

RESEARCH ARTICLE

Meta-transcriptomics Reveals Dysbiosis of the Respiratory Microbiome in Older Adults with Long COVID

Meng'en Liao^{1†}, Jianpeng Cai^{1†}, Feng Zhu^{1,2†}, Yuanbo Lan^{1,3†}, Tianqi Xu¹, Jingxin Guo¹, Quanlin Xue¹, Yilong Wen³, Fan Zou³, Yu Zhang⁴, Shiliang Zhang⁵, Yan Yan⁶, Jingwen Ai^{1*}, Jie Cui^{1,7*}, and Wenhong Zhang¹

¹Shanghai Sci-Tech Inno Center for Infection & Immunity, National Medical Center for Infectious Diseases, Huashan Hospital, Institute of Infection and Health, Fudan University, Shanghai, China. ²Department of Respiratory and Critical Care Medicine, Affiliated Wuxi Fifth Hospital of Jiangnan University, The Fifth People's Hospital of Wuxi, Wuxi, China. ³Department of Respiratory and Critical Care Medicine, Affiliated Hospital of Zunyi Medical University, Zunyi, China. ⁴Department of Respiratory, Wuxi Ninth Affiliated Hospital of Soochow University, Wuxi, China. ⁵Center of Clinical Laboratory, Affiliated Wuxi Fifth Hospital of Jiangnan University, The Fifth People's Hospital of Wuxi, Wuxi, China. ⁶Shanghai Institute of Immunity and Infection, Chinese Academy of Sciences, Shanghai, China. ⁷Laboratory for Marine Biology and Biotechnology, Qingdao Marine Science and Technology Center, Qingdao, China.

*Address correspondence to: jingwenai1990@126.com (J.A.); jiecui@fudan.edu.cn (J.C.)

†These authors contributed equally to this work.

Limited research has investigated the connection between long COVID (LC) and the respiratory microbiome, particularly in older adults. This study aimed to characterize the respiratory microbiome of older LC patients (with an average age of 65 years old), through meta-transcriptomic sequencing of 201 individual samples. Marked differences in microbial diversity were observed between LC and non-LC patients, including disruptions in both pathogenic bacteria and fungi. Importantly, viral taxa, such as *Herpes simplex virus type 1* and *Human coronavirus 229E*, were more frequently detected in LC patients, indicating the vulnerability of LC patients to viral infections. Functional annotation at the expression level revealed notable differences in microbial metabolism with alterations observed in pathways related to tryptophan–serotonin metabolism in LC patients. These findings underscore the altered microbial landscape, especially in older adults who developed LC, and fill the gap for the potentially clinical roles played by the respiratory microbiome.

Introduction

The ongoing impact of coronavirus disease 2019 (COVID-19) has brought great attention to the long-lasting and often incapacitating sequelae that follow in the wake of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, a condition now formally recognized as “long COVID” (LC) [1,2]. Typically, among older adults, the impact of this condition not only increases mortality and hospitalization rates [3–5] but also results in an elevated susceptibility to more severe and persistent LC symptoms compared to those in younger individuals, further complicating recovery and leading to a more challenging postinfection course [6]. LC is a heterogeneous and multisystemic condition with a diverse array of symptoms—including persistent fatigue, cognitive impairment, breathing difficulties, gastrointestinal disturbances, autonomic dysfunction, and musculoskeletal issues—and mental health challenges

[5]. It is widely acknowledged that older adults with LC symptoms face distinct and prolonged health threats [7,8], and therefore, understanding the pathophysiology related to LC is crucial.

The human respiratory tract, spanning from the nostrils to the lung alveoli, hosts niche-specific multikingdom microbial communities that likely act as gatekeepers against pathogens and support respiratory physiology and immune homeostasis [9,10]. As immune function declines with age [11,12], the respiratory microbiota undergoes significant changes [13,14], exemplified by the nasopharyngeal microbiome losing its niche specificity and becoming increasingly dominated by *Streptococcus*, *Prevotella*, and *Veillonella*, which may elevate the risk of infection [15]. Age-related microbial dysbiosis is linked to chronic inflammation, increased vulnerability to infections, and other respiratory diseases [16–21]. These issues are particularly relevant for LC patients. However, the dysregulation of key microbial taxa

Citation: Liao M, Cai J, Zhu F, Lan Y, Xu T, Guo J, Xue Q, Wen Y, Zou F, Zhang Y, et al. Meta-transcriptomics Reveals Dysbiosis of the Respiratory Microbiome in Older Adults with Long COVID. *Research* 2025;8:Article 0720. <https://doi.org/10.34133/research.0720>

Submitted 29 December 2024

Revised 30 April 2025

Accepted 7 May 2025

Published 2 June 2025

Copyright © 2025 Meng'en Liao et al. Exclusive licensee Science and Technology Review Publishing House. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License (CC BY 4.0).

within the respiratory microbiome of elderly patients, particularly how these imbalances may exacerbate conditions associated with LC, remains a critical yet insufficiently explored area of investigation.

The human microbiome is intricately linked to LC. Emerging evidence based on animal models suggests that dysbiosis may have a causal role in LC [22]. LC patients exhibit gut microbiome dysbiosis characterized by a reduction in beneficial short-chain fatty acid-producing bacteria, such as *Faecalibacterium prausnitzii* [23–25] and *Bifidobacterium* [26,27], along with an increase in opportunistic pathogens like *Ruminococcus gnavus* [23,28,29] and *Flavonifractor* [27,30]. These microbial alterations correlate with systemic inflammation and immune dysregulation, potentially perpetuating chronic symptoms [31–34]. Specific LC symptoms have also been linked to gut microbiome changes; for instance, respiratory dysfunction is associated with an enrichment of *Veillonella* and *Flavonifractor* [30], while fatigue and neuropsychiatric symptoms correlate with *Clostridium innocuum* and *Actinomyces naeslundii* [23]. Given these connections, microbiome-based interventions, including probiotics [35–38], dietary modifications [39], and fecal microbiota transplantation [40], are being explored as potential therapies for mitigating LC symptoms.

Few studies have investigated the relationship between respiratory microbiota and LC. A study demonstrated that LC patients exhibited a markedly increased presence of pro-inflammatory and lipopolysaccharide-producing microbiota, including *Prevotella* and *Veillonella* [41]. Critical gaps remain in understanding the association between LC and the respiratory microbiome alteration, especially in older patients. Additionally, nonbacterial microorganisms, including viruses and fungi, which may also contribute to persistent LC symptoms, have been largely overlooked.

This study aims to investigate individuals' respiratory-tract-resident microbiome, that is the collection of viruses, bacteria, and fungi, identified by unbiased total RNA sequencing, of middle-aged and older adults with LC symptoms and compare them to those of former patients who did not develop LC. We hypothesize that the respiratory microbiome alteration in LC patients is associated with the persistence of symptoms and could serve as a potential microbial biomarker for the condition. To fill such a gap, we performed unbiased meta-transcriptomic sequencing on throat swab and sputum samples, conducted cross-kingdom microbiome analyses, and explored functional changes in microbial communities and associations between microbiome composition and relevant clinical features.

Results

Subject characteristics and samples

To investigate the changes in the respiratory tract microbiome of middle-aged and older adult patients with LC, we collected oropharyngeal swab (SW) and sputum (SP) samples from individuals with a history of at least one SARS-CoV-2 infection. The study center was not a major determinant of respiratory microbiome alterations in LC patients, as shown by principal coordinates analysis (Fig. S1). Demographic and clinical information was collected simultaneously by using a structured questionnaire (Table S1). A total of 122 participants completed the questionnaire, and 201 paired samples from 101 participants—comprising 101 sputum specimens and 100 throat swab specimens—were subsequently processed for sequencing and analysis. Of

these participants, 43.6% ($n = 44$) reported experiencing persistent symptoms following COVID-19 infection (Fig. 1A). Nonspecific symptoms, such as sleep disorders and memory loss, along with cardiopulmonary symptoms including cough and difficulty breathing, are among the most frequently reported symptoms in patients with LC (Fig. 1B). Coexistence of multiple symptom types was commonly observed. The average age for patients with LC symptoms was 65.2 years, compared to 60.8 years for those without LC (Fig. 1C). Forty-one (40.6%) participants experienced multiple infections, and 88 (87.1%) participants received at least one COVID-19 vaccine (Fig. 1D). Overall, the cohort design and sampling strategies offer a valuable opportunity to investigate age-related changes in the respiratory tract microbiome specific to post-COVID-19 sequelae.

Distinct taxonomic and diversity shifts in the respiratory microbiome of LC patients

To optimize the recovery of cross-kingdom microbiome composition, we analyzed the microbiomes of the upper respiratory tract (URT) and lower respiratory tract (LRT) in LC and non-long COVID (NLC) patients using unbiased meta-transcriptomic sequencing (Fig. 2A). After removing low-quality reads, human-genome-derived reads, and ribosomal RNA (rRNA) sequences, the clean reads were processed using the Kraken2 annotation workflow. To minimize false-positive hits, the following criteria were prioritized: (a) the confidence cutoff of Kraken2 was set as 0.1; (b) the reads per million (RPM) threshold was 10 for bacteria and 1 for fungi and viruses; (c) viruses with a known host of human, bacterial, or fungal origin were retained; and (d) bacterial and fungal microorganisms present in multiple samples were retained. Only microorganisms that met these requirements were included in the subsequent analysis. These strategies provided stringent characterization of respiratory tract microbiome for patients with a COVID-19 infection history. It was observed that URT and LRT exhibited similar bacterial compositions at the phylum level, with Bacillota, Pseudomonadota, and Bacteroidetes as the dominant phyla, along with fungal phyla Ascomycota and Basidiomycota (Fig. 2B). Notably, the viral communities differed significantly between these 2 niches. In SP samples, the top 4 viral phyla were *Pisuviricota*, *Uroviricota*, *Negarnaviricota*, and *Peploviricota*, whereas in SW samples, they were *Uroviricota*, *Artverviricota*, *Negarnaviricota*, and *Pisuviricota*.

To investigate the respiratory tract microbiome between LC and NLC groups in 2 niches (i.e., SP and SW samples), we profiled the family-level microbial composition of each individual (Fig. S2A). The analysis of microbial taxa distribution across different sample types and patient groups revealed similar patterns with distinct microbial signatures that may be associated with disease states. Specifically, most bacterial taxa can be detected in 4 groups (Fig. S2B). The family-level analysis revealed significant differences between LC and NLC groups, with Burkholderiaceae (bacteria), Pichiaceae (fungi), and *Redondoviridae* (viruses) enriched in LC across both sputum and swab samples (Fig. S3). In sputum, LC samples showed increased Microbacteriaceae and Aspergillaceae, while commensal families like Staphylococcaceae were depleted [42–44]. Swab samples exhibited similar dysbiosis, with Treponemataceae enriched and health-associated Gemellaceae depleted [45–48].

At the genus level, the respiratory microbiome composition showed distinct alterations across the 4 groups, with *Streptococcus*

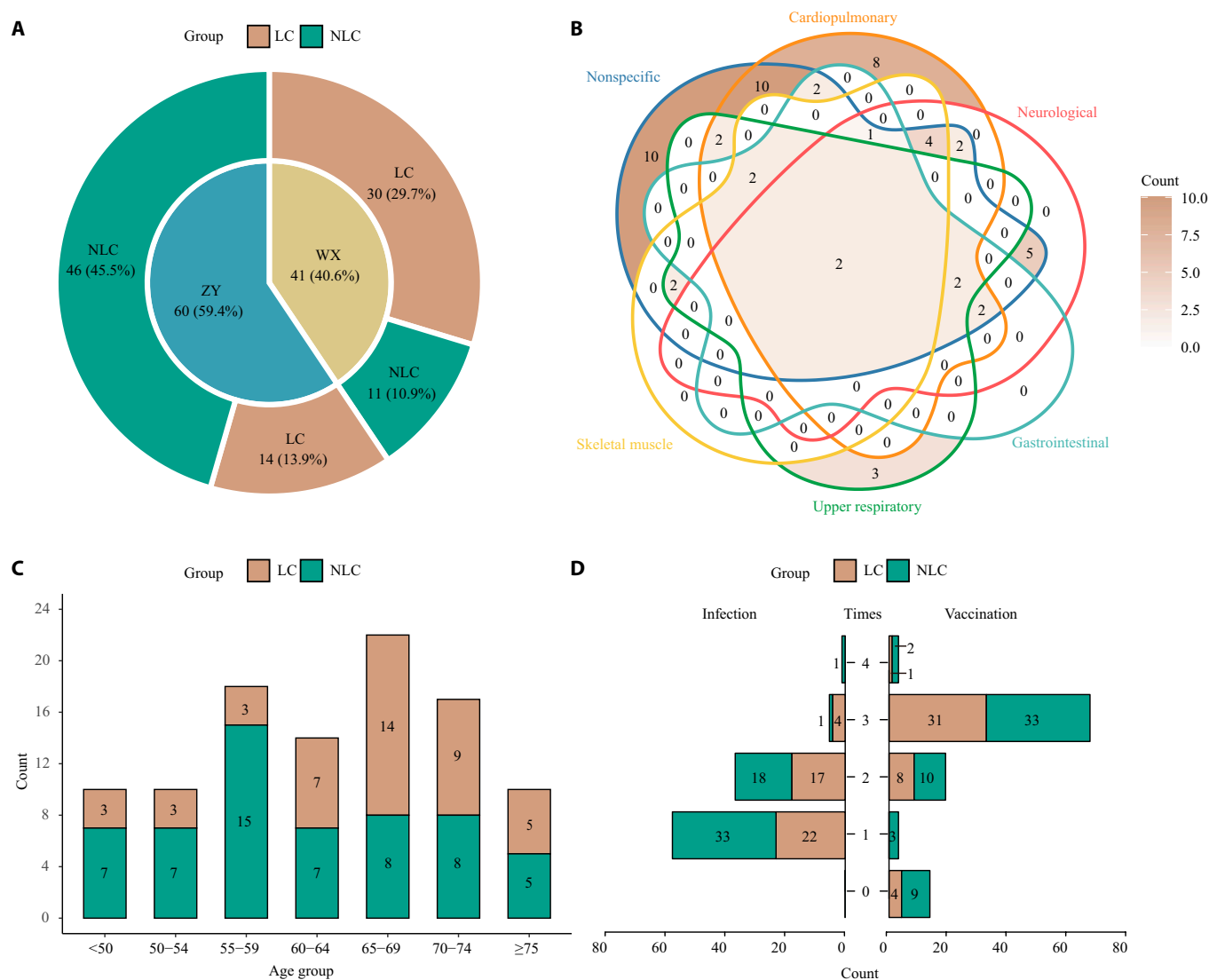


Fig. 1. Overview of the study cohort and clinical characteristics. (A) Donut plot illustrating the proportion of long COVID (LC) and non-long COVID (NLC) patients across different centers (WX for Wuxi Fifth People’s Hospital and ZY for Affiliated Hospital of Zunyi Medical University). The inner circle represents patient recruitment sources, while the outer segments indicate the proportion of LC (brown) and NLC (green) patients within each center. (B) Venn diagram showing the distribution of clinical symptoms among LC patients, classified into 6 categories: cardiopulmonary, neurological, gastrointestinal, upper respiratory, skeletal/muscle, and nonspecific. Overlapping regions indicate the number of LC patients presenting with multiple categories of symptoms. (C) Age distribution of LC and NLC patients. Bars represent the count of patients in each age group, stratified by LC and NLC groups. The numbers of patients are shown. (D) Comparison of history severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection frequency (left) and COVID-19 vaccination status (right) between LC and NLC groups. Bars indicate the number of individuals in each group.

consistently dominating in all groups (Fig. 2C). Notable differences were observed in the relative abundances of *Neisseria*, *Prevotella*, and *Rothia*, between LC and NLC groups. Specifically, LC groups demonstrated a reduction in *Streptococcus* and *Prevotella*, alongside an increase in *Neisseria*, suggesting potential microbial dysbiosis associated with LC. *Rothia mucilaginosa* emerged as the most dominant species in all groups and showed higher relative abundance in LC patients than in NLC groups (Fig. 2D).

To identify the respiratory microbial signatures associated with LC, we conducted comparative analyses of community-level diversity. A significant decrease in alpha diversity, measured by the Shannon and Gini-Simpson indices, was observed in sputum samples from LC patients compared to those from NLC patients ($P < 0.05$, Wilcoxon test, Fig. 2E). In terms of

sampling sites, LC patients showed comparable alpha diversity between the 2 sites ($P > 0.05$, Wilcoxon test, Fig. 2E). In contrast, NLC patients showed higher alpha diversity in sputum samples than swabs ($P < 0.001$, Wilcoxon test, Fig. 2E). Consistently, significant differences in microbial composition were observed between groups stratified by symptom status and sample types, as shown by the result of the permutational multivariate analysis of variance (ANOVA) test (Adonis $R^2 = 0.03$, $P = 0.004$, Fig. 2F). However, these factors accounted for only a small portion of the total variation. The first principal coordinate (PCoA 1) explained 20.5% of the total variance observed, while the second principal coordinate (PCoA 2) accounted for 15.1% of the variance. Taken together, these results revealed altered microbial diversity associated with LC in older adults, especially in the LRT.

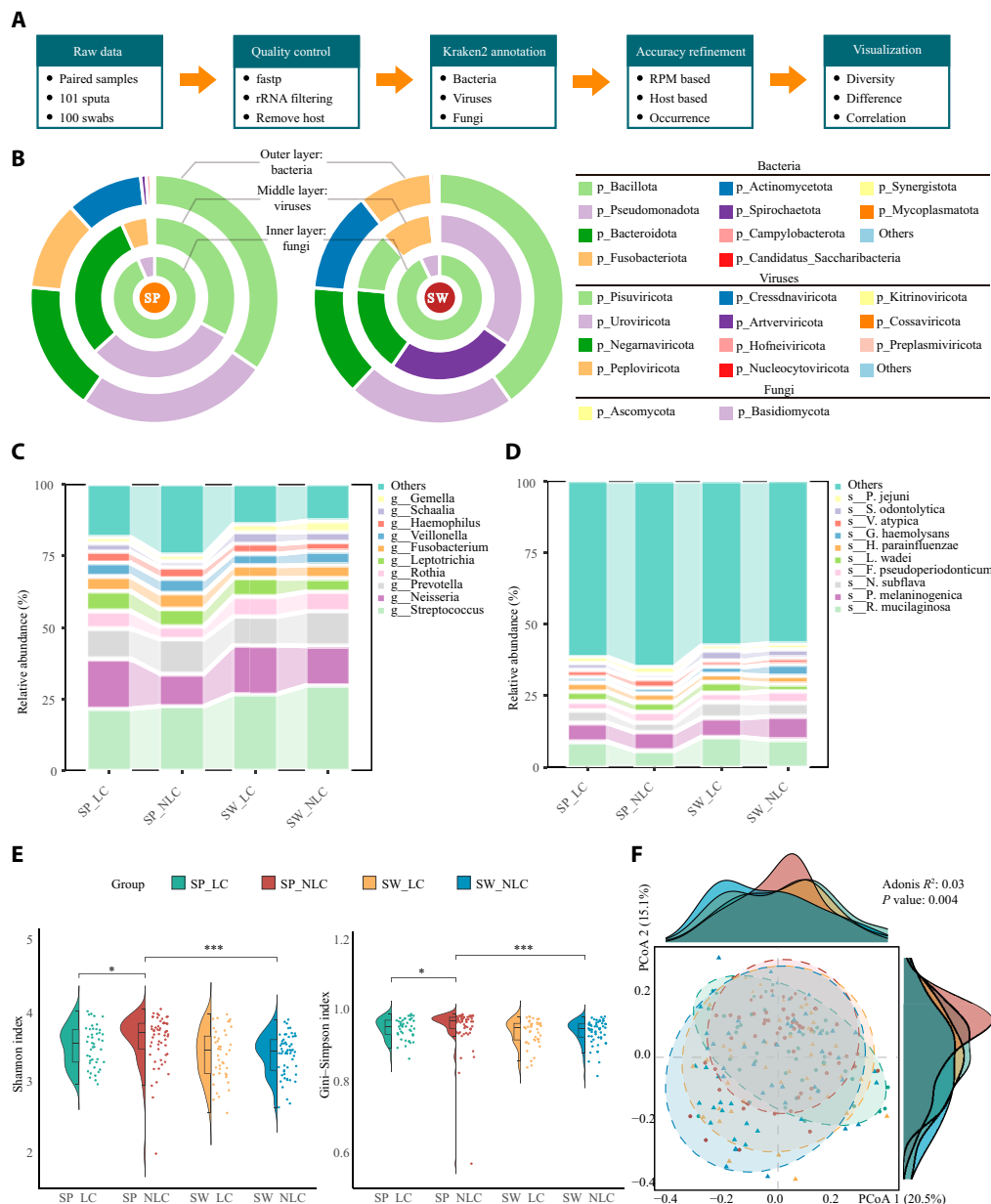


Fig. 2. Microbiome profiling of upper and lower respiratory tract samples. (A) Schematic diagram of the analytical workflow for meta-transcriptomic data processing. (B) Phylum-level composition of the respiratory microbiome in sputum (SP) and swab (SW) samples. Donut plots depict the relative abundance of bacterial (outer ring), viral (middle ring), and fungal (inner ring) phyla. For clarity, the same color suite was used for different layers. Colors within each ring correspond to distinct phyla listed on the right. (C) Genus- and (D) species-level community composition presented by the relative abundance of bacterial taxa across 4 groups. The top 10 dominant taxa at each taxonomic level are highlighted. *s__R. mucilaginosa*, *s__Rothia_mucilaginosa*; *s__P. melaninogenica*, *s__Prevotella_melaninogenica*; *s__N. subflava*, *s__Neisseria_subflava*; *s__F. pseudoperiodonticum*, *s__Fusobacterium_pseudoperiodonticum*; *s__L. wadei*, *s__Leptotrichia_wadei*; *s__H. parainfluenzae*, *s__Haemophilus_parainfluenzae*; *s__G. haemolysans*, *s__Gemella_haemolysans*; *s__V. atypica*, *s__Veillonella_atypica*; *s__S. odontolytica*, *s__Schaalia_odontolytica*; *s__P. jejuni*, *s__Prevotella_jejuni*. (E) Violin plots showing the alpha diversity indices (Shannon index and Gini-Simpson index) for sputum (SP) and swab (SW) samples in 4 groups: SP_LC, SP_NLC, SW_LC, and SW_NLC. Statistical comparisons were performed using the Wilcoxon rank-sum test. (F) Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarity, illustrating beta diversity differences among the SP_LC, SP_NLC, SW_LC, and SW_NLC groups. Each point represents a sample, colored and shaped according to group. The variance explained by PCoA 1 and PCoA 2 is shown. Statistical significance was assessed using permutational multivariate analysis of variance (PERMANOVA). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. rRNA, ribosomal RNA; RPM, reads per million.

Patients with LC showed dysregulated pathogenic microbes

To explore potential pathogenicity from infecting organisms, we compared the abundance and occurrence of pathogenic and nonpathogenic microbes identified in this cohort. This was achieved by incorporating the Global Catalogue of Pathogens, a comprehensive resource that provides strong evidence on

human pathogens [49]. In sputum samples, 2 genera (*Bacteroides* and *Stenotrophomonas*) showed a decreased detection rate, although no pathogenic genera showed significant differences in terms of relative abundance (Fig. 3A and B). By contrast, a genus containing pathogenic fungi, *Candida*, was more frequently detected in the sputum of patients with LC symptoms. Specifically, 2 species from the family Prevotellaceae, *Prevotella bivia* and

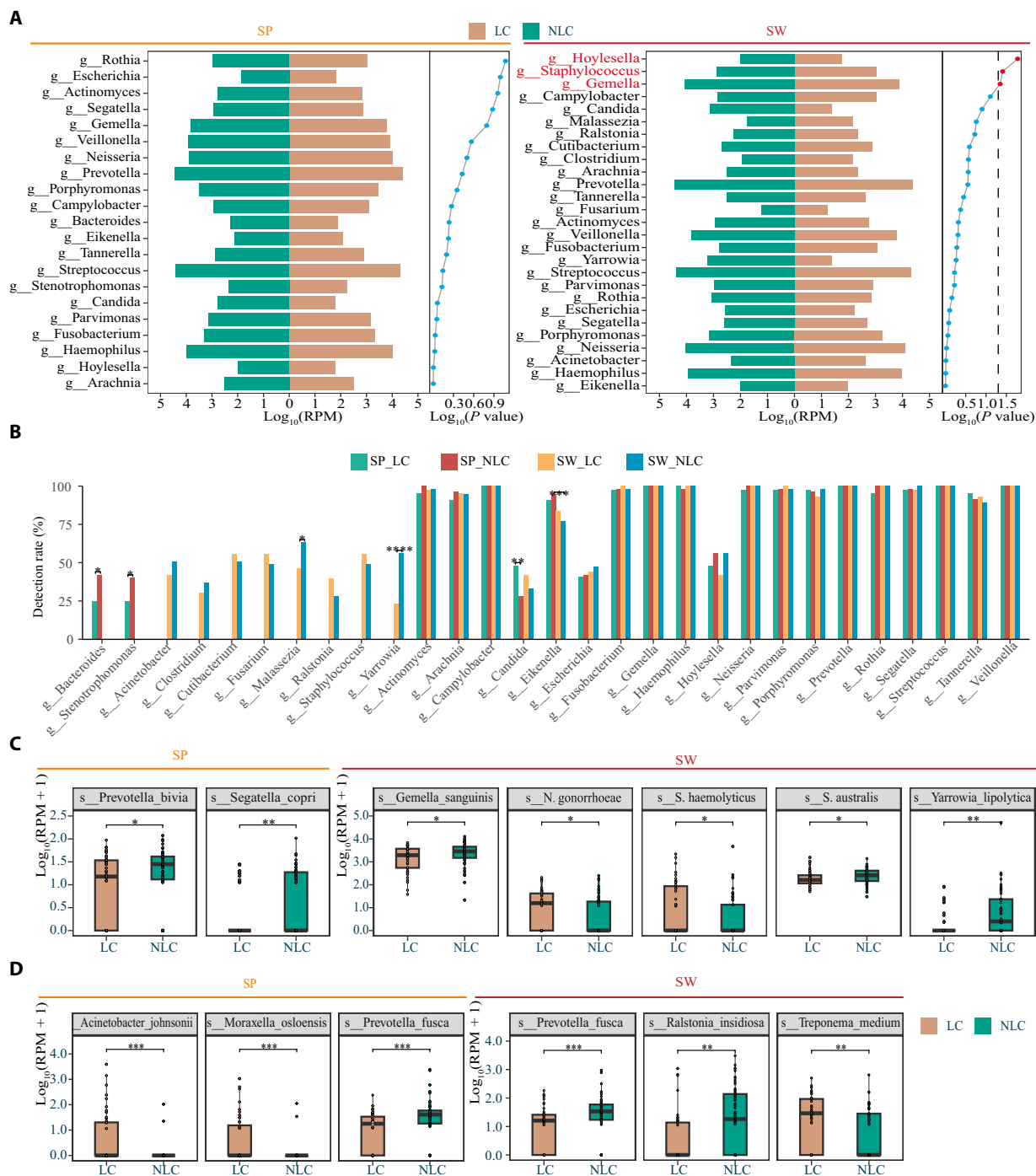


Fig. 3. Differential analysis of pathogenic and nonpathogenic taxa between LC and NLC groups. (A) Bidirectional bar plots showing the genus-level relative abundance of pathogenic microbes annotated using the Global Catalogue of Pathogens. The left panel displays results for sputum (SP) samples, and the right panel for swab (SW) samples. Brown bars indicate LC groups, and green bars indicate NLC groups. Wilcoxon rank-sum test results are shown on the right side of each plot. (B) Detection rate (%) of genera containing pathogenic microorganisms between groups. Significant differences were assessed using the chi-square test, with P values indicated for each genus. (C) Boxplots comparing the log-transformed relative abundances calculated as $\log_{10}(\text{RPM} + 1)$ of significantly different pathogenic species between LC and NLC groups in both sample types. (D) Boxplots showing nonpathogenic species with significant differences in relative abundance between LC and NLC in SP and SW samples. The top 3 species are shown in each sample type. Statistical tests were performed using the Wilcoxon rank-sum test. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. See also Fig. S3.

Segatella copri, had lower average abundance in LC sputum than in NLC (Fig. 3C). In swab samples, 3 detected bacterial genera, *Hoylella*, *Staphylococcus*, and *Gemella*, showed significantly changed abundance between LC patients and NLC patients (Fig. 3A). Another 2 genera (i.e., *Malassezia* and *Yarrowia*) showed decreased detection rates in patients with LC compared

to those in NLC patients (Fig. 3B). Consistently, most detected pathogens showed reduced relative abundance in LC patients, including *Gemella sanguinis*, *Streptococcus australis*, and *Yarrowia lipolytica* (Fig. 3C). Conversely, only *Neisseria gonorrhoeae* showed increased average abundance in these patients (Fig. 3C). We also identified significantly different nonpathogenic microbial

species between LC and NLC patients (Fig. S4). The most differential microbes were *Acinetobacter johnsonii*, *Moraxella osloensis*, and *Prevotella fusca* in the LRT and *P. fusca*, *Ralstonia insidiosus*, and *Treponema medium* in the URT (Fig. 3D). Collectively, these results demonstrated a reduction in most differential pathogenic microorganisms in older patients with LC, as reflected in both relative abundance and detection rates. However, the findings necessitate validation in larger cohorts to ensure consistency and further investigation into its implications for LC and the potential underlying mechanisms.

Relationship between viral detection and LC in older adults

Given the fact that SARS-CoV-2 interaction with human viral communities may serve as a potential biological contributor to post-acute sequelae of COVID-19 [50], we analyzed the species-level virome characteristics related to LC. Six viruses from the genus *Litunavirus* were detected in patients with LC (Fig. 4A and B), suggesting a possible connection between this genus of bacteriophage and LC symptoms. Among them, *Litunavirus LIT1* and *PA26* showed similar increasing trends in both sputum and swab samples, while the other 4 viruses, *Litunavirus Mag4*, *Pap02*, *Yh6*, and *Ab09* increased only in throat swabs. Similarly, several additional viruses were found to be uniquely present in LC patients, including *Simplexvirus humanalpha1* and *Staphylococcus phage StB12* and *StB27* in the LRT. *Simplexvirus humanalpha1* was known as *Herpes simplex virus type 1* (HSV-1) and has been associated with critically ill COVID-19 patients [51]. *Vientovirus* of genus *Torbevirus* showed an increased detection rate in the LRT and URT of LC patients. *Human coronavirus 229E* and *Rothia phage Spartoi* were detected in the sputum of both groups of patients with and without LC with a significant increase in LC patients. *Pseudomonas phage VB PaeP VLI* showed a similar trend in swabs. Several other viruses exhibited distinct patterns in their occurrence. *Cepunavirus* was not detected in LC patients; however, 2 of its members, *Cepunavirus Cp1* and *CP7*, were identified in 8.77% of sputum samples and 10.53% of swabs from NLC patients, respectively. *Streptococcus phage EJ-1* was much less frequently detected in LC patients. Collectively, most differentially detected viral species had increased detection rates in older patients undergoing LC, indicating that an exploratory study to investigate the underlying biological relationship is required.

A noteworthy human respiratory-associated virus was *Vientovirus* sp. (Fig. S5A and B), a circular Rep-encoding single-stranded DNA virus known to be elevated in the oro-respiratory samples of patients with periodontitis, acute illness, and severe COVID-19 [52]. Sequence analysis revealed that the *Vientovirus* genome contains a capsid protein, a replication-associated protein, and an open reading frame 3 protein (Fig. S5C), whose function remains unclear. The *Vientovirus* Rep protein (Fig. S5C) contains 2 domains found in many small DNA viruses: one involved in rolling-circle replication (InterPro: PS52020) and a second helicase domain within the P-loop NTPase superfamily (Pfam: PF00910) [53]. Based on the abundance difference of *Vientovirus* in patients with LC, the Rep protein of *Vientovirus* may serve as a potential biomarker for LC detection. *Human coronavirus 229E*, a human coronavirus typically associated with mild upper respiratory infections, may experience alterations in its prevalence and virulence potentially due to the influence of SARS-CoV-2 [54]. Considering the unique occurrence of *Human coronavirus 229E* in patients suffering from LC, we

assembled contigs from meta-transcriptomic reads and identified a nearly complete genome of *Human coronavirus 229E* that showed 99.905% nucleotide similarity with *Human coronavirus 229E* strain LC22 (GenBank: ON791801.1) (Fig. S5D).

To link identified bacteriophages with possible host interactions, we conducted correlation network analyses in sputum and swab samples. The analysis in SP samples revealed strong correlations between specific bacteriophages and both bacterial and fungal genera (Fig. 4C, left panel). Notably, phages like *Litunavirus PA26* and *Staphylococcus phage STB27* showed extensive connections with multiple bacterial hosts, suggesting their potential role in modulating microbial dynamics. In the SW samples, the correlation network displayed a more intricate pattern, highlighting that phages such as *Pseudomonas phage VB PaeP VLI* and members in *Litunavirus* were prominently linked to genera like *Pichia* and *Fusarium* (Fig. 4C, right panel).

These findings suggested relatively differential viral members and phage–host interaction dynamics between the URT and the LRT, which may influence the persistence and symptomatology of LC.

Functional potential differences in the expression of genetic materials of LC and NLC respiratory microbiomes

One advantage of meta-transcriptomics is the capacity to analyze the activated genetic elements expressed by microbial communities. We analyzed the functional potential of the respiratory microbiome in LC and NLC patients using MetaCyc pathway data inferred through HUMAnN 3.0 [55]. Differential abundance analysis identified 26 MetaCyc pathways significantly associated with LC or NLC in sputum samples and 4 in swab samples (Fig. 5A and C). Pathways enriched in the NLC group included those involved in core metabolic functions, such as fatty acid salvage, adenosine nucleotide degradation, phospholipid metabolism, and amino acid metabolism (Fig. 5B and D). In LC patients, the most enriched pathway was ectoine biosynthesis, along with pathways related to aromatic biogenic amine degradation, serotonin degradation, and L-tryptophan degradation (Fig. 5B). These findings highlight distinct functional differences in microbial metabolism between LC and NLC groups.

Association between the respiratory tract microbiome and clinical features

To explore the potential link between the respiratory tract microbiome and specific clinical features, we performed correlation analyses to identify significant microbial patterns. Firstly, the Mantel test was used to evaluate the correlation between microbial diversity at the family, genus, and species levels, and several factors, including age, gender, smoking status, vaccination, treatment, and the presence of LC symptoms (Fig. S6). The Mantel test revealed significant correlations between LC symptoms and clinical factors such as age, gender, and treatment status. These factors, especially at the species level, displayed stronger associations with microbial composition, consistent with our previous observations.

To elucidate the correlations between specific microbial species and clinical parameters in middle-aged and elderly patients with LC, we conducted correlation analysis in sputum and swab samples from both LC and NLC groups (Fig. 6A). In SP samples, *Porphyromonas* sp. oral taxon 275 and *Curtobacterium*

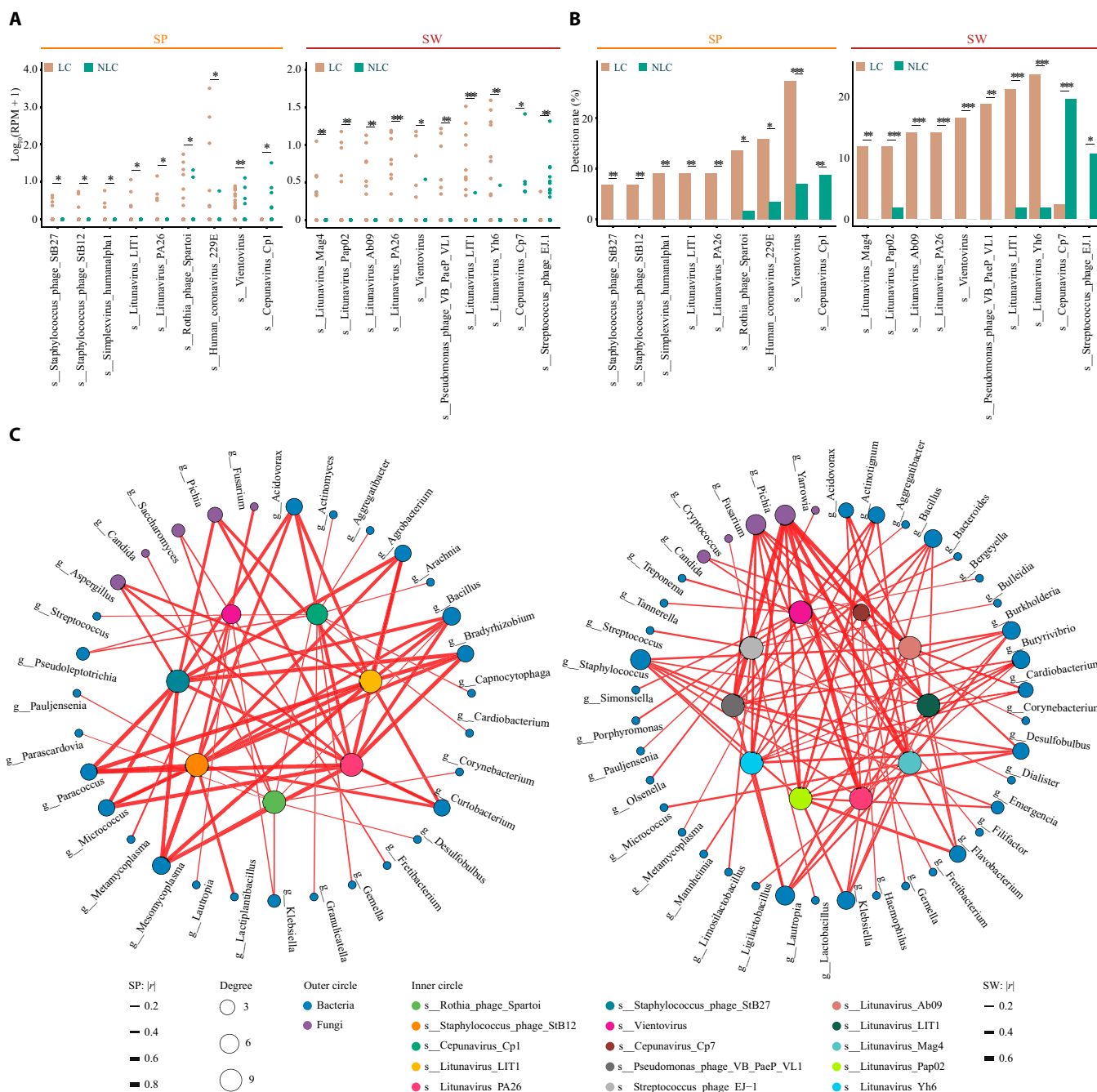


Fig. 4. Comparative analysis of virus differences and host-microbe interactions in LC and NLC groups. (A) Boxplots showing the relative abundance calculated as $\log_{10}(\text{RPM} + 1)$ of viral species with significant differences between LC and NLC groups in sputum (SP) and swab (SW) samples. Statistical significance was determined using the Wilcoxon rank-sum test. (B) Detection rates (%) of key viral species in SP and SW samples between LC and NLC groups. Chi-square test results are shown for each species. (C) Network analysis of virus-host interactions in SP (left) and SW (right) samples. Inner circles represent viral species with significant differences between LC and NLC groups, color coded by viral species. Outer circles display the top 30 bacterial (blue) and fungal (purple) genera that show significant positive correlations with these viruses. Circle size corresponds to interaction degree, and edge thickness represents the strength of correlation (Spearman correlation coefficient, $|r| > 0.2$). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. See also Fig. S4.

flaccumfaciens exhibited positive correlations with infection frequency (Fig. 6A and B). Similarly, in SW samples, *Capnocytophaga leadbetteri* and *Metamycoplasma orale* showed positive correlations with infection times (Fig. 6A and C). Conversely, *Streptococcus suis* and *Litunavirus Yh6*, in the LRT and URT, respectively, negatively correlated with the times of previous SARS-CoV-2 infections (Fig. 6A to C). The correlation analysis between microbial species in swab samples from

LC patients and the number of vaccinations revealed predominantly negative correlations, as seen in species like *Desulfomicrobium orale* and *Neisseria dumasiana* and the bacteriophage *Streptococcus phage phiARI0131-2* (Fig. 6A and D). Interestingly, *Paraburkholderia fungorum* exhibited a moderate positive correlation ($R = 0.41$, $P = 0.0068$) with vaccination frequency, contrasting with the general trend and indicating a particular response to vaccination among the analyzed microbes

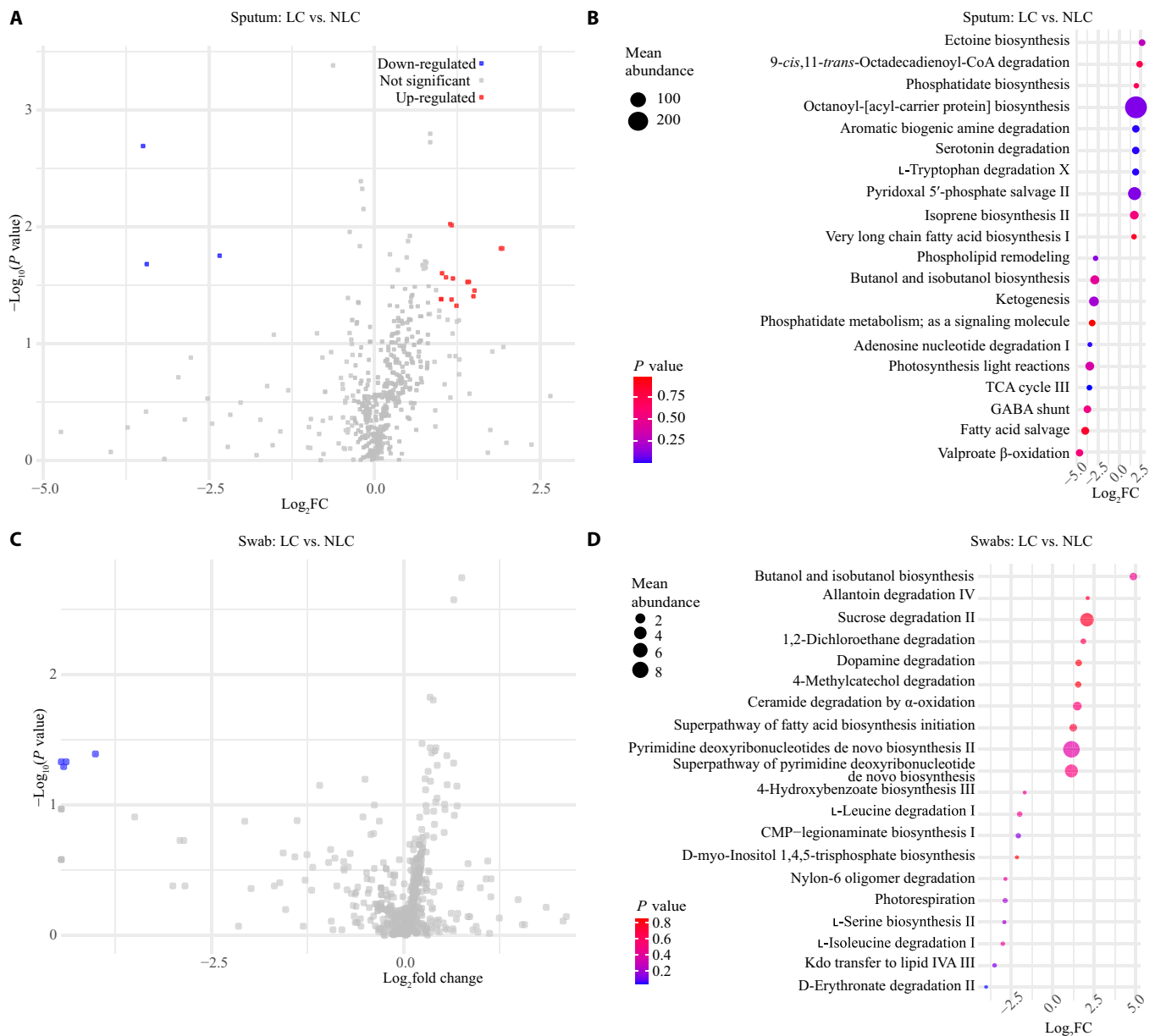


Fig. 5. Differential abundance of respiratory microbial MetaCyc pathways between LC and NLC patients. (A) Volcano plot showing the differential MetaCyc pathway abundance between LC and NLC patients in sputum samples. (B) The top 10 LC associated and NLC associated MetaCyc pathways in sputum samples based on the Wilcoxon signed-rank test. The y axis lists the pathway names, and the x axis represents the log-transformed pathway abundance ratio (LC versus NLC). Bubble size corresponds to the mean abundance of the pathway in the respective group, while color indicates the *P* value. (C) Volcano plot showing the differential MetaCyc pathway abundance between LC and NLC patients in swab samples. (D) The top 10 LC associated and NLC associated MetaCyc pathways in swab samples based on the Wilcoxon signed-rank test. The y axis lists the pathway names, and the x axis represents the log-transformed pathway abundance ratio (LC versus NLC). Bubble size corresponds to the mean abundance of the pathway in the respective group, while color indicates the *P* value. FC, fold change; TCA, tricarboxylic acid; GABA, γ -aminobutyric acid; CMP, cytidine monophosphate.

(Fig. 6A and D). Overall, these results revealed a complex interplay between microbial dynamics and clinical parameters, where samples from different respiratory sites exhibited heterogeneous microbial responses to infection frequency and vaccination.

Discussion

LC remains a critical public health issue following the COVID-19 pandemic, particularly on older adults. While previous studies have examined gut microbiome changes in LC patients, the respiratory microbiome, especially in older adults, remains

underexplored. This study addressed this gap by characterizing the respiratory microbiome of middle-aged and elderly LC patients using meta-transcriptomic sequencing of 201 individual samples. Our findings revealed microbial dysbiosis in LC patients, marked by reduced microbial diversity and disruptions in both pathogenic bacteria and fungi. Additionally, functional annotation at the expression level highlighted distinct alterations in microbial metabolism between LC and NLC groups, with LC patients showing enrichment in pathways related to ectoine biosynthesis, aromatic biogenic amine degradation, and L-tryptophan degradation. Importantly, LC patients exhibited increased susceptibility to viral infections, with higher abundances

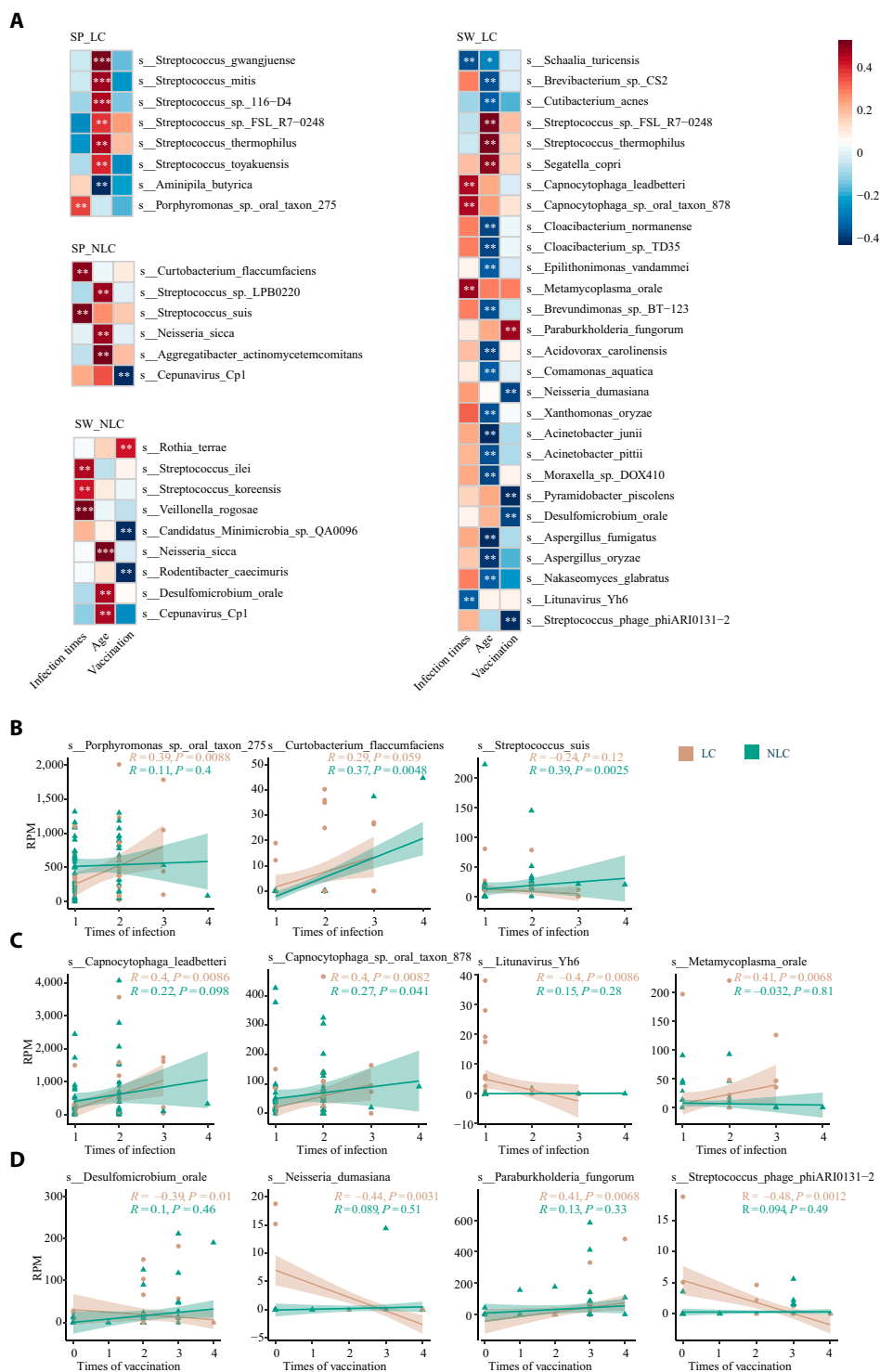


Fig. 6. Correlation analysis between microbial species and clinical characteristics in LC and NLC groups. (A) Heatmap showing correlations between microbial species and 3 clinical characteristics, including infection times, age, and vaccination times, in SP and SW samples for LC and NLC groups. The top-ranked species were shown in each group. (B) Scatter plots showing the relationships between infection times and microbial species with significant correlations in SP samples. (C) Scatter plots showing the relationships between infection times and microbial species with significant correlations in SW samples. (D) Scatter plots showing the relationships between vaccination times and microbial species with significant correlations in SW samples. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. See also Fig. S5.

of viral taxa such as *Herpes simplex virus type 1*, *Human coronavirus 229E*, and *Vientovirus*. These metabolic shifts, alongside microbial and viral imbalances, suggest a complex interplay between bacteria, viruses, and fungi that could contribute to the prolonged symptoms observed in LC.

Previous studies have demonstrated reduced alpha and changed beta diversity of the gut microbiota in LC patients compared to those in NLC individuals, with an increase in opportunistic pathogens and a decrease in beneficial bacteria [56,57]. These findings align with our results, which also show

reduced alpha diversity in the LRT. Some studies have suggested that opportunistic bacteria, such as members of the chronic-respiratory-condition-associated Burkholderiaceae family [58], are more enriched in COVID-19 patients [59]. At the genus level, significant increases in *Staphylococcus* [60,61] and *Gemella* [62] have been reported in both COVID-19 and LC patients, whereas *Bacteroides* and *Stenotrophomonas* showed a marked decrease [61,63]. Additionally, *Candida* has frequently been detected in COVID-19 intensive care unit patients and is often associated with severe infections [64]. At the species level, *P. bivia* has been negatively correlated with monocyte levels in COVID-19 patients [65], while the abundance of *Staphylococcus haemolyticus* has appeared to decrease [66]. Our study also observed similar variations in bacterial abundance or detection rates, suggesting that increased opportunistic pathogen levels may be associated with sustained immune dysregulation and chronic inflammation in LC patients [62,67,68]. Functionally, the enrichment of the tryptophan–serotonin metabolic pathway (e.g., aromatic biogenic amine degradation, serotonin degradation, and L-tryptophan degradation X) in the LC group suggests a functional implication in chronic inflammation and sleep disturbances, as serotonin plays a crucial role in modulating immune responses, reducing pro-inflammatory cytokines [69], and serving as a precursor for melatonin, which is essential for circadian rhythm regulation and sleep maintenance [70]. The highest enrichment of the ectoine biosynthesis pathway suggests a key role in anti-inflammatory protection and cellular stress resistance by respiratory microbiota.

The role of bacteriophages in human diseases remains underexplored. However, several studies suggest that bacteriophages may influence disease resistance or susceptibility through mechanisms independent of their bacterial hosts [71,72]. While adaptive phage therapy has been shown to reduce antibiotic use in the rehabilitation of post-COVID-19 patients in intensive care units [73], direct experimental studies investigating the contribution of bacteriophages to LC pathogenesis are needed. In our study, we found significant differences in bacteriophage abundance and detection rates between LC and NLC patients. These differences may be associated with bacterial hosts and the reactivation of latent infection following immunosuppression [74]. However, the specific role of bacteriophages in the progression of LC requires further investigation. A similar mechanism may explain the observed differences in the abundance and detection rates of respiratory viruses such as *HSV*, *Vientovirus*, and *Human coronavirus 229E*. *HSV-1* reactivation may contribute to LC through immune modulation or indirect effects on the respiratory microbiome [75,76]. While *Vientovirus* is generally considered a commensal virus and does not appear to influence SARS-CoV-2 infection [77], its potential role in respiratory health and immune interaction in LC remains unclear and requires further investigation. Current studies on the relationship between redondovirus and SARS-CoV-2 remain inconclusive. Some research has reported higher abundance and detection rates of redondovirus in SARS-CoV-2-positive samples [78,79], while others found no significant association [80]. In our study, the differences in redondovirus abundance and detection between LC and NLC patients were notable but insufficient to establish a causal role in the progression of LC. Further research is needed to clarify the complex interactions between these viruses and the disease.

Our study identifies microbial and viral dysbiosis in older adults with LC, yet the cross-sectional design limits causal

inference. While these imbalances may reflect associative consequences of post-COVID immune dysregulation, plausible mechanisms suggest bidirectional interactions. For example, *HSV-1* [81] and *Human coronavirus 229E* [82,83] enrichment could drive chronic inflammation, while dysregulated commensals like *Gemella* may impair mucosal immunity [84,85], perpetuating dysbiosis. Longitudinal studies tracking microbial dynamics from acute infection to recovery, combined with mechanistic models, are needed to establish causality.

Several limitations of this study should be acknowledged. One limitation of this study is the absence of a healthy control group; however, prior studies have shown that the respiratory microbiome diversity in COVID-19-recovered patients without LC (NLC) tends to normalize toward healthy baselines [86–88], supporting our focus on the LC vs. NLC comparison to specifically identify microbial signatures associated with LC. Saliva contamination in throat swabs and sputum samples is difficult to eliminate entirely, but in this study, efforts were made to minimize this issue during sampling in this. While metatranscriptomics offers a snapshot of cross-kingdom microbial activity, it lacks the genomic context of microorganisms with DNA as genetic materials needed to capture their full metabolic potential and complex interkingdom interactions. Future studies could benefit from integrating metagenomics to provide a more comprehensive view. Reliance on self-reported symptoms introduces potential recall bias and limits the precision of symptom characterization. Incorporating objective clinical measures, such as inflammatory markers (e.g., TIMP-1) that can be triggered by *HSV-1* [89] and neurocognitive tests, could improve the accuracy of symptom–microbiome associations and help contextualize immune responses in LC. This study is cross-sectional in nature, which limits our ability to assess causal relationships and the temporal progression of microbial dysbiosis in LC. Future longitudinal studies are essential to better understand how microbial changes evolve over time and their role in the persistence of symptoms.

Methods

Study design and ethics approval

Participants in this study were patients with a history of at least one COVID-19 infection, recruited from 2 hospitals in China: 42 participants from Wuxi Fifth People's Hospital with samples collected between 2023 November 19 and 2024 January 6 and 80 participants from the Affiliated Hospital of Zunyi Medical University with samples collected between 2024 January 20 and 2024 February 3. This cohort comprised individuals with real-time polymerase chain reaction-confirmed primary SARS-CoV-2 infection, followed up at a median of 12 months postdiagnosis. While subsequent reinfections were captured through self-reports, no virological data were available for these events. Symptoms and clinical characteristics were collected through a standardized questionnaire during visits. The follow-up was conducted in person with face-to-face visits between physicians and participants. The standardized questionnaire was validated through preliminary testing, where a small group of non-research individuals provided feedback, leading to adjustments in wording and format for clarity and comprehension. The questionnaire, described in our previous study [90], included details such as sex, age, smoking history, chronic disease history, number of COVID-19 infections, vaccination status, and prior treatments during infection.

Based on self-reported symptoms collected via questionnaires, 53 participants reported LC symptoms, such as persistent fatigue, respiratory difficulties, and cognitive impairments, and were assigned to the LC group. Individuals who were diagnosed with COVID-19 at the initial stage (SARS-CoV-2 positive) but did not report any LC symptoms during the follow-up period were referred to the NLC group ($n = 65$). The diagnosis of whether symptoms were related to LC was made by trained physicians. Ethical approval for this study was granted by the Ethics Committee of Wuxi Fifth People's Hospital (approval no. 2022-019-1). All procedures performed were in accordance with the ethical standards of the committee. Written informed consent was obtained from all participants prior to enrollment in the study.

Sample collection

Throat swabs were collected by trained personnel, and the swabs were then placed into the collection tube immediately. The method for induced sputum collection is as follows: subjects inhaled 4% sterile nebulized saline through a mouthpiece for 15 min. During the induction procedure, participants' condition was assessed by the trained personnel. Subjects were instructed to rinse their mouths with clean water before sample collection and to spit saliva into a saliva container before coughing sputum into a sterilized sputum container. The inhalation period was interrupted every 2 min to allow subjects to cough up sputum. All samples are immediately placed in dry ice for storage and transported to the laboratory.

RNA extraction and quality control

Total RNA extraction was performed using the RNA PowerSoil Total RNA Isolation Kit (MoBio, Cat. No. 12866-25), following the manufacturer's protocols. To assess RNA quality, a 1.5% agarose gel electrophoresis was used, ensuring intact RNA bands without signs of degradation. RNA concentration and purity were measured using a NanoDrop NC2000 ultraviolet-visible spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Samples with an RNA concentration greater than 50 ng/ μ l and a total RNA quantity exceeding 1 μ g were considered adequate for library preparation. RNA purity was confirmed by an A_{260}/A_{280} ratio >1.8 , although slight deviations were acceptable based on sample-specific characteristics. Additionally, RNA integrity number was assessed using an Agilent 2100 bioanalyzer, and samples with RNA integrity number ≥ 5.5 were used for further analyses.

Library construction and sequencing

rRNA depletion was performed using rRNA-specific probes that hybridize to total RNA, followed by magnetic bead separation to remove rRNA/probe complexes. The remaining messenger RNA was further purified using ethanol precipitation. Complementary DNA (cDNA) libraries were constructed using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. Fragmentation of RNA was carried out using divalent cations under elevated temperature conditions (94 °C for 8 min). This was followed by first-strand cDNA synthesis using the Superscript III enzyme in the presence of Act D Mix, which was incubated at 25 °C for 10 min, 42 °C for 50 min, and then 85 °C for 15 min. Second-strand cDNA synthesis was performed using Second Strand Marking Master Mix, incubated at 16 °C for 1 h. Purified cDNA was recovered using AMPure

XP beads. After second-strand synthesis, A-tailing was performed to add a single adenosine (A) base to the 3' ends of cDNA fragments, preventing self-ligation. This was followed by the ligation of indexed adapters. Adapter-ligated cDNA fragments were amplified by polymerase chain reaction, followed by final purification and size selection using AMPure XP beads. Library quality was assessed using Agilent High Sensitivity DNA Kit on Agilent 2100 Bioanalyzer, where a single peak indicated proper library construction. Library quantification was performed using Quant-iT PicoGreen dsDNA Assay Kit on a Promega QuantiFluor instrument. Libraries with a concentration above 2 nM were deemed suitable for sequencing. Qualified libraries were sequenced on the Illumina NovaSeq 6000 platform at Personalbio (Shanghai, China), using NovaSeq 6000 S4 Reagent Kit (300 cycles) to generate paired-end reads of 2×150 bp.

Sequencing data preprocessing

For the processing of raw sequencing reads, fastp (version 0.23.4) [91] was employed to filter out low-quality reads, including those that were too short or contained excessive ambiguous bases (N), as well as to trim low-quality bases, remove adapter sequences, and drop duplicated reads/pairs. Default parameters were used except for an accuracy level of 6 to calculate duplication. To remove rRNA sequences in silico, we used SortMeRNA (version 4.3.6) [92] with the `smr_v4.3_default_db` database built from the Silva 138 SSURef NR99 (16S, 18S), Silva 132 LSURef (23S, 28S), and RFAM v14.1 (5S, 5.8S) source databases with default parameters. The sequences were clustered at 90% identity for bacterial 16S rRNA, with 5S and 5.8S rRNA as seed sequences, and the remaining sequences clustered at 95%, achieving a benchmark accuracy of 99.899%. Bowtie 2 (version 2.5.2) [93] was then used to align and remove reads derived from the host (human) genome, utilizing the GRCh38.p13 reference genome obtained from GenCode. The alignment was performed in the very-sensitive-local mode to maximize accuracy.

Microbiome detection and quantification

The clean reads from each sample were annotated using Kraken2 (version 2.1.3) [94], a taxonomic classification system that employs a k -mer-based method to assign taxonomic labels quickly and accurately by matching k -mers to the lowest common ancestor of genomes. The PlusPF reference database, which includes combined annotations of archaeal, bacterial, viral, protozoal, and fungal sequences from the RefSeq databases, was used (downloaded on 2024 August 12). The confidence parameter was set as 0.1 for higher annotation accuracy. To enable intersample comparisons, microbial abundance was quantified as RPM. This was calculated by dividing the number of reads assigned to each taxon by the total number of reads in the library and multiplying the result by 10^6 . To minimize false positives, we applied stringent filtering criteria: bacterial taxa with an RPM below 10, and viral and fungal taxa with an RPM below 1, were excluded from further analysis. This filtering process was automated using a custom stand-alone script.

Virus host curation

Full NCBI Virus metadata were obtained (2024 August 17) from the National Center for Biotechnology Information (NCBI) FTP site (<https://ftp.ncbi.nlm.nih.gov/genomes/Viruses/>)

AllNuclMetadata/AllNuclMetadata.csv.gz). The taxonomic lineage of each virus was manually curated using taxonkit (version 0.17.0) [95]. For each viral family, we curated the host range and the number of virus entries. Only viruses with hosts including “Human”, “Bacteria”, or “Fungi” were included for further analysis.

Community diversity analysis

Alpha diversity was assessed to quantify microbial richness and diversity across samples. The analysis was conducted using the vegan R package (version 2.6-6.1) based on the filtered RPM table as mentioned above. For each sample, microbial richness was calculated as the number of taxa with nonzero abundance. Shannon and Gini–Simpson indices were used as the diversity measures, computed using the diversity function with the “shannon” method and the “simpson” method, respectively. For beta-diversity analysis, Bray–Curtis dissimilarity matrices were calculated at the genus, family, and species levels using the vegan package (version 2.6-6.1) in R. Principal coordinates analysis was then performed on the Bray–Curtis dissimilarity matrices, retaining the first 2 principal coordinates for visualization. Eigenvalues were computed to determine the proportion of variance explained by each coordinate.

Pathogenic and nonpathogenic microorganism analysis

Metadata for human pathogens, including 510 bacterial taxa, 407 fungal taxa, and 226 viral taxa, were obtained from the gcPathogen database (downloaded on 2024 August 17, <https://nmdc.cn/gcpathogen/>) [49]. The taxonomic lineage of each pathogen was curated manually using taxonkit (version 0.17.0) [95]. Microbes identified from this study were classified as pathogenic or nonpathogenic at the genus and species levels according to the organized information. A genus was considered pathogenic if it contained at least one pathogenic species. The relative abundance, and detection rate, calculated as the number of patients positive for the taxon divided by the total number of individuals in each group, of pathogenic genera and species were then compared across groups.

Correlation network analysis

Correlation networks between viruses and microbial taxa were constructed using igraph (2.0.3) and ggraph (version 2.2.1) in R. Spearman correlations between viral, bacterial, and fungal genera in sputum and swab samples were calculated with false-discovery-rate-adjusted P values ($P \leq 0.05$). The top 10 bacterial or fungal interactions per virus were retained for network visualization.

Viral genome assembly, abundance estimation, and sequence annotation

Viral contigs were assembled based on clean reads by using Megahit (version 1.2.9) [96] with default parameters. The assembled contigs were annotated via BLASTp against the nonredundant protein (NR) database (downloaded on 2023 October 23) by using DIAMOND (version 2.1.9.163) [97]. Only those with viral sequences as the top hit were retained. Contigs were further evaluated using CheckV (version 1.0.3) [98], and only complete, high-quality, and medium-quality contigs containing at least one viral gene were kept as potential viral fragments. To estimate viral abundance, reads were mapped to the assembled

contigs using Bowtie 2 (version 2.5.4) [93], and the normalized relative abundance of each contig was calculated as RPM, as previously described.

Phylogenetic analysis and contig annotation of specific viruses

Prodigal with default parameters was used to predict proteins from viral contigs. Protein functions were annotated using BLASTp (version 2.15.0+) [99] against the NR database and CD-search against the Conserved Domain Database (version 3.21) [100]. To select reference sequences for phylogenetic analysis, all replication proteins from *Redondoviridae* and RNA-dependent RNA polymerases from *Human coronavirus 229E* in the NCBI database were chosen. For *Vientovirus* and *Human coronavirus 229E*, the respective replication proteins and RNA-dependent RNA polymerases were selected. These proteins were aligned with reference sequences using MAFFT (version 7.525) [101] with default settings. Ambiguously aligned regions were removed using TrimAl (version 1.5) [102]. Phylogenetic trees were constructed using FastTree (version 2.1.11) [103] with the LG model and visualized in iTOL [104]. The viral genomic structure was visualized using Geneious Prime (version 2024.0.5), and conserved domains of the replication proteins were predicted with InterPro using default parameters.

Functional annotation and differential pathway analysis

Functional profiling of microbial communities was conducted using HUMAnN 3.0 to annotate the functional expression of the microbiome based on MetaCyc pathway definitions on sputum and swab samples, respectively. Pathway abundances were normalized to copies per million units using the humann_renorm_table tool. Differential abundance analysis of MetaCyc pathways between LC and NLC groups was performed using the Wilcoxon signed-rank test.

Correlation between microbial diversity and clinical features

Spearman correlation analysis was conducted to assess associations between microbial species in sputum and swab samples and the clinical features of patients in LC and NLC groups. Correlations were calculated using the psych package (version 2.4.6.26), and significance was determined with false-discovery-rate-adjusted P values ($P \leq 0.05$). Significant correlations were visualized using pheatmap (version 1.0.12). Heatmaps were generated for each group, and significant species correlations were highlighted by using scatter plots.

Statistical analysis

Given the non-normal distribution of continuous variables in this study, as determined by normality testing, the Wilcoxon signed-rank test was used for group comparisons of numerical variables. The chi-square test was employed to assess differences in microbial detection rates across groups. Differences in microbial community composition between groups were evaluated using permutational multivariate ANOVA via the adonis2 function of the vegan package. Spearman rank correlation analysis was performed to assess correlations between microbial abundance and clinical characteristics. All statistical tests were 2-tailed, and P values < 0.05 were considered statistically significant.

Ethical Approval

Ethical approval for this study was granted by the Ethics Committee of Wuxi Fifth People's Hospital (approval no. 2022-019-1). All procedures performed were in accordance with the ethical standards of the committee. Written informed consent was obtained from all participants prior to enrollment in the study.

Acknowledgments

Funding: This work was supported by the Shanghai Municipal Science and Technology Major Project (ZD2021CY001); National Natural Science Foundation of China grants (nos. 32470151 and 32470054); the Talent Plan of Shanghai Branch, Chinese Academy of Sciences (CASSHB-QNPD-2023-019); and the National Key Research and Development Program of China (2023YFC3043501 and 2024YFC3044600).

Author contributions: M.L. and T.X. conducted the experiments. M.L., T.X., F. Zhu, Y.L., J.G., Q.X., Y.W., F. Zou, Y.Z., S.Z., Y.Y., and J.A. collected and analyzed the data. M.L. and J. Cui wrote the main text of the manuscript. J. Cui, J.A., and W.Z. supervised the study. All authors reviewed the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data Availability

The raw sequence data generated in this paper have been deposited in the Genome Sequence Archive in the National Genomics Data Center, China National Center for Bioinformatics/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA020324), which are publicly accessible at <https://ngdc.cnbc.ac.cn/gsa> [105,106]. This study did not generate new unique reagents.

Supplementary Materials

Figs. S1 to S6
Table S1

References

- Su S, Zhao Y, Zeng N, Liu X, Zheng Y, Sun J, Zhong Y, Wu S, Ni S, Gong Y, et al. Epidemiology, clinical presentation, pathophysiology, and management of long COVID: An update. *Mol Psychiatry*. 2023;28(10):4056–4069.
- Harris E. Review calls for standardized long COVID definition. *JAMA*. 2024;332(2):97.
- Perlis RH, Santillana M, Ognyanova K, Safarpour A, Lunz Trujillo K, Simonson MD, Green J, Quintana A, Druckman J, Baum MA, et al. Prevalence and correlates of long COVID symptoms among US adults. *JAMA Netw Open*. 2022;5(10):Article e2238804.
- Katz GM, Bach K, Bobos P, Cheung A, Décary S, Goulding S, Herridge MS, McNaughton CD, Palmer KS, Razak FA, et al. Understanding how post-COVID-19 condition affects adults and health care systems. *JAMA Health Forum*. 2023;4(7):Article e231933.
- Greenhalgh T, Sivan M, Perlowski A, Nikolich J. Long COVID: A clinical update. *Lancet*. 2024;404(10453):707–724.
- Mansell V, Hall Dykgraaf S, Kidd M, Goodyear-Smith F. Long COVID and older people. *Lancet Healthy Longev*. 2022;3(12):E849–E854.
- Li J, Zhou Y, Ma J, Zhang Q, Shao J, Liang S, Yu Y, Li W, Wang C. The long-term health outcomes, pathophysiological mechanisms and multidisciplinary management of long COVID. *Signal Transduct Target Ther*. 2023;8(1):416.
- Russell SJ, Parker K, Lehoczi A, Lieberman D, Partha IS, Scott SJ, Phillips LR, Fain MJ, Nikolich JZ. Post-acute sequelae of SARS-CoV-2 infection (long COVID) in older adults. *Geroscience*. 2024;46(6):6563–6581.
- Perdijk O, Azzoni R, Marsland BJ. The microbiome: An integral player in immune homeostasis and inflammation in the respiratory tract. *Physiol Rev*. 2023;104(2):835–879.
- Özçam M, Lynch SV. The gut–airway microbiome axis in health and respiratory diseases. *Nat Rev Microbiol*. 2024;22(8):492–506.
- Liu Z, Liang Q, Ren Y, Guo C, Ge X, Wang L, Cheng Q, Luo P, Zhang Y, Han X. Immunosenescence: Molecular mechanisms and diseases. *Signal Transduct Target Ther*. 2023;8(1):200.
- Santoro A, Bientinesi E, Monti D. Immunosenescence and inflammaging in the aging process: Age-related diseases or longevity? *Ageing Res Rev*. 2021;71:Article 101422.
- Miller SJ, Zhang F, Taylor S, Woodman R, Shoubridge AP, Papanicolaou LE, Rogers GB. Oropharyngeal *Staphylococcus aureus* is linked to higher mortality in long-term aged care residents. *Age Ageing*. 2025;54(3):afaf042.
- Liu J, He X-Y, Yang K-L, Zhao Y, Dai E-Y, Chen W-J, Raj AK, Li D, Zhuang M, Yin XH, et al. Oropharyngeal microbiome profiling and its association with age and heart failure in the elderly population from the northernmost province of China. *Microbiol Spectr*. 2024;12(10):e0021624.
- Whelan FJ, Verschoor CP, Stearns JC, Rossi L, Luinstra K, Loeb M, Smieja M, Johnstone J, Surette MG, Bowdish DME. The loss of topography in the microbial communities of the upper respiratory tract in the elderly. *Ann Am Thorac Soc*. 2014;11(4):513–521.
- Kolhe R, Sahajpal NS, Vyavahare S, Dhanani AS, Adusumilli S, Ananth S, Mondal AK, Patterson GT, Kumar S, Rojiani AM, et al. Alteration in nasopharyngeal microbiota profile in aged patients with COVID-19. *Diagnostics*. 2021;11(9):1622.
- van den Munckhof EHA, Hafkamp HC, de Kluijver J, Kuijper EJ, de Koning MNC, Quint WGV, Knetsch CW. Nasal microbiota dominated by *Moraxella* spp. is associated with respiratory health in the elderly population: A case control study. *Respir Res*. 2020;21(1):181.
- Lee JJ, Kim SH, Lee MJ, Kim BK, Song WJ, Park HW, Cho SH, Hong SJ, Chang YS, Kim BS. Different upper airway microbiome and their functional genes associated with asthma in young adults and elderly individuals. *Allergy*. 2019;74(4):709–719.
- Jacobs JP, Spencer EA, Helmus DS, Yang JC, Lagishetty V, Bongers G, Britton G, Gettler K, Reyes-Mercedes P, Hu J, et al. Age-related patterns of microbial dysbiosis in multiplex inflammatory bowel disease families. *Gut*. 2024;73(12):1953–1964.
- Zhao M, Chu J, Feng S, Guo C, Xue B, He K, Li L. Immunological mechanisms of inflammatory diseases caused by gut microbiota dysbiosis: A review. *Biomed Pharmacother*. 2023;164:Article 114985.
- Evenepoel P, Stenvinkel P, Shanahan C, Pacifici R. Inflammation and gut dysbiosis as drivers of CKD–MBD. *Nat Rev Nephrol*. 2023;19(10):646–657.
- Mendes de Almeida V, Engel DF, Ricci MF, Cruz CS, Lopes ÍS, Alves DA, d' Auriol M, Magalhães J, Machado EC, Rocha VM, et al. Gut microbiota from patients with COVID-19 cause

- alterations in mice that resemble post-COVID symptoms. *Gut Microbes*. 2023;15(2):2249146.
23. Liu Q, Mak JWY, Su Q, Yeoh YK, Lui GC, Ng SSS, Zhang F, Li AYL, Lu W, Hui DSC, et al. Gut microbiota dynamics in a prospective cohort of patients with post-acute COVID-19 syndrome. *Gut*. 2022;71(3):544–552.
 24. Carneiro VL, Littlefield KM, Watson R, Palmer BE, Lozupone C. Inflammation-associated gut microbiome in postacute sequelae of SARS-CoV-2 points towards new therapeutic targets. *Gut*. 2024;73(2):376–378.
 25. Zhou Y, Zhang J, Zhang D, Ma WL, Wang X. Linking the gut microbiota to persistent symptoms in survivors of COVID-19 after discharge. *J Microbiol*. 2021;59(10):941–948.
 26. Ishizaka A, Koga M, Mizutani T, Yamayoshi S, Iwatsuki-Horimoto K, Adachi E, Suzuki Y, Kawaoka Y, Yotsuyanagi H. Association of gut microbiota with the pathogenesis of SARS-CoV-2 infection in people living with HIV. *BMC Microbiol*. 2024;24(1):6.
 27. Ferreira-Junior AS, Borgonovi TF, De Salis LVV, Leite AZ, Dantas AS, De Salis GVV, Cruz GNF, De Oliveira Lfv, Gomes E, Penna ALB, et al. Detection of intestinal dysbiosis in post-COVID-19 patients one to eight months after acute disease resolution. *Int J Environ Res Public Health*. 2022;19(16):10189.
 28. Su Q, Lau RI, Liu Q, Chan FKL, Ng SC. Post-acute COVID-19 syndrome and gut dysbiosis linger beyond 1 year after SARS-CoV-2 clearance. *Gut*. 2023;72(6):1230–1232.
 29. Tkacheva ON, Klimenko NS, Kashtanova DA, Tyakht AV, Maytesyan LV, Akopyan AA, Koshechkin SI, Strazhesko ID. Gut microbiome in post-COVID-19 patients is linked to immune and cardiovascular health status but not COVID-19 severity. *Microorganisms*. 2023;11(4):1036.
 30. Vestad B, Ueland T, Lerum TV, Dahl TB, Holm K, Barratt-Due A, Kåsine T, Dyrhol-Riise AM, Stiksrud B, Tonby K, et al. Respiratory dysfunction three months after severe COVID-19 is associated with gut microbiota alterations. *J Intern Med*. 2022;291(6):801–812.
 31. Barichello T, Kluwe-Schiavon B, Borba LA, Pedro LC, Niero FS, Santos LND, Leonardo LM, Ignácio ZM, Morales R, Ceretta LB, et al. Alterations in gut microbiome composition and increased inflammatory markers in post-COVID-19 individuals. *Mol Neurobiol*. 2025;62(6):8038–8047.
 32. Lau RI, Su Q, Ng SC. Long COVID and gut microbiome: Insights into pathogenesis and therapeutics. *Gut Microbes*. 2025;17(1):2457495.
 33. Ancona G, Alagna L, Alteri C, Palomba E, Tonizzo A, Pastena A, Muscatello A, Gori A, Bandera A. Gut and airway microbiota dysbiosis and their role in COVID-19 and long-COVID. *Front Immunol*. 2023;14:1080043.
 34. Hromić-Jahjefendić A, Mahmutović L, Sezer A, Bećirević T, Rubio-Casillas A, Redwan EM, Uversky VN. The intersection of microbiome and autoimmunity in long COVID-19: Current insights and future directions. *Cytokine Growth Factor Rev*. 2025;82:43–54.
 35. Lau RI, Su Q, Lau ISF, Ching JYL, Wong MCS, Lau LHS, Tun HM, Mok CKP, Chau SWH, Tse YK, et al. A synbiotic preparation (SIM01) for post-acute COVID-19 syndrome in Hong Kong (RECOVERY): A randomised, double-blind, placebo-controlled trial. *Lancet Infect Dis*. 2024;24(3):256–265.
 36. Docampo MJ, Batruch M, Oldrati P, Berenjano-Correa E, Hilty M, Leventhal G, Lutterotti A, Martin R, Sospedra M. Clinical and immunologic effects of paraprobiotics in long-COVID patients: A pilot study. *Neurol Neuroimmunol Neuroinflamm*. 2024;11(5):Article e200296.
 37. Ranisavljev M, Stajer V, Todorovic N, Ostojic J, Cvejic JH, Steinert RE, Ostojic SM. The effects of 3-month supplementation with synbiotic on patient-reported outcomes, exercise tolerance, and brain and muscle metabolism in adult patients with post-COVID-19 chronic fatigue syndrome (STOP-FATIGUE): A randomized placebo-controlled clinical trial. *Eur J Nutr*. 2024;64(1):28.
 38. Gutiérrez-Castrellón P, Gandara-Martí T, Abreu Y AT, Nieto-Rufino CD, López-Orduña E, Jiménez-Escobar I, Jiménez-Gutiérrez C, López-Velazquez G, Espadaler-Mazo J. Probiotic improves symptomatic and viral clearance in Covid19 outpatients: A randomized, quadruple-blinded, placebo-controlled trial. *Gut Microbes*. 2022;14(1):2018899.
 39. Wang Y, Wu G, Zhao L, Wang W. Nutritional modulation of gut microbiota alleviates severe gastrointestinal symptoms in a patient with post-acute COVID-19 syndrome. *MBio*. 2022;13(2):Article e0380121.
 40. Lau RI, Su Q, Ching JYL, Lui RN, Chan TT, Wong MTL, Lau LHS, Wing YK, Chan RNY, Kwok HYH, et al. Fecal microbiota transplantation for sleep disturbance in post-acute COVID-19 syndrome. *Clin Gastroenterol Hepatol*. 2024;22(12):2487–96.e6.
 41. Haran JP, Bradley E, Zeamer AL, Cincotta L, Salive MC, Dutta P, Mutaawe S, Anya O, Meza-Segura M, Moormann AM, et al. Inflammation-type dysbiosis of the oral microbiome associates with the duration of COVID-19 symptoms and long COVID. *JCI Insight*. 2021;6(20):e152346.
 42. Hilty M, Qi W, Brugger SD, Frei L, Agyeman P, Frey PM, Aebi S, Mühlemann K. Nasopharyngeal microbiota in infants with acute otitis media. *J Infect Dis*. 2012;205(7):1048–1055.
 43. Kumpitsch C, Koskinen K, Schöpf V, Moissl-Eichinger C. The microbiome of the upper respiratory tract in health and disease. *BMC Biol*. 2019;17(1):87.
 44. Chonmaitree T, Jennings K, Golovko G, Khanipov K, Pimenova M, Patel JA, McCormick DP, Loeffelholz MJ, Fofanov Y. Nasopharyngeal microbiota in infants and changes during viral upper respiratory tract infection and acute otitis media. *PLOS ONE*. 2017;12(7):Article e0180630.
 45. Dodi G, Attanasi M, Di Filippo P, Di Pillo S, Chiarelli F. Virome in the lungs: The role of anelloviruses in childhood respiratory diseases. *Microorganisms*. 2021;9(7):1357.
 46. Cao L, Ma Y, Wan Z, Li B, Tian W, Zhang C, Li Y. Longitudinal anellome dynamics in the upper respiratory tract of children with acute respiratory tract infections. *Virus Evol*. 2023;9(2):vead045.
 47. Taylor LJ, Keeler EL, Bushman FD, Collman RG. The enigmatic roles of *Anelloviridae* and *Redondoviridae* in humans. *Curr Opin Virol*. 2022;55:Article 101248.
 48. Liang W, Yang Y, Gong S, Wei M, Ma Y, Feng R, Gao J, Liu X, Tu F, Ma W, et al. Airway dysbiosis accelerates lung function decline in chronic obstructive pulmonary disease. *Cell Host Microbe*. 2023;31(6):1054–1070.e9.
 49. Guo C, Chen Q, Fan G, Sun Y, Nie J, Shen Z, Meng Z, Zhou Y, Li S, Wang S, et al. gcPathogen: A comprehensive genomic resource of human pathogens for public health. *Nucleic Acids Res*. 2024;52(D1):D714–D723.
 50. Sherif ZA, Gomez CR, Connors TJ, Henrich TJ, Reeves WB, RECOVER Mechanistic Pathway Task Force. Pathogenic mechanisms of post-acute sequelae of SARS-CoV-2 infection (PASC). *eLife*. 2023;12:e86002.

51. Meyer A, Buetti N, Houhou-Fidouh N, Patrier J, Abdel-Nabey M, Jaquet P, Presente S, Girard T, Sayagh F, Ruckly S, et al. HSV-1 reactivation is associated with an increased risk of mortality and pneumonia in critically ill COVID-19 patients. *Crit Care*. 2021;25(1):417.
52. Keeler EL, Merenstein C, Reddy S, Taylor LJ, Cobián-Güemes AG, Zankharia U, Collman RG, Bushman FD. Widespread, human-associated redondoviruses infect the commensal protozoan *Entamoeba gingivalis*. *Cell Host Microbe*. 2023;31(1):58–68.e5.
53. Zhang Y, Wang C, Feng X, Chen X, Zhang W. Redondoviridae and periodontitis: A case–control study and identification of five novel redondoviruses from periodontal tissues. *Virus Evol*. 2021;7(1):veab033.
54. Hasöksüz M, Kiliç S, Saraç F. Coronaviruses and SARS-CoV-2. *Turk J Med Sci*. 2020;50(SI-1):549–556.
55. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, Mailyan A, Manghi P, Scholz M, Thomas AM, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife*. 2021;10:e65088.
56. Chen Y, Gu S, Chen Y, Lu H, Shi D, Guo J, Wu WR, Yang Y, Li Y, Xu KJ, et al. Six-month follow-up of gut microbiota richness in patients with COVID-19. *Gut*. 2022;71(1):222–225.
57. Zuo T, Zhang F, Lui GCY, Yeoh YK, Li AYL, Zhan H, Wan Y, Chung ACK, Cheung CP, Chen N, et al. Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. *Gastroenterology*. 2020;159(3):944–55.e8.
58. Motta H, Reuwsaat JCV, Lopes FC, Viezzer G, Volpato FCZ, Barth AL, de Tarso Roth Dalcin P, Staats CC, Vainstein MH, Kmetzsch L. Comparative microbiome analysis in cystic fibrosis and non-cystic fibrosis bronchiectasis. *Respir Res*. 2024;25(1):211.
59. Bose T, Wasimuddin, Acharya V, Pinna NK, Kaur H, Ranjan M, Krishna JS, Nagabandi T, Varma B, Tallapaka KB, et al. A cross-sectional study on the nasopharyngeal microbiota of individuals with SARS-CoV-2 infection across three COVID-19 waves in India. *Front Microbiol*. 2023;14:1238829.
60. Wang P, Zhang S, Qi C, Wang C, Zhu Z, Shi L, Cheng L, Zhang X. Blood microbial analyses reveal long-term effects of SARS-CoV-2 infection on patients who recovered from COVID-19. *Comput Biol Med*. 2024;168:Article 107721.
61. Romani L, Del Chierico F, Pane S, Ristori MV, Pirona I, Guarrasi V, Cotugno N, Bernardi S, Lancella L, Perno CF, et al. Exploring nasopharyngeal microbiota profile in children affected by SARS-CoV-2 infection. *Microbiol Spectr*. 2024;12(3):Article e0300923.
62. Gupta A, Karyakarte R, Joshi S, Das R, Jani K, Shouche Y, Sharma A. Nasopharyngeal microbiome reveals the prevalence of opportunistic pathogens in SARS-CoV-2 infected individuals and their association with host types. *Microbes Infect*. 2022;24(1):Article 104880.
63. Ignatyeva O, Gostev V, Taraskina A, Tsvetkova I, Pavlova P, Sulian O, Ageevets V, Likholetova D, Chulkova P, Nikitina E, et al. General dynamics of the URT microbiome and microbial signs of recovery in COVID-19 patients. *Benef Microbes*. 2024;15(2):145–164.
64. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *Lancet*. 2020;395(10223):507–513.
65. Li S, Yang S, Zhou Y, Disoma C, Dong Z, Du A, Zhang Y, Chen Y, Huang W, Chen J, et al. Microbiome profiling using shotgun metagenomic sequencing identified unique microorganisms in COVID-19 patients with altered gut microbiota. *Front Microbiol*. 2021;12:Article 712081.
66. Rosas-Salazar C, Kimura KS, Shilts MH, Strickland BA, Freeman MH, Wessinger BC, Gupta V, Brown HM, Rajagopala SV, Turner JH, et al. SARS-CoV-2 infection and viral load are associated with the upper respiratory tract microbiome. *J Allergy Clin Immunol*. 2021;147(4):1226–33.e2.
67. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: Major findings, mechanisms and recommendations. *Nat Rev Microbiol*. 2023;21(3):133–146.
68. Phetsouphanh C, Darley DR, Wilson DB, Howe A, Munier CML, Patel SK, Juno JA, Burrell LM, Kent SJ, Dore GJ, et al. Immunological dysfunction persists for 8 months following initial mild-to-moderate SARS-CoV-2 infection. *Nat Immunol*. 2022;23(2):210–216.
69. Haq S, Grondin JA, Khan WI. Tryptophan-derived serotonin-kynurenine balance in immune activation and intestinal inflammation. *FASEB J*. 2021;35(10):Article e21888.
70. Leu-Semenescu S, Arnulf I, Decaix C, Moussa F, Clot F, Boniol C, Touitou Y, Levy R, Vidailhet M, Roze E. Sleep and rhythm consequences of a genetically induced loss of serotonin. *Sleep*. 2010;33(3):307–314.
71. Virgin HW. The virome in mammalian physiology and disease. *Cell*. 2014;157(1):142–150.
72. Gogokhia L, Buhrke K, Bell R, Hoffman B, Brown DG, Hanke-Gogokhia C, Ajami NJ, Wong MC, Ghazaryan A, Valentine JF, et al. Expansion of bacteriophages is linked to aggravated intestinal inflammation and colitis. *Cell Host Microbe*. 2019;25(2):285–99.e8.
73. Beloborodova NV, Grechko AV, Zurabov AY, Zurabov FM, Kuzovlev AN, Petrova MV, Chernevskaya EA, Yakovlev AA. Prospects of using adaptive phage therapy in the rehabilitation of post-COVID-19 patients. *PhysRehab*. 2021;3(3):254–259.
74. Fishman JA. Opportunistic infections—Coming to the limits of immunosuppression? *Cold Spring Harb Perspect Med*. 2013;3(10):Article a015669.
75. Liu Y, Wen Z, Fang Y, Wang T, Wu F, Zhang H, Chen D, Liu J. Herpesvirus reactivation in respiratory tract is associated with increased mortality of severe pneumonia patients and their respiratory microbiome dysbiosis. *Front Cell Infect Microbiol*. 2023;13:1294142.
76. Liu Z, Qi CJ, Shi Y, Li T, Fang Y, Zhang Q. Active herpesviruses are associated with intensive care unit (ICU) admission in patients pulmonary infection and alter the respiratory microbiome. *Front Microbiol*. 2024;15:1441476.
77. Charlys da Costa A, Mendes-Correa MC, Tozetto-Mendoza TR, Villas-Boas LS, de Paula AV, Paiao HGO, Leal FE, Ferreira NE, Honorato L, Leal E, et al. Detailed characterization of Redondovirus in saliva of SARS-CoV-2-infected individuals in Sao Paulo, Brazil. *PLOS ONE*. 2023;18(8):Article e0291027.
78. Spezia PG, Baj A, Drago Ferrante F, Boutahar S, Azzi L, Genoni A, Dalla Gasperina D, Novazzi F, Dentali F, Focosi D, et al. Detection of Torquetenovirus and Redondovirus DNA in saliva samples from SARS-CoV-2-positive and -negative subjects. *Viruses*. 2022;14(11):e0291027.
79. Merenstein C, Liang G, Whiteside SA, Cobián-Güemes AG, Merlino MS, Taylor LJ, Glascock A, Bittinger K, Tanes C,

- Graham-Wooten J, et al. Signatures of COVID-19 severity and immune response in the respiratory tract microbiome. *mBio*. 2021;12(4):Article e0177721.
80. Mohebbi A, Kiani SJ, Khanaliha K, Donyavi T, Emtiazi N, Sharifian K, Mohebbi M, Gholami A, Behnezhad F, Abbasi-Kolli M, et al. Dental complications as a potential indicator of Redondovirus infection: A cross-sectional study. *BMC Infect Dis*. 2024;24(1):673.
 81. Sivasubramanian MK, Monteiro R, Harrison Kelly S, Plakkot B, Subramanian M, Jones C. Herpes simplex virus type 1 preferentially enhances neuro-inflammation and senescence in brainstem of female mice. *J Virol*. 2022;96(17):e0108122.
 82. Sell EA, Tan LH, Renner DM, Douglas J, Lee RJ, Kohanski MA, Bosso JV, Kennedy DW, Palmer JN, Adappa ND, et al. Common cold coronavirus 229E induces higher interferon stimulating gene responses in human nasal epithelial cells from patients with chronic rhinosinusitis with polyposis. *Am J Rhinol Allergy*. 2024;39(1):13–20.
 83. Gaspar-Rodríguez A, Padilla-González A, Rivera-Toledo E. Coronavirus persistence in human respiratory tract and cell culture: An overview. *Braz J Infect Dis*. 2021;25(5): Article 101632.
 84. Torres-Morales J, Mark Welch JL, Dewhurst FE, Borisy GG. Site-specialization of human oral *Gemella* species. *J Oral Microbiol*. 2023;15(1):2225261.
 85. Park YC, Choi SY, Cha Y, Yoon HW, Son YM. Microbiome-mucosal immunity nexus: Driving forces in respiratory disease progression. *J Microbiol*. 2024;62(9):709–725.
 86. Wei N, Zhu G, Zhao T, Wang Y, Lou H, Li H, Yang Z, Zhang Z, Wang Q, Han M, et al. Characterization of oral bacterial and fungal microbiome in recovered COVID-19 patients. *BMC Microbiol*. 2023;23(1):123.
 87. Xie L, Luo G, Yang Z, Wu WC, Chen J, Ren Y, Zeng Z, Ye G, Pan Y, Zhao WJ, et al. The clinical outcome of COVID-19 is strongly associated with microbiome dynamics in the upper respiratory tract. *J Infect*. 2024;88(3):Article 106118.
 88. Ren L, Wang Y, Zhong J, Li X, Xiao Y, Li J, Yang J, Fan G, Guo L, Shen Z, et al. Dynamics of the upper respiratory tract microbiota and its association with mortality in COVID-19. *Am J Respir Crit Care Med*. 2021;204(12):1379–1390.
 89. Miyazaki D, Haruki T, Takeda S, Sasaki S-i, Yakura K, Terasaka Y, Komatsu N, Yamagami S, Touge H, Touge C, et al. Herpes simplex virus type 1-induced transcriptional networks of corneal endothelial cells indicate antigen presentation function. *Invest Ophthalmol Vis Sci*. 2011;52(7):4282–4293.
 90. Lin K, Cai J, Guo J, Zhang H, Sun G, Wang X, Zhu K, Xue Q, Zhu F, Wang P, et al. Multi-omics landscapes reveal heterogeneity in long COVID patients characterized with enhanced neutrophil activity. *J Transl Med*. 2024;22(1):753.
 91. Chen S, Zhou Y, Chen Y, Gu J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–i890.
 92. Kopylova E, Noé L, Touzet H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*. 2012;28(24):3211–3217.
 93. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–359.
 94. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20(1):257.
 95. Shen W, Ren H. TaxonKit: A practical and efficient NCBI taxonomy toolkit. *J Genet Genomics*. 2021;48(9):844–850.
 96. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31(10):1674–1676.
 97. Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods*. 2021;18(4):366–368.
 98. Nayfach S, Camargo AP, Schulz F, Eloë-Fadrosh E, Roux S, Kyrpidis NC. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol*. 2021;39(5):578–585.
 99. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: Architecture and applications. *BMC Bioinformatics*. 2009;10:421.
 100. Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, et al. The conserved domain database in 2023. *Nucleic Acids Res*. 2023;51(D1):D384–D388.
 101. Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. MAFFT-DASH: Integrated protein sequence and structural alignment. *Nucleic Acids Res*. 2019;47(W1):W5–w10.
 102. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25(15):1972–1973.
 103. Price MN, Dehal PS, Arkin AP. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE*. 2010;5(3):Article e9490.
 104. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res*. 2024;52(W1):W78–W82.
 105. Chen T, Chen X, Zhang S, Zhu J, Tang B, Wang A, Dong L, Zhang Z, Yu C, Sun Y, et al. The Genome Sequence Archive family: Toward explosive data growth and diverse data types. *Genomics Proteomics Bioinformatics*. 2021;19(4):578–583.
 106. CNCB-NGDC Members and Partners. Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2024. *Nucleic Acids Res*. 2024;52(D1):D18–d32.