

<https://doi.org/10.1038/s41522-025-00762-2>

Benefits and challenges of host depletion methods in profiling the upper and lower respiratory microbiome



Chun Wang^{1,2,6}, Li Zhang^{1,2,6}✉, Chuanwen Kan^{1,2}, Junyan He³, Weijie Liang⁴, Ruiyan Xia^{1,2}, Linyu Zhu^{1,2}, Jing Yang^{1,2}, Xuan Jiang^{1,2}, Wentai Ma^{1,2}, Zhihang Liang³, Zhiqing Xiao⁵, Juanjuan Zhang³, Jiaxin Zhong^{1,2}, Xiao Sun^{1,2}, De Chang⁵, Zhang Wang⁴, Guoliang Zhang³✉ & Mingkun Li^{1,2}✉

Metagenomic sequencing for respiratory pathogen detection faces two challenges: efficient host DNA depletion and the representativeness of upper respiratory samples for lower tract infections. In this study, we benchmarked seven host depletion methods, including a new method (F_{ase}), using bronchoalveolar lavage fluid (BALF), oropharyngeal swab (OP), and mock samples. All methods significantly increased microbial reads, species richness, genes richness, and genome coverage while reduced bacterial biomass, introduced contamination, and altered microbial abundance. Some commensals and pathogens, including *Prevotella* spp. and *Mycoplasma pneumoniae*, were significantly diminished. F_{ase} demonstrated the most balanced performance. High-resolution microbiomes profiling revealed distinct microbial niche preferences and microbiome disparities between the upper and lower respiratory tract. In pneumonia patients, 16.7% of high-abundance species (>1%) in BALF were underrepresented (<0.1%) in OP, highlighting OP's limitations as lower respiratory proxies. This study underscores both the potential and challenges of metagenomic sequencing in advancing microbial ecology and clinical research.

The respiratory microbiome plays a crucial role in maintaining ecological and immunological homeostasis within the respiratory tract^{1,2}. It is intimately linked with the development and progression of respiratory diseases, and response to treatment³. Metagenomic next-generation sequencing (mNGS) has revolutionized the study of the respiratory microbiome, enabling the characterization of microbial communities with unprecedented coverage and depth. Specifically, mNGS exhibits enhanced sensitivity and specificity while reducing turnaround time compared to conventional diagnostic methods in the detection of respiratory pathogens^{4–6}. mNGS has the capability to not only identify novel or challenging-to-culture pathogens, but also provide information on antibiotic resistance, enabling effective treatments and curtailing the spread of antibiotic resistance^{7–9}. However, the accuracy and sensitivity of metagenomic sequencing are hampered by the overwhelming amount of host-derived nucleic acids that overshadow microbial signals in respiratory tract samples^{10–12}.

Host DNA depletion has emerged as a promising solution to increase the yield of microbial sequences from metagenomic sequencing. The methods can be categorized into two categories: pre-extraction and post-extraction methods¹⁰. The pre-extraction methods involve a two-step procedure that eliminates mammalian cells and cell-free DNA, leaving only intact microbial cells for downstream DNA extraction; the post-extraction method selectively eliminates host DNA based on the theory that methylated nucleotides are more prevalent in the human genome compared to microbial genomes. Host DNA depletion has been applied on multiple samples types including saliva^{10,13}, cerebrospinal fluid^{14–17}, tissue biopsy^{13,18–21}, food (milk, chicken, pork, prawns)^{22–25}, ocular surface²⁶, skin swabs²⁷, and respiratory samples (bronchoalveolar lavage fluid (BALF), sputum, and nasopharyngeal aspirate specimens)^{12,14,17,21,28–33}.

A widely-used pre-extraction method for respiratory samples is saponin lysis of human cells followed by endonuclease digestion of cell-free DNA. The

¹Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, and China National Center for Bioinformation, Beijing, China. ²University of Chinese Academy of Sciences, Beijing, China. ³National Clinical Research Center for Infectious Diseases, Shenzhen Third People's Hospital, Southern University of Science and Technology, Shenzhen, China. ⁴Institute of Ecological Sciences, School of Life Sciences, South China Normal University, Guangzhou, China. ⁵Department of Pulmonary and Critical Care Medicine at The Seventh Medical Center, College of Pulmonary and Critical Care Medicine of The Eighth Medical Center, Chinese PLA General Hospital, Beijing, China. ⁶These authors contributed equally: Chun Wang, Li Zhang.

✉ e-mail: zhangli@big.ac.cn; szdsyy@aliyun.com; limk@big.ac.cn

concentration of saponin in previous studies varied from 0.025% to 2.50%^{14,29–31}. However, comparing their effectiveness is difficult due to differences in sample types and evaluation criteria. Other methods, including hypotonic lysis of human cells followed by endonuclease or PMA degradation of cell-free DNA^{17,28}, microfluidic separation of microbial cells followed by endonuclease digestion of cell-free DNA³², as well as commercial kits, including QIAamp DNA Microbiome kit, Zymo HostZERO Microbial DNA Kit and Molzym MolYsis Basic kit, also show varying effectiveness in host DNA removal for respiratory samples^{12,28,30,33}. In contrast, post-extraction methods, specifically using the NEBNext Microbiome DNA Enrichment Kit^{21,30}, have shown poor performance in removing host DNA from respiratory samples, consistent with findings from other sample types^{10,34}. Furthermore, the host removal process may cause varying extents of damage to microorganisms depending on the degree of cell wall fragility¹⁴, potentially leading to alterations in microbiota composition. However, the taxonomic biases of these methods are rarely investigated. Overall, although varied methods are available for respiratory samples, a comprehensive comparison of host DNA depletion effects across different methods and respiratory sample types is lacking, making it difficult to choose the optimal method.

The upper and lower respiratory tract microbiomes are interconnected, with some studies demonstrating good concordance between them. This supports the use of upper airway samples as surrogates for studying the lung microbiome, particularly when lower airway sampling is challenging^{35–38}. However, other studies have reported significant differences between these microbiomes^{39,40}. This discrepancy may be attributed to variations in health status, disease type and pathogenesis, age groups, or methodologies used for microbiome profiling. Additionally, previous studies on respiratory microbiomes were primarily conducted at lower resolution levels (genus and species). A finer-scale study at the strain and SNP level is needed to provide deeper insights into the concordance between the upper and lower respiratory microbiomes, enabling the establishment of a more realistic ecological model of the respiratory microbiome.

This study aimed to benchmark host DNA depletion methods for BALF and OP, which are representative sample types of the upper and lower respiratory tract, respectively. We applied seven pre-extraction host DNA depletion methods, including one developed in this study, four existing methods from the literature (with optimization of experimental conditions), and two commercial kits. We evaluated the performance of each method on the two sample types using systematic metrics, including effectiveness, fidelity, and contamination. Additionally, we identified the taxonomic bias for each method and further confirmed them using a mock microbial community. Finally, the enrichment of microbial DNA after host depletion enables an unprecedented comparison of the upper and lower respiratory tract microbiomes at the strain and SNP levels.

Results

Experiment design and respiratory sample characteristics

Thirty-five BALF samples and 34 paired OP samples were collected from 35 patients, including 26 with pneumonia and 9 with other diseases (Supplementary Table 1). Negative controls, consisting of saline water through the bronchoscope, unused flocking swabs, and deionized water, were collected and processed following the same experimental protocol. The OP samples had a median bacterial load of 24.37 ng/swab (interquartile range (IQR) 2.16–224.76), a median host DNA content of 50.20 ng/swab (IQR 6.45–235.21), and a median microbe-to-host read ratio of 1:7 in metagenomic data (IQR 1:14–1:2) (Fig. 1A–C). In contrast, BALF samples had a lower bacterial load of 1.28 ng/ml (IQR 0.32–16.90), higher host DNA content of 4446.16 ng/ml (IQR 1443.90–14289.62), and a microbe-to-host read ratio of 1:5263 (IQR 1:9091–1:667), which is consistent with previous reports⁴. Notably, both sample types contained a large proportion of cell-free microbial DNA, representing 68.97% of total microbial DNA in BALF and 79.60% in OP. These cell-free microbial DNAs cannot be captured by any of the pre-extraction host DNA depletion methods.

The seven pre-extraction host DNA depletion methods employed in this study included nuclease digestion (R_{ase}), osmotic lysis followed by

propidium monoazide (PMA) degradation (O_{pma}) or nuclease digestion (O_{ase}), saponin lysis followed by nuclease digestion (S_{ase}), 10 µm filtering followed by nuclease digestion (F_{ase}, a new method developed in this study), as well as two commercial kits: the QIAamp DNA Microbiome kit (K_{qia}) and the HostZERO Microbial DNA Kit (K_{zym}) (Fig. 1D, cost and turnaround time of each method are provided in Supplementary Table 2).

Experimental conditions were optimized to enhance host DNA depletion efficiency and minimize bacterial DNA loss. This included: 1) Testing PMA concentrations (10 µM, 30 µM, and 50 µM) with selection of 10 µM in O_{pma}; 2) Testing saponin concentrations (0.025%, 0.10%, and 0.50%) with selection of 0.025% in S_{ase}; 3) Testing sample cryopreservation methods (with glycerol and without glycerol), and selecting the addition of 25% glycerol to samples (Supplementary Fig. 1, Supplementary Materials).

In total, 483 host-depleted respiratory samples, 69 respiratory samples without host depletion (Raw), and 63 negative controls (NCs) underwent shotgun DNA sequencing. The median sequencing read number was 14.07 million for BALF samples (IQR 12.00–16.68 million), and 12.99 million for OP samples (IQR 10.16–16.47 million).

The host DNA depletion enhanced the taxonomic and genomic resolution of the respiratory microbiome

Host DNA load that measured by qPCR was significantly decreased by all methods by one to four orders of magnitudes (*p*-adj values < 0.05, Wilcoxon matched-pairs tests, Fig. 2A, B). Samples processed with the S_{ase} and K_{zym} methods exhibited the highest host DNA removal efficiency. For BALF, S_{ase} samples had a median of 493.82 pg/mL human DNA (1.1‰ of the original concentration) and K_{zym} samples had a median of 396.60 pg/mL human DNA (0.9‰). For OP, 82.35% of S_{ase} samples and 70.59% of K_{zym} samples had human DNA concentration below the detection limit (8.34 pg/swab) (Fig. 2A, B). Regarding the change in bacterial DNA loads, the R_{ase} method resulted in the highest bacterial retention rate in BALF (median 31%, IQR 5%–100%), and R_{ase} and K_{qia} showed the highest bacterial retention rate in OP (median 20%, IQR 9%–34% for R_{ase}, and 21%, IQR 11%–72% for K_{qia}, respectively) (Fig. 2C, D).

From the perspective of sequencing results, the K_{zym} method showed the best performance in increasing microbial reads (2.66% of total reads after host DNA depletion, 100.3-fold of that in Raw) in BALF samples, followed by S_{ase} (1.67%, 55.8-fold), F_{ase} (1.57%, 65.6-fold), K_{qia} (1.39%, 55.3-fold), O_{ase} (0.67%, 25.4-fold), and R_{ase} (0.32%, 16.2-fold), whereas O_{pma} showed the least effectiveness (0.09%, 2.5-fold) (Fig. 2E, Supplementary Fig. 2A). In OP samples, the S_{ase} method was the most effective in increasing microbial reads (65.60%, 5.9-fold), followed by K_{qia} (63.00%, 4.2-fold), K_{zym} (61.68%, 3.5-fold), F_{ase} (56.64%, 3.1-fold), O_{ase} (53.46%, 2.8-fold), R_{ase} (52.64%, 2.8-fold), and O_{pma} (44.92%, 2.1-fold) (Fig. 2F, Supplementary Fig. 2B). Collectively, the S_{ase}, F_{ase}, and K_{qia} methods performed relatively better in BALF samples, while the S_{ase} method showed superior performance in OP samples. Although K_{zym} removed the most host DNA and obtained the highest microbial reads, it resulted in the most significant loss of microbial DNA in both sample types (Fig. 2C, D), potentially causing severe contamination.

To assess the benefits of host DNA depletion, we compared the number of detected microbial species and functional gene families. In BALF samples, all methods, except O_{pma}, significantly increased the number of detected species compared to the Raw method. The highest increase was observed with K_{zym} (1.5-fold), followed by S_{ase} (1.5-fold), F_{ase} (1.4-fold), K_{qia} (1.4-fold), O_{ase} (1.4-fold), and R_{ase} (1.3-fold, *p*-adj values < 0.05 for all methods, Wilcoxon matched-pairs tests, Fig. 2G). K_{zym} detected the greatest number of gene families, with a 41.4-fold increase compared to Raw method, followed by F_{ase} (37.4-fold), Kit_{qia} (29.8-fold), S_{ase} (19.6-fold), O_{ase} (13.0-fold), R_{ase} (10.8-fold), and O_{pma} (2.4-fold) (Fig. 2H). In contrast, in OP samples, there was no significant increase in the number of detected species after host DNA depletion (Fig. 2I), suggesting that the species number had already reached saturation in Raw samples. Meanwhile, the numbers of detected gene families were moderately improved by all

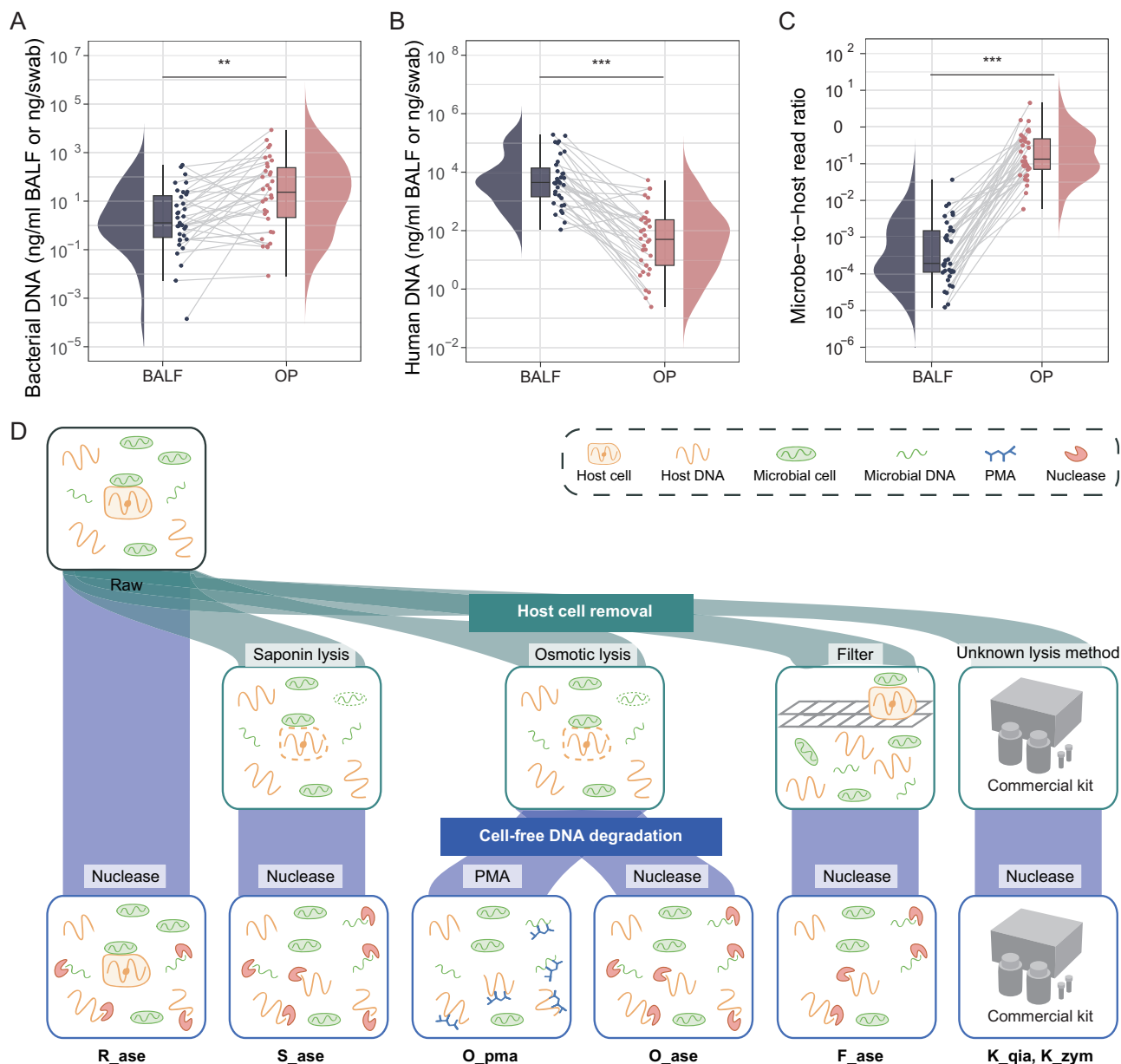


Fig. 1 | Respiratory sample characteristics and the schematic overview of host DNA depletion methods. **A** The amount of bacterial DNA. **B** The amount of host DNA. **C** The ratio of microbe-to-host read numbers in metagenomic data. Reads from contaminated microbes were discarded. In (A–C), each dot represents a sample, and samples from the same individuals were connected with lines; the center

line of the boxplot represents the median, box limits represent the upper and lower quartiles, and whiskers represent the 1.5x interquartile range. ** $p < 0.01$, *** $p < 0.001$, Wilcoxon matched pairs tests. $n = 35$ for BALF and $n = 34$ for OP. **D** Schematic overview of host DNA depletion methods evaluated in this study.

methods (1.1–1.4 fold) except O_pma (Fig. 2J). We further examined the enhancements in genome coverage (the percentage of the genome covered by at least one sequencing read) after host depletion. In BALF samples, genome coverage for 34 common respiratory bacteria substantially increased, with the highest median fold change observed for K_qia (37.6-fold), followed by K_zym (29.1-fold), F_ase (28.4-fold), S_ase (20.3-fold), O_ase (11.0-fold), R_ase (10.8-fold), and the least for O_pma (1.9-fold, p_{adj} values < 0.001 for all methods, Wilcoxon matched-pairs tests). In contrast, in OP samples, only a small number of the common bacteria exhibited moderately increased genome coverage following host depletion (Supplementary Fig. 3A, B).

We noted that better performance was achieved in samples with lower levels of microbial nucleic acids. When the microbial proportion in the original sample reached 10% or higher, no improvement in species richness was observed. Similarly, when the microbial proportion reached 30% or

higher, gene richness did not improve (Fig. 2K, L). The suboptimal performance in samples with a high microbial proportion might be due to the damage of microbial nucleic acids during the treatment.

Impact of host DNA depletion on the fidelity of microbiome profiling

We further investigated the impact of host DNA depletion on microbial composition, i.e., fidelity. First, we found that patient identity explained the highest variation in microbial composition ($R^2 = 49.64\%$, $p < 0.001$, PERMANOVA), whereas host DNA depletion methods accounted for 3.29% of the variance ($p < 0.001$). When clustering the samples based on microbiome similarities, they primarily clustered by patient identity rather than by the specific host depletion methods employed (Supplementary Fig. 4), suggesting that microbial characteristics were largely retained following the host DNA depletion treatments.

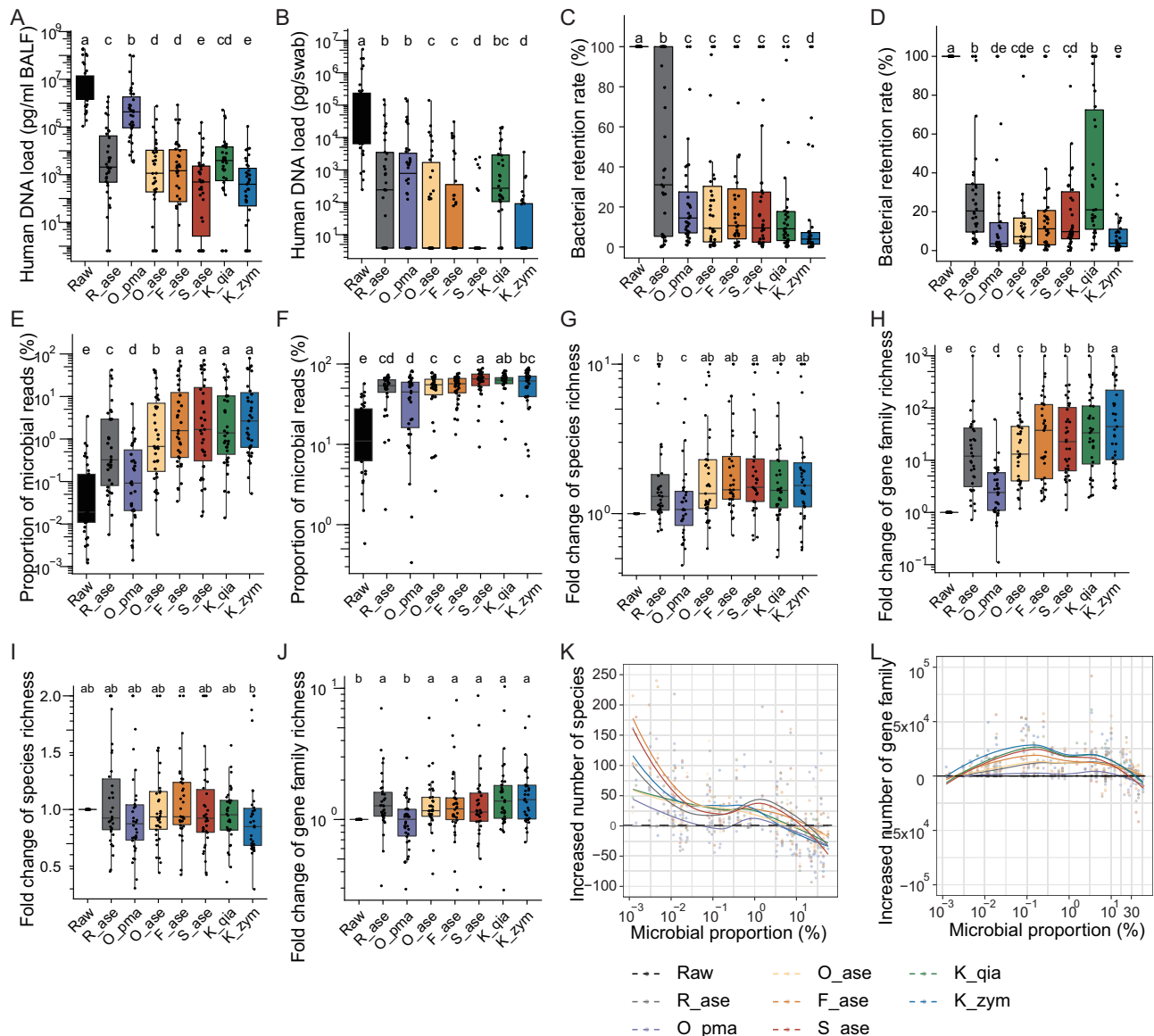


Fig. 2 | Comparison of effectiveness between host DNA depletion methods. Host DNA concentration (pg/ml BALF or pg/swab) in BALF (A) and OP (B) samples treated with different host DNA depletion methods. Bacterial retention rate in BALF (C) and OP (D) samples after host removal. Data in (A–D) were measured using qPCR assays. The proportion of non-contaminant microbial reads in raw sequencing data for BALF (E) and OP (F). Fold change of species richness (G) and gene family richness (H) between host-depleted and raw BALF samples. Fold change of species richness (I) and gene family richness (J) between host-depleted and raw OP samples. Species richness was represented by the number of species. Gene family richness was represented by the number of gene families. Correlation between the

proportion of microbial reads in Raw samples and the improvement in species richness (K) and gene family richness (L) after host depletions. The y-axis represents the difference in richness between host-depleted and Raw samples. Fitting curves of loess regression are indicated for each method. In (A–J), the center line of the boxplot represents the median, box limits represent upper and lower quartiles, and whiskers represent the 1.5x interquartile range. Different letters on the top indicate statistically significant differences among methods (Wilcoxon matched pairs tests followed by FDR adjustment). In (G–L), data from eight samples from the same patient were rarefied to the lowest sequencing depth among them. $n = 35$ for BALF and $n = 34$ for OP.

Notably, the clustering analysis revealed several distinct branches that were clustered based on the depletion methods, specifically involving the K_zym and K_qia methods. Samples within these branches had a lower bacterial load and included NC, suggesting a significant influence of kit-specific contamination in samples with low microbial biomass. Even after applying the 5-fold NC algorithm (details in Methods) to eliminate potential contaminating microbes, the two methods still accounted for a higher variance in microbiome composition compared to other methods ($R^2 = 1.85\%$ for K_qia, $R^2 = 1.61\%$ for K_zym, $p < 0.05$; $p > 0.05$ for other methods, PERMANOVA). In contrast, the F_ase method displayed the highest similarity in microbiota composition to the R_ase method (median Jensen-Shannon distance (JSD) = 0.24 for BALF and 0.12 for OP), which

selectively removed cell-free microbial DNA, thereby reflecting the composition of microbes with intact cells. Following F_ase, the next closest methods were O_ase (0.25 for BALF, and 0.19 for OP), S_ase (0.29 for BALF and 0.20 for OP), and O_pma (0.32 for BALF and 0.29 for OP) (Fig. 3A, B). These distances fell within the range observed between technical replicates (0.09–0.33) of OP samples processed in our lab (Fig. 3A, B).

The impact of host removal methods was further explored on the microbial load and the abundance of common respiratory microbial species, which comprised 74.32% of the microbial reads in BALF samples and 82.73% in OP samples. Despite the significant reduction in bacterial loads following host depletion as compared to the R_ase method, notable alterations in relative abundances were observed in both directions

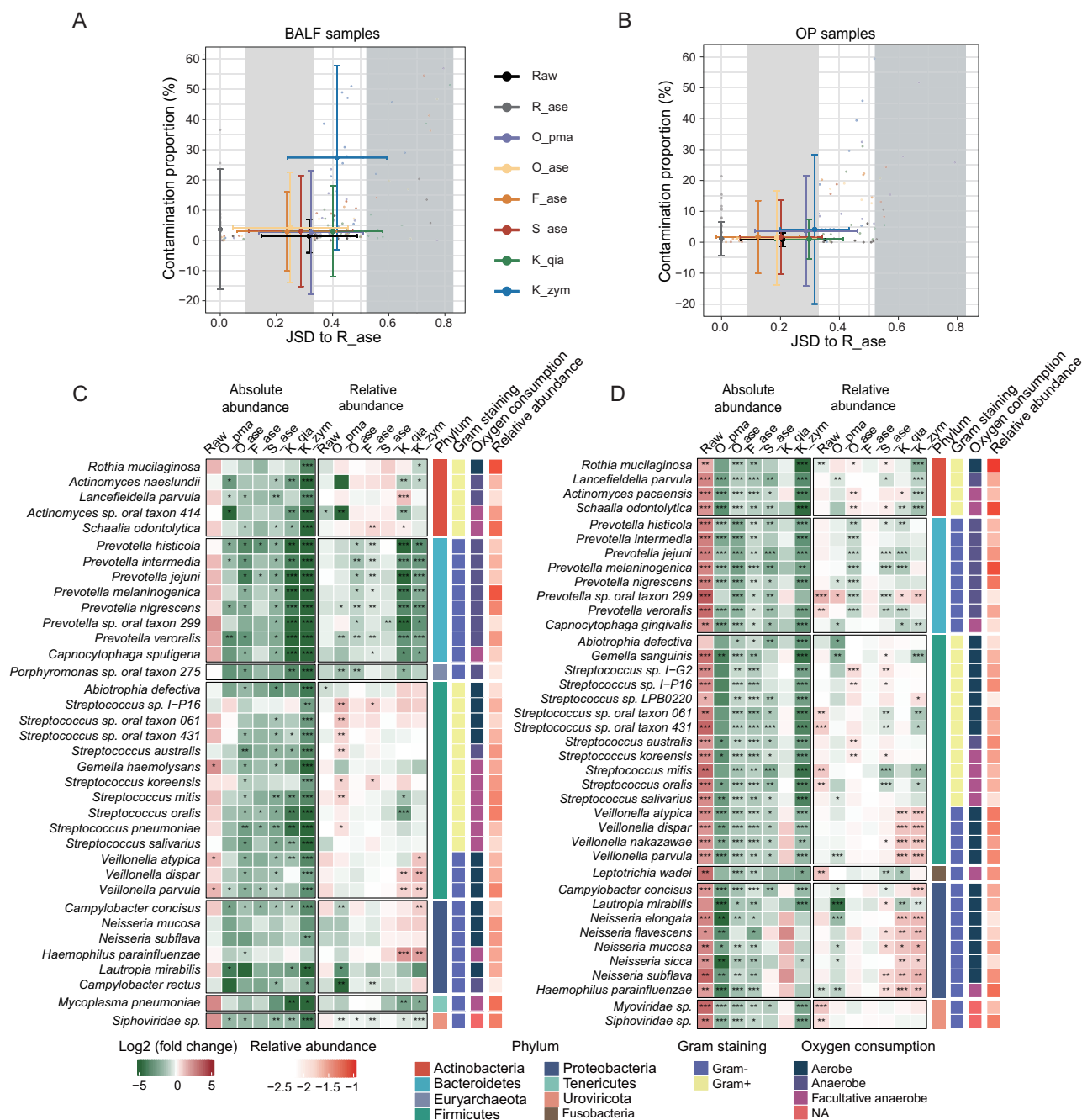


Fig. 3 | Influence of host DNA depletion on the fidelity of microbiome profiling. Similarity in microbial composition between R_ase and other methods, and proportion of contaminating microbial reads in total microbial reads in BALF (A) and OP (B) samples. Dots indicate median values of different methods, and error bars indicate standard deviations. JSD (Jensen-Shannon distance) was calculated excluding identified contaminating components. The gray area on the left indicates the range of JSD between technical replicates, while the gray area on the right

indicates the 5th to 95th percentile range of JSD among samples from different individuals. Absolute and relative abundance changes of common species in BALF (C) and OP (D) samples. Fold changes in abundances were calculated between different methods and R_ase. Asterisks indicate significant differences in abundances between different methods and R_ase, * p .adj < 0.05, ** p .adj < 0.01, *** p .adj < 0.001, Wilcoxon matched pairs tests. $n = 35$ for BALF and $n = 34$ for OP.

(Fig. 3C, D). F_ase resulted in the fewest species with significant alterations in relative abundances (12 in BALF and 0 in OP), followed by O_ase (9 in BALF and 14 in OP) and S_ase (1 in BALF and 24 in OP). It is noteworthy that most host DNA depletion methods led to a decrease in the relative abundance of seven *Prevotella* species, which are Gram-negative and anaerobic. Additionally, the pathogenic microorganism *Mycoplasma pneumoniae* exhibited a significant decline following treatments with the K_qia and K_zym. This decline could pose a significant challenge in pathogen detection.

To measure the influence of exogenous contaminant microorganisms introduced by different host depletion methods, species with relative abundances less than five-fold of those in negative controls were labeled as contaminants. K_zym introduced the highest level of contaminants. The median percentage of contaminant reads was 27.39% in BALF samples (IQR 16.19–50.00%), and 4.22% in OP samples (IQR 1.01–21.13%). The other methods exhibited a median proportion of microbial contaminants below 4.21% in BALF samples and 3.67% in OP samples (Fig. 3A, B). This was likely due to two factors. First, potential contamination from the kit itself

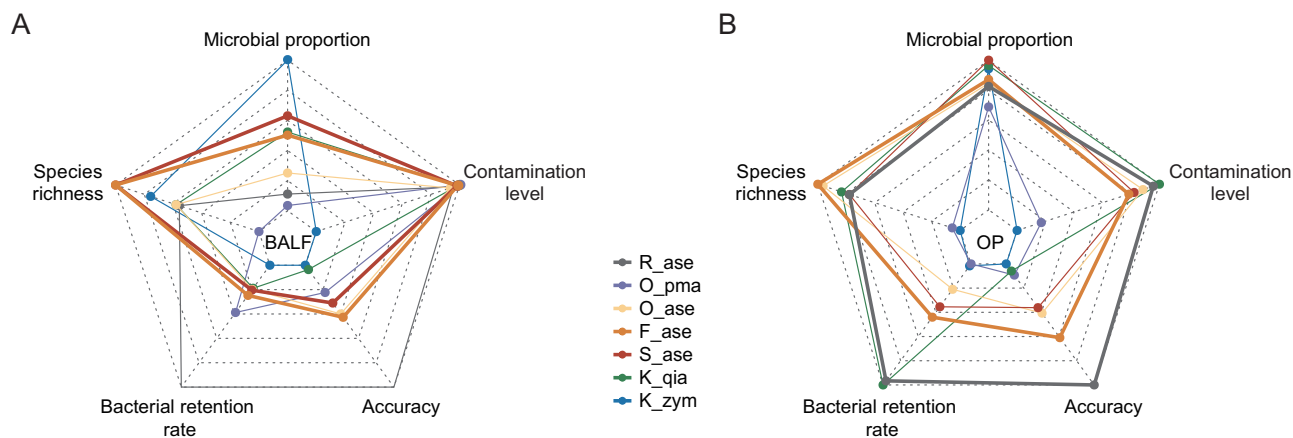


Fig. 4 | Integrated evaluation of host DNA depletion methods. Radar charts showing the performance of host DNA depletion methods on five evaluation metrics in BALF (A) and OP (B) samples. Accuracy was defined as 1 subtracted by the JSD between microbial compositions of different methods and *R_ase*. The contamination level was defined as 1 subtracted by the proportion of contaminants in total

microbial reads. Microbial proportion and accuracy were estimated excluding exogenous contamination. Maximum-minimum measures were used for data scaling. The two methods with the best performance (largest radar area) are labeled in bold. $n = 35$ for BALF and $n = 34$ for OP.

may have introduced microbial nucleic acids during sample processing. Second, significant microbial loss occurred during processing (Fig. 2C, D). This loss increased the proportion of contaminants in the final sample composition. Meanwhile, the microbial composition of negative control samples varied significantly among host DNA depletion methods ($R^2 = 0.30$, $p < 0.001$, PERMANOVA). The contaminants introduced by *K_qia* and *K_zym* exhibited relatively homogeneous compositions, forming distinct clusters on the PCoA plot, while the other methods showed heterogeneous contaminants compositions (Supplementary Fig. 5A, B).

Integrated evaluation of host DNA depletion methods

Five performance metrics were used for the comprehensive evaluation of host DNA depletion methods: the proportion of non-contaminant microbial reads, species richness, bacterial retention rate, compositional similarity to the *R_ase* (i.e., accuracy), and the proportion of contaminants in total microbial reads (i.e., contamination level). The *F_ase* and *S_ase* methods exhibited superior performance for BALF samples, demonstrating a well-balanced outcome across all metrics. Other methods exhibited significant defects in various metrics, including low effectiveness (microbial proportion and species richness) for *O_ase*, *O_pma*, and *R_ase*, low accuracy for the two commercial kits, and substantial bacterial loss and contaminants for *K_zym* (Fig. 4). For OP samples, *R_ase* demonstrated the largest radar chart area, indicating it is the optimal method. Despite *R_ase* showing lower effectiveness compared to top-performing methods, the performance difference was minimal or not statistically significant. Other viable options include the *F_ase*, *O_ase*, and *S_ase* methods, whereas *O_pma*, *K_qia*, and *K_zym* exhibited notable limitations in at least one aspect of their performance.

Assessment of host DNA depletion methods using a mock microbial community

To further assess the performance of different host depletion methods, we constructed a mock community comprising 15 respiratory commensal bacteria and pathogens (Supplementary Table 3). All methods resulted in varying levels of bacterial DNA loss. The *O_ase* method exhibited the most substantial loss (Supplementary Fig. 6), whereas in respiratory samples, *K_zym* showed the greatest reduction in bacterial DNA (Fig. 2C, D). This difference may indicate variations in microbial characteristics and status among the samples. The relative and absolute abundances of 15 species obtained with *F_ase* most closely resembled the reference profile (obtained with *R_ase*), followed by *S_ase*, *O_ase*, and *O_pma*. In contrast, low concordance was observed between the two commercial kits and the reference profile (Fig. 5A, B), validating findings using respiratory samples.

Comparisons of upper and lower respiratory tract microbiomes

Host DNA depletion enabled higher-resolution analysis of the upper and lower respiratory microbiota. Using data from the *F_ase* method, which exhibited optimal performance in host DNA depletion as a representative, we observed similar Shannon diversity indices at the species level between OP and BALF samples (Supplementary Fig. 7A). The JSD distances between OP and BALF samples from the same individuals was notably smaller than distances among OP samples, among BALF samples, and among unpaired OP-BALF samples (Fig. 6A). Despite these similarities, microbial species composition significantly differed between OP and BALF samples ($R^2 = 3.42\%$, $p = 0.001$, PERMANOVA, Fig. 6A). While species abundances were correlated between the two sites (Pearson correlation coefficient 0.60, $p < 0.001$, Fig. 6B), some species, such as several *Veillonella* species, showed distinct abundance patterns between OP and BALF samples (Fig. 6C). Among species with relative abundances above 0.01, shared species accounted for 52.0% (IQR 26.6–66.7%) and 55.3% (IQR 18.8–77.1%) of the number of species, and 79.7% (IQR 32.2–92.5%) and 69.0% (IQR 19.5–83.7%) of the abundance in OP and BALF samples, respectively (Fig. 6D).

Notably, pneumonia cases exhibited higher numbers and abundances of shared species between OP and BALF compared to non-pneumonia diseases (Fig. 6E, F). Moreover, pneumonia cases showed a stronger correlation in species abundances between the two sites compared to non-pneumonia cases (Pearson correlation coefficient 0.67 vs. 0.41, Supplementary Fig. 7B). However, several pneumonia outliers showed low concordance, characterized by an exceptionally higher abundance of potential pathogens, such as *M. pneumoniae*, *Haemophilus influenzae*, and *Mycobacterium tuberculosis* in BALF than in OP (Fig. 6D, Supplementary Fig. 7C).

We further conducted a strain-level comparison of the OP and BALF microbiomes using the strainGE results⁴¹. In most cases, strains detected in BALF were a subset of those found in OP, probably due to higher read numbers obtained from OP samples. An exception was observed with *M. pneumoniae*, where strains in OP were a subset of those in BALF, with higher *M. pneumoniae* read numbers in BALF (Fig. 6G). Of note, the strain diversity of *V. parvula* and *V. dispar* was higher in BALF compared to OP in a substantial number of individuals (6 of 34, and 7 of 24, respectively), despite the *Veillonella* reads being more abundant in OP samples in all but two individuals (Fig. 6G). This suggests a potential niche preference for specific strains. Specifically, 20.0–64.4% (median: 46.8%) of strains in OP and 15.0–72.2% (52.2%) of strains in BALF were shared between the two sites across different species.

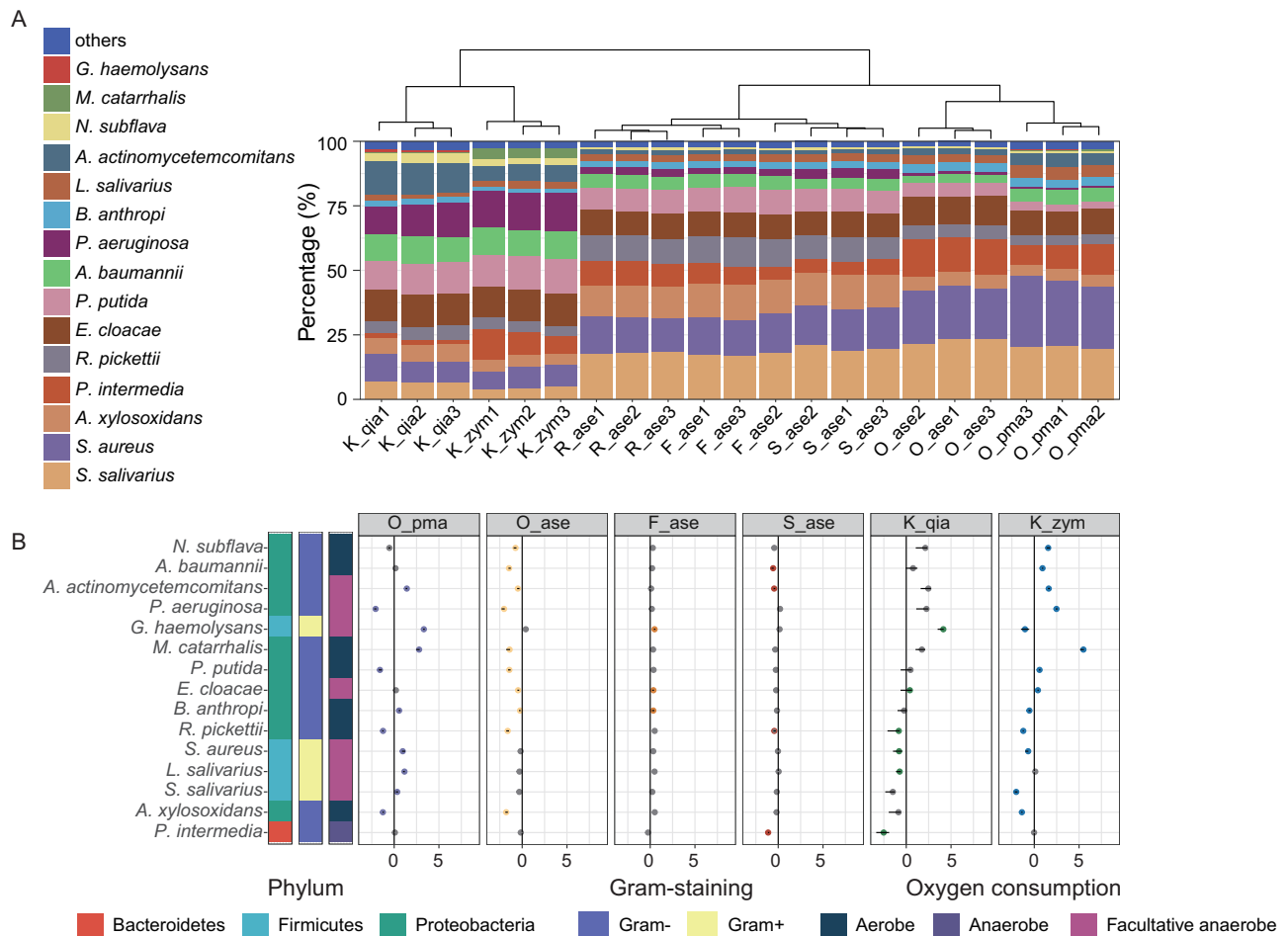


Fig. 5 | Evaluation of host DNA depletion methods using a mock microbial community. **A** Microbial composition of the mock community obtained with different methods. Hierarchical clustering was performed based on JSD. **B** Fold changes

in microbial absolute abundances between different methods and R_ase. Significant differences between different methods and R_ase are indicated with colored dots (paired Student's t-tests followed by FDR adjustment). $n = 3$ replicates.

At the sequence level, *V. parvula*, the predominant *Veillonella* species in nearly half (20/41) of the samples, showed greater genomic similarity between BALF and OP samples from the same individuals compared to those from different individuals, suggesting transmission between the upper and lower respiratory tracts (4% vs. 13%, $p = 1.9 \times 10^{-7}$, Wilcoxon signed-rank test, Supplementary Fig. 7D). Notably, 63 SNPs showed significant differences in prevalence between the two sites ($p < 0.01$, Fisher's exact test), including 14 nonsynonymous mutations in genes involved in vitamin B2 and B12 biosynthesis (Supplementary Fig. 7E).

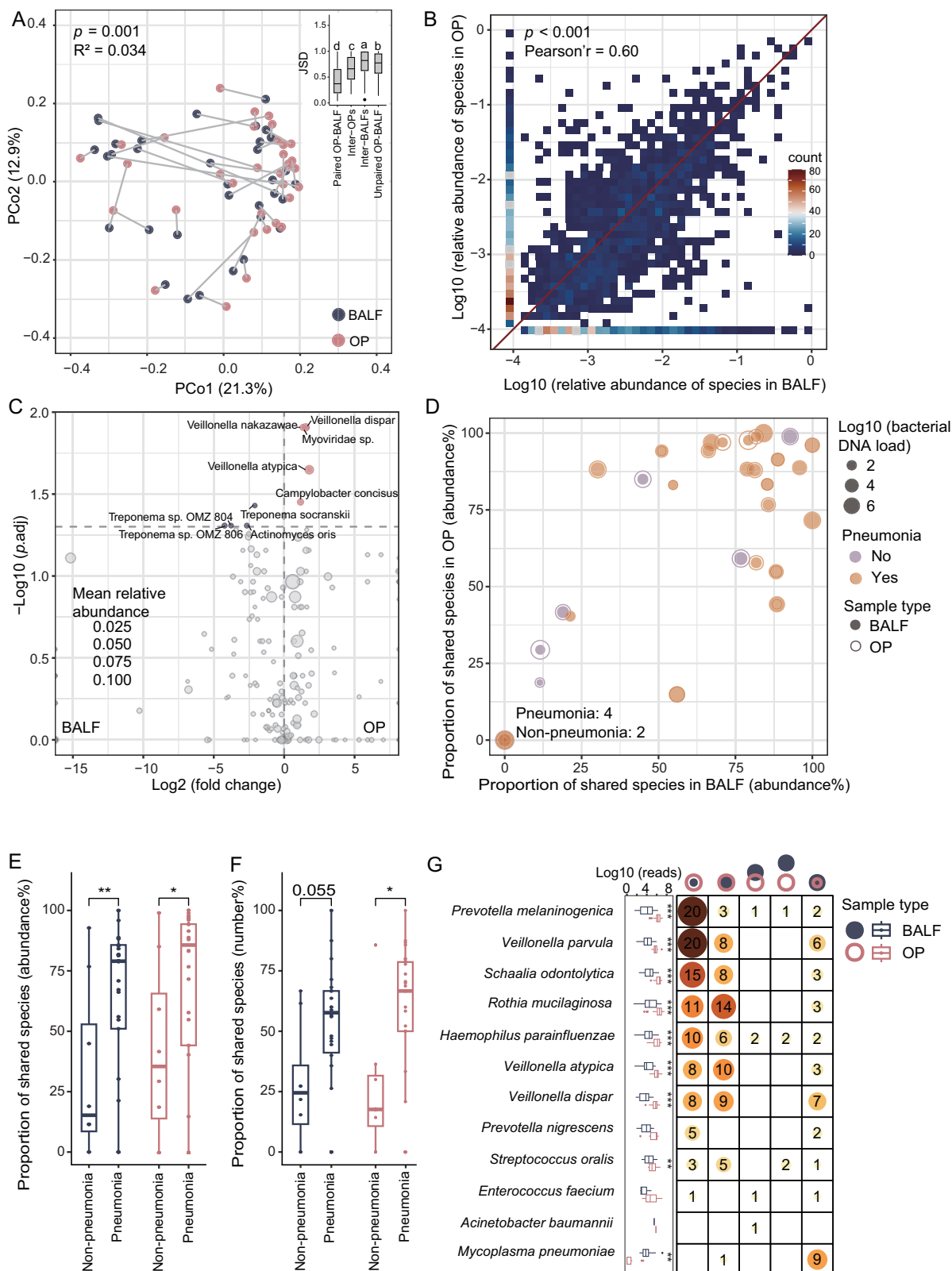
In contrast, in samples without host depletion, no differential species were detected between OP and BALF samples. Moreover, the significant differences previously observed in the number of OP-BALF-shared abundant species between pneumonia and non-pneumonia diseases were no longer evident. Additionally, the detected number of strains and SNPs were insufficient to discern meaningful differences (Supplementary Fig. 8).

Discussion

Respiratory infections are the leading cause of infectious disease-related death globally. Despite the promise of clinical metagenomics—a cutting-edge diagnostic method capable of identifying both known and novel pathogens—the sensitivity for detecting respiratory pathogens varies from 50.7% to 88.3%^{42–46}, indicating a substantial need for improvement. These challenges are partially due to the low microbial biomass and high host DNA content in respiratory samples, which hampers obtaining sufficient high-quality data. By benchmarking seven pre-extraction host DNA depletion methods for BALF and OP samples, we identified optimal methods and

experimental conditions tailored to each sample type, which can substantially promote the efficiency of metagenomics in unveiling respiratory microbiomes.

We developed a new host DNA depletion method, F_ase, which combines 10 μm filtration with nuclease digestion. This method demonstrated superior performance in BALF and mock samples, and was the second-best option after R_ase for OP samples. F_ase showed well-balanced performance in the host removal, resulting in an increased microbial proportion, a higher number of detected microbial species, good retention of bacterial DNA, low levels of exogenous contamination, and notably low bias in profiling endogenous microbes. The high accuracy of F_ase can be attributed to its use of cell size-based selection for eliminating host cells, rather than chemical or osmotic lysis, which causes varying degrees of microbial cell damage. We noted that a previous study applied filtration with a 5 μm filter to saliva¹⁰, deeming the method ineffective. This may be due to the existence of the cell-free host DNA, which can constitute a significant portion (mean 96% in our study) of total host DNA and dominate sequencing reads. This highlights the essential need for digesting cell-free host DNA after filtration. In addition, digesting cell-free DNA can also reduce noise from dead microbes in subsequent analyses. Additionally, three other cell size-based strategies are available for selectively removing host cells: microfluidic separation, flow cytometry cell sorting, and mechanical shearing with beads^{13,32}. However, the first two strategies have limitations: higher facility requirements, increased time consumption, and elevated contamination risk. These limitations arise from complex procedures in microfluidic separation (e.g., chip degassing, BSA incubation to



prevent cell adherence⁴⁷, and pre-filtration to avoid clogging) and the need for specialized equipment and protocols to prevent biological hazards when processing infectious agents with flow cytometry. Both methods result in substantial microbial component loss: microfluidic separation leads to a ~50% loss rate in sputum³², while flow cytometry often loses viral particles. In contrast, a recently developed method, mechanical shearing with 1.4 mm

beads (MEM method), is simpler and has shown promising performance across various sample types, though it has not yet been tested with respiratory samples¹³. Here, we compared the performance of MEM, S_{ase}, and F_{ase} against R_{ase} on eight BALF samples from pneumonia samples. All three methods showed better performance in microbial proportion and species richness than R_{ase} (Supplementary Fig. 9, Supplementary results).

Fig. 6 | Comparison between the upper and lower respiratory tract microbiomes. **A** Principal Co-ordinates Analysis (PCoA) plot based on JSD of microbial compositions. R^2 and p values from PERMANOVA are shown. BALF and OP samples from the same individuals were connected with lines. Inserted boxplots show the JSD between different sample types, and different letters on the top indicate statistically significant differences between groups. **B** 2D density plot showing the correlation of species abundances between BALF and OP. A total of 209 species with relative abundances higher than 0.01 in at least one sample were included. Pearson correlation coefficient and p value are shown. **C** Volcano plot showing species with differential relative abundances between BALF and OP samples. The horizontal dash line indicates the adjusted p value of 0.05. Differential species are color-coded and labeled with their names. In (A, B), Wilcoxon matched pairs test followed by FDR adjustment was used. **D** The proportion of OP-BALF-shared species abundance to the total abundance of species. Comparison of the proportion of OP-BALF-shared

species between pneumonia and non-pneumonia cases, as per abundance (E) and number (F) of species. In (D–F), only abundant species with relative abundances higher than 0.01 were included. **G** The relationship between strains detected in the upper and lower respiratory tracts. The number as well as the color and size of dots indicate the number of strains. Symbols on the top indicate the following relationships: strains in BALF were a subset of those in OP, BALF and OP had the same strains, BALF and OP had partially overlapping strains, BALF and OP had completely different strains, strains in OP were a subset of those in BALF. The boxplots on the left indicate the read number of species. In (E–G), asterisks indicate significant differences, * $p_{\text{adj}} < 0.05$, ** $p_{\text{adj}} < 0.01$, *** $p_{\text{adj}} < 0.001$, Wilcoxon matched pairs tests. In (A) and (E–G), the center line of the boxplot represents the median, box limits represent upper and lower quartiles, and whiskers represent 1.5x interquartile range. F_{ase} treated samples were used. $n = 34$ for BALF and OP.

However, S_{ase} performed worse in accuracy and contamination control compared to the others. MEM outperformed the F_{ase} method in terms of microbial proportion and contamination control but showed comparable accuracy. Notably, MEM discards the supernatant after centrifugation, which may result in the loss of viral particles, warranting caution and further improvements.

The performance of host DNA depletion methods varied between BALF, OP, and mock samples. This variability can be attributed to the intrinsic properties of the samples, including host cell integrity, microbial cell integrity and viability, microbial biomass and proportion, and the ratio of extracellular to intracellular nucleic acids. Each method's effectiveness in host removal is significantly influenced by the working concentrations of chemicals used. While our study found optimal concentrations of PMA and saponin were similar for BALF and OP samples, caution is warranted when generalizing these findings to other respiratory samples. Meanwhile, host removal procedures are not always necessary. Samples with more than 10% DNA attributed to microbes did not significantly benefit from host removal compared to samples with lower microbial DNA amounts in our study. Therefore, the choice of host DNA depletion method and experimental conditions should be tailored to specific sample types and study objectives.

All host removal methods resulted in the loss of microbial DNA and introduced bias in elucidating the microbial composition due to the difference in microbial physiological characteristics. For instance, Gram-negative bacteria and bacteria without cell walls are more susceptible to damage caused by host removal methods⁴⁸. This aligns with our finding, where the absolute and relative abundance of *Prevotella* spp. and *M. pneumoniae* notably decreased after host removal. Conversely, the loss of some Gram-negative bacteria, such as *Veillonella* spp., *Neisseria* spp., and *H. influenzae* was less severe, with slight decreases in absolute abundances and significant increases in relative abundances. This suggests that cell wall structure is not the sole influencing factor; oxygen tolerance may also play a role. Anaerobic bacteria, including *Prevotella* spp. and *M. pneumoniae*, being more vulnerable to cell damage under ambient conditions, might be more affected during the host removal process⁴⁹. The degree of bias in microbiota composition varied among different methods, with the two commercial kits, K_{qia} and K_{zym} exhibiting the most significant changes in the relative abundance of respiratory microbes (Figs. 3, 5). This might be due to greater damage to indigenous microbes and the introduction of more extraneous microbial DNA by these kits. Therefore, comparisons of the microbial composition should be made among samples subjected to the same host depletion methods, and correction methods for microbiota composition biases after host depletion should be developed when sufficient data are available.

The proportion of microbial reads increased by 66-fold for BALF samples and 3-fold for OP samples after host DNA depletion, enabling us a more detailed characterization of the upper and lower respiratory tract microbiomes. We found a significant correlation between the abundance of species in OP and BALF samples, with a stronger correlation in pneumonia samples compared to non-pneumonia samples. Furthermore, similarities in microorganisms between OP and BALF from the same individual were

observed at both the strain and genome level. This supports the rationale for using OP samples as a proxy for lower respiratory tract samples in pathogen detection and microbiota profiling in previous studies^{36,50,51}. These correlations are likely due to microbial aspiration, which introduces microorganisms from the upper to the lower respiratory tract, a process that can be exacerbated during respiratory infections^{51,52}. However, 16.7% of the high-abundance species (relative abundance >1%) in BALF were undetectable (<0.1%) in OP in pneumonia patients, and some potential pathogens, such as *M. pneumoniae*, *Haemophilus influenzae*, and *Mycobacterium tuberculosis* exhibited significantly reduced abundance in OP samples, indicating a potentially high false-negative rate in pathogen detection. Additionally, we found colonization preferences among different species and strains, underscoring limitations in using OP as a proxy for profiling the lower respiratory tract microbiome. Notably, other low-abundance species may exhibit similar phenomena, but their low abundance or limited prevalence prevented them from being detected in our study. The colonization preferences might be attributed to environmental factors such as oxygen levels, nutrient availability, immune pressure, and pH, which select for specific microbial communities⁵³. For instance, we observed that differential SNPs between OP and BALF *V. parvula* strains could impact microbial functions like vitamin synthesis. Developing a comprehensive map of the lower respiratory microbiome and isolating these microorganisms for functional analysis are further needed to enhance the understanding of their specific roles in respiratory diseases.

There are several limitations in this study. First, we used frozen but not fresh respiratory samples, which had been stored for over a year prior to the experiment. This may have led to reduced microbial viability, higher cell-free host DNA proportion, and consequently poorer method performance. While this scenario is suitable for microbiome studies, it is not ideal for clinical diagnostics, where using fresh samples may increase the effectiveness of host DNA depletion methods. Second, sputum samples were not included in our assessment as they are not feasible for some patients. While results from sticky BALF samples can offer some insights, optimal experiment parameters for sputum still need to be determined. Third, despite a significant increase in microbial proportion observed in BALF samples after host DNA depletion, human reads still constituted more than 90% of total reads. This underscores the necessity for novel host depletion methods for challenging samples with very low microbial biomass and proportion.

Nevertheless, this benchmarking study provides a foundational basis for selecting host removal methods for different types of respiratory samples. The adoption of these methods could facilitate the accumulation of high-quality microbial genomes from the respiratory tract, enhance understanding of microbial composition and function across various respiratory tract locations, and ultimately advance pathogen detection and microbiota-targeted prevention and treatment of respiratory diseases.

Methods

Sample collection and ethics

BALF and OP samples were collected from patients with pneumonia (26) or other diseases (including two chronic obstructive pulmonary disease

patients, two bronchitis patients, one bronchiectasis patient, one lung cancer patient, and three patients without pulmonary diseases) at Longgang People's Hospital and Huizhou First People's Hospital in Shenzhen, China. Another eight BALF samples were collected from patients with pneumonia at the Chinese PLA General Hospital in Beijing. Each BALF sample was mixed with an equal volume of saline with 50% glycerol, and each OP swab was soaked in 1 mL saline with 25% glycerol after collection. The samples were left at room temperature for 20 min to let glycerol diffuse across the cell wall and membrane, and then transferred to -80°C freezer. One negative control sample for BALF and one for OP were collected each day whenever respiratory samples were collected, by injecting saline through the bronchoscope or using plain swabs. For the evaluation of the repeatability of technical replicates, five healthy volunteers were enrolled, and one OP samples were taken from each volunteer. Each sample was processed and analyzed twice under identical conditions to assess the extent of technical bias.

The protocol of this study was approved by the Ethics Committee of Shenzhen Third People's Hospital (No.2018-013). The written informed consent was obtained from the patient or legal guardian of the included patient.

Construction of mock community

Fifteen bacterial species representing respiratory commensals and pathogens were used to construct a mock community to test the performance of host depletion methods. The mock community included 4 Gram-positive and 11 Gram-negative bacteria, with 7 aerobes, 7 facultative anaerobes, and 1 anaerobe, ensuring a balanced representation of Gram-positive and Gram-negative bacteria, as well as different oxygen utilization types. The growth medium for each species was as follows: brain-heart infusion medium for *Aggregatibacter actinomycetemcomitans*, *Gemella haemolysans*, *Moraxella catarrhalis*, *Neisseria subflava*, and *Streptococcus salivarius*; Lysogeny medium for *Achromobacter xylosoxidans*, *Acinetobacter baumannii*, *Enterobacter cloacae*, *Ochrobactrum anthropi*, *Pseudomonas putida*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus*; Man-Rogosa-Sharpe medium for *Lactobacillus salivarius*; nutrient Broth with 0.5% NaCl for *Ralstonia pickettii*; Columbia blood medium with 5% defibrinated sheep blood for *Prevotella intermedia*. Bacterial cells were harvested at the late log phase, rinsed twice with saline, resuspended using saline with 25% glycerol, and stored at -80°C . For *P. intermedia*, blood cells from the culture medium were removed by filtering through 20 μm pre-separation filters (Miltenyi, 130-101-812) before use. Cells from all 15 species were pooled at equal colony-forming units, resulting in 20 million colony-forming units in 200 μL saline for downstream analysis (Supplementary Table 3).

Host DNA depletion methods

Respiratory and mock samples were thawed at room temperature and vortexed for 30 s, and aliquots of 200 μL were used for host DNA depletion as described below.

1. Method Raw: samples without host removal.
2. Method R_ase: nuclease digestion of cell-free DNA. Fourteen units of TURBO DNase (Invitrogen, AM2239) and 50 units of Benzonase (Merck Millipore, 70664-3) were added to each sample, and the mixture was incubated at 37°C for 3 h with gentle agitation. Afterwards, the nucleases were quenched by adding 20 μL proteinase K (QIAGEN, 19133) and incubating at 56°C for 15 min with gentle agitation.
3. Method O_pma: osmotic lysis followed by PMA degradation of cell-free DNA¹⁰. The sample was centrifuged at 10,000 g for 8 min to pellet any prokaryotic and eukaryotic cells. The supernatant was transferred to a new 1.5 mL microcentrifuge tube and stored at 4°C , while the pellet was resuspended in 500 μL sterile H_2O and incubated at room temperature for 10 min with gentle agitation. After incubation, the sample was centrifuged at 10,000 g for 8 min, and the supernatant was discarded. The previously collected supernatant was added to the pellet, and the sample was pipetted at least ten times to fully resuspend

all cells, resulting in the osmotic lysis-treated sample. A final concentration of 10 μM PMAxx[™] (Biotium, 40069) was added, and the sample was briefly vortexed and incubated in the dark at room temperature for 5 min. The sample was then laid horizontally on ice, positioned less than 20 cm from a tubular fluorescent lamp (Philips TL 20 W/52 SLV/25) for 20 min with gentle agitation.

4. Method O_ase: osmotic lysis followed by nuclease digestion of cell-free DNA⁵⁴. Osmotic lysis was performed as described for O_pma, and nuclease digestion of cell-free DNA was performed as described for R_ase.
5. Method F_ase: 10 μm filtering followed by nuclease digestion of cell-free DNA. The sample was filtered through a 10 μm syringe-type filter (BD Biosciences, 340728) to remove host cells. Next, nuclease digestion of cell-free DNA was performed as described for R_ase.
6. Method S_ase, saponin lysis followed by nuclease digestion of cell-free DNA¹⁴. A final concentration of 0.025% saponin (Sigma, 47036-50G-F) was added to the sample. The sample was then briefly vortexed, and incubated at room temperature for 10 min. Next, nuclease digestion of cell-free DNA was performed as described for R_ase.
7. Method K_qia: QIAamp DNA Microbiome kit (QIAGEN, 51704). Samples were processed according to the manufacturer's instructions.
8. Method K_zym: HostZERO Microbial DNA Kit (Zymo, D4310). Samples were processed according to the manufacturer's instructions.
9. Method MEM: mechanical shear lysis followed by nuclease digestion of cell-free DNA as described previously¹³. Briefly, 200 μL of BALF aliquot was placed into 2-ml 1.4-mm ceramic bead-beating tubes (MP, 116913050-CF) and homogenized using FastPrep-24 (MP, 116004500) for 30 s at 4.5 m/s. The supernatant was transferred to a clean microfuge tube, followed by the addition of 10 μL buffer, 33 μL PBS, 2 μL Benzonase Nuclease (Merck Millipore, 70664-3), and 5 μL Proteinase K (QIAGEN, 19133). The mixture was incubated at 37°C for 15 minutes with gentle agitation. Subsequently, the samples were centrifuged at 10,000 g for 2 min, the supernatant was discarded, and the pellets were resuspended in 200 μL PBS.

DNA extraction and metagenomic sequencing

DNA from respiratory and mock samples was extracted using DNeasy PowerSoil Pro Kit (QIAGEN, 47016). Metagenomic libraries were constructed using QIAseq FX DNA Library Kit (QIAGEN, 180479). Shotgun sequencing was carried out on the Illumina Novaseq 6000 PE150 platform, generating approximately six gigabytes of data for each sample.

Quantitative PCR

The human DNA load and bacterial DNA load were measured using quantitative PCR targeting genes encoding bacterial 16S rRNA and human β -actin, respectively. The primers and probes were: bACTIN_F, TCAACACCCAGCCATGTAC; bACTIN_R, AGGGCATACCCCTCGTAGAT; bACTIN_probe, VIC-CTGTGCTATCCCTGTACGCCTCTGGC-BHQ; Bact_341f, CCTACGGGNGGCWGCAG; U1052-1071-reverse, GARCTGRCGRRCRCATGCA; 515F(Parada)_probe, FAM-TGYCAGC MGCCGCGGTAA-BHQ. The multiplex PCR reactions were conducted with 0.25 μM forward/reverse primers, 0.15 μM probes, 10 μL TaqMan Advanced Master Mix (Thermo Fisher, 4444557), and less than 10 ng template DNA in a total volume of 20 μL . The assays were performed on the Bio-Rad CFX96 Touch System under the following PCR conditions: 50°C for 2 min; 95°C for 3 min; 43 cycles of 95°C for 10 sec, 50°C for 30 s, 60°C for 40 sec. Standard curves were developed using serial dilutions of *E. coli* genomic DNA and HEP-2 cell genomic DNA. It is important to note that the quantification of bacteria in this study represents an approximation, reflecting a generalized estimation.

Sequencing data analysis

Raw reads were processed by trimming adapters and low-quality sequences and merging paired-end reads with fastp v0.20.0⁵⁵. Reads of low complexity were removed with komplexity v0.3.6 (-f -k -t 0.2)⁵⁶, and human reads were

removed with BMTagger⁵⁷. The reads were then aligned against the non-redundant nucleotide database (2021.07) from NCBI using megablast⁵⁸ with parameters `eval = 1e-10`, `max_target_seqs = 1500`, `max_hsp = 1`, `qcov_hsp_perc = 60` and `perc_identity = 60`. MEGAN v6.20.14⁵⁹ with the NCBI nucleotide to taxonomy mapping database (July 2021 version) was used for taxonomic classification. Species with relative abundances less than fivefold of that in negative controls were labeled as contaminations. HUMAnN2 v2.8.1⁶⁰ was used to obtain gene family profiles. The compositional data was normalized using total sum scaling (i.e., relative abundance), and the absolute abundances of microbial species were obtained by multiplying the bacterial load and relative abundances. Shannon index was calculated using the vegan R package v2.5.7⁶¹. Jensen-Shannon divergence of microbial composition was calculated using the philentropy R package v0.7.0⁶² and further square rooted to obtain the Jensen-Shannon distance.

Species with relative abundances higher than 0.01 in at least 10% of the BALF or OP samples were referred to as common respiratory microbial species. To estimate their genome coverage, reads were mapped to their representative genomes using BWA v0.7.12-r1039⁶³ and analyzed using Samtools v1.8⁶⁴. Before comparing microbial richness and genome coverage between host removal methods, data from eight samples of the same individual was rarefied to the lowest sequencing depth among them. Overall, the median sequencing depth of raw data was 8,644,428 reads per individual, with a range of 6,061,353 to 9,810,144 reads. When comparing OP and BALF samples at the species level, data was rarefied to 10,000 microbial reads per sample prior to the calculation of the Shannon index and JSD; for the rest analyses, data from OP and BALF samples from the same individual was rarefied to the lower microbial reads of the two. Strains were identified for species with mean relative abundances greater than 0.01 using StrainGE v1.3.8 with default parameters, and species with only one strain (clustered with genome similarity 0.99) detected were filtered out. To compare SNPs on *V. parvula* genome between BALF and OP samples, sequencing reads were aligned to the reference genome NZ_LT906445.1 using BWA v0.7.12-r1039. Sequencing depth and pileup bases were obtained from the BAM files using Samtools (mpileup -OsBa) and Varscan v2.3.9 (--min-coverage 0)⁶⁵. Only samples with more than 80% genome covered by at least five reads were included in the subsequent analyses. A mutation was identified if the frequency of the mutant allele exceeded 90% at positions covered by at least 100 reads. The phylogenetic tree of *V. parvula* genome consensus sequences was built using JolyTree V2.1⁶⁶.

Statistical analysis

PERMANOVA was performed using the `adonis2` function of vegan R package v2.5.7, and the patient identity was used as the strata argument when comparing paired samples⁶¹. Hierarchical clustering of samples was performed using the `hclust` function of stats R package v4.3.1⁶⁷ with JSD and the “ward.D2” method. Wilcoxon matched pairs tests were used to detect differences in the performance metrics between host DNA depletion methods as well as between OP and BALF samples. The Benjamini-Hochberg FDR algorithm was used to correct multiple comparisons with a significance level of 0.05 (*p*.adj value) unless otherwise specified.

Data availability

Sequencing data used in this study were deposited in the Genome Sequence Archive in National Genomics Data Center, China National Center for Bioinformatics, under accession CRA015070 that are publicly accessible at <https://bigd.big.ac.cn/gsa>.

Code availability

The scripts used to analyzed the data is shared at https://github.com/WangChun2019/Host_depletion_methods_data_analysis.

Received: 17 November 2024; Accepted: 20 June 2025;

Published online: 07 July 2025

References

- Pérez-Cobas, A. E., Rodríguez-Beltrán, J., Baquero, F. & Coque, T. M. Ecology of the respiratory tract microbiome. *Trends Microbiol.* **31**, 972–984 (2023).
- Natalini, J. G., Singh, S. & Segal, L. N. The dynamic lung microbiome in health and disease. *Nat. Rev. Microbiol.* **21**, 222–235 (2023).
- Di Simone, S. K., Rudloff, I., Nold-Petry, C. A., Forster, S. C. & Nold, M. F. Understanding respiratory microbiome-immune system interactions in health and disease. *Sci. Transl. Med.* **15**, eabq5126 (2023).
- He, Y. et al. Enhanced DNA and RNA pathogen detection via metagenomic sequencing in patients with pneumonia. *J. Transl. Med.* **20**, 195 (2022).
- Qu, J. et al. Aetiology of severe community acquired pneumonia in adults identified by combined detection methods: a multi-centre prospective study in China. *Emerg. Microbes Infect.* **11**, 556–566 (2022).
- Zinter, M. S. et al. Pulmonary metagenomic sequencing suggests missed infections in immunocompromised children. *Clin. Infect. Dis.* **68**, 1847–1855 (2019).
- Wang, J., Wang, P., Wang, X., Zheng, Y. & Xiao, Y. Use and prescription of antibiotics in primary health care settings in China. *JAMA Intern. Med.* **174**, 1914–1920 (2014).
- Antimicrobial Resistance, C Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).
- Serpa, P. H. et al. Metagenomic prediction of antimicrobial resistance in critically ill patients with lower respiratory tract infections. *Genome Med.* **14**, 74 (2022).
- Marotz, C. A. et al. Improving saliva shotgun metagenomics by chemical host DNA depletion. *Microbiome* **6**, 42 (2018).
- Pereira-Marques, J. et al. Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. *Front. Microbiol.* **10**, 1277 (2019).
- Rajar, P. et al. Microbial DNA extraction of high-host content and low biomass samples: Optimized protocol for nasopharynx metagenomic studies. *Front. Microbiol.* **13**, 1038120 (2022).
- Wu-Woods, N. J. et al. Microbial-enrichment method enables high-throughput metagenomic characterization from host-rich samples. *Nat. Methods*, <https://doi.org/10.1038/s41592-023-02025-4> (2023).
- Hasan, M. R. et al. Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. *J. Clin. Microbiol.* **54**, 919–927 (2016).
- Simner, P. J. et al. Development and optimization of metagenomic next-generation sequencing methods for cerebrospinal fluid diagnostics. *J. Clin. Microbiol.* **56**, <https://doi.org/10.1128/JCM.00472-18> (2018).
- Oechslein, C. P. et al. Limited correlation of shotgun metagenomics following host depletion and routine diagnostics for viruses and bacteria in low concentrated surrogate and clinical samples. *Front. Cell Infect. Microbiol.* **8**, 375 (2018).
- Bellehumeur, C. et al. Propidium monoazide (PMA) and ethidium bromide monoazide (EMA) improve DNA array and high-throughput sequencing of porcine reproductive and respiratory syndrome virus identification. *J. Virol. Methods* **222**, 182–191 (2015).
- Heravi, F. S., Zakrzewski, M., Vickery, K. & Hu, H. Host DNA depletion efficiency of microbiome DNA enrichment methods in infected tissue samples. *J. Microbiol. Methods* **170**, 105856 (2020).
- Marchukov, D., Li, J., Juillerat, P., Misselwitz, B. & Yilmaz, B. Benchmarking microbial DNA enrichment protocols from human intestinal biopsies. *Front. Genet.* **14**, 1184473 (2023).
- Bruggeling, C. E. et al. Optimized bacterial DNA isolation method for microbiome analysis of human tissues. *Microbiologyopen* **10**, e1191 (2021).
- Wiscovitch-Russo, R., Singh, H., Oldfield, L. M., Fedulov, A. V. & Gonzalez-Juarbe, N. An optimized approach for processing of frozen

- lung and lavage samples for microbiome studies. *PLoS One* **17**, e0265891 (2022).
22. Ganda, E. et al. DNA extraction and host depletion methods significantly impact and potentially bias bacterial detection in a biological fluid. *mSystems* **6**, e0061921 (2021).
23. Bloomfield, S. J. et al. Determination and quantification of microbial communities and antimicrobial resistance on food through host DNA-depleted metagenomics. *Food Microbiol.* **110**, 104162 (2023).
24. Pathirana, E., McPherson, A., Whittington, R. & Hick, P. The role of tissue type, sampling and nucleic acid purification methodology on the inferred composition of Pacific oyster (*Crassostrea gigas*) microbiome. *J. Appl. Microbiol.* **127**, 429–444 (2019).
25. Li, Q. et al. DNA enrichment methods for microbial symbionts in marine bivalves. *Microorganisms* **10**, <https://doi.org/10.3390/microorganisms10020393> (2022).
26. Delbeke, H., Casteels, I. & Joossens, M. DNA extraction protocol impacts ocular surface microbiome profile. *Front. Microbiol.* **14**, 1128917 (2023).
27. Amar, Y. et al. Pre-digest of unprotected DNA by Benzonase improves the representation of living skin bacteria and efficiently depletes host DNA. *Microbiome* **9**, 123 (2021).
28. Nelson, M. T. et al. Human and extracellular DNA depletion for metagenomic analysis of complex clinical infection samples yields optimized viable microbiome profiles. *Cell Rep.* **26**, 2227–2240. e2225 (2019).
29. Charalampous, T. et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat. Biotechnol.* **37**, 783–792 (2019).
30. Fong, W., Rockett, R., Timms, V. & Sintchenko, V. Optimization of sample preparation for culture-independent sequencing of *Bordetella pertussis*. *Microb. Genom.* **6**, <https://doi.org/10.1099/mgen.0.000332> (2020).
31. Kok, N. A. et al. Host DNA depletion can increase the sensitivity of *Mycobacterium* spp. detection through shotgun metagenomics in sputum. *Front. Microbiol.* **13**, 949328 (2022).
32. Shi, X. et al. Microfluidics-based enrichment and whole-genome amplification enable strain-level resolution for airway metagenomics. *mSystems* **4**, <https://doi.org/10.1128/mSystems.00198-19> (2019).
33. Gan, M. et al. Combined nanopore adaptive sequencing and enzyme-based host depletion efficiently enriched microbial sequences and identified missing respiratory pathogens. *BMC Genom.* **22**, 732 (2021).
34. Yap, M. et al. Evaluation of methods for the reduction of contaminating host reads when performing shotgun metagenomic sequencing of the milk microbiome. *Sci. Rep.* **10**, 21665 (2020).
35. Boutin, S. et al. Comparison of microbiomes from different niches of upper and lower airways in children and adolescents with cystic fibrosis. *PLoS One* **10**, e0116029 (2015).
36. Charlson, E. S. et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am. J. Respir. Crit. Care Med.* **184**, 957–963 (2011).
37. Duvallet, C. et al. Aerodigestive sampling reveals altered microbial exchange between lung, oropharyngeal, and gastric microbiomes in children with impaired swallow function. *PLoS One* **14**, e0216453 (2019).
38. Liu, H. Y. et al. Oropharyngeal and sputum microbiomes are similar following exacerbation of chronic obstructive pulmonary disease. *Front. Microbiol.* **8**, 1163 (2017).
39. Prevaes, S. M. et al. Concordance between upper and lower airway microbiota in infants with cystic fibrosis. *Eur. Respir. J.* **49**, <https://doi.org/10.1183/13993003.02235-2016> (2017).
40. Baranova, E. et al. Comparison of sputum and oropharyngeal microbiome compositions in patients with non-small. *Cell Lung Cancer* **6**, 1–23 (2022).
41. van Dijk, L. R. et al. StrainGE: a toolkit to track and characterize low-abundance strains in complex microbial communities. *Genome Biol.* **23**, 74 (2022).
42. Huang, J. et al. Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions. *Infect. Drug Resistance* **13**, 567–576 (2020).
43. Chen, S., Kang, Y., Li, D. & Li, Z. Diagnostic performance of metagenomic next-generation sequencing for the detection of pathogens in bronchoalveolar lavage fluid in patients with pulmonary infections: systematic review and meta-analysis. *Int. J. Infect. Dis.* **122**, 867–873 (2022).
44. Zhang, D., Yang, X., Wang, J., Xu, J. & Wang, M. J. J. O. I. M. R. Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in patients with lower respiratory tract infections. **50**, 03000605221089795 (2022).
45. Xiao, Y.-H. et al. Clinical efficacy and diagnostic value of metagenomic next-generation sequencing for pathogen detection in patients with suspected infectious diseases: a retrospective study from a large tertiary hospital. 1815–1828 (2023).
46. Li, N. et al. High-throughput metagenomics for identification of pathogens in the clinical settings. **5**, 2000792 (2021).
47. Wu, T. et al. Streamline-based purification of bacterial samples from liquefied sputum utilizing microfluidics. *Lab Chip* **17**, 3601–3608 (2017).
48. Ahannach, S. et al. Microbial enrichment and storage for metagenomics of vaginal, skin, and saliva samples. *iScience* **24**, 103306 (2021).
49. Li, Z., Fallon, J., Mandeli, J., Wetmur, J. & Woo, S. L. A genetically enhanced anaerobic bacterium for oncopathic therapy of pancreatic cancer. *J. Natl. Cancer Inst.* **100**, 1389–1400 (2008).
50. Mammen, M. J., Scannapieco, F. A. & Sethi, S. Oral-lung microbiome interactions in lung diseases. *Periodontology 2000* **83**, 234–241 (2020).
51. Bassis, C. M. et al. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *mBio* **6**, e00037 (2015).
52. Claassen-Weitz, S., Lim, K. Y. L., Mullally, C., Zar, H. J. & Nicol, M. P. The association between bacteria colonizing the upper respiratory tract and lower respiratory tract infection in young children: a systematic review and meta-analysis. *Clin. Microbiol. Infect.* **27**, 1262–1270 (2021).
53. Man, W. H., de Steenhuijsen Piters, W. A. & Bogaert, D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat. Rev. Microbiol.* **15**, 259–270 (2017).
54. Nelson, M. T. et al. Human and extracellular DNA depletion for metagenomic analysis of complex clinical infection samples yields optimized viable microbiome profiles. *Cell Rep.* **26**, 2227–2240. e2225 (2019).
55. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
56. Clarke, E. L. et al. Sunbeam: an extensible pipeline for analyzing metagenomic sequencing experiments. *Microbiome* **7**, 46 (2019).
57. Rotmistrovsky, K. & Agarwala, R. *BMTagger: Best Match Tagger for Removing Human Reads from Metagenomics Datasets* (NCBI/NLM, National Institutes of Health, 2011).
58. Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**, 203–214 (2000).
59. Huson, D. H. et al. MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* **12**, e1004957 (2016).
60. Franzosa, E. A. et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* **15**, 962–968 (2018).
61. Oksanen, J. et al. The vegan package. *Community Ecol. package* **10**, 719 (2007).

62. Drost, H.-G. Philentropy: information theory and distance quantification with R. *J. Open Source Softw.* **3**, 765 (2018).
63. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
64. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, <https://doi.org/10.1093/gigascience/giab008> (2021).
65. Koboldt, D. C. et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* **25**, 2283–2285 (2009).
66. Criscuolo, A. J. R. I. & Outcomes. A fast alignment-free bioinformatics procedure to infer accurate distance-based phylogenetic trees from genome assemblies. **5**, e36178 (2019).
67. Team, R. C. *R: A Language and Environment for Statistical Computing* Vol. 1 (R Foundation for Statistical Computing, 2013).

Acknowledgements

This work was supported by the National Key R&D Program of China (grant no. 2022YFA1304300 to M.L.), the Beijing Nova Program (grant no. Z211100002121135 to L.Z.), the National Natural Science Foundation of China (grant no. 32100098 to L.Z.), the Youth Innovation Promotion Association of Chinese Academy of Sciences (grant no. 2021097 to L.Z.), the Shenzhen Scientific and Technological Foundation (grant no. KCXFZ20211020163545004, RCJC20221008092726022 to G.Z.), the Sanming Project of Medicine in Shenzhen (grant no. SZZYSM202311009 to G.Z.), and the Shenzhen High-level Hospital Construction Fund (to G.Z.).

Author contributions

M.L. and L.Z. took the lead in conceptualizing the study. G.Z., Z.L., J.Z., D.C., Z.X. and J.H. were responsible for cohort management and sample collection. C.W., L.Z., W.L., Z.W., R.X., X.J. and W.M. carried out the wet-lab experiments. C.W., C.K., L.Z., M.L., J.Y., J.Z. and X.S. handled data analysis and interpretation. L.Z., C.W. and M.L. wrote the main manuscript text. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41522-025-00762-2>.

Correspondence and requests for materials should be addressed to Li Zhang, Guoliang Zhang or Mingkun Li.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025