

RESEARCH

Open Access



The virome composition of respiratory tract changes in school-aged children with *Mycoplasma pneumoniae* infection

Dianqi Zhang^{1†}, Yang Cao^{2†}, Biao Dai³, Teng Zhang⁵, Xing Jin¹, Qingyue Lan¹, Chaoying Qian^{1*}, Yumin He^{4*} and Yi Jiang^{3*}

Abstract

Background *Mycoplasma pneumoniae* (MP) is a common pathogen for respiratory infections in children. Previous studies have reported respiratory tract microbial disturbances associated with MP infection (MPI); however, since the COVID-19 pandemic, respiratory virome data in school-aged children with MPI remains insufficient. This study aims to explore the changes in the respiratory virome caused by MPI after the COVID-19 pandemic to enrich local epidemiological data.

Methods Clinical samples from 70 children with MPI (70 throat swab samples and 70 bronchoalveolar lavage fluid (BALF) samples) and 78 healthy controls (78 throat swab samples) were analyzed using viral metagenomics. Virus reads were calculated and normalized using MEGAN.6, followed by statistical analysis.

Results Principal Coordinate Analysis (PCoA) showed that viral community diversity is a significant difference between disease cohorts and healthy controls. After MPI, the number of virus species in the upper respiratory tract (URT) increased obviously, and the abundance of families *Poxviridae*, *Retroviridae*, and *Iridoviridae*, which infect vertebrates, rose evidently, particularly the species *BeAn 58,085 virus* (BAV). Meanwhile, phage alterations in the disease cohorts were predominantly characterized by increased *Myoviridae* and *Ackermannviridae* families and decreased *Siphoviridae* and *Salasmaviridae* families ($p < 0.01$). In addition, some new viruses, such as rhinovirus, respirovirus, dependoparvovirus, and a novel gemykibvirus, were also detected in the BALF of the disease cohort.

Conclusions This cross-sectional research highlighted the respiratory virome characteristics of school-aged children with MPI after the COVID-19 outbreak and provided important epidemiological information. Further investigation into the impact of various microorganisms on diseases will aid in developing clinical treatment strategies.

[†]Dianqi Zhang and Yang Cao contributed equally to this work.

*Correspondence:

Chaoying Qian
staff1516@yxph.com
Yumin He
hheyumin@126.com
Yi Jiang
staff1763@yxph.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **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/>.

Keywords School-aged children, Respiratory infection, *Mycoplasma pneumoniae*, Viral metagenomics, Respiratory virome

Introduction

Acute respiratory infections (ARIs), as a global health problem closely linked to high morbidity and mortality, have become the fourth leading cause of death in the world [1], especially acute lower respiratory infections (ALRIs), which have become the leading cause of hospitalization and death in young children [2, 3]. Previous studies have reported that *Mycoplasma pneumoniae* infection (MPI) in Chinese children decreased during the COVID-19 outbreak [4, 5]. However, since mid-2023, the atypical rise of MPI in Jiangsu Province (China) has gradually shown the trend of large-scale epidemics. The reason for the early onset of this MPI is still unclear, and there is still a lack of post-epidemic respiratory virome data related to MPI. Given the heavy burden of hospitals, further efforts should be made to elucidate the current region's unique epidemiological characteristics and potential pathogenic factors.

Mycoplasma pneumoniae (MP), an atypical bacterium, is one of the primary pathogens causing community-acquired pneumonia (CAP) and can cause severe respiratory infections in children [6, 7]. Clinical symptoms include fever, cough (dry or productive), sore throat, headache, and other common respiratory symptoms [8]. In severe cases, it can develop into lung abscess, bronchiectasis, pleural effusion, bronchiolitis obliterans, and respiratory distress syndrome [9]. The pathogenic mechanism of MP is a complex process as the main adhesins (P1 and P30) and accessory proteins (P90 and P40) in MP bind respiratory epithelial cells, either directly or indirectly, leading to host cell metabolism and structural changes, causing severe host damage [10–12]. He et al. demonstrated that standard antigenic components between MP and host cells can evade the host immune surveillance system and thus achieve long-term infection [13]. MP can also cause multi-system immune damage to the host, which may further increase the risk of disease and co-infection [14, 15]. Li et al. reported that the co-infection rate of MP and adenovirus (ADV) was as high as 92.3% (84/91) in bronchoalveolar lavage fluid (BALF) in 91 pediatric severe community-acquired pneumonia (SCAP) patients [16]. Similarly, Gao et al. found that children with MP and ADV infections had more severe clinical symptoms, as shown by a longer duration of fever and hospitalization, a higher proportion of dyspnea, and oxygen treatment [17]. Previous research has shown that MPI may induce dysregulation of the respiratory microbiota. *Firmicutes* and *Bacteroidetes* phyla were most common in healthy lungs [18], whereas, in MPI, *Tenericutes* abundance was significantly increased, exceeding 65% of

the entire microbiota [19]. One study reported that compared with healthy controls, microbial diversity in the oropharyngeal and nasopharyngeal regions decreased in MPI, with significant enrichment of *streptococcus* in the oropharyngeal and *Staphylococcus* in the nasopharyngeal area [20]. Dai et al. found that the lung microbial community was more similar to that in the nasopharyngeal under MPI conditions and that *Mycoplasma* dominated the BALF microbiota, while its abundance was lower in URT [21]. However, bacterial dysbiosis, including richness and diversity, cannot effectively explain virome changes in disease status [22].

At present, many studies have scrutinized the respiratory viral community. However, the primary focus is on discovering novel viruses or the absence of corresponding control groups and distinct pathological states [23, 24], which still possess some limitations in elucidating the potential pathogenic causes of diseases. Over the past few years, non-pharmaceutical interventions (NPIs) have effectively reduced the spread of SARS-CoV-2 during the COVID-19 epidemic [25]. However, protective measures inevitably affect the development of children's immune systems, especially those of current school age (>6 years). An epidemiological survey showed that school-aged children (aged 6–11) were more likely to develop severe conditions [26], possibly due to the relatively immature immune system and a more complex social environment. The host immune response was confirmed to be closely related to the severity of MPI [27, 28]. The respiratory microbiota is an essential factor in regulating and shaping the pulmonary immune response, and slight variations in this composition may impact host immunity. At the same time, microbe-microbe interactions can also significantly affect the process of respiratory diseases [29, 30].

This study used viral metagenomics to analyze the viral composition and respiratory tract differences in school-aged children with MPI. It identified potential pathogens and provided clues as to whether a co-infection in the current MP epidemic was possible. This study will not only enrich local epidemiological data but also contribute to providing empirical treatment strategies for these infections.

Materials and methods

Participants and sample collection

From July to September 2023, a total of 140 clinical specimens (70 throat swab specimens and 70 BALF specimens) from 70 school-aged children clinically diagnosed with MPI were collected, and 78 throat swab specimens from healthy school-aged children without clinical

symptoms were used as controls in Wuxi city, Jiangsu Province. The inclusion criteria for the disease group were children aged 6–12 years with fever, cough, runny nose, headache, and sore throat, positive for *Mycoplasma pneumoniae* and negative for other clinical diagnostic tests, including respiratory viruses (influenza A virus, influenza B virus, human parainfluenza virus, adenovirus, respiratory syncytial virus), *Legionella pneumophila*, *Chlamydia pneumoniae*, and *Coxiella burnetii*. Professional pediatricians collected throat swab samples from healthy children in this city's three experimental primary schools (6–12 years old). The healthy control had no obvious clinical symptoms of respiratory infection. Samples were tested for nine respiratory pathogens (influenza A virus, influenza B virus, human parainfluenza virus, adenovirus, respiratory syncytial virus, *Legionella pneumophila*, *Chlamydia pneumoniae*, *Coxiella burnetii*, and *Mycoplasma pneumoniae*) by the laboratory department of the hospital. Those with negative results were included in the study cohort.

All samples were collected by professional clinicians, stored in sterile centrifuge tubes or sputum collectors, and then transported by cold chain. Throat swab samples were vigorously vortexed in 1 mL Dulbecco's phosphate-buffered saline (DPBS) for 10 min before being freeze-thawed three times in dry ice. After centrifugation (10 min, 15,000 g, 4 °C), each sample's supernatants were collected in a new 1.5 mL centrifuge tube and stored at -80 °C. After filtration through sterile gauze, BALF samples were centrifuged at 400 g for 10 min at 4 °C, and then the supernatant was transferred to a high-speed centrifuge tube. After centrifugation (10 min, 15000 g, 4 °C), the supernatant was discarded, and 1 mL of DPBS was added to resuspend the precipitate and store it at -80 °C.

Viral nucleic acid extraction and library construction

The 218 supernatants were randomly and evenly combined into 15 sample pools (500 µL/pool) according to the health status and sample type. Eukaryotic and some bacterial cell-sized particles were removed from each sample pool using a 0.45 µm filter (5 min, 8,000 g, 4 °C), and DNase (Turbo DNase from Thermo Fisher; Baseline-ZERO from Epicentre; benzonase from Novagen) and RNase (Promega, Madison, WI, USA) kits were used to digest unprotected nucleic acid in filtrates at 37 °C for 60 min. QIAamp viral RNA Minikit (Qiagen, HQ, Germany) was used to extract the remaining viral nucleic acids according to the manufacturer's protocol.

The enriched viral RNA from the respective pools was subjected to reverse transcription reactions using SuperScript IV Reverse Transcriptase (Invitrogen, CA, USA) and six-base random primers, and then double-strand DNA synthesis using Klenow fragment polymerase (New England Biolabs, Ipswich, MA, USA). Using the Nextera

XT DNA Sample Preparation Kit (Illumina, CA, USA), 15 libraries were constructed and sequenced on the MiSeq Illumina platform with 250 bases paired-ends with dual barcoding for each sample library.

Bioinformatics analysis

The 250-bp paired-end reads with barcodes generated by the Illumina MiSeq platform were trimmed using vendor software from Illumina. Data was processed using an in-house analysis pipeline running on a 32-node Linux cluster, and reads were treated as duplicates if bases 5 to 55 were identical, with only one random copy of duplicates kept. The Phred quality score of 10 as the threshold was used to trim tails with low sequencing quality, and adaptors were trimmed using the default parameters of VecScreen in the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/tools/vecsreen/>). Bacterial nucleotide sequences were mapped from the BLAST NT database and subtracted using Bowtie2, and cleaned reads were *de-novo* assembled by SOAPdenovo2 (version r240) with a K-mer size of 63 with default settings [31, 32]. Contigs and singlets were then aligned to an in-house viral proteome database using BLASTx (<https://ftp.ncbi.nih.gov/refseq/release/viral/>) with an E-value cutoff of $<10^{-5}$. Candidate viral hits were then compared to an in-house non-virus non-redundant (NVNR) database (based on annotation taxonomy excluding viruses) with an E-value cutoff of $<10^{-5}$ to eliminate false-positive viral hits. MEGAN (v6.22.2) was used to assign each sequence present in metagenomic data to different taxa using the NCBI taxonomic database.

Viral sequences extension and annotation

The generated individual viral contig was used as a reference to obtain a longer contig or nearly complete genome using the Low Sensitivity/Fastest parameter in Geneious Prime v2020.0.5, and then predicted putative viral open reading frames (ORFs) with self-defined parameters (genetic code: Standard; minimum size: 300; start codons: ATG) [33], and the predicted ORFs were further compared against the NCBI database using BLASTp.

Phylogenetic analysis

Detailed methods were described in the previous research [34, 35]. Roughly as follows, the amino acid (aa) sequences of viral reference strains were downloaded from the NCBI GenBank database, and sequence alignment was performed using Cluster X with the default parameters in MEGA (v10.1.8) [36]. Sites with more than 50% gaps were removed from the alignment results, and Bayesian inference trees were constructed using MrBayes v3.2.7 [37]. For MrBayes analysis, "lset nst=6 rates=invgamma" were used for phylogenetic analysis based on

nucleotide sequences, which set the evolutionary model to the GTR substitution model with gamma-distributed rate variation across sites and a proportion of invariable sites (“GTR+I+ Γ ”), while “prset aamodelpr = mixed” were set for the phylogenetic analysis using aa sequences, which allows the program to utilize the 10 built-in aa models, and the number of generations was increased to one million until the standard deviation of split frequencies is below 0.01, sampled every 50 generations, and the first 25% of Markov chain Monte Carlo (MCMC) samples were discarded as burn-in. Meanwhile, all Bayesian inference trees were further validated using the maximum likelihood trees constructed by the MEGA.

Statistical analysis

The virus composition analyses were normalized using MEGAN.6, and the viral community analyses were compared by the R v4.3.1 package (vegan, ggpubr, ggsignif, dplyr, and ggplot2). The Kruskal-Wallis test compared the alpha diversity. The Mann-Whitney U test was used to compare groups. Statistical Analysis of Metagenomic Profiles (STAMP) analysis was constructed by the R v4.3.1 package (tidyverse and boot), and the Wilcoxon test was calculated to compare groups. $P < 0.05$ was considered statistically significant.

Nucleotide sequence accession number

The raw sequence reads generated in the study are available at the China National Center for Bioinformatics (CNCB) database under the BioProject accession PRJCA027224 (<https://ngdc.cncb.ac.cn/gsub/>), and the Biosample accession can be found in Supplementary Table. S1. All viral sequences identified in this study were deposited in the CNCB database (<https://ngdc.cncb.ac.cn/genbase/>), and the accession numbers were provided in the Supplementary Table. S2.

Result

Overview of viral metagenomic

Fifteen libraries were divided into three cohorts according to participants' health status and sample types, including Healthy Control (5 libraries, WXN1-WXN5, HC), MP-infected throat swab (5 libraries, WXD1-WXD5, MT), and MP-infected BALF (5 libraries, WXD1-WXD5, MB).

After next-generation sequencing with the Illumina Miseq platform, the 15 libraries generated 193,279,390 raw sequence reads. With an E-value cutoff of $< 10^{-5}$, 762,858 sequences demonstrated significant sequence identities for known viruses using MEGAN.6 (18,874–134,174 reads per library), most of which were assigned to phage sequences, mainly belonging to *Siphoviridae*, *Myoviridae*, *Podoviridae*, and *Autographiviridae*. Besides, 301,576 sequences from the following families

of viruses that could infect vertebrates were detected, including *Poxviridae*, *Picornaviridae*, *Retroviridae*, *Coronaviridae*, *Anelloviridae*, and *Paramyxoviridae*, among others. Of these, *Poxviridae* and *Picornaviridae* accounted for the majority of the viruses observed infecting vertebrates in this study. Detailed information on each library can be found in the Supplementary Table. S3.

Diversity analysis of viral communities

Virus reads were calculated and normalized using MEGAN.6, followed by composition analysis. The flattening of each library curve in the Rarefaction Curve indicates that the biodiversity of community predictions can be made within the range of sequencing depth (Fig. 1A). Assessing alpha diversity using Shannon, Simpson, and ACE indices to reflect virus diversity and richness. The results showed significant differences in alpha diversity among the three cohorts (Fig. 1B). The Shannon and Simpson indices decreased slightly in the MT cohort compared to the HC cohort. The differences were significant in the MB cohort ($p < 0.001$), while the ACE index increased significantly in the MT cohort ($p < 0.05$). These results indicated that MPI resulted in changes in viral community composition, including diversity and richness. On the other hand, the above results also revealed apparent differences in the composition of upper and lower respiratory virus communities (MT vs. MB).

We further explored whether the overall viral phenotype differed among the three groups. PCoA was performed using the Bray Curtis distance matrix to compare differences in community composition (Fig. 1C), and the results showed that viral signatures among the three groups were significantly distinct (ANOSIM: $R = 0.928$, $P < 0.01$). The difference in virus community structure among the three groups was significant, and the five libraries in each group could be effectively clustered, indicating that the difference within each group was slight and the sampling was reasonable. Meanwhile, the Partial Least Squares Discrimination Analysis (PLS-DA) model showed a similar result (Fig. 1D).

Composition of the viral communities

Viral reads were analyzed at the family and species levels to reveal community composition. At the family level, viral reads in the three cohorts were classified into 81 known viral families, which contain 24 viral families that can infect vertebrates, including 12 RNA virus families and 12 DNA virus families (Fig. 2A). Among the viral reads that infect vertebrates, *Picornaviridae* (70.41%), *Coronaviridae* (7.67%), and *Picobirnaviridae* (7.06%) showed a relatively high proportion in the HC cohort. However, *Picornaviridae* (39.01%, 20.06%), *Poxviridae* (31.19%, 45.81%), and *Retroviridae* (9.28%, 10.35%) were the main viruses in the MT and MB cohorts, respectively.

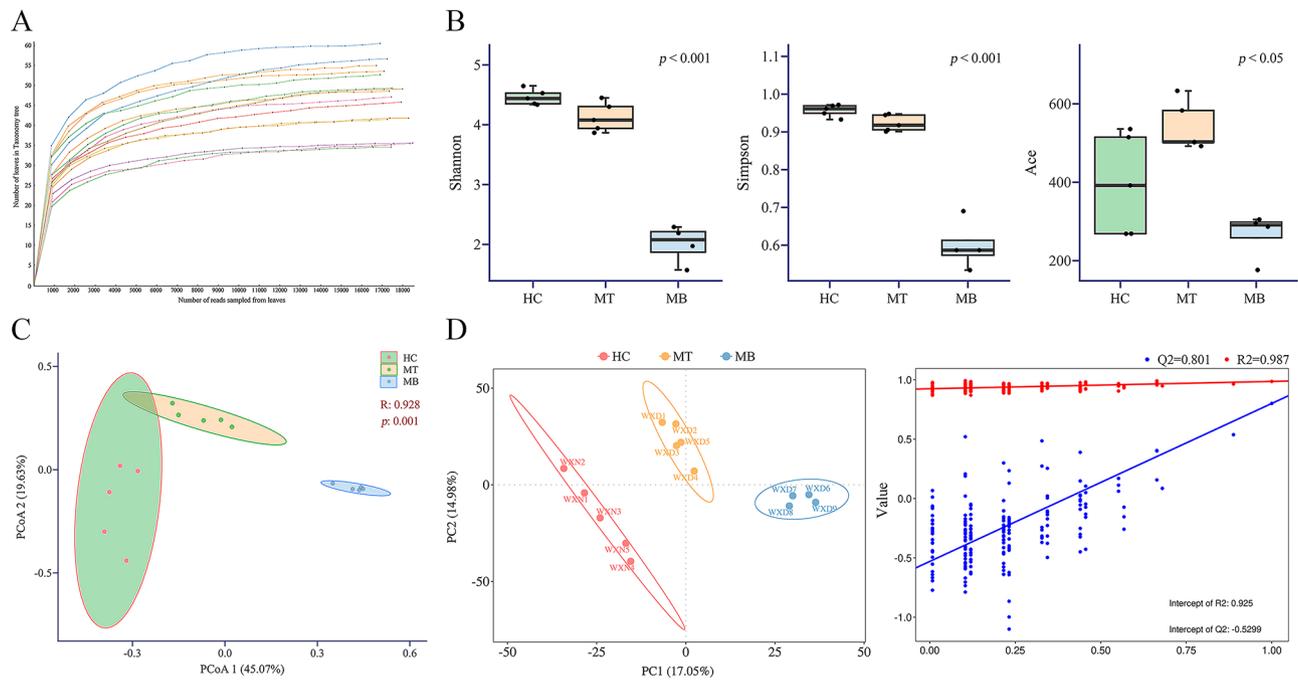


Fig. 1 Comparison of viral communities among the three cohorts. **(A)** Rarefaction curves of viral species in each library. **(B)** Alpha diversity was assessed with Shannon, Simpson, and ACE indices in the three cohorts using the Kruskal-Wallis test. **(C)** PCoA was performed based on Bray-Curtis distance for the overall differences and similarities of all samples at the specie level. ANOSIM was utilized to assess the similarity among high-dimensional data sets and provide a foundation for evaluating the statistical significance of differences. **(D)** The supervised PLS-DA model was further used to assess group differences, and the model was calculated using seven-fold cross-validations and 200 response permutation tests to prevent overfitting

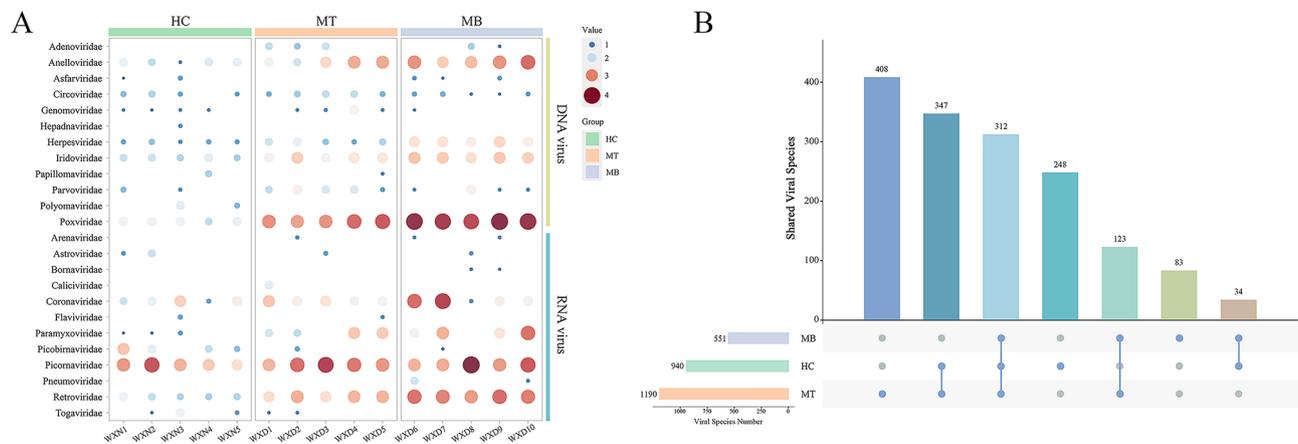


Fig. 2 Composition of viral communities. **(A)** Viral reads in the three cohorts were calculated and normalized by Megan6, and then the reads covered on a log10 scale. The circle size and color represent the relative abundance of different virus families. **(B)** UpSet plot showing the numbers of shared species in the three cohorts. The horizontal bars represent the total number of viral species in each group

Notably, the composition of viral communities in the URT was significantly altered following MPI (HC vs. MT), with a significant increase in the proportion of *Poxviridae* (4.35% vs. 31.19%) and *Retroviridae* (2.00% vs. 9.28%). In addition, *Adenoviridae*, *Pneumoviridae*, *Caliciviridae*, *Arenaviridae*, and *Bornaviridae* were only detected in the disease cohort. At the species level, 1554 species were identified in virus reads across all cohorts. The MT cohort contained the most virus species, with

408 viruses unique to this group, while only 248 viruses were unique to the HC group (Fig. 2B). This indicates that *Mycoplasma pneumoniae* infection had a significant impact on the respiratory virus community.

Differences in virus communities

In the HC and MT cohorts, 6 viral families displayed significant differences in abundance. The families *Herpesviridae*, *Iridoviridae*, *Parvoviridae*, *Poxviridae*, and

Retroviridae were significantly higher than the HC cohort in the MT cohort. At the same time, the number of *Picobirnaviridae* was clearly lower than the HC cohort (Fig. 3A). In the state of MPI, 5 viral families were significantly different between the upper and lower respiratory tracts (MT vs. MB), and viral reads in the families of *Herpesviridae*, *Iridoviridae*, *Poxviridae*, and *Retroviridae* showed visibly higher abundance in the MB cohort compared to the MT cohort (Fig. 3B). Although the abundance of *Anelloviridae* was overtly increased in the disease group, it did not show a statistical difference. At the genus level, *Enterovirus* read abundance was higher in the upper respiratory tract virus community of the healthy control group. In contrast, abundance of the genus *unclassified Chordopoxvirinae* and *Enterovirus* was the main domain in the infection cohort (Fig. 3C).

In addition, phage, as the main component of the viral community, also changed significantly after infection with MP. Phage *Siphoviridae*, *Salasmaviridae*, and *Microviridae* showed an obvious lower abundance in the

MT cohort compared to HC, while *Ackermannviridae* and *Myoviridae* were significantly more abundant. Meanwhile, the abundance of phage reads in the LRT (23,115 reads, MB) was obviously lower than that in the URT (99,290 reads, MT). Phage *Myoviridae*, *Microviridae*, and *Autographiviridae* showed significant differences (Fig. 3D).

Subsequently, STAMP was used to investigate further differences in the relative abundance of viral reads at the species level. STAMP provides extensive hypothesis testing and confidence intervals for identifying of biologically relevant differences [38]. The results showed that 13 virus species were significantly different between the HC and MT cohorts (Fig. 4A), most of which (*Mouse mammary tumor virus*, *Bovine retrovirus CH15*, *Feline leukemia virus*, *Harvey murine sarcoma virus*, *Reticuloendotheliosis virus*, *Baboon endogenous virus*, and *Atlantic salmon swim bladder sarcoma virus*) belonged to the family *Retroviridae*. In the MT and MB cohorts (Fig. 4B), there were 15 distinct virus species, most of which (*Murmansk*

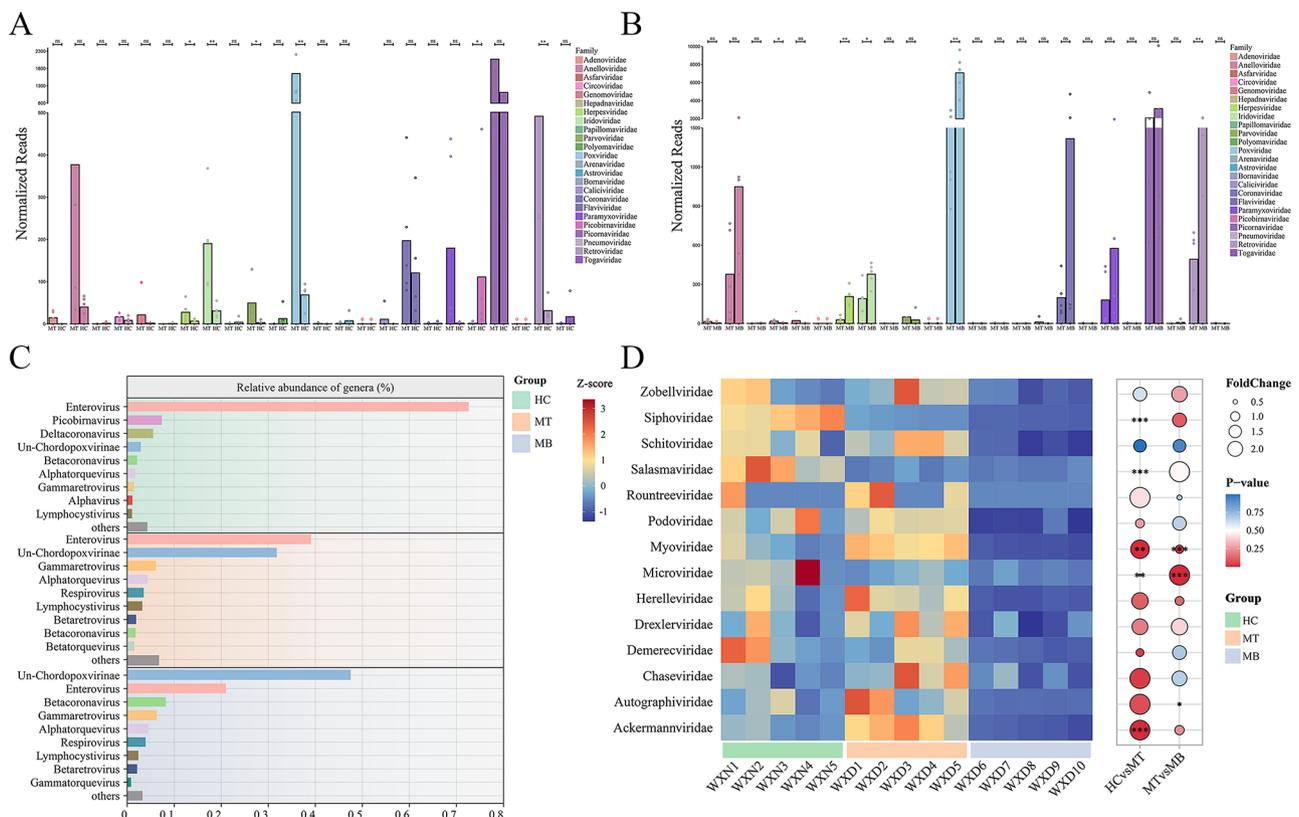


Fig. 3 Differences in viral community composition. **(A-B)** Bar plots showed the normalized viral reads of vertebrate infection in HC vs. MT cohorts or MT vs. MB cohorts. Each viral family was indicated in a different colored rectangle, and each point represented a library. Differences between the two groups were calculated using the Mann-Whitney U test. 0 represented the absence of viral reads between the two cohorts. **(C)** Bar plots showing the top 9 most abundant viral genera in the three cohorts. The same viral genera between each cohort were indicated with consistent color, and the horizontal axis represented the relative abundance of reads assigned to each genus. Un-Chordopoxvirinae represented the unclassified Chordopoxvirinae. **(D)** Heatmap representing the normalized viral reads of the family *Microviridae* and the order *Caudovirales*. Differences between the two groups were measured by the Mann-Whitney U test. The bubble size and color depth represented the Fold Change value and significance, respectively. (ns indicates a nonsignificant difference, * $P < 0.05$, ** $P < 0.01$)

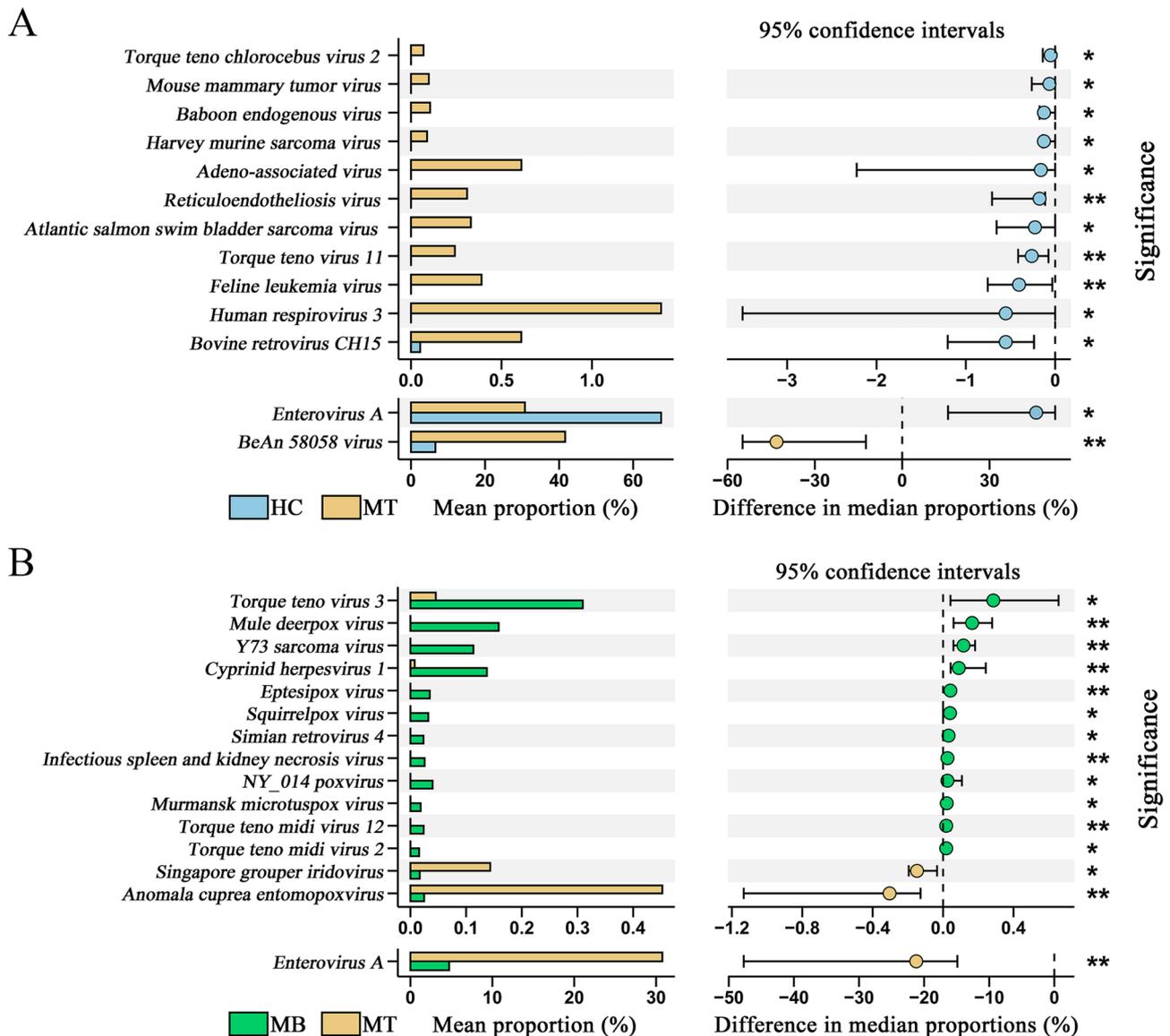


Fig. 4 STAMP analyses of the respiratory virome at the species level (A-B) The difference in relative viral abundance between cohorts was analyzed by STAMP. The differences were calculated using Wilcoxon test. Significance was shown on the right. (* $P < 0.05$, ** $P < 0.01$)

microtus pox virus, *NY_014 poxvirus*, *Mule deerpox virus*, *Squirrelpox virus*, *Eptesipox virus*, and *Anomala cuprea entomopoxvirus*) belonged to the family *Poxviridae*.

Analysis of new viruses

Ultimately, to compensate for the lack of conventional clinical methods, this study utilized viral metagenomic technology to examine possible vertebrate viruses in all libraries and obtained seven relatively complete virus sequences.

Sequence reads corresponding to the family *Picornaviridae* were identified from all libraries. Four nearly complete genomes of the virus were obtained and named WX/HC1/Picor2023 (mean coverage: 26.1), WX/HC2/Picor2023 (mean coverage: 37.7), WX/BALF8/Picor2023

(mean coverage: 44.4), and WX/BALF10/Picor2023 (mean coverage: 155.1), respectively (Figure. S1A). Phylogenetic analysis was performed based on VP1 nucleotide (nt) sequences, including reference sequences from the genus *Enterovirus* [39]. The results showed that WX/HC2/Picor2023, WX/BALF8/Picor2023, and WX/BALF10/Picor2023 clustered significantly with the species *Rhinovirus A* forming a clade, while WX/HC1/Picor2023 clustered with the species *Enterovirus A* forming a clade (Figure. S1B). Sequence analysis showed that WX/HC1/Picor2023 shared the 97.7% nt sequence identity with *Coxsackievirus A6 strain* (OR394975), WX/HC2/Picor2023 shared 97.0% with *Rhinovirus A21 strain* (OL638406), WX/BALF8/Picor2023 shared 97.1% with *Rhinovirus A21 strain* (MZ835609), and WX/BALF10/

Picor2023 shared 98.8% with *Rhinovirus A54 strain* (MZ540950).

One nearly complete respirovirus was discovered in library #WXD10 and named WX/BALF10/Param2023 (mean coverage: 55.6). The genome was 15,492 bp and contained a complete coding sequence (CDS), including N, P, M, F, HN, and L (Fig. S2A). Phylogenetic analysis based on N protein and L protein construction showed that WX/BALF10/Param2023 belonged to a member of the species *Human respirovirus 3* and shared 99.6% nt sequence identity with *Human respirovirus 3 isolated NICU/BJ/CHN/2022* (OQ785280) (Figure. S2B).

A nearly complete genome of the family *Parvoviridae* was detected in library #WXD8 and named WX/BALF8/Parv2023 (mean coverage: 21.6). The virus was 4,529 bp in length and contained two complete ORFs encoding a nonstructural protein (NS1, 648aa) and a structural protein (VP1, 726aa), respectively. Phylogenetic analysis was performed based on VP1 aa sequences, including reference sequences from the genus *Dependoparvovirus*, and results showed that WX/BALF8/Parv2023 was clustered with avian adeno-associated viruses (MT138226, MT138246, MT138249) but formed a separate branch (Figure. S3A). According to the International Committee on Taxonomy of Viruses (ICTV) classification criteria for species in the family *Parvoviridae*, viruses within a species usually encode NS1 protein that exhibits >85% aa sequence identity. Analysis of the NS1 aa sequence showed that the virus shared the identity of 91.5–93.7% with the above three virus strains. On the basis of the above evidence, WX/BALF8/Parv2023 belongs to a member of the genus *Dependoparvovirus*.

A complete genome of *Genomoviridae* was obtained from library #WXD10 and named WX/BALF10/Genom2023 (mean coverage: 8.9). The virus was 2150 bp in length, and the genome organization was typical of the family *Genomoviridae*. WX/BALF10/Genom2023 was examined for genetic relationships with family references using the replication-associated protein. The phylogenetic tree showed that the virus detected in the study was clustered significantly with the genus *Gemykibvirus* (Figure. S3B). Further analysis showed that WX/BALF10/Genom2023 shared only 57.4% and 50.2% genome-wide identity with closely related *faecal-associated gemycircularvirus 9* (NC_025731) and *human gemycircularvirus GeTz1* (NC_038497). According to the 78% genome-wide pairwise identity as a species demarcation threshold [40], WX/BALF10/Genom2023 is a new member of the genus *Gemykibvirus*.

Discussion

The respiratory tract, as the communication channel between the human body and the external environment, is also the main route for invading many pathogens.

Previous studies have shown that changes in respiratory microecology can affect host respiratory health and the development of disease processes [41, 42]. This study compared the respiratory virus group between MPI children and healthy controls, revealing differences in viral community changes due to infection. Small changes in differences in the diversity of the URT virome may depend on common environmental exposures, as well as the effects of inherently colonizing viruses in the respiratory tract, such as the *Picornaviridae* that we have observed. Subsequently, we further explored the differences in the viral groups of the upper and lower respiratory tract after MPI. The similarity in viral community composition between the two cohorts may be attributed to two explanations: (1) Viral metastasis due to hyper-ventilation; (2) Infectious agents cause viral reactivation.

Special bacterial communities that colonize the respiratory tract are thought to play a major role in maintaining human health [43]. For most pathogens, colonization on the surface of the respiratory tract is a necessary precondition for infection, and the “colonization resistance” effect generated by the resident microbiota may be important for maintaining respiratory health [44]. In addition, the respiratory microbiota also plays an essential role in the formation of local immunity [45]. The complex pathogenic mechanism of MP often leads to varying degrees of respiratory system damage. MP interacts with the host bronchial epithelium through adhesion proteins, inducing intracellular metabolism and ultrastructural changes in infected cells, leading to the destruction of the integrity of the airway lumen surface membranes and further affecting the self-purification ability of the respiratory tract, which may provide a prerequisite for other pathogens to invade the host [12]. A report of pathogen detection using the TaqMan Array Card (TAC) in MP-infected patients found that at least one other bacterial or viral codetection was identified in 59.8% (125/209) patients, and bacterial and viral codetection was found in 19% (34/209) patients, and was observed only in pediatric patients [14]. Moreover, several studies have demonstrated that MPI can cause an imbalance in the respiratory microbiota [19, 21]. As shown in our study, the relative abundance of many viruses increases significantly in the disease state, such as *Anelloviridae*, *Iridoviridae*, *Poxviridae*, *Coronaviridae*, *Paramyxoviridae*, and *Retroviridae*. In a previous study investigating unexplained acute respiratory infections, Mao et al. found that *Picornaviridae*, *Parvoviridae*, *Paramyxoviridae*, *Coronaviridae*, and *Anelloviridae* were the top five virus families with the highest relative abundance in the URT [23]. Our results have partially overlapped with theirs. Due to the lack of a healthy control group, their results do not reflect changes in the viral community under disease status. As we observed in this study, the family *Picornaviridae* in

the healthy state still has a high proportion, indicating that it may be a resident virus in the respiratory tract. The family *Anelloviridae* has been confirmed as a widely prevalent family of viruses throughout the respiratory tract [46]. Although highly prevalent in humans, no studies have shown that it is directly involved in any pathogenic process [47]. However, the relative abundance of *Anelloviridae* is increasing under MPI, a trend consistent with previous research results [23, 48]. Several studies have shown that anellovirus can be used as a potential clinical biomarker for immune function; as observed in organ transplant patients, blood levels of anellovirus increase with immunosuppressive therapy before the onset of acute organ rejection showed a relative decline [49, 50]. Changes in anellovirus levels may indicate the body's immune status. It is suggested that *Anelloviridae* replication may have a potential role in disease progression, and the underlying mechanism remains to be explored.

As major members of the virome, phages are generally thought to maintain microbiota homeostasis by regulating and modifying bacterial colonies [51]. In this study, DNA phages of *Caudovirales* were dominated by the bacteriophage component and mainly concentrated in the URT, which also reflected the relatively clean environment of the LRT. The MP infection caused changes in the phage community. The phage *Siphoviridae* was the dominant virus in the healthy state, while the abundance of *Myoviridae* increased significantly and became dominant after MPI. Intriguingly, one study reported that pathogen-induced inflammation catalyzes phage induction and transfer during infection [52], and immune activity and phage-killing effects are synergistic in combating bacterial infection [53]. Given recent evidence, we propose a conjecture as to whether MPI leads to changes in the phage community by destabilizing the microbiota or whether infection catalyzes phage induction and promotes rebalancing of dysregulated microbiota. Therefore, exploring phage-microbiota and phage-immune system interactions may have important biological implications and broaden the horizons of clinical therapeutics.

It is worth noting that a large number of *poxviridae* reads were observed in the respiratory tract of MPI patients in this study, particularly the BAV. BAV has been characterized and recommended to belong to the genus *Orthopoxvirus*. The genus belongs to double-stranded DNA (dsDNA) viruses and are zoonotic pathogen [54]. In a study of the lung microbiome of COPD patients in Tshwane, South Africa, the team reported a high abundance of BAV in all sputum samples in the study [55], consistent with what was observed in the disease cohort in our study. Several immature theories point to possible reasons for the high abundance of BAV. First, BAV was an ancient virus that had become part of the human genome over time [56]. Second, the virus is a DNA artifact of the

smallpox vaccine received years earlier because BAV has a high degree of homology with the vaccinia virus (used for smallpox vaccination) [54]. Finally, participants were infected with the virus via environmental exposure [57]. However, even if the above theories can explain the existence of BAV in humans, the potential significance and mechanism of the apparently elevated abundance of BAV in disease states have yet to be elucidated. To the best of our knowledge, this is the first cross-sectional study to report the differential performance of BAV in the human respiratory tract.

Another significant difference in the infection cohort was the *Retroviridae*, for which more useful biological information was not obtained because of the short viral sequences. Notably, when we performed BLASTx analysis of longer sequences, some of which showed high sequence similarity with human endogenous retroviruses (HERVs), we speculated whether MPI caused the reactivation of HERVs and then delivery to the extracellular space by vesicles. Previous studies have verified that Environmental stimulation can activate partial HERV expression, such as infectious agents, exogenous viruses, radiation, or cytokines [58]. Several studies have confirmed that HERV reactivation produces of viral transcripts and proteins, which can affect important host biological functions and promote disease development [59, 60]. Therefore, it is necessary to explore further the role of the abnormal increase of *Retroviridae* in respiratory tract infections.

Conventional clinical methods, including microbial culture, serology tests, and nucleic acid amplification technology (NAAT), for the detection of respiratory pathogens still have certain limitations. Microbial culture not only takes a long time due to its characteristics but is also susceptible to antimicrobial therapy, especially for viruses. Serology tests are often subject to thresholds and population variations, such as antibodies that persist long after infection has cleared in young children [61]. Although NAAT is highly sensitive and specific, it requires effective prediction of the pathogen. Meanwhile, pulmonary infection, due to its universality, complexity, and heterogeneity in clinical diagnosis, is unclear in about 19–62% of aetiology. Using unbiased viral metagenomics techniques, we investigated viral community composition under MPI. As expected, some common ARI viruses were detected in the disease group, such as *Adenoviridae*, *Paramyxoviridae*, and *Pneumoviridae*. It is worth noting that the inclusion criteria for the disease group in this study already excluded subjects who tested positive for routine clinical respiratory viruses. Obviously, conventional methods still have a certain degree of missed detection. A retrospective study of pneumonia infection showed that the positive rate of metagenomic next-generation sequencing (mNGS) for virus detection

was as high as 92.31% compared to conventional methods (7.69%) [62]. Therefore, for each regional or large-scale outbreak of infectious disease, comprehensive testing is needed to explore the potential causes of the epidemic trend. At the same time, with cost reduction and process simplification, mNGS technology is expected to become a conventional clinical method for microbial detection and diagnosis.

The emergence of SARS-CoV-2 has reshaped our understanding of respiratory viruses, highlighting the challenges of etiological diagnosis due to symptom similarity. As we observed, *Mycoplasma pneumoniae* infection leads to changes in the respiratory virome. The clinical significance of the significantly increased abundance of *Poxviridae* and *Retroviridae* is not clear, and it may be involved in the host immune regulation process. The detection of *Adenoviridae*, *Coronaviridae*, and *Paramyxoviridae* suggests the omission of clinical routine detection. Therefore, for severe and persistent MPI, coinfection and superinfection should be considered, and clinical medication should be guided. In addition, the changes in phage composition also reflect the imbalance of respiratory flora and can provide some evidence for phage therapy for refractory *Mycoplasma pneumoniae* pneumonia.

Several limitations of this study should also be considered. First, considering the need to test as many clinical samples as possible, sample pooling obscured the resolution of sample-specific diversity to some extent. It may affect the detection of rare species. Meanwhile, the inevitable loss of nucleic acid fragments in the library enrichment process has led to the inability to obtain complete genetic information for some viruses, such as *Coronaviridae* and *Anelloviridae*. Subsequent studies can obtain more complete sequencing information by single sample sequencing and optimizing the experimental process. Second, MP was not sequenced to compare the effect of resistance gene mutation on changes in respiratory viral communities. Further research is needed to investigate the effect of this aspect. Third, as an invasive procedure, BALF collection is relatively limited and requires future large-scale cohort studies to validate the findings of this study. In addition, a single cross-sectional study cannot monitor the temporal dynamic changes of the microbiome. Therefore, it is necessary to form a continuous surveillance state by establishing a multi-center joint to provide scientific support for developing prevention and control strategies for *Mycoplasma pneumoniae* pneumonia.

Conclusions This cross-sectional research highlighted the respiratory virome characteristics of school-aged children with MPI after the COVID-19 outbreak, with obvious increases in *Poxviridae*, *Retroviridae*, *Iridoviridae*

families, and the phages *Siphoviridae* and *Ackermannviridae*. Further investigation into the differentially altered viral communities may help to develop new clinical treatment strategies.

Abbreviations

ADV	Adenovirus
ALRIs	Acute lower respiratory infections
ARIs	Acute respiratory infections
BAV	BeAn 58085 virus
BALF	Bronchoalveolar lavage fluid
CAP	Community-acquired pneumonia
CDS	Complete coding sequence
CNCB	China national center for bioinformation
DPBS	Dulbecco's phosphate-buffered saline
HC	Healthy control
LRT	Lower respiratory tract
MB	MP-infected balf
MCMC	Markov chain monte carlo
MP	<i>Mycoplasma pneumoniae</i>
MPI	<i>Mycoplasma pneumoniae</i> infection
MT	MP-infected throat swab
NAAT	Nucleic acid amplification technology
NCBI	National center for biotechnology information
NPIs	Non-pharmaceutical interventions
NVNR	Non-virus non-redundant
ORFs	Open reading frames
PCoA	Principal Coordinate Analysis
PLS-DA	Partial least squares discrimination analysis
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SCAP	Severe community acquired pneumonia
STAMP	Statistical analysis of metagenomic profiles
TAC	TaqMan array card
URT	Upper respiratory tract

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12985-025-02626-9>.

Supplementary Material 1: Table S1: Detailed information of 15 libraries constructed in this study.

Supplementary Material 2: Table S2: Detailed information of viral sequences identified in this study.

Supplementary Material 3: Table S3: The number of viral reads and contigs in each library.

Supplementary Material 4: Figure. S1: Genome organization and phylogenetic analysis of enterovirus.

Supplementary Material 5: Figure. S2: Genome organization and phylogenetic analysis of the respirovirus.

Supplementary Material 6: Figure. S3: The phylogenetic analysis of parvovirus and Gemykibivirus.

Acknowledgements

We thank Prof. Shixing Yang and Prof. Wen Zhang for their guidance on this research.

Author contributions

Conceptualization: Y.J, Y.M.H., and D.Q.Z. Methodology: D.Q.Z., B.D., T.Z., X.J., Q.Y.L., and C.Y.Q. Investigation: D.Q.Z., Y.J., Y.M.H., and C.Y.Q. Funding acquisition: D.Q.Z, X.J, and Y.M.H. Writing—original draft: D.Q.Z, and Y.C. Writing—review and editing: D.Q.Z, Y.J, Y.M.H, and C.Y.Q.

Funding

This research was supported by National Natural Science Foundation of China No. 32201990; Natural Science Foundation of Jiangsu Province of China No.

BK20210461; Scientific Research Program of Wuxi Health Commission No. Q202342; Hospital-level project of Northern Jiangsu People's Hospital No. SBQN22013.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics statements

Sample collection and all experiments in the present were performed with Ethical Approval given by the Ethics Committee of The Affiliated Yixing People's Hospital.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Radiology, the Affiliated Yixing Hospital of Jiangsu University, Wuxi, China

²Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, USA

³Department of Pediatric, the Affiliated Yixing Hospital of Jiangsu University, Wuxi, China

⁴Medical research center, Northern Jiangsu People's Hospital, Yangzhou, China

⁵School of Medicine, Qilu Institute of Technology, Jinan, China

Received: 30 October 2024 / Accepted: 6 January 2025

Published online: 13 January 2025

References

- Global burden. Of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of Disease Study 2019. *Lancet*. 2020;396:1204–22.
- Mortality morbidity. Hospitalisations due to influenza lower respiratory tract infections, 2017: an analysis for the global burden of Disease Study 2017. *Lancet Respir Med*. 2019;7:69–89.
- Bryce J, Boschi-Pinto C, Shibuya K, Black RE. WHO estimates of the causes of death in children. *Lancet*. 2005;365:1147–52.
- Xu M, Liu P, Su L, Cao L, Zhong H, Lu L, Jia R, Xu J. Comparison of respiratory pathogens in Children with Lower Respiratory Tract infections before and during the COVID-19 pandemic in Shanghai, China. *Front Pediatr*. 2022;10:881224.
- Chen J, Zhang J, Lu Z, Chen Y, Huang S, Li H, Lin S, Yu J, Zeng X, Ji C, et al. *Mycoplasma pneumoniae* among Chinese outpatient children with mild respiratory tract infections during the Coronavirus Disease 2019 Pandemic. *Microbiol Spectr*. 2022;10:e0155021.
- Cillóniz C, Torres A, Niederman M, van der Eerden M, Chalmers J, Welte T, Blasi F. Community-acquired pneumonia related to intracellular pathogens. *Intensive Care Med*. 2016;42:1374–86.
- Kutty PK, Jain S, Taylor TH, Bramley AM, Diaz MH, Ampofo K, Arnold SR, Williams DJ, Edwards KM, McCullers JA, et al. *Mycoplasma pneumoniae* among children hospitalized with community-acquired Pneumonia. *Clin Infect Dis*. 2019;68:5–12.
- Waites KB, Xiao L, Liu Y, Balish MF, Atkinson TP. *Mycoplasma pneumoniae* from the respiratory tract and Beyond. *Clin Microbiol Rev*. 2017;30:747–809.
- Gao LW, Yin J, Hu YH, Liu XY, Feng XL, He JX, Liu J, Guo Y, Xu BP, Shen KL. The epidemiology of paediatric *Mycoplasma pneumoniae* pneumonia in North China: 2006 to 2016. *Epidemiol Infect*. 2019;147:e192.
- Chang HY, Jordan JL, Krause DC. Domain analysis of protein P30 in *Mycoplasma pneumoniae* cytoadherence and gliding motility. *J Bacteriol*. 2011;193:1726–33.
- Hu J, Ye Y, Chen X, Xiong L, Xie W, Liu P. Insight into the pathogenic mechanism of *Mycoplasma pneumoniae*. *Curr Microbiol*. 2022;80:14.
- Jiang Z, Li S, Zhu C, Zhou R, Leung PHM. *Mycoplasma pneumoniae* infections: Pathogenesis and Vaccine Development. *Pathogens*. 2021;10(2):119.
- He J, Liu M, Ye Z, Tan T, Liu X, You X, Zeng Y, Wu Y. Insights into the pathogenesis of *Mycoplasma pneumoniae* (review). *Mol Med Rep*. 2016;14:4030–6.
- Diaz MH, Cross KE, Benitez AJ, Hicks LA, Kutty P, Bramley AM, Chappell JD, Hymas W, Patel A, Qi C, et al. Identification of bacterial and viral codections with *Mycoplasma pneumoniae* using the TaqMan array card in patients hospitalized with community-acquired pneumonia. *Open Forum Infect Dis*. 2016;3:ofw071.
- Ledford JG, Mukherjee S, Kiskan MM, Nugent JL, Hollingsworth JW, Wright JR. Surfactant protein-A suppresses eosinophil-mediated killing of *Mycoplasma pneumoniae* in allergic lungs. *PLoS ONE*. 2012;7:e32436.
- Li F, Zhang Y, Shi P, Cao L, Su L, Fu P, Abuduxikuer K, Wang L, Wang Y, Lu R, et al. *Mycoplasma pneumoniae* and Adenovirus Coinfection cause Pediatric severe community-acquired pneumonia. *Microbiol Spectr*. 2022;10:e0002622.
- Gao J, Xu L, Xu B, Xie Z, Shen K. Human adenovirus Coinfection aggravates the severity of *Mycoplasma pneumoniae* pneumonia in children. *BMC Infect Dis*. 2020;20:420.
- Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG. Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit Care Med*. 2011;184:957–63.
- Wang Y, Yu X, Liu F, Tian X, Quan S, Jiao A, Yang X, Zeng X, Jiao W, Qi H, et al. Respiratory microbiota imbalance in children with *Mycoplasma pneumoniae* pneumonia. *Emerg Microbes Infect*. 2023;12:2202272.
- Zhou Q, Xie G, Liu Y, Wang H, Yang Y, Shen K, Dai W, Li S, Zheng Y. Different nasopharynx and oropharynx microbiota imbalance in children with *Mycoplasma pneumoniae* or influenza virus infection. *Microb Pathog*. 2020;144:104189.
- Dai W, Wang H, Zhou Q, Feng X, Lu Z, Li D, Yang Z, Liu Y, Li Y, Xie G, et al. The concordance between upper and lower respiratory microbiota in children with *Mycoplasma pneumoniae* pneumonia. *Emerg Microbes Infect*. 2018;7:92.
- Norman JM, Handley SA, Baldrige MT, Droit L, Liu CY, Keller BC, Kambal A, Monaco CL, Zhao G, Flesher P, et al. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell*. 2015;160:447–60.
- Mao Q, Sun G, Qian Y, Qian Y, Li W, Wang X, Shen Q, Yang S, Zhou C, Wang H, Zhang W. Viral metagenomics of pharyngeal secretions from children with acute respiratory diseases with unknown etiology revealed diverse viruses. *Virus Res*. 2022;321:198912.
- Zhou Y, Fernandez S, Yoon IK, Simasathien S, Watanaveeradej V, Yang Y, Marte-Salcedo OA, Shuck-Lee DJ, Thomas SJ, Hang J, Jarman RG. Metagenomics Study of Viral Pathogens in undiagnosed respiratory specimens and Identification of Human Enteroviruses at a Thailand Hospital. *Am J Trop Med Hyg*. 2016;95:663–9.
- Leung K, Wu JT, Liu D, Leung GM. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *Lancet*. 2020;395:1382–93.
- Xu M, Li Y, Shi Y, Liu H, Tong X, Ma L, Gao J, Du Q, Du H, Liu D, et al. Molecular epidemiology of *Mycoplasma pneumoniae* pneumonia in children, Wuhan, 2020–2022. *BMC Microbiol*. 2024;24:23.
- Zhang Z, Dou H, Tu P, Shi D, Wei R, Wan R, Jia C, Ning L, Wang D, Li J, et al. Serum cytokine profiling reveals different immune response patterns during general and severe *Mycoplasma pneumoniae* pneumonia. *Front Immunol*. 2022;13:1088725.
- Zhu Y, Luo Y, Li L, Jiang X, Du Y, Wang J, Li H, Gu H, Li D, Tang H, et al. Immune response plays a role in *Mycoplasma pneumoniae* pneumonia. *Front Immunol*. 2023;14:1189647.
- Wu BG, Segal LN. The Lung Microbiome and its role in Pneumonia. *Clin Chest Med*. 2018;39:677–89.
- Budden KF, Shukla SD, Rehman SF, Bowerman KL, Keely S, Hugenholtz P, Armstrong-James DPH, Adcock IM, Chotirmall SH, Chung KF, Hansbro PM. Functional effects of the microbiota in chronic respiratory disease. *Lancet Respir Med*. 2019;7:907–20.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience*. 2012;1:18.
- Li H, Wang H, Ju H, Lv J, Yang S, Zhang W, Lu H. Comparison of gut viral communities in children under 5 years old and newborns. *Viol J*. 2023;20:52.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–9.
- Zhang D, Wang Y, Chen X, He Y, Zhao M, Lu X, Lu J, Ji L, Shen Q, Wang X, et al. Diversity of viral communities in faecal samples of farmed red foxes. *Heliyon*. 2023;9:e12826.

35. Zhang D, Wang Y, He Y, Ji L, Zhao K, Yang S, Zhang W. Identification of avihepadnaviruses and circoviruses in an unexplained death event in farmed ducks. *Arch Virol*. 2023;168:85.
36. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing platforms. *Mol Biol Evol*. 2018;35:1547–9.
37. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012;61:539–42.
38. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics*. 2014;30:3123–4.
39. Brown BA, Maher K, Flemister MR, Naraghi-Arani P, Uddin M, Oberste MS, Pallansch MA. Resolving ambiguities in genetic typing of human enterovirus species C clinical isolates and identification of enterovirus 96, 99 and 102. *J Gen Virol*. 2009;90:1713–23.
40. Varsani A, Krupovic M. Family Genomoviridae: 2021 taxonomy update. *Arch Virol*. 2021;166:2911–26.
41. de Steenhuijsen Piters WA, Huijskens EG, Wyllie AL, Biesbroek G, van den Bergh MR, Veenhoven RH, Wang X, Trzciński K, Bonten MJ, Rossen JW, et al. Dysbiosis of upper respiratory tract microbiota in elderly pneumonia patients. *Isme j*. 2016;10:97–108.
42. Hasegawa K, Linnemann RW, Mansbach JM, Ajami NJ, Espinola JA, Petrosino JF, Piedra PA, Stevenson MD, Sullivan AF, Thompson AD, Camargo CA Jr. Nasal Airway Microbiota Profile and severe bronchiolitis in infants: a case-control study. *Pediatr Infect Dis J*. 2017;36:1044–51.
43. Narendrakumar L, Ray A. Respiratory tract microbiome and pneumonia. *Prog Mol Biol Transl Sci*. 2022;192:97–124.
44. Man WH, de Steenhuijsen Piters WA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol*. 2017;15:259–70.
45. Gollwitzer ES, Saglani S, Trompette A, Yadava K, Sherburn R, McCoy KD, Nicod LP, Lloyd CM, Marsland BJ. Lung microbiota promotes tolerance to allergens in neonates via PD-L1. *Nat Med*. 2014;20:642–7.
46. Wang Y, Zhu N, Li Y, Lu R, Wang H, Liu G, Zou X, Xie Z, Tan W. Metagenomic analysis of viral genetic diversity in respiratory samples from children with severe acute respiratory infection in China. *Clin Microbiol Infect*. 2016;22:e458451–459.
47. Taylor LJ, Keeler EL, Bushman FD, Collman RG. The enigmatic roles of Anelloviridae and Redondoviridae in humans. *Curr Opin Virol*. 2022;55:101248.
48. Young JC, Chehoud C, Bittinger K, Bailey A, Diamond JM, Cantu E, Haas AR, Abbas A, Frye L, Christie JD, et al. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. *Am J Transpl*. 2015;15:200–9.
49. Maggi F, Focosi D, Statzu M, Bianco G, Costa C, Macera L, Spezia PG, Medici C, Albert E, Navarro D, et al. Early post-transplant Torquetenovirus Viremia predicts Cytomegalovirus Reactivations in solid organ transplant recipients. *Sci Rep*. 2018;8:15490.
50. Görzer I, Jaksch P, Kundi M, Seitz T, Klepetko W, Puchhammer-Stöckl E. Pre-transplant plasma Torque Teno virus load and increase dynamics after lung transplantation. *PLoS ONE*. 2015;10:e0122975.
51. Blanco-Picazo P, Fernández-Orth D, Brown-Jaque M, Miró E, Espinal P, Rodríguez-Rubio L, Muniesa M, Navarro F. Unravelling the consequences of the bacteriophages in human samples. *Sci Rep*. 2020;10:6737.
52. Diard M, Bakkeren E, Cornuault JK, Moor K, Hausmann A, Sellin ME, Loverdo C, Aertsen A, Ackermann M, De Paepe M, et al. Inflammation boosts bacteriophage transfer between *Salmonella* spp. *Science*. 2017;355:1211–5.
53. Roach DR, Leung CY, Henry M, Morello E, Singh D, Di Santo JP, Weitz JS, Debarbieux L. Synergy between the host Immune System and bacteriophage is essential for successful phage therapy against an Acute Respiratory Pathogen. *Cell Host Microbe*. 2017;22:38–e4734.
54. Fonseca FG, Lanna MC, Campos MA, Kitajima EW, Peres JN, Golgher RR, Ferreira PC, Kroon EG. Morphological and molecular characterization of the poxvirus BeAn 58058. *Arch Virol*. 1998;143:1171–86.
55. Goolam Mahomed T, Peters RPH, Allam M, Ismail A, Mtshali S, Goolam Mahomed A, Ueckermann V, Kock MM, Ehlers MM. Lung microbiome of stable and exacerbated COPD patients in Tshwane, South Africa. *Sci Rep*. 2021;11:19758.
56. Oliveira GP, Rodrigues RAL, Lima MT, Drummond BP, Abrahão JS. Poxvirus host range genes and virus-host spectrum: a critical review. *Viruses*. 2017;9(11):331.
57. Haller SL, Peng C, McFadden G, Rothenburg S. Poxviruses and the evolution of host range and virulence. *Infect Genet Evol*. 2014;21:15–40.
58. Küry P, Nath A, Créange A, Dolei A, Marche P, Gold J, Giovannoni G, Hartung HP, Perron H. Human endogenous retroviruses in neurological diseases. *Trends Mol Med*. 2018;24:379–94.
59. Yu C, Lei X, Chen F, Mao S, Lv L, Liu H, Hu X, Wang R, Shen L, Zhang N, et al. ARID1A loss derepresses a group of human endogenous retrovirus-H loci to modulate BRD4-dependent transcription. *Nat Commun*. 2022;13:3501.
60. Jakobsson J, Vincendeau M. SnapShot: human endogenous retroviruses. *Cell*. 2022;185:400–e400401.
61. Meyer Sauter PM, Trück J, van Rossum AMC, Berger C. Circulating antibody-secreting cell response during *Mycoplasma pneumoniae* Childhood Pneumonia. *J Infect Dis*. 2020;222:136–47.
62. Jin X, Li J, Shao M, Lv X, Ji N, Zhu Y, Huang M, Yu F, Zhang C, Xie L, et al. Improving suspected pulmonary infection diagnosis by Bronchoalveolar Lavage Fluid Metagenomic Next-Generation sequencing: a Multicenter Retrospective Study. *Microbiol Spectr*. 2022;10:e0247321.

Publisher's note

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