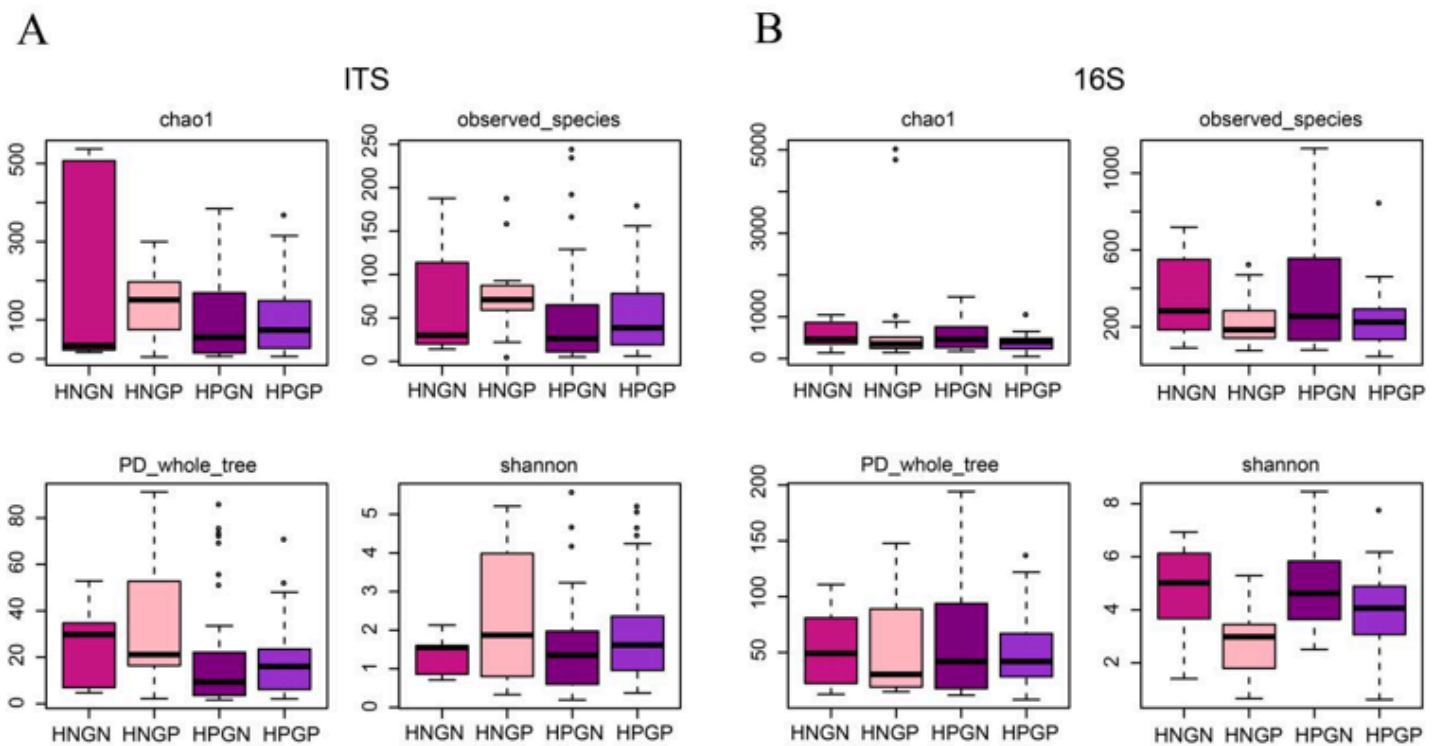


Figure 2

Diversity of microbe in HIV and non-HIV individuals.

A. The index statistics (ITS) of chao1, observed species, PD whole tree and shannon among groups.

B. The index statistics (16S) of chao1, observed species, PD whole tree and shannon among groups.

HNGN: HIV-negative and GM testing negative, HNGP: HIV-negative and GM testing positive, HPGN: HIV-positive and GM testing negative, HPGP: HIV-positive and GM testing positive.

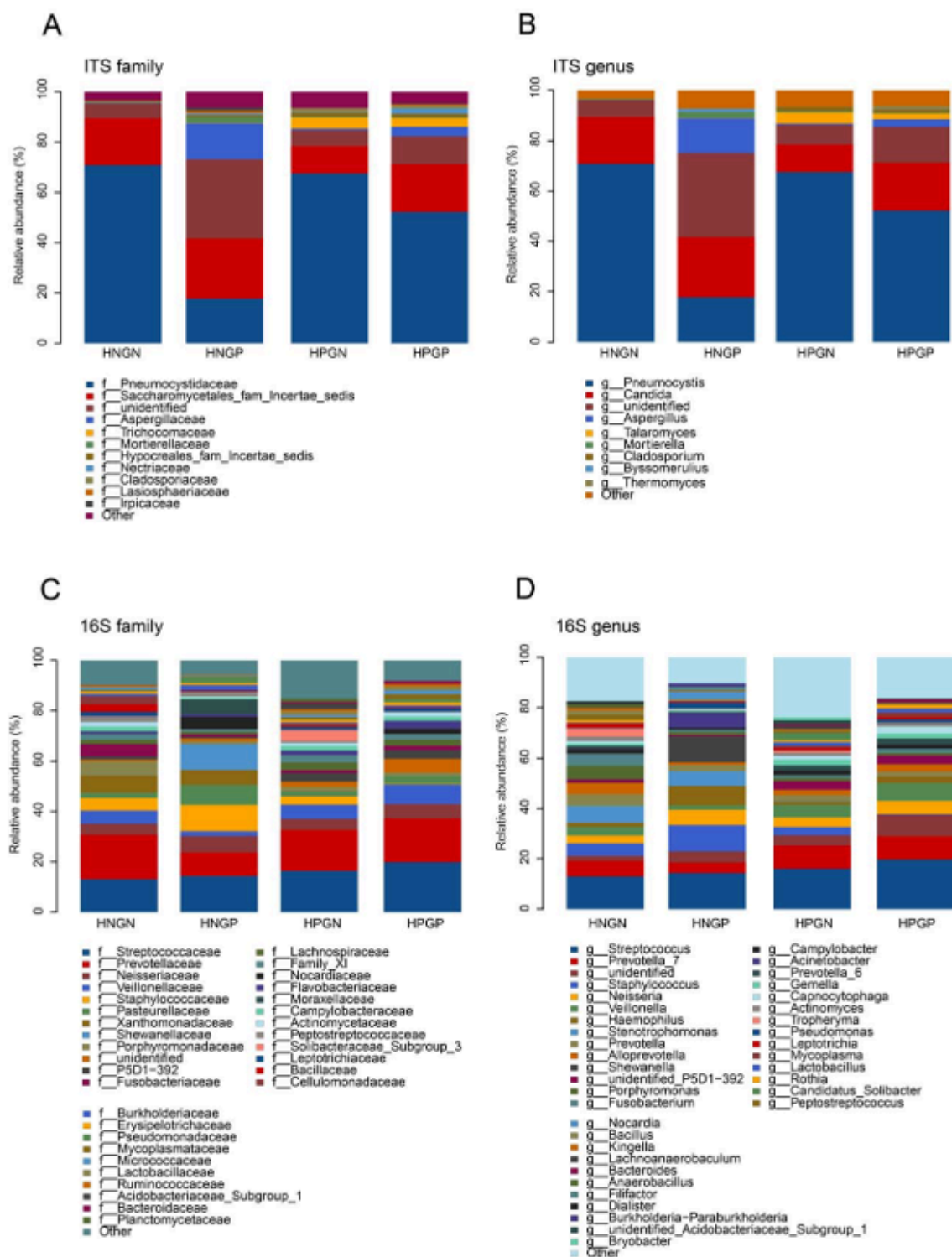


Figure 3

Landscape of microbial composition for enrolled individuals.

A. Bar plot of microbes at family level among groups for ITS.

B. Bar plot of microbes at genus level among groups for ITS.

C. Bar plot of microbes at family level among groups for 16 S.

D. Bar plot of microbes at genus level among groups for 16 S.

HNGN: HIV-negative and GM testing negative, HNGP: HIV-negative and GM testing positive, HPGN: HIV-positive and GM testing negative, HPGP: HIV-positive and GM testing positive.

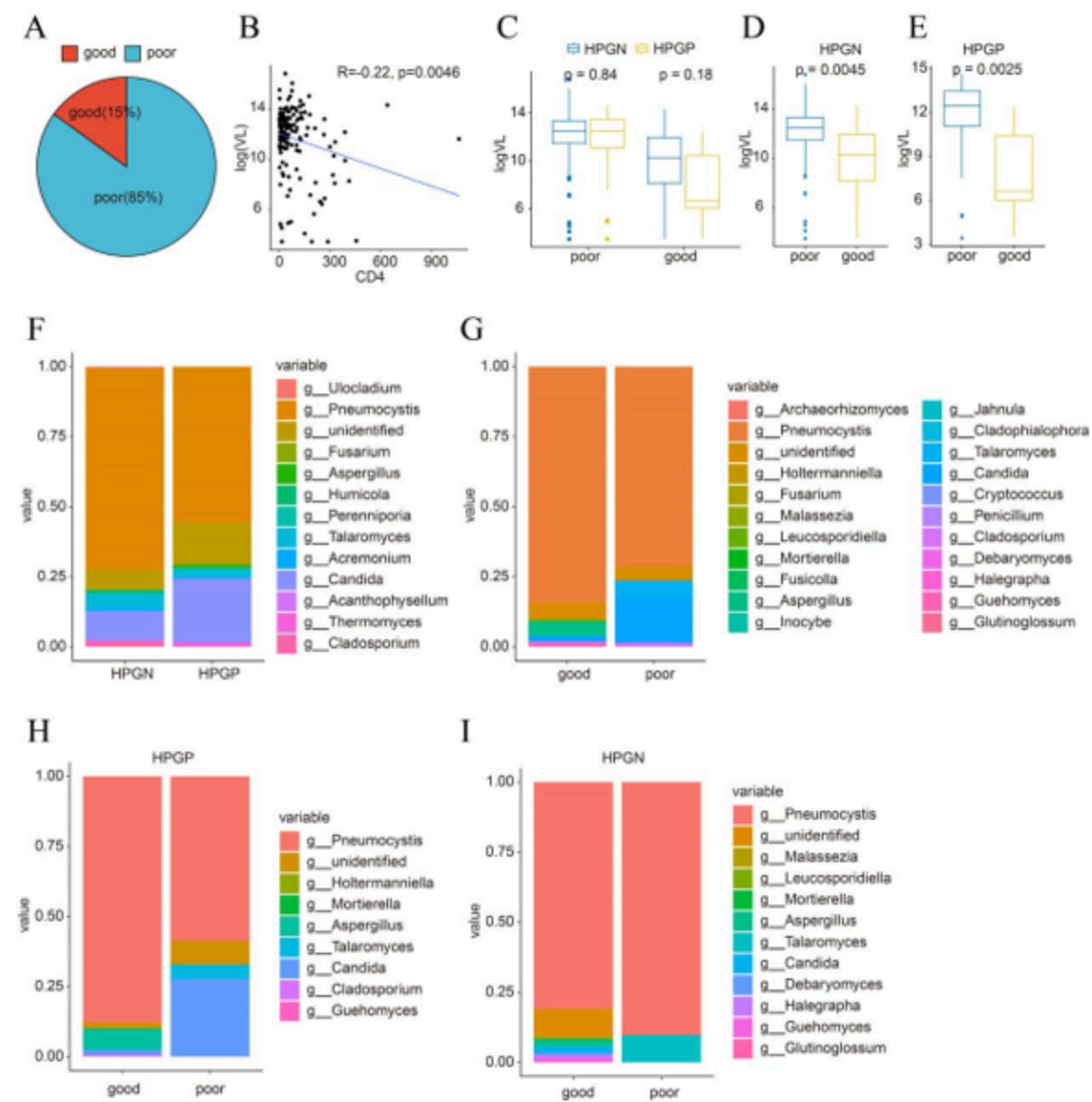


Figure 4

The correlation of percentage of microbe, CD4⁺ T Cell level and virus load.

A. Pie plot of HIV individuals with good or poor immune status.

B. Correlation map between CD4⁺ T Cell levels and virus load (log-transformed).

C. Box plot between HPGN and HPGP.

D. Box plot between poor and good immune status in HPGN individuals.

E. Box plot between poor and good immune status in HPGP individuals.

F. Bar plot of microbes at genus level between HPGN and HPGP in HIV individuals.

G. Bar plot of microbes at genus level between poor and good immune status in HIV individuals.

H. Bar plot of microbes at genus level between poor and good immune status in HPGN HIV individuals.

I. Bar plot of microbes at genus level between poor and good immune status in HPGP HIV individuals.

HNGN: HIV-negative and GM testing negative, HNGP: HIV-negative and GM testing positive, HPGN: HIV-positive and GM testing negative, HPGP: HIV-positive and GM testing positive.