The respiratory tract virome: unravelling the role of viral dark matter in respiratory health and disease

Martha Purcell [1,2], Jodie Ackland [1], Karl J. Staples [1,2], Anna Freeman[1,2] and Tom M.A. Wilkinson[1,2]

[1]Department of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK. [2]NIHR Southampton Biomedical Research Centre, University Hospital Southampton, Southampton, UK.

Corresponding author: Martha Purcell (martha.purcell@soton.ac.uk)

Shareable abstract (@ERSpublications)
Metagenomic sequencing reveals the respiratory virome as a diverse ecosystem. Beyond pathogens, commensal viruses and phages may shape immunity and microbial balance, offering potential as biomarkers and modulators of respiratory disease. https://bit.ly/45Agh5z

Cite this article as: Purcell M, Ackland J, Staples KJ, *et al*. The respiratory tract virome: unravelling the role of viral dark matter in respiratory health and disease. *Eur Respir Rev* 2025; 34: 240284 [DOI: 10.1183/16000617.0284-2024].

Received: 20 Dec 2024
Accepted: 23 July 2025

Abstract

The human respiratory tract virome is an underexplored component of the microbiome that includes eukaryotic viruses, bacteriophages and archaeal viruses. The respiratory virome represents a dynamic and heterogeneous ecosystem, shaped by host, environmental and microbial factors. Advances in metagenomic sequencing have expanded our understanding of virome composition, dynamics and potential roles in health and disease. Despite increasing interest, virome research remains fragmented and often secondary to bacteriome studies. Challenges in study design, genomic characterisation and interpretation limit consistent conclusions. This review summarises current knowledge of the respiratory virome in health and across acute and chronic respiratory diseases, including acute respiratory infection, asthma, COPD, cystic fibrosis and bronchiectasis. While each condition is distinct, they share features of airway inflammation and immune dysregulation where the virome may act as a modifier or marker. Across these syndromes, emerging evidence highlights the consistent detection of respiratory viruses including potential commensals, such as *Anelloviridae*, and the often-overlooked role of bacteriophages. We also discuss the concept of viral dark matter, where large proportions of sequence data remain unclassified, potentially representing novel viral taxa. Technical and conceptual challenges are evaluated, alongside recent methodological innovations such as meta-transcriptomics and viral enrichment protocols. We outline how standardised, multi-omic and longitudinal approaches are urgently needed to clarify the virome's functional role, interactions with immunity and microbial communities and its utility as a biomarker or therapeutic target.

Literature review methodology

Literature searches were conducted in PubMed to include articles available up until September 2024. Search terms included "metagenomic next generation sequencing", "viral metagenomics", "microbiome", "virome", "virobiota", "lung", "respiratory tract", "nasopharyngeal", "sputum", "bronchoalveolar lavage", "chronic obstructive pulmonary disease", "asthma", "cystic fibrosis", "pneumonia", "bronchiectasis", "respiratory tract infection" and "acute respiratory infection". Once identified, articles were reviewed to determine relevance and bibliographies were searched for additional studies.

Introduction

The human microbiome refers to the diverse community of microorganisms, including bacteria, fungi, archaea and viruses, which inhabit different anatomical sites of the body. Along with fungi (mycobiome) and bacteria (bacteriome), the viral component of this ecosystem is known as the virome, see glossary of terms, table 1. The virome includes eukaryotic viruses that infect human cells, viruses that infect bacterial, fungal or archaeal hosts, and integrated retroviral elements, as well as viruses with nonhuman hosts, such as plant viruses, typically acquired through environmental exposure or diet [1–3]. The human virome is established from birth, is temporally dynamic [4, 5] and exhibits heterogeneity in abundance, diversity and

composition between anatomical sites [6, 7]. The respiratory tract virome is impacted by environmental factors including diet, geography, age, birth mode and breastfeeding [8, 9], see figure 1 [8–48].

Despite rapid advancements in lung microbiome research, studies of the respiratory virome remain scarce, leaving gaps in our understanding of how viral populations interact with host immunity and microbial communities in respiratory health and disease. Study of the virome in other body systems, such as the gut, has preceded the respiratory tract. Alterations in virome composition or virome dysbiosis has been linked to multiple diseases such as diabetes, inflammatory bowel disease and malignancy [49–51].

Although the concept of "sterile lungs" has long been disproven [52], our understanding of the respiratory virome still lags behind that of the bacteriome. Traditional PCR technologies revolutionised our ability to detect specific, known respiratory viruses, yet have largely reinforced a narrow focus on pathogenic viruses such as influenza, respiratory syncytial virus (RSV) and coronaviruses [53, 54]. These viruses are clinically important, contributing to both acute respiratory infection (ARI) and viral exacerbations of chronic lung disease (CLD), but they represent only a fraction of the viral diversity present in the respiratory tract [8, 10, 55–61].

Metagenomic next-generation sequencing (mNGS) enables untargeted detection of the full spectrum of viral genetic material within a sample, revealing not only established pathogens but also commensal, transient or latent viruses present during health and disease [60–63]. This broader lens has uncovered novel and unexpected virome members, such as *Anelloviridae*, which are highly prevalent in healthy individuals but of uncertain clinical significance [64–67]. Distinguishing between commensal and pathogenic microorganisms is increasingly recognised as complex, given their ecological overlap and context-dependent roles in health and disease [68, 69]. Metagenomic approaches have also highlighted potential cross-kingdom interactions between viruses, other microbes and the host immune system, with growing interest in how such interactions may modulate disease risk and outcomes [1, 11, 70–74].

The severe acute respiratory syndrome coronavirus 2 pandemic catalysed global investment in viral surveillance and sequencing infrastructure [75–77], underscoring the importance of understanding not just

TABLE 1 Glossary of terms used in the text relating to the characterisation of the respiratory tract virome

Term	Definition
Respiratory microbiome	The community of all microorganisms, including bacteria, fungi, archaea and viruses, and their genetic material, found in the respiratory tract
Virome	The viral component of the microbiome The collection of all viruses and their viral genomes within a specific host or environment
Bacteriome	The collection of all bacterial species and their genomes present in a specific environment or host, forming a component of the overall microbiome
Dysbiosis	An imbalance in the composition and metabolic activities of the components of the microbiome at a specific body site, often associated with health or disease
Eukaryotic virus	A virus that infects a eukaryotic host cell in order to replicate
Prokaryotic virus	A virus that infects prokaryotic organisms, such as bacteria or archaea Those that infect bacteria are referred to as *bacteriophages*, while viruses that infect archaea are known as *archaeal viruses*
Metagenomics	An untargeted sequencing approach that captures the total genetic material within a sample, enabling analysis of the full microbial community, including bacteria, viruses, fungi and archaea
Meta-transcriptomics	A technique used to sequence the total RNA (rather than DNA) from a sample, allowing simultaneous analysis of actively expressed genes from viruses, bacteria and the host, and offering insight into functional activity and host–microbe interactions
Viral reads	Short nucleotide sequences obtained from high-throughput sequencing that originate from viral genomes and are used to detect and analyse viral presence in a sample
Viral contigs	Longer sequences of viral genetic material assembled from overlapping sequencing reads, used to reconstruct parts of viral genomes and identify specific viruses in metagenomic datasets
Bioinformatics	A field that combines biology, computer science and statistics to analyse and interpret biological data, particularly large datasets generated by high-throughput sequencing methods such as metagenomics

individual viruses but also viral interactions, co-infections and their impact on immune responses and disease severity [2, 73, 78–80]. However, much of the respiratory virome, especially in adults with CLD, remains uncharted and virome studies still trail behind those of the bacteriome in both quantity and depth.

Moving beyond a pathogen-centric model, there is growing interest in conceptualising the respiratory virome as a dynamic ecological system, shaped by host, microbial and environmental factors, with the potential to influence immune tone, disease susceptibility and therapeutic response [2, 29, 69, 81]. This review explores the potential of metagenomic approaches to characterise the virome, with a particular focus on the respiratory tract virome in health and in respiratory diseases including ARI, COPD, asthma, cystic fibrosis (CF) and bronchiectasis. We evaluate current insights, clinical applications and the technical and conceptual challenges inherent to virome research, while also highlighting emerging opportunities and future directions for discovery.

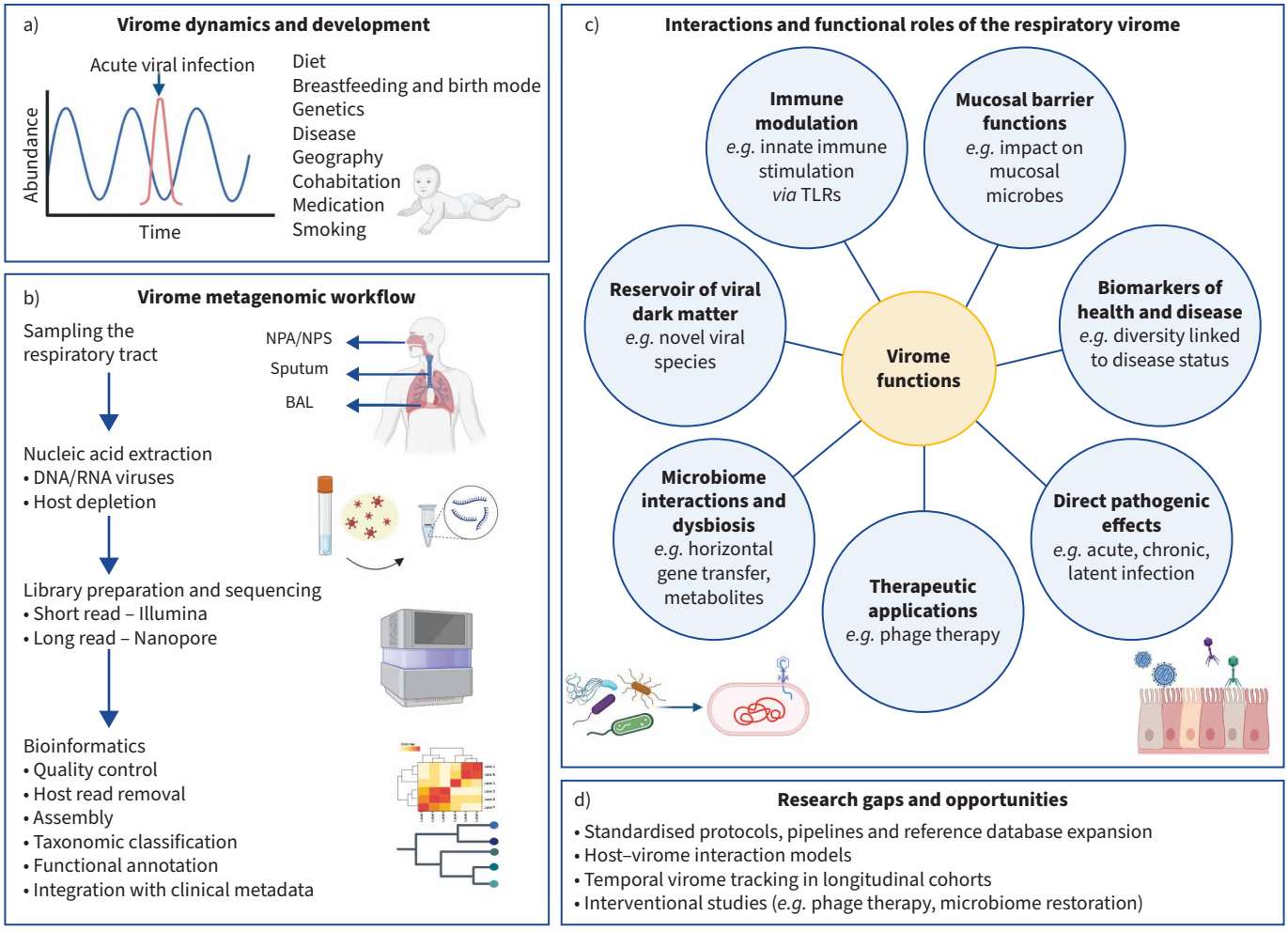

FIGURE 1 Conceptual framework of respiratory virome dynamics, analysis workflow, functions and future research directions. This figure outlines key components in understanding the respiratory virome. a) Virome dynamics and development: the virome forms from birth and is shaped by host, environmental and lifestyle factors. Viral abundance fluctuates over time, especially during acute infections. b) Virome metagenomic workflow: common steps in respiratory virome analysis, from sampling (*e.g.* nasopharyngeal aspirate (NPA)/nasopharyngeal swab (NPS), sputum and bronchoalveolar lavage (BAL)) and nucleic acid extraction to sequencing (short- and long-read) and bioinformatics, including taxonomic classification and integration with clinical data. c) Functional roles and interactions: the virome may influence health through immune modulation, disruption of mucosal barriers, direct pathogenicity and microbiome interactions (*e.g.* gene transfer). It may also serve as a reservoir of viral dark matter and provide biomarkers or therapeutic targets. d) Research gaps and opportunities: priorities include standardised methods, better reference databases, host–virome interaction models, longitudinal studies and interventional approaches, such as phage therapy [8–48]. TLR: Toll-like receptor. Figure created in BioRender (https://BioRender.com/jsxl2wc).

Viral dark matter and bacteriophages

Characterising the virome remains challenging as a substantial proportion of viral sequences recovered from metagenomic studies cannot be matched to any known reference genomes, a phenomenon termed "viral dark matter" [82–85]. Depending on the sample type, methodology and analysis pipeline used, this unclassified fraction can represent 40–90% of total viral reads [8, 82, 84–87].

Viral dark matter arises from several factors, including the vast genetic diversity of viruses, high mutation and recombination rates, recovery of short or fragmented sequences, and the limited representation of environmental and host-associated viruses in public databases [84, 85, 88]. Additional contributors include the lack of universal marker genes (*e.g.* 16S rRNA in bacteria), under-representation of RNA viruses due to technical challenges and poor assembly of low-abundance viral genomes [85, 89]. Contamination from host genetic material can further obscure detection, while some viruses integrate into host genomes as prophages, making them difficult to distinguish [89–92]. Many unclassified sequences may also contain genes of unknown function or auxiliary metabolic genes, which mimic host processes and are challenging to identify with standard databases [87]. Far from being a trivial artefact, viral dark matter likely reflects a biologically important reservoir of viral diversity, encompassing both novel eukaryotic viruses and previously uncharacterised bacteriophages [84–86, 90, 93, 94].

A large portion of viral dark matter is presumed to be composed of bacteriophages. These phages are not merely passive entities; they are genetically diverse and may play pivotal roles in shaping microbial ecology, modulating host immunity and mediating inter-microbial interactions [8, 36, 82, 83, 90, 95]. While direct evidence in the respiratory tract is still emerging, studies from other body sites support the hypothesis that phages can contribute to both microbial homeostasis and dysbiosis (table 2), underscoring their potential relevance in respiratory health and disease.

Challenges for metagenomic respiratory virome profiling

mNGS characterises all genetic material, including human, bacterial and viral, within a sample [86, 103]. Although this comprehensive and unbiased approach offers unprecedented insight into complex ecosystems, it also introduces significant technical and analytical challenges. Key decisions made during study design, such as sampling strategy, sample processing, sequencing platform selection and analytical pipeline, can influence a range of factors, including potential biases, data quality, reproducibility, and the accuracy of downstream interpretation, see table 3 [121, 123].

TABLE 2 Summary of proposed bacteriophage functions relevant to the respiratory tract virome

Regulatory mechanism	Description/effect	References
Regulation of bacterial populations	Phages drive bacterial turnover *via* lytic infection and may promote the expansion of lysogenised bacteria This shapes community structure and contributes to microbial homeostasis	[33–37]
Modulation of bacterial fitness and virulence	Phages influence bacterial gene expression and can carry fitness-enhancing or virulence-related genes, impacting both pathogenic and commensal strains	[30–32, 38]
Interaction with host immune system	Phages engage with host immunity through TLR signalling, cytokine induction (*e.g.* IL-1β, TNF-α) and generation of adaptive responses	[39–41, 43]
Barrier and mucosal defence	Some phages bind to mucins *via* immunoglobulin-like domains reducing bacterial colonisation at epithelial surfaces *in vitro*	[47, 48]
Influence on bacterial metabolism	Phages can alter bacterial metabolic pathways or carry metabolic genes, potentially affecting host–microbe metabolic interactions	[37, 96, 97]
Therapeutic and biomarker potential, including RNA phages	Due to their specificity and abundance, phages are being explored as therapeutics (*e.g.* for drug-resistant infections) and biomarkers of microbial dynamics	[42, 44–46, 98]
Role in disease tolerance and immune modulation	In some contexts, phage exposure correlates with protective immune-metabolic states and improved clinical outcomes	[99–102]

This table outlines the hypothesised and emerging roles of bacteriophages in modulating respiratory microbial ecology, host immunity and disease outcomes. While most evidence, to date, is derived from gut or environmental virome studies, these functions are increasingly recognised as relevant to respiratory health and disease. IL-1β: interleukin-1β; TNF-α: tumour necrosis factor-α; TLR: Toll-like receptor.

TABLE 3 Technical challenges facing virome studies across the workflow

Step	Challenge
Sampling	Low biomass limits detection and increases contamination sensitivity [104] Temporal and spatial virome heterogeneity [105] Inconsistent sampling protocols across studies or sites [106]
Nucleic acid extraction	Viral loads often below detection thresholds in clinical samples [107] Sample pre-treatment (*e.g.* centrifugation and filtration) to remove host or bacteria influence detection of larger viruses or intracellular viral particles [108] RNA instability leads to degradation of RNA viral genomes [109] Extraction method influences virus detection and genome coverage [110]
Library preparation	Biases introduced during amplification (*e.g.* MDA, SISPA and LASL) can distort community composition and reduce genome coverage [108] Reverse transcription efficiency varies, limiting consistent detection of RNA viruses [110]
Sequencing	Platform-specific biases Illumina-generated short reads make it difficult to assemble complex and large genomes, but generation of longer reads using ONT have higher error rates [111] Contamination through index hopping or sequencing controls [91, 112] Batch effects and run-to-run variability impact data comparability [113] Insufficient sequencing depth may limit detection of low-abundance or rare viruses but can also increase the potential of contaminant detection [114]
Data analysis	Choice of assembly tool and analysis pipeline can bias analytical outcome [115] Viral genomic material may comprise <1% of sequenced reads [24] Incomplete, sparse, contaminated or taxonomically biased viral reference databases limit accurate identification [116] Prophages or integrated viral elements can confound separation between bacterial and viral reads [117] High mutation rates, recombination and genome plasticity make alignment-based approaches less effective for novel virus detection [118] *De novo* assembly is challenged by low genome coverage, short contigs or complex genome structures [119]
Interpretation and validation	Contaminated reference databases can lead to poor annotation, misclassification and misleading biological inferences [120] Lack of standardised analysis pipelines and reporting guidelines complicates cross-study comparisons [106, 121, 122]

Virome studies are limited by low viral biomass, high host background and variability in sampling and processing protocols. Each step, from extraction to sequencing and analysis, can introduce biases, affect genome recovery and hinder virus detection, especially for RNA viruses. Sparse and contaminated reference databases, combined with high viral diversity, further complicate classification and interpretation. This table outlines key technical limitations across the virome workflow. LASL: linker amplification shotgun libraries; MDA: multiple displacement amplification; ONT: Oxford Nanopore Technologies; SISPAL: sequence-independent, single-primer amplification.

The virome specifically presents unique methodological challenges due to the diverse array of viral genome types (double-stranded DNA, single-stranded DNA, double-stranded RNA and single-stranded RNA) and structures (linear, circular or segmented). Due to the anatomical and microbial complexity of the respiratory tract, the sampling strategy and site selection can profoundly affect the quality, interpretability and contextualisation of respiratory virome data. These challenges are heightened by the fact that despite considerable efforts by researchers, there remains a lack of standardised approaches, not only in virome-specific studies [121, 122], but also in respiratory microbiome research [106].

Sampling the respiratory tract – where, when and how?

Compounding the technical and analytical challenges surrounding mNGS is the complex and dynamic nature of the respiratory tract, which is composed of distinct anatomical compartments. It is divided into the upper respiratory tract (URT) (nasopharynx and oropharynx) and lower respiratory tract (LRT) (trachea, lower airways and lungs), which are continuously exposed to microorganisms from the environment [124]. Colonisation of the LRT occurs *via* micro-aspiration and mucosal migration, balanced through eradication by host immunity and the muco-ciliary escalator [125, 126]. As a result, the lung microbiome is transient, shaped by dynamic interactions between microbial immigration, elimination and growth. Growth and persistence of microbes in the LRT also depend on local environmental conditions, including temperature, oxygen availability, pH and nutrient levels, as well as host factors such as cough reflex and airway architecture [127, 128]. Microbial community composition varies across anatomical regions, with both site-specific and overlapping populations, which adds further complexity to the selection of sampling method, timing and site [129–131]. Whether similar temporal and spatial patterns exist for viral populations is less clear.

Sampling the distinct regions of the airway can be challenging. Differences in viral tropism may impact which viruses are detected throughout the respiratory tract depending on sampling location [132]. Even within a single anatomical site, distinct ecological niches exist and the choice of sampling technique can influence the microbial diversity captured [105]. Noninvasive methods, such as nasopharyngeal swabs (NPS) or washes and sputum induction, are commonly used in clinical settings, but may not accurately represent the LRT in disease. Whereas invasive sampling, including bronchoalveolar lavage (BAL) and protected specimen brushes, provide more direct access to the lower airways, this sampling is resource-intensive and invasive. Given the continuity of the airway, nasopharyngeal (NP) sampling has been considered a proxy for the LRT and many respiratory virome studies to date have been conducted using this sample type [129, 133].

Contamination challenges in metagenomic studies

Respiratory tract samples generally yield a low microbial biomass in comparison to other body sites [134]. Community abundance and diversity of microbial populations diminish further distally in the respiratory tract resulting in lower yield despite invasive sampling [134]. These factors can amplify the risk of contamination, presenting a major obstacle to generating robust and reproducible virome data. Contaminants can lead to the misidentification of novel viruses. For example, studies have reported cases where what were initially thought to be novel viral discoveries were later traced back to contamination from extraction kits [135, 136].

Contamination of samples can occur across the course of a metagenomic study, see table 4. During sampling, lower airway specimens may be contaminated by upper airway microbes as instruments traverse the higher-biomass URT [129]. During processing, contaminants may originate from DNA extraction or enrichment kits (known as the "kitome"), reagents or consumables, handling by technicians or machines, or exposure to ambient air [137–139]. These contaminants can be addressed analytically, typically through *in silico* removal using negative controls, *e.g.* bronchoscope washes [140]. However, cross-contamination between wells and tubes can result in contamination of negative controls (termed the "splashome"), which may result in removal of genuine low-abundance microbes that may be key ecological features of the sampled niche [137]. Thus, studies need to not only consider carefully designed controls, but also experimental design throughout the processing and sample handling stage to reduce contamination. To address and mitigate the impact of contamination, Eisenhofer et al. [141] proposed the RIDE checklist, a minimum standards guidance specifically for low microbial biomass studies.

Innovative approaches for studying the virome

Unlike bacteria, viruses lack universally conserved genetic signatures, such as 16S rRNA, used to target organisms for ease of molecular analysis, and hence virome characterisation has lagged [10, 63]. PCR assays enable detection of a targeted panel of pathogenic DNA and RNA RVs, but this approach relies on prior knowledge. The advent of metagenomics allows for both RNA and DNA viruses within a single sample to be amplified, capturing a more accurate reflection of viral content and reducing potential enrichment biases [63, 84]. This is achieved without requiring prior knowledge of the organisms and represents a tool in novel virus discovery, genome assembly and functional gene analysis [86].

TABLE 4 Potential sources of contamination in metagenomic studies

Contamination source	Examples
Anatomical	Contamination introduced by sampling tools passing through high-biomass URT regions (*e.g.* bronchoscopy) Inherent contamination in sample types such as sputum that contain both upper and lower airway material Gastro-oesophageal reflux or microaspiration
Background	Nucleic acid extraction kits, enrichment kits, lab reagents (water, master mixes) and consumables (plasticware)
Handling	Automated lab equipment, human user (researchers, technicians, *etc.*) and air
Cross-contamination	Sample-to-sample contamination (*e.g.* wells/tubes), tag switching and index hopping
Analytical	Poor quality genomes, alignment artefacts and contaminants present in reference databases

Contaminants can arise from sampling, reagents, handling, cross-contamination or analytical artefacts, and must be carefully controlled to ensure accurate virome characterisation. URT: upper respiratory tract.

Recent technological and analytical innovations in the metagenomics field have significantly advanced the field of metagenomic virome profiling, see table 5 [27, 81, 92, 106, 111, 120, 122, 142–147, 149–172]. A recent article has comprehensively outlined the technical and analytical tools available for metagenomic studies [173]. Long-read sequencing technologies now enable the assembly of complete viral genomes and the resolution of within-host diversity, while targeted enrichment strategies improve the detection of low-abundance and diverse viruses [111, 157–160, 142–147]. These approaches enhance the clinical utility of mNGS by enabling the identification of clinically relevant genome variations and viral subtypes from a single sample, demonstrating the role of mNGS as an invaluable tool in viral surveillance, diagnostics and development of therapeutics [111, 157–160, 142–147]. Beyond viral detection and characterising community composition, resolving host–virus interactions is critical to understanding viral pathogenesis. Spatial transcriptomics allows localisation of viral presence within tissues, while single cell viromics can directly link viruses to specific host cells and detect intra-host diversity, such as quasispecies [149–156].

As technology has advanced, data analysis tools have evolved to keep pace. Integrative multi-omic approaches now enable exploration of the host–virus interactome, while ongoing improvements in viral assembly algorithms and machine learning applications have enhanced the detection and annotation of novel viruses, overcoming the need for homologous matching to known sequences in current reference databases [81, 120, 164–166, 161–163]. Furthermore, publicly available viral sequencing data from previous studies is more freely available and provides an opportunity to retrospectively perform virome analysis, offering a cost-effective means to extract new insights from existing data [167–169, 174]. Together, these approaches can provide deeper insights into interkingdom interactions, enhancing our understanding of the role of the virome in health and disease.

However, despite these advances, the field still lacks standardised protocols and reporting guidelines for metagenomic virome studies. Establishing consensus frameworks which cover the entire metagenomic workflow including sampling, sequencing, analysis and metadata reporting, is essential for ensuring robust and reproducible insights across diverse study settings. Although there has been significant effort within the research community to develop standardised guidelines, they are not consistently implemented in practice [92, 122, 141, 170]. This lack of standardisation not only hinders the reproducibility and interpretability of virome studies, but also limits their clinical applicability, impeding the integration of virome data into diagnostic workflows, therapeutic strategies and efforts to improve patient outcomes.

Despite these complexities and limitations, increased availability of mNGS, improved bioinformatic pipelines and expansion of viral genome reference libraries have resulted in the characterisation of the virome across a variety of sample types, including blood, BAL, sputum and faeces [7, 25, 63, 175], offering critical insights into its potential role in both health and disease.

The respiratory tract virome in health

Much of our current understanding of the healthy respiratory virome derives from clinically well or asymptomatic individuals recruited to studies of acute and CLD [12, 17, 25–27, 60, 174, 176–180]. In these healthy populations, the virome includes eukaryotic respiratory viruses, non-respiratory eukaryotic viruses and bacteriophages. This resident viral community is thought to impact host immunity and may modulate susceptibility to disease [60].

A recent meta-analysis of publicly available metagenomic sequencing datasets demonstrated greater virome diversity in healthy controls than in individuals with respiratory disease, suggesting the existence of disease-associated shifts in virome composition [174]. Several studies also support the existence of a core respiratory virome, comprising stable populations of both eukaryotic viruses and bacteriophages, [177]. Among eukaryotic viruses, *Anelloviridae* are frequently detected in healthy airways and have been proposed as possible commensals, though they are also linked to immunosuppression and disease states [176, 178, 179].

Pathogenic viral families, such as *Picornaviridae* (human rhinovirus (HRV)), *Adenoviridae*, *Coronaviridae*, *Orthomyxoviridae* (influenza), *Paramyxoviridae* (RSV) and *Parvoviridae* (human bocavirus), have all been detected in the nasopharynx of healthy children [17, 25–27]. DNA viruses, including *Papillomaviridae*, *Polyomaviridae*, *Herpesviridae* and *Adenoviridae*, are also common in the nasal virome of healthy adults and children [12, 176]. Their presence may reflect transient exposure or asymptomatic carriage, with clearance mediated by the innate immune system.

Bacteriophages form a substantial proportion of the healthy virome, often dominating the sequenced viral reads. Commonly detected phage families include, *Myoviridae*, *Podoviridae* and *Siphoviridae* [26, 27, 60,

TABLE 5 Overview of innovative approaches and associated challenges driving progress in the field of viral metagenomics

Approach	What is this approach?	Challenges	Impact on the field	Tools	References
Technology					
Targeted enrichment	Uses bait-capture or hybridisation to selectively enrich a targeted panel of viral nucleic acids before sequencing	Requires prior knowledge of target viruses May miss highly divergent or novel sequences	Overcomes the limitation of a low proportion of viral reads in complex samples Enhances detection of rare or low-abundance viruses Reduces background contamination Allows for increased genome coverage	ViroCap, VirCapSeq-VERT, Twist Bioscience Respiratory Virus Research Panel, Illumina Respiratory Virus Oligos	[142–148]
Single-cell viromics	Captures genomic material from individual host cells along with any associated virus	Currently limited mostly to *in vitro* samples or RNA viruses Technically challenging and low throughput	Direct linking of viruses to specific host cells Uncovers intracellular infection dynamics Resolves micro-diversity (*e.g.* quasispecies) and uncovers rare or novel viral entities	kallisto, bustools, PalmDB viscRNA-Seq DART-seq, STRS, Viral-Track	[149–154]
Spatial mapping	Adapts spatial transcriptomics to localise viruses within tissues or environments	Low sensitivity Limited resolution in complex tissues and expensive	Enables spatial mapping of virus–host interactions Adds ecological and pathophysiological context	10x Genomics (Visium, Xenium)	[155, 156]
Long-read sequencing	Sequencing technologies that read longer fragments of DNA/RNA	Higher error rates than short-read methods	Better resolution of variants and quasispecies Portable platforms can facilitate rapid viral identification and real-time surveillance during outbreaks Can detect RNA modifications	Oxford Nanopore Technologies, PacBio	[109, 152, 157–160]
Analytical					
Machine learning	Uses machine learning models to predict viral sequences, functions and host interactions from sequence data	Dependent on quality of training data Bias from contaminated or under-represented genomes	Bypasses limitations of current reference databases Can annotate novel viruses	DeepVirFinder, MArVD2	[161, 162]
Interactome	Integrates metagenomics with transcriptomics, proteomics and metabolomics	Complex to integrate and interpret multi-layered datasets	Combination of multi-omic data can improve functional analysis to better understand viral pathogenesis and reveal targets for novel therapeutic development		[27, 81, 163]
Database curation	Development and refinement of curated metagenome and metadata databases	Lag in database updates, incomplete metadata requiring lengthy curation, contamination present in reference databases	Enhance annotation and identification of viral sequences, enable global comparative viromics	Integrated Microbial Genomes, Genomes OnLine Database, MG-RAST	[120, 164–166]
Reanalysis of existing data	Applying new computational approaches to previously generated data	Original data may be low-quality, contaminated or poorly annotated	Identifies of new virus types Finds novel disease–virus associations Cost-effective, but may be hampered by data quality (*e.g.* contaminated sequence data, incomplete metadata)		[167–169]
Standardisation and development of best practice guidelines	Development of consensus protocols and reporting frameworks to encompass the entire metagenomic workflow	Inconsistent practices across studies Lack of standardised guidance implementation	Improves reproducibility and comparability across studies Enhances reliability of viral discovery and quantification Enables data integration and meta-analysis Lays groundwork for clinical translation and regulatory acceptance	MIxS, MIUViG, MISAG, MIMAG, RIDE	[92, 106, 120, 170–172]

Emerging technological and analytical approaches in virome research. This table summarises recent developments aimed at improving viral detection, characterisation and interpretation in metagenomic studies. Each approach is described alongside its challenges, potential impact, examples of some key tools, where applicable, and representative references.

177, 180]. While the precise roles of these viruses in respiratory health remain unclear, their presence suggests an influence on bacterial community structure and host–microbe interactions, as described in table 2.

These findings suggest that the healthy respiratory virome is diverse and includes both potentially pathogenic and commensal viruses. However, further work is needed to define the functional roles of these viral communities and their contributions to respiratory homeostasis.

Anelloviridae – *respiratory tract commensals?*

Anelloviruses are small, diverse, single-stranded DNA viruses first identified in 1997, with no confirmed pathogenicity [181]. The *Anelloviridae* family includes torque teno virus (TTV) and they are frequently detected in respiratory samples, sometimes representing the most abundant eukaryotic DNA viruses [14, 15, 21, 182, 183]. Longitudinal sampling in children has shown individuals carry a unique and persistent "anellome" in the URT [182]. Anelloviruses are commonly detected in children with and without ARI or pneumonia, but no causal link has been demonstrated [14, 15, 21, 24, 26]. TTV abundance in nasal and plasma samples has been associated with airflow obstruction and disease severity in children with asthma and bronchiectasis [184, 185]. In COPD, TTV is prevalent in sputum during exacerbations [186] and increased TTV load has been linked to severe COVID-19 outcomes [187].

Emerging evidence suggests *Anelloviridae* may modulate host immunity and contribute to chronic disease trajectories [65, 66, 188]. TTV has been shown to stimulate inflammatory cytokine production *in vitro* [189] and circulating TTV DNA is increasingly explored as a biomarker of immune competence [179, 190]. *Redondoviridae*, another family of novel, circular DNA viruses, have also been detected at high abundance in critical illness, correlating with changes in bacterial communities and disease severity [183, 187]. Further temporal metagenomic studies in health and respiratory disease will determine the role these novel DNA viruses play in lung health and pathogenesis of respiratory diseases.

The virome in respiratory tract disease

This section explores current evidence from metagenomic studies of the respiratory tract virome across four clinical contexts, namely ARI, COPD, asthma and CF/bronchiectasis. While these conditions differ in aetiology and clinical trajectory, they are unified by shared features of airway inflammation, immune modulation and recurrent infection, settings in which the virome may play a modifying or contributory role. A visual summary of the key virome features identified across each condition is provided in figure 2 [12, 13, 17, 19–21, 23, 25–28, 60, 65, 81, 131, 176–178, 180, 191–200]. These findings highlight the potential role of the virome in shaping lung health and here we discuss these studies in more detail.

A list of included studies is provided in supplementary tables S1 [12–26, 182, 192, 193, 195, 201–208] and S2 [28, 65, 81, 131, 177, 180, 191, 194, 196–200, 209–212]. Several additional studies have demonstrated the value of mNGS for pathogen detection in the clinical management of respiratory infections, particularly in severe pneumonia [213–220]. While these highlight important clinical applications, they are primarily focused on diagnostics rather than ecological characterisation of the virome and are therefore not discussed in detail here.

The virome in ARI

ARI encompasses infections of the URT and LRT caused by a range of pathogens, including viruses, bacteria and fungi. Across the studies reviewed, ARI was variably defined, spanning diagnostic clinical syndromes such as rhinitis, pharyngitis, tonsillitis, bronchitis, bronchiolitis, lower respiratory tract infection (LRTI) and pneumonia [13, 15, 21, 24, 206, 207], as well as public health case definitions such as influenza-like illness and severe acute respiratory infection (SARI) [17, 19, 26], see supplementary table S1. To add complexity, some studies focused on infections of unknown aetiology, while others included samples with detected bacterial or viral pathogens using standard tests [15, 18, 20, 195, 207]. This heterogeneity in clinical definitions, pathogen focus and illness severity complicates direct comparisons across studies and limits the ability to draw consistent conclusions about the respiratory virome in ARI (supplementary table S1).

The studies also differed in methodological approaches, including the use of various sequencing platforms, enrichment protocols and bioinformatic pipelines. Respiratory tract samples included NPS, NP aspirates, sputum and BAL, each representing distinct anatomical compartments with different microbial and viral communities (supplementary table S1). These technical and anatomical differences likely contribute to the diversity in virome profiles reported across studies and constrain the generalisability of individual findings.

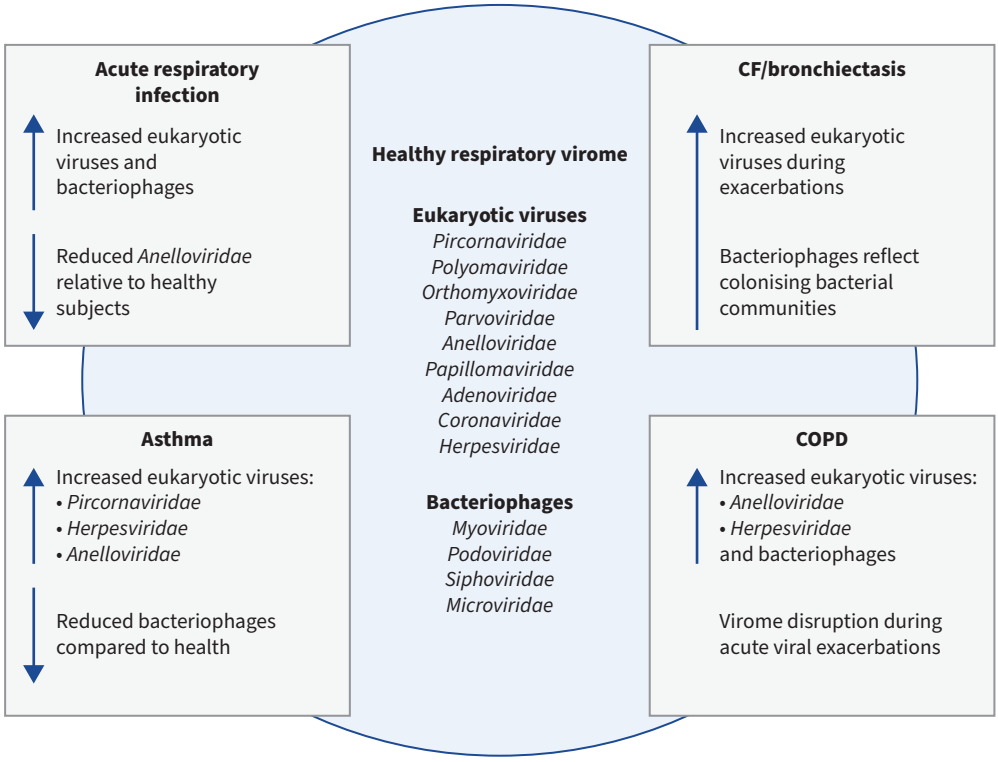

FIGURE 2 Respiratory virome changes in health and disease. The healthy respiratory virome includes diverse eukaryotic viruses and bacteriophages. In acute infections, both increase while *Anelloviridae* decrease. Asthma shows increased *Picornaviridae*, *Herpesviridae* and *Anelloviridae*, with fewer bacteriophages. In COPD and cystic fibrosis (CF)/bronchiectasis, exacerbations are linked to increased eukaryotic viruses and virome disruption, with phage communities reflecting bacterial colonisation [12, 13, 17, 19–21, 23, 25–28, 60, 65, 81, 131, 176–178, 180, 191–200]. Figure created in BioRender (https://BioRender.com/l88z285).

Eukaryotic viruses

Metagenomic studies in ARI consistently identify a high diversity and richness of eukaryotic viral families in both URT and LRT samples, particularly among patients with severe infection syndromes such as pneumonia and SARI. Frequently detected families include *Paramyxoviridae* (*e.g.* RSV and human metapneumovirus), *Picornaviridae* (*e.g.* rhinoviruses), *Orthomyxoviridae* (influenza), *Adenoviridae*, *Anelloviridae*, *Polyomaviridae*, *Parvoviridae*, *Coronaviridae*, *Herpesviridae*, *Astroviridae* and *Papillomaviridae* [15, 17–19, 23, 25, 26, 195, 207, 221, 222]. These viral families have been observed in both NP and LRT samples (*e.g.* sputum and BAL), with some studies showing correlations between viral abundance and disease severity [13, 207, 222]. The severity of pneumonia in immunosuppressed paediatric intensive care unit subjects was positively correlated with abundance of human mastadenovirus B, TTV (*Anelloviridae* family) in BAL [207]. A recent study by Cui *et al.* [13] demonstrated that children with pneumonia exhibit distinct respiratory virome profiles compared to healthy controls, with specific viral taxa, such as human adenoviruses and polyomaviruses, being more prevalent in severe cases.

Across multiple studies, the respiratory tract virome appears more diverse and abundant in pathogenic eukaryotic viruses compared to that of asymptomatic or healthy controls [17, 25, 26]. In contrast to subjects with SARI, healthy controls, had significantly reduced abundance and diversity of pathogenic viruses. In controls, the virome was instead dominated by the possibly commensal *Anelloviridae* family, particularly TTVs [26]. This pattern persists in children with recurrent respiratory tract infections (RTIs), where greater pathogenic viral richness and diversity was observed in those with multiple RTI episodes compared to single infections, but with *Anelloviridae* abundance remaining stable between groups [21, 182]. These findings suggest that acute viral infection may disrupt the virome, introducing transient pathogenic viruses, while a more stable "core virome" may exist across health and disease.

Metagenomic studies of the virome in COVID-19 further support trends in other ARI cohorts. Despite a marked reduction in circulating seasonal RVs during the pandemic, a diverse range of eukaryotic viral

families, including *Picornaviridae*, *Adenoviridae*, *Herpesviridae*, *Papillomaviridae*, *Orthomyxoviridae* and *Anelloviridae*, remained detectable in COVID-19-positive samples [20, 204, 205, 208], but no consistent associations between virome composition and disease severity could be established. However, *Anelloviridae* were found in greater abundance in hospitalised and deceased patients compared to ambulatory cases, suggesting a potential link between host immune status and *Anelloviridae* dynamics, consistent with their proposed role as markers of immune modulation [205]. Complementing these findings, Merenstein *et al.* [187] observed that *Anelloviridae* and *Redondoviridae* were more frequently detected and present at higher titres in patients with severe COVID-19, suggesting a potential link between these commensal viruses and disease severity.

Several studies have also detected low-abundance eukaryotic viruses in airway samples using metagenomic approaches, including bocavirus, picobirnavirus, measles virus, polyomaviruses and retroviruses, with uncertain clinical significance [17, 20, 23, 24, 204]. In parallel, metagenomic approaches have demonstrated utility in novel virus discovery and zoonotic surveillance, underscoring their potential for emerging pathogen detection [16, 17, 23, 24, 202].

Collectively, current studies underscore the complexity and variability of the respiratory virome across ARI syndromes. The consistent detection of diverse eukaryotic viruses across different patient groups and clinical presentations highlights the dynamic nature of the virome and its potential relevance to disease processes. Despite emerging associations with disease severity, causal links and mechanistic insights into the virome's role remain underexplored. The detection of viral genetic material in asymptomatic individuals also raises important questions about persistent colonisation and potential viral commensalism within the respiratory tract. Emerging tools, such as meta-transcriptomics, which enable simultaneous profiling of viral, bacterial and host gene expression, offer a promising route to explore these interactions in greater depth [27, 206]. However, the field remains constrained by fragmented study designs, inconsistent integration of nonviral microbiota data and a lack of longitudinal approaches. Moving forward, more standardised, multi-omic and temporally resolved studies will be essential to define both the structure and function of the respiratory virome in ARI, clarify its role in symptom onset and progression, and understand its broader interplay with microbial ecology and host immune responses.

Bacteriophages

Bacteriophages have been identified in multiple metagenomic studies of the respiratory tract virome during ARI [13, 15, 21, 24]. The variety, diversity and abundance of bacteriophages in respiratory tract samples is thought to mirror the composition of the lung bacterial flora [24, 26]. Families of phage identified in respiratory samples include, *Microviridae*, *Myoviridae*, *Podoviridae* and *Siphoviridae* [21, 193, 208]. The corresponding bacterial hosts for these phages in the respiratory tract include *Enterobacteria*, *Pseudomonas*, *Klebsiella*, *Burkholderia*, *Streptococcus*, *Cronobacter*, *Actinomyces*, *Escherichia*, *Haemophilus*, *Staphylococcus*, *Yersinia* and *Lactobacillus* [12, 15, 21, 24, 26].

Bacteriophages may represent biomarkers of microbiome dysbiosis and disease severity. A study of recurrent RTI in children demonstrated that abundance of *Propionibacterium* phage was elevated in recurrent RTI compared to those with single RTI over a 6-year period, which may indicate ongoing microbial dysbiosis [21]. Increased serum levels of factors associated with tissue inflammation and airway remodelling, such as tissue inhibitor metalloproteinase 1 and platelet-derived growth factor subunits BB, were also associated with recurrent RTI. In combination with *Propionibacterium* phage, these factors predicted recurrent RTI in the paediatric study population [21]. In COVID-19, phage abundances in the nasopharynx were higher in subjects with severe COVID-19 relative to those with mild disease [193]. Bacteriophage profiles also varied depending on the age of subjects, highlighting the dynamism of the virome and impact of environmental factors. The airway "phageome" may hold value as a biomarker of disease severity and should be included in virome research alongside eukaryotic viruses [193].

Given the diverse roles of phages in other niches (table 2), it is likely they also influence respiratory health. Barr *et al.* [47] showed that phages bind airway mucins and reduce bacterial colonisation, supporting a potential barrier-protective role in the lung. Conversely, some phages may negatively modulate host immunity; for example, filamentous Pf phages produced by *Pseudomonas aeruginosa* have been shown in murine and *in vitro* models to suppress antibacterial responses *via* Toll-like receptor-3-mediated type I interferon signalling, impairing phagocytosis and promoting chronic infection [43]. These contrasting roles highlight the need to explore phage function, not just composition, in ARI. Building on extensive research in gut phage therapy, emerging studies indicate that bacteriophage applications in respiratory infections hold significant promise, both as direct antimicrobial agents and as modulators of host immunity, paving the way for innovative therapeutic strategies [45, 98, 102].

The virome in CF and bronchiectasis

RVIs are associated with reduced pulmonary function and increased illness severity during CF viral exacerbations compared to RVI in non-CF individuals [223]. The CF airway represents a distinct ecological niche, requiring microbial adaptation to persistent inflammation, thick mucus and antibiotic exposure [210, 223]. Bacteriophages shape the microbial ecosystem by lysing colonising bacterial hosts and facilitating horizontal gene transfer, including antibiotic resistance genes, thereby contributing to microbial persistence and adaptation in the inflamed airway environment [33, 38, 42, 209, 210].

Virome profiling in CF has primarily used sputum samples, where bacterial sequences dominate, and oropharyngeal contamination may confound interpretation [199]. In a study comparing CF and non-CF individuals, non-CF subjects exhibited a relatively conserved core eukaryotic virome, while CF patients showed highly variable profiles, with viruses such as *Herpesviridae* and *Retroviridae* dominating in some cases [177]. A similar divergence was observed in bacteriophage communities, CF viromes contained phages associated with colonising pathogens such as *Pseudomonas*, *Burkholderia*, *Enterobacteria* and *Streptococcus*, whereas non-CF samples contained more stable phage populations [177, 191, 199]. Further studies have identified eukaryotic DNA viruses in CF sputum, including *Anelloviridae*, *Adenoviridae* and *Papillomaviridae*, although the exclusion of RNA viruses limits interpretability [131, 199]. Geographically distinct regions of the lung may host different viral communities, eukaryotic viruses dominate upper lobes, while bacteriophages predominate in the lower lobes [131]. Functional metagenomic analyses have revealed that CF airway viromes exhibit distinct metabolic profiles, reflecting adaptations to the unique conditions of the CF lung environment and could indicate the influence of phages on metabolism of bacterial populations [177].

In bronchiectasis, RVs are detected year-round, including in clinically stable patients [81, 224]. DNA viruses are enriched in BAL during exacerbations with *Herpesviridae* a dominant family across disease states and *Anelloviridae* more common in stable disease [200]. Sputum-based virome studies have shown similar bacteriophage profiles to CF, with >70% of viral sequences mapping to *Siphoviridae*, *Caudovirales*, *Myoviridae* and *Phycodnaviridae* families [81]. These communities appear geographically conserved and may contribute to microbial network dynamics during exacerbations [81]. Specific phage signatures have also been linked to exacerbation frequency, supporting a potential role for bacteriophages as modulators of clinical outcomes in bronchiectasis [81]. The high burden of bacterial colonisation in bronchiectasis suggests that bacteriophages play a central role in shaping airway microbial ecology and contributing to exacerbation dynamics. As exacerbations may be viral, bacterial, or both, phages warrant investigation not only as biomarkers but also as therapeutic agents [45, 209, 225]. However, case reports have shown that adaptive immune responses, including the development of neutralising antibodies, can emerge during phage therapy and potentially reduce its effectiveness in some individuals [226]. A deeper understanding of host–phage immune interactions will be key to advancing virome-informed strategies in CLD.

The virome in asthma

Eukaryotic viruses are well-established triggers of asthma exacerbations, particularly HRV and RSV, both in early life and adulthood [53, 55–60, 227, 228]. These viruses are also linked to asthma development and severity. Early-life airway microbiome disruption, including virus–bacteria interactions involving *Haemophilus* and *Moraxella*, may further contribute to asthma risk and progression, though causal mechanisms remain unclear [55, 227–229]. In line with observations from ARI studies, the virome during asthma exacerbation is dominated by eukaryotic respiratory viruses from six families, namely *Picornaviridae*, *Parvoviridae*, *Paramyxoviridae*, *Anelloviridae*, *Orthomyxoviridae* and *Herpesviridae* [28].

Recent metagenomic studies have profiled the virome during stable asthma using NP and sputum samples [65, 180]. In adults, reduced bacteriophage abundance and increased presence of *Herpesviridae* (including cytomegalovirus and Epstein–Barr virus) in sputum were associated with asthma severity and exacerbation risk [180]. Similarly, children with stable asthma displayed virome profiles distinct from healthy controls, characterised by a lower diversity and abundance of bacteriophages, and a higher prevalence of *Anelloviridae* and HRVs [65]. Clustering of virome profiles correlated with clinical severity, suggesting that virome composition may reflect underlying disease endotypes. Reduced phage presence could contribute to impaired microbial regulation and airway dysbiosis, although this may also reflect concurrent loss of microbial diversity [65].

Emerging work has begun exploring how virome dynamics relate to host immune responses in asthma. Antiviral cytokine responses from peripheral blood mononuclear cells have been shown to correlate with virome profiles in children with stable asthma, suggesting bidirectional interactions between host immunity and virome composition [211]. Whether these virome shifts are a cause or consequence of immune

dysregulation remains uncertain. Ongoing studies aim to leverage virome–bacteriome interactions to restore balanced lung microbial communities [5, 65, 230].

The virome in COPD

In addition to bacteria, RVs are a common cause of acute exacerbations of COPD (AECOPD), contributing to healthcare burden, chronic inflammation and deteriorating lung function [231, 232]. Recent interest has turned toward the role of viral persistence in stable COPD and how virome–microbe interactions may influence disease severity and progression [56, 233]. A multi-omic and ecological approach is increasingly recognised as essential to understand COPD pathogenesis and identify potential therapeutic avenues [234].

The virome in healthy smokers, a potential model for early COPD, is dominated by bacteriophages, particularly *Propionibacterium* phages, alongside co-detection of *Herpesviridae*, *Adenoviridae* and *Papillomaviridae* [235]. Smoking was associated with reduced viral diversity and enrichment of Proteobacteria-targeting phages, alongside distinct metabolomic shifts, suggesting that virome alterations may reflect smoke-induced inflammation and metabolic change [235]. In established COPD, the LRT virome (BAL and sputum) contains abundant bacteriophage communities (*Siphoviridae*, *Podoviridae*, *Myoviridae*) and DNA viruses (*Anelloviridae*, *Retroviridae*, *Herpesviridae*), with diversity declining alongside COPD severity [194, 196].

Gut virome profiling in COPD has also demonstrated reduced viral diversity and altered phage composition, with changes linked to disease severity and markers of systemic inflammation [236]. These findings complement respiratory virome studies and support the hypothesis that virome shifts across mucosal sites may contribute to COPD progression, potentially through immune modulation or effects on bacterial metabolism. Notably, bacterial metabolites have been linked to clinical outcomes in COPD [237], raising the possibility that phages, through their impact on bacterial metabolism, may contribute to host–microbe interactions relevant to disease progression.

Metagenomic studies of NPS from COPD patients hospitalised with LRTI and acute exacerbation have identified a range of eukaryotic viruses, such as *Picornaviridae*, *Orthomyxoviridae*, *Coronaviridae*, *Herpesviridae*, *Paramyxoviridae*, *Papillomaviridae*, *Adenoviridae* and *Anelloviridae*, as well as bacteriophages [197, 198]. During AECOPD, patients with PCR-confirmed viral infections showed reduced overall viral diversity and a contraction of the phage community, while the bacterial microbiota remained relatively stable [197]. This contrasts with the more diverse phages and DNA virus-rich viromes observed in stable COPD, where diversity declined with disease severity [194, 196]. While these findings highlight a potential role for virome disruption in virus-associated exacerbations, results should be interpreted with caution given small sample sizes and methodological variability across studies.

Conclusions

The respiratory virome remains an under-characterised component of the airway microbiome, with emerging evidence suggesting it plays important roles in both health and disease. While studies consistently detect diverse eukaryotic viruses and bacteriophages in airway samples, major gaps remain in our understanding of their function, host interactions and temporal dynamics. The concept of viral dark matter underscores the vast unclassified fraction of the virome and the limitations of current databases. Bacteriophages may shape microbial ecology, immune responses and host metabolism, yet remain poorly studied in the respiratory context. The potential for phages to serve as biomarkers or therapeutic agents, especially in settings of chronic colonisation and recurrent exacerbation, warrants closer investigation. However, findings to date are limited by small sample sizes, variation in methods and inconsistent reporting, making standardisation of sampling, sequencing and analytical pipelines essential to enable comparison between studies.

Emerging technologies such as targeted enrichment, single-cell viromics, long-read sequencing and spatial mapping, alongside machine learning and multi-omic integration, offer promising avenues to deepen our understanding of virome function and dynamics. Going forward, multi-omic approaches, including meta-transcriptomics and host response profiling, offer opportunities to uncover mechanistic insights into the virome's role in disease. Importantly, these efforts should be embedded within a broader ecological framework that integrates the virome with bacterial, fungal, and host factors to fully capture the complexity of respiratory health and disease. A shift from pathogen-focused detection to ecological characterisation of the virome could open new avenues for diagnostics, risk stratification and precision therapeutics in respiratory medicine.

Questions for future research

- How can metagenomic approaches be standardised and optimised to comprehensively characterise the respiratory virome, including both DNA and RNA viruses, and integrate it with broader microbiome analyses, across diverse populations and clinical settings?
- What roles do commensal viruses and bacteriophages play in shaping host immunity and influencing the pathogenesis of respiratory infection and CLDs?
- Can longitudinal, multi-omic studies establish causal links between virome dynamics and the onset, progression or exacerbation of respiratory diseases?
- How can systems biology and functional genomics be applied to model virome–host–microbiome interactions and uncover novel therapeutic targets?

Provenance: Submitted article, peer reviewed.

Author contributions: M. Purcell and J. Ackland performed the literature search, conceptualised and drafted the manuscript. K.J Staples, A. Freeman and T.M.A. Wilkinson provided critical revisions and expertise on the review content. All authors reviewed and approved the final version of the manuscript.

Conflict of interest: M. Purcell, J. Ackland and A. Freeman have nothing to disclose. K.J. Staples reports research grants from AstraZeneca and Epiendo, and payment or honoraria for lectures, presentations, manuscript writing or educational events from AstraZeneca. T.M.A. Wilkinson reports support for the present study from the National Institute for Health and Care Research, Medical Research Council, BerGenBio, AstraZeneca, UCB and Janssen, consultancy fees from AstraZeneca, Valneva, Olam Pharma, Janssen, My mHealth and Synairgen, payment or honoraria for lectures, presentations, manuscript writing or educational events from AstraZeneca, Boehringer Ingelheim and Roche, participation on a data safety monitoring board or advisory board with Valneva, and stock (or stock options) with My mHealth.

References

1 Duerkopf BA, Hooper LV. Resident viruses and their interactions with the immune system. *Nat Immunol* 2013; 14: 654–659.
2 Virgin HW. The virome in mammalian physiology and disease. *Cell* 2014; 157: 142–150.
3 Cao Z, Sugimura N, Burgermeister E, *et al*. The gut virome: a new microbiome component in health and disease. *EBioMedicine* 2022; 81: 104113.
4 Shkoporov AN, Clooney AG, Sutton TDS, *et al*. The human gut virome is highly diverse, stable, and individual specific. *Cell Host Microbe* 2019; 26: 527–541.
5 Mageiros L, Megremis S, Papadopoulos NG. The virome in allergy and asthma: a nascent, ineffable player. *J Allergy Clin Immunol* 2023; 152: 1347–1351.
6 Abeles SR, Robles-Sikisaka R, Ly M, *et al*. Human oral viruses are personal, persistent and gender-consistent. *ISME J* 2014; 8: 1753–1767.
7 Reyes A, Haynes M, Hanson N, *et al*. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 2010; 466: 334–338.
8 Liang G, Bushman FD. The human virome: assembly, composition and host interactions. *Nat Rev Microbiol* 2021; 19: 514–527.
9 Bosch AATM, Levin E, van Houten MA, *et al*. Development of upper respiratory tract microbiota in infancy is affected by mode of delivery. *EBioMedicine* 2016; 9: 336–345.
10 Porto BN. Insights into the role of the lung virome during respiratory viral infections. *Front Immunol* 2022; 13: 885341.
11 Flynn M, Lyall Z, Shepherd G, *et al*. Interactions of the bacteriome, virome, and immune system in the nose. *FEMS Microbes* 2022; 3: xtac020.
12 Sandybayev N, Beloussov V, Strochkov V, *et al*. Metagenomic profiling of nasopharyngeal samples from adults with acute respiratory infection. *R Soc Open Sci* 2024; 11: 240108.
13 Cui S, Guo R, Chen C, *et al*. Next-generation sequencing for characterizing respiratory tract virome and improving detection of viral pathogens in children with pneumonia. *Influenza Other Respir Viruses* 2024; 18: 13362.
14 Bal A, Destras G, Sabatier M, *et al*. Metagenomic analysis reveals high abundance of torque teno mini virus in the respiratory tract of children with acute respiratory illness. *Viruses* 2022; 14: 955.
15 Guo Q, Li L, Wang C, *et al*. Comprehensive virome analysis of the viral spectrum in paediatric patients diagnosed with *Mycoplasma pneumoniae* pneumonia. *Virol J* 2020; 19: 181.

16 Kha Tu NT, Thu Hong NT, Han Ny NT, *et al*. The virome of acute respiratory diseases in individuals at risk of zoonotic infections. *Viruses* 2020; 12: 960.
17 Ogunbayo AE, Mogotsi MT, Sondlane H, *et al*. Metagenomic analysis of respiratory RNA virome of children with and without severe acute respiratory infection from the Free State, South Africa during COVID-19 pandemic reveals higher diversity and abundance in summer compared with winter period. *Viruses* 2022; 14: 2516.
18 Liu Y, Wang H, Yang J, *et al*. Virome of respiratory secretion from children with unknown etiological acute respiratory disease revealed recombinant human parechovirus and other significant viruses. *Virol J* 2021; 18: 122.
19 Zhou Y, Fernandez S, Yoon IK, *et al*. Metagenomics study of viral pathogens in undiagnosed respiratory specimens and identification of human enteroviruses at a Thailand hospital. *Am J Trop Med Hygiene* 2016; 95: 663–669.
20 Ferravante C, Sanna G, Melone V, *et al*. Nasopharyngeal virome analysis of COVID-19 patients during three different waves in Campania region of Italy. *J Med Virol* 2022; 94: 2275–2283.
21 Li Y, Fu X, Ma J, *et al*. Altered respiratory virome and serum cytokine profile associated with recurrent respiratory tract infections in children. *Nat Commun* 2019; 10: 2288.
22 Yang J, Yang F, Ren L, *et al*. Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach. *J Clin Microbiol* 2011; 49: 3463–3469.
23 Lysholm F, Wetterbom A, Lindau C, *et al*. Characterization of the viral microbiome in patients with severe lower respiratory tract infections, using metagenomic sequencing. *PLoS One* 2012; 7: e30875.
24 Xu L, Zhu Y, Ren L, *et al*. Characterization of the nasopharyngeal viral microbiome from children with community-acquired pneumonia but negative for Luminex xTAG respiratory viral panel assay detection. *J Med Virol* 2017; 89: 2098–2107.
25 Wylie KM, Mihindukulasuriya KA, Sodergren E, *et al*. Sequence analysis of the human virome in febrile and afebrile children. *PLoS One* 2012; 7: e27735.
26 Wang Y, Zhu N, Li Y, *et al*. Metagenomic analysis of viral genetic diversity in respiratory samples from children with severe acute respiratory infection in China. *Clin Microbiol Infect* 2016; 22: 458.
27 Rajagopala SV, Bakhoum NG, Pakala SB, *et al*. Metatranscriptomics to characterize respiratory virome, microbiome, and host response directly from clinical samples. *Cell Rep Methods* 2021; 1: 100091.
28 Romero-Espinoza JA, Moreno-Valencia Y, Coronel-Tellez RH, *et al*. Virome and bacteriome characterization of children with pneumonia and asthma in Mexico City during winter seasons 2014 and 2015. *PLoS One* 2018; 13: e0192878.
29 Wallace MA, Wille M, Geoghegan J, *et al*. Making sense of the virome in light of evolution and ecology. *Proc Biol Sci* 2025; 292: 20250389.
30 Taylor VL, Fitzpatrick AD, Islam Z, *et al*. The diverse impacts of phage morons on bacterial fitness and virulence. *Adv Virus Res* 2019; 103: 1–31.
31 Kurilovich E, Geva-Zatorsky N. Effects of bacteriophages on gut microbiome functionality. *Gut Microbes* 2025; 17: 2481178.
32 Brüssow H, Canchaya C, Hardt WD. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* 2004; 68: 560.
33 Howard-Varona C, Hargreaves KR, Abedon ST, *et al*. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J* 2017; 11: 1511–1520.
34 Fortier LC, Sekulovic O. Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence* 2013; 4: 354–365.
35 Marchi J, Zborowsky S, Debarbieux L, *et al*. The dynamic interplay of bacteriophage, bacteria and the mammalian host during phage therapy. *iScience* 2023; 26: 106004.
36 Shkoporov AN, Hill C. Bacteriophages of the human gut: the “known unknown” of the microbiome. *Cell Host Microbe* 2019; 25: 195–209.
37 Naureen Z, Dautaj A, Anpilogov K, *et al*. Bacteriophages presence in nature and their role in the natural selection of bacterial populations. *Acta Biomed* 2020; 91: Suppl. 13, e2020024.
38 Borodovich T, Shkoporov AN, Ross RP, *et al*. Phage-mediated horizontal gene transfer and its implications for the human gut microbiome. *Gastroenterol Rep* 2022; 10: goac012.
39 Van Belleghem JD, Dąbrowska K, Vaneechoutte M, *et al*. Interactions between bacteriophage, bacteria, and the mammalian immune system. *Viruses* 2019; 11: 10.
40 Chatterjee A, Duerkop BA. Beyond bacteria: bacteriophage–eukaryotic host interactions reveal emerging paradigms of health and disease. *Front Microbiol* 2018; 9: 1394.
41 Kan L, Barr JJ. A mammalian cell’s guide on how to process a bacteriophage. *Annu Rev Virol* 2023; 10: 183–198.
42 Federici S, Nobs SP, Elinav E. Phages and their potential to modulate the microbiome and immunity. *Cell Mol Immunol* 2021; 18: 889–904.

43 Sweere JM, Van Belleghem JD, Ishak H, *et al*. Bacteriophage trigger antiviral immunity and prevent clearance of bacterial infection. *Science* 2019; 363: eaat9691.
44 Subramanian A. Emerging roles of bacteriophage-based therapeutics in combating antibiotic resistance. *Front Microbiol* 2024; 15: 1384164.
45 Khosravi A, Chen Q, Echterhof A, *et al*. Phage therapy for respiratory infections: opportunities and challenges. *Lung* 2024; 202: 223–232.
46 Kirk D, Costeira R, Visconti A, *et al*. Bacteriophages, gut bacteria, and microbial pathways interplay in cardiometabolic health. *Cell Rep* 2024; 43: 113728.
47 Barr JJ, Auro R, Furlan M, *et al*. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc Natl Acad Sci USA* 2013; 110: 10771–10776.
48 Tzani-Tzanopoulou P, Skliros D, Megremis S, *et al*. Interactions of bacteriophages and bacteria at the airway mucosa: new insights into the pathophysiology of asthma. *Front Allergy* 2020; 1: 617240.
49 Zhao G, Vatanen T, Droit L, *et al*. Intestinal virome changes precede autoimmunity in type I diabetes-susceptible children. *Proc Natl Acad Sci USA* 2017; 114: E6166–E6175.
50 Hannigan GD, Duhaime MB, Ruffin MT, *et al*. Diagnostic potential and interactive dynamics of the colorectal cancer virome. *mBio* 2018; 9: e02248-18.
51 Norman JM, Handley SA, Baldridge MT, *et al*. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 2015; 160: 447–460.
52 Dickson RP, Erb-Downward JR, Martinez FJ, *et al*. The microbiome and the respiratory tract. *Annu Rev Physiol* 2016; 78: 481–504.
53 Hewitt R, Farne H, Ritchie A, *et al*. The role of viral infections in exacerbations of chronic obstructive pulmonary disease and asthma. *Ther Adv Respir Dis* 2016; 10: 158–174.
54 Singanayagam A, Joshi PV, Mallia P, *et al*. Viruses exacerbating chronic pulmonary disease: the role of immune modulation. *BMC Med* 2012; 10: 27.
55 Ptaschinski C, Lukacs NW. Early life respiratory syncytial virus infection and asthmatic responses. *Immunol Allergy Clin North Am* 2019; 39: 309–319.
56 Linden D, Guo-Parke H, Coyle PV, *et al*. Respiratory viral infection: a potential “missing link” in the pathogenesis of COPD. *Eur Respir Rev* 2019; 28: 180063.
57 Mallia P, Footitt J, Sotero R, *et al*. Rhinovirus infection induces degradation of antimicrobial peptides and secondary bacterial infection in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012; 186: 1117–1124.
58 Molyneaux PL, Mallia P, Cox MJ, *et al*. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013; 188: 1224–1231.
59 Gern JE. Rhinovirus and the initiation of asthma. *Curr Opin Allergy Clin Immunol* 2009; 9: 73–78.
60 Jankauskaitė L, Misevičienė V, Vaidelienė L, *et al*. Lower airway virology in health and disease–from invaders to symbionts. *Medicina (Lithuania)* 2018; 54: 72.
61 Wylie KM. The virome of the human respiratory tract. *Clin Chest Med* 2017; 38: 11–19.
62 Mitchell AB, Glanville AR. Introduction to techniques and methodologies for characterizing the human respiratory virome. *Methods Mol Biol* 2018; 1838: 111–123.
63 Sandybayev N, Beloussov V, Strochkov V, *et al*. Next generation sequencing approaches to characterize the respiratory tract virome. *Microgorganisms* 2022; 10: 2327.
64 Jaksch P, Kundi M, Görzer I, *et al*. Torque teno virus as a novel biomarker targeting the efficacy of immunosuppression after lung transplantation. *J Infect Dis* 2018; 218: 1922–1928.
65 Megremis S, Constantinides B, Xepapadaki P, *et al*. Respiratory eukaryotic virome expansion and bacteriophage deficiency characterize childhood asthma. *Sci Rep* 2023; 13: 8319.
66 Dodi G, Attanasi M, Di Filippo P, *et al*. Virome in the lungs: the role of anelloviruses in childhood respiratory diseases. *Microorganisms* 2021; 23: 1357.
67 Maggi F, Pifferi M, Tempestini E, *et al*. Correlation between torque tenovirus infection and serum levels of eosinophil cationic protein in children hospitalized for acute respiratory diseases. *J Infect Dis* 2004; 190: 971–974.
68 Burrows K, Ngai L, Chiaranunt P, *et al*. A gut commensal protozoan determines respiratory disease outcomes by shaping pulmonary immunity. *Cell* 2024; 188: 316–330.
69 Maier L, Stein-Thoeringer C, Ley RE, *et al*. Integrating research on bacterial pathogens and commensals to fight infections–an ecological perspective. *Lancet Microbe* 2024; 5: 100843.
70 Neu U, Mainou BA. Virus interactions with bacteria: partners in the infectious dance. *PLoS One* 2020; 16: e1008234.
71 Brealey JC, Sly PD, Young PR, *et al*. Viral bacterial co-infection of the respiratory tract during early childhood. *FEMS Microbiol Lett* 2015; 362: fnv062.
72 Hanada S, Pirzadeh M, Carver KY, *et al*. Respiratory viral infection-induced microbiome alterations and secondary bacterial pneumonia. *Front Immunol* 2018; 9: 2640.
73 Cadwell K. The virome in host health and disease. *Immunity* 2015; 42: 805–813.

74 Kikkert M. Innate immune evasion by human respiratory RNA viruses. *J Innate Immun* 2020; 12: 4–20.
75 Zhou P, Yang X Lou, Wang XG, *et al*. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020; 579: 270–273.
76 Wu F, Zhao S, Yu B, *et al*. A new coronavirus associated with human respiratory disease in China. *Nature* 2020; 579: 265–269.
77 Zhu N, Zhang D, Wang W, *et al*. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020; 382: 727–733.
78 Wu H, Zhou HY, Zheng H, *et al*. Towards understanding and identification of human viral co-infections. *Viruses* 2024; 16: 673.
79 Carvalho T, Krammer F, Iwasaki A. The first 12 months of COVID-19: a timeline of immunological insights. *Nat Rev Immunol* 2021; 21: 245–256.
80 Houldcroft CJ, Beale MA, Breuer J. Clinical and biological insights from viral genome sequencing. *Nat Rev Microbiol* 2017; 15: 183–192.
81 Mac Aogáin M, Kumar Narayana J, Yee Tiew P, *et al*. Integrative microbiomics in bronchiectasis exacerbations. *Nat Med* 2021; 27: 688–699.
82 Reyes A, Semenkovich NP, Whiteson K, *et al*. Going viral: next generation sequencing applied to human gut phage populations. *Nat Rev Microbiol* 2012; 10: 607.
83 Pedulla ML, Ford ME, Houtz JM, *et al*. Origins of highly mosaic mycobacteriophage genomes. *Cell* 2003; 113: 171–182.
84 Santiago-Rodriguez TM, Hollister EB. Unraveling the viral dark matter through viral metagenomics. *Front Immunol* 2022; 13: 1005107.
85 Krishnamurthy SR, Wang D. Origins and challenges of viral dark matter. *Virus Res* 2017; 239: 136–142.
86 Roux S, Matthijnssens J, Dutilh BE. Metagenomics in virology. *Encyclopedia Virol* 2021; 1: 133–140.
87 Brum JR, Sullivan MB. Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol* 2015; 13: 147–159.
88 Roux S, Hallam SJ, Woyke T, *et al*. Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. *eLife* 2015; 4: e08490.
89 Garmaeva S, Sinha T, Kurilshikov A, *et al*. Studying the gut virome in the metagenomic era: challenges and perspectives. *BMC Biol* 2019; 17: 84.
90 Aggarwala V, Liang G, Bushman FD. Viral communities of the human gut: metagenomic analysis of composition and dynamics. *Mob DNA* 2017; 8: 12.
91 Jurasz H, Pawłowski T, Perlejewski K. Contamination issue in viral metagenomics: problems, solutions, and clinical perspectives. *Front Microbiol* 2021; 12: 745076.
92 Conceição-Neto N, Zeller M, Lefrère H, *et al*. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 2015; 5: 16532.
93 Mokili JL, Rohwer F, Dutilh BE. Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2012; 2: 63–77.
94 Dutilh BE, Cassman N, McNair K, *et al*. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun* 2014; 5: 4498.
95 Breitbart M, Salamon P, Andresen B, *et al*. Genomic analysis of uncultured marine viral communities. *Proc Natl Acad Sci USA* 2002; 99: 14250–14255.
96 Tiamani K, Luo S, Schulz S, *et al*. The role of virome in the gastrointestinal tract and beyond. *FEMS Microbiol Rev* 2022; 46: fuac027.
97 Thompson LR, Zeng Q, Kelly L, *et al*. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci USA* 2011; 108: E757–E764.
98 Li L, Zhong Q, Zhao Y, *et al*. First-in-human application of double-stranded RNA bacteriophage in the treatment of pulmonary *Pseudomonas aeruginosa* infection. *Microb Biotechnol* 2023; 16: 862.
99 Gogokhia L, Buhrke K, Bell R, *et al*. Expansion of bacteriophages is linked to aggravated intestinal inflammation and colitis. *Cell Host Microbe* 2019; 25: 285–299.
100 Fluckiger A, Daillère R, Sassi M, *et al*. Cross-reactivity between tumor MHC class I–restricted antigens and an enterococcal bacteriophage. *Science* (1979) 2020; 369: 936–942.
101 Champagne-Jorgensen K, Luong T, Darby T, *et al*. Immunogenicity of bacteriophages. *Trends Microbiol* 2023; 31: 1058–1071.
102 de Souza EB, Pinto AR, Fongaro G. Bacteriophages as potential clinical immune modulators. *Microorganisms* 2023; 11: 2222.
103 van Boheemen S, van Rijn AL, Pappas N, *et al*. Retrospective validation of a metagenomic sequencing protocol for combined detection of RNA and DNA viruses using respiratory samples from pediatric patients. *J Mol Diagn* 2020; 22: 196–207.
104 Austin GI, Korem T. Planning and analyzing a low-biomass microbiome study: a data analysis perspective. *J Infect Dis* 2024; 22: jiae378.

105 Pérez-Losada M, Crandall KA, Freishtat RJ. Two sampling methods yield distinct microbial signatures in the nasopharynges of asthmatic children. *Microbiome* 2016; 4: 25.
106 Carney SM, Clemente JC, Cox MJ, *et al*. Methods in lung microbiome research. *Am J Respir Cell Mol Biol* 2020; 62: 283–299.
107 Kohl C, Brinkmann A, Dabrowski PW, *et al*. Protocol for metagenomic virus detection in clinical specimens. *Emerg Infect Dis* 2015; 21: 48–57.
108 Parras-Moltó M, Rodríguez-Galet A, Suárez-Rodríguez P, *et al*. Evaluation of bias induced by viral enrichment and random amplification protocols in metagenomic surveys of saliva DNA viruses. *Microbiome* 2018; 6: 199.
109 Ginocchio CC, Wang XP, Kaplan MH, *et al*. Effects of specimen collection, processing, and storage conditions on stability of human immunodeficiency virus type 1 RNA levels in plasma. *J Clin Microbiol* 1997; 35: 2886–2893.
110 Li L, Deng X, Mee ET, *et al*. Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. *J Virol Methods* 2015; 213: 139–146.
111 Cook R, Telatin A, Hsieh SY, *et al*. Nanopore and Illumina sequencing reveal different viral populations from human gut samples. *Microb Genom* 2024; 10: 001236.
112 Mukherjee S, Huntemann M, Ivanova N, *et al*. Large-scale contamination of microbial isolate genomes by Illumina Phix control. *Stand Genomic Sci* 2015; 10: 18.
113 Leek JT, Scharpf RB, Bravo HC, *et al*. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat Rev Genet* 2010; 11: 733–739.
114 Buddle S, Forrest L, Akinsuyi N, *et al*. Evaluating metagenomics and targeted approaches for diagnosis and surveillance of viruses. *Genome Med* 2024; 16: 111.
115 Sutton TDS, Clooney AG, Ryan FJ, *et al*. Choice of assembly software has a critical impact on virome characterisation. *Microbiome* 2019; 7: 12.
116 Chorlton SD. Ten common issues with reference sequence databases and how to mitigate them. *Front Bioinform* 2024; 4: 1278228.
117 Glickman C, Hendrix J, Strong M. Simulation study and comparative evaluation of viral contiguous sequence identification tools. *BMC Bioinformatics* 2021; 22: 329.
118 Sanjuán R, Nebot MR, Chirico N, *et al*. Viral mutation rates. *J Virol* 2010; 84: 9733–9748.
119 Quince C, Walker AW, Simpson JT, *et al*. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol* 2017; 35: 833–844.
120 Steinegger M, Salzberg SL. Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biol* 2020; 21: 155.
121 Chang WS, Harvey E, Mahar JE, *et al*. Improving the reporting of metagenomic virome-scale data. *Commun Biol* 2024; 7: 1687.
122 Roux S, Adriaenssens EM, Dutilh BE, *et al*. Minimum information about an uncultivated virus genome (MIUVIG). *Nat Biotechnol* 2019; 37: 29–37.
123 Charon J, Olendraite I, Forgia M, *et al*. Consensus statement from the first RdRp Summit: advancing RNA virus discovery at scale across communities. *Front Virol* 2024; 4: 1371958.
124 de Steenhuijsen Piters WAA, Binkowska J, Bogaert D. Early life microbiota and respiratory tract infections. *Cell Host Microbe* 2020; 28: 223–232.
125 Bassis CM, Erb-Downward JR, Dickson RP, *et al*. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *mBio* 2015; 6: e00037.
126 Dickson RP, Erb-Downward JR, Freeman CM, *et al*. Spatial variation in the healthy human lung microbiome and the adapted island model of lung biogeography. *Ann Am Thorac Soc* 2015; 12: 821–830.
127 Man WH, De Steenhuijsen Piters WAA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol* 2017; 15: 259–270.
128 Dickson R, Erb-Downward J, Huffnagle G. Towards an ecology of the lung: new conceptual models of pulmonary microbiology and pneumonia pathogenesis. *Lancet Respir Med* 2014; 2: 238–246.
129 Goddard AF, Staudinger BJ, Dowd SE, *et al*. Direct sampling of cystic fibrosis lungs indicates that DNA-based analyses of upper-airway specimens can misrepresent lung microbiota. *Proc Natl Acad Sci USA* 2012; 109: 13769–13774.
130 Dickson R, Erb-Downward J, Freeman C, *et al*. Bacterial topography of the healthy human lower respiratory tract. *mBio* 2017; 8: e02287-16.
131 Willner D, Haynes MR, Furlan M, *et al*. Case studies of the spatial heterogeneity of DNA viruses in the cystic fibrosis lung. *Am J Respir Cell Mol Biol* 2012; 46: 127–131.
132 Rijsbergen LC, van Dijk LLA, Engel MFM, *et al*. *In vitro* modelling of respiratory virus infections in human airway epithelial cells – a systematic review. *Front Immunol* 2021; 12: 683002.
133 Man WH, van Houten MA, Mérelle ME, *et al*. Bacterial and viral respiratory tract microbiota and host characteristics in children with lower respiratory tract infections: a matched case–control study. *Lancet Respir Med* 2019; 7: 417–426.
134 Natalini JG, Singh S, Segal LN. The dynamic lung microbiome in health and disease. *Nat Rev Microbiol* 2023; 21: 222–235.

135 Smuts H, Kew M, Khan A, *et al*. Novel hybrid parvovirus-like virus, NIH-CQV/PHV, contaminants in silica column-based nucleic acid extraction kits. *J Virol* 2014; 88: 1398–1398.

136 Naccache SN, Greninger AL, Lee D, *et al*. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. *J Virol* 2013; 87: 11966–11977.

137 Minich JJ, Sanders JG, Amir A, *et al*. Quantifying and understanding well-to-well contamination in microbiome research. *mSystems* 2019; 4: 186–205.

138 Salter SJ, Cox MJ, Turek EM, *et al*. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014; 12: 87.

139 Asplund M, Kjartansdóttir KR, Mollerup S, *et al*. Contaminating viral sequences in high-throughput sequencing viromics: a linkage study of 700 sequencing libraries. *Clin Microbiol Infect* 2019; 25: 1277–1285.

140 Davis NM, Proctor DiM, Holmes SP, *et al*. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018; 6: 226.

141 Eisenhofer R, Minich JJ, Marotz C, *et al*. Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol* 2019; 27: 105–117.

142 Xiao M, Xiao M, Liu X, *et al*. Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. *Genome Med* 2020; 12: 57.

143 Wen S, Sun C, Zheng H, *et al*. High-coverage SARS-CoV-2 genome sequences acquired by target capture sequencing. *J Med Virol* 2020; 92: 2221–2226.

144 Koskela von Sydow A, Lindqvist CM, Asghar N, *et al*. Comparison of SARS-CoV-2 whole genome sequencing using tiled amplicon enrichment and bait hybridization. *Sci* 2023; 13: 6461.

145 Wylie TN, Wylie KM, Herter BN, *et al*. Enhanced virome sequencing using targeted sequence capture. *Genome Res* 2015; 25: 1910–1920.

146 Mourik K, Sidorov I, Carbo EC, *et al*. Comparison of the performance of two targeted metagenomic virus capture probe-based methods using reference control materials and clinical samples. *J Clin Microbiol* 2024; 62: e0034524.

147 Briese T, Kapoor A, Mishra N, *et al*. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio* 2015; 6: e01491-15.

148 Ceballos-Garzon A, Comtet-Marre S, Peyret P. Applying targeted gene hybridization capture to viruses with a focus to SARS-CoV-2. *Virus Res* 2024; 340: 199293.

149 Saikia M, Burnham P, Keshavjee SH, *et al*. Simultaneous multiplexed amplicon sequencing and transcriptome profiling in single cells. *Nat Methods* 2019; 16: 59–62.

150 Luebbert L, Sullivan DK, Carilli M, *et al.* Detection of viral sequences at single-cell resolution identifies novel viruses associated with host gene expression changes. *Nat Biotechnol* 2025; in press [http://doi.org/10.1038/s41587-025-02614-y].

151 Russell AB, Trapnell C, Bloom JD. Extreme heterogeneity of influenza virus infection in single cells. *eLife* 2018; 7: e32303.

152 Russell AB, Elshina E, Kowalsky JR, *et al*. Single-cell virus sequencing of influenza infections that trigger innate immunity. *J Virol* 2019; 93: e00500-19.

153 Bost P, Giladi A, Liu Y, *et al*. Host-viral infection maps reveal signatures of severe COVID-19 patients. *Cell* 2020; 181: 1475–1488.

154 McKellar DW, Mantri M, Hinchman MM, *et al*. Spatial mapping of the total transcriptome by in situ polyadenylation. *Nat Biotechnol* 2022; 41: 513–520.

155 Barrozo ER, Seferovic MD, Castro ECC, *et al*. SARS-CoV-2 niches in human placenta revealed by spatial transcriptomics. *Med* 2023; 4: 612–634.

156 Yu X, Gong Q, Yu D, *et al*. Spatial transcriptomics reveals a low extent of transcriptionally active hepatitis B virus integration in patients with HBsAg loss. *Gut* 2023; 73: 797–809.

157 Pichler I, Schmutz S, Ziltener G, *et al*. Rapid and sensitive single-sample viral metagenomics using Nanopore Flongle sequencing. *J Virol Methods* 2023; 320: 114784.

158 Imai K, Tamura K, Tanigaki T, *et al*. Whole genome sequencing of influenza A and B viruses with the MinION sequencer in the clinical setting: a pilot study. *Front Microbiol* 2018; 9: 417312.

159 Jiao X, Imamichi H, Sherman BT, *et al*. QuasiSeq: profiling viral quasispecies *via* self-tuning spectral clustering with PacBio long sequencing reads. *Bioinformatics* 2022; 38: 3192–3199.

160 Tan L, Guo Z, Wang X, *et al*. Utilization of nanopore direct RNA sequencing to analyze viral RNA modifications. *mSystems* 2024; 9: e0116323.

161 Ren J, Song K, Deng C, *et al*. Identifying viruses from metagenomic data using deep learning. *Quant* 2020; 8: 64–77.

162 Vik D, Bolduc B, Roux S, *et al*. MArVD2: a machine learning enhanced tool to discriminate between archaeal and bacterial viruses in viral datasets. *ISME Commun* 2023; 3: 87.

163 Narayana JK, Mac Aogáin M, Ali NABM, *et al*. Similarity network fusion for the integration of multi-omics and microbiomes in respiratory disease. *Eur Respir J* 2021; 58: 2101016.

164 Chen IMA, Chu K, Palaniappan K, *et al*. The IMG/M data management and analysis system v.7: content updates and new features. *Nucleic Acids Res* 2023; 51: D723–D732.
165 Mukherjee S, Stamatis D, Tianqing Li C, *et al*. Genomes OnLine Database (GOLD) v. 10: new features and update. *Nucleic Acids Res* 2025; 53: D989.
166 Meyer F, Paarmann D, D'Souza M, *et al*. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 2008; 9: 386.
167 Callanan J, Stockdale SR, Shkoporov A, *et al*. Expansion of known ssRNA phage genomes: from tens to over a thousand. *Sci Adv* 2020; 6: eaay5981.
168 Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, *et al*. Uncovering Earth's virome. *Nature* 2016; 536: 425–430.
169 Edgar RC, Taylor J, Lin V, *et al*. Petabase-scale sequence alignment catalyses viral discovery. *Nature* 2022; 602: 142–147.
170 Yilmaz P, Kottmann R, Field D, *et al*. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat Biotechnol* 2011; 29: 415–420.
171 Adriaenssens EM, Roux S, Brister JR, *et al*. Guidelines for public database submission of uncultivated virus genome sequences for taxonomic classification. *Nat Biotechnol* 2023; 41: 898–902.
172 Bowers RM, Kyrpides NC, Stepanauskas R, *et al*. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 2017; 35: 725–731.
173 Liu S, Rodriguez JS, Munteanu V*, et al.* Analysis of metagenomic data. *Nat Rev Methods Primers* 2025; 5: 5.
174 Conradie T, Caparros-Martin JA, Egan S, *et al*. Exploring the complexity of the human respiratory virome through an *in silico* analysis of shotgun metagenomic data retrieved from public repositories. *Viruses* 2024; 16: 953.
175 Mitchell AB, Mourad B, Morgan LC, *et al*. Transplanting the pulmonary virome: dynamics of transient populations. *J Heart Lung Transplant* 2018; 37: 1111–1118.
176 Wylie KM, Mihindukulasuriya KA, Zhou Y, *et al*. Metagenomic analysis of double-stranded DNA viruses in healthy adults. *BMC Biol* 2014; 12: 71.
177 Willner D, Furlan M, Haynes M, *et al*. Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. *PLoS One* 2009; 4: e7370.
178 Young JC, Chehoud C, Bittinger K, *et al*. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. *Am J Transplant* 2015; 15: 200–209.
179 Redondo N, Navarro D, Aguado JM, *et al*. Viruses, friends, and foes: the case of torque teno virus and the net state of immunosuppression. *Transpl Infect Dis* 2022; 24: e13778.
180 Choi S, Sohn KH, Jung JW, *et al*. Lung virome: new potential biomarkers for asthma severity and exacerbation. *J Allergy Clin Immunol* 2021; 148: 1007–1015.
181 Kaczorowska J, Van Der Hoek L. Human anelloviruses: diverse, omnipresent and commensal members of the virome. *FEMS Microbiol Rev* 2020; 44: 305–313.
182 Cao L, Ma Y, Wan Z, *et al*. Longitudinal anellome dynamics in the upper respiratory tract of children with acute respiratory tract infections. *Virus Evol* 2023; 9: vead045.
183 Abbas AA, Taylor LJ, Dothard MI, *et al*. Redondoviridae, a family of small, circular DNA viruses of the human oro-respiratory tract associated with periodontitis and critical illness. *Cell Host Microbe* 2019; 25: 719–729.
184 Pifferi M, Maggi F, Andreoli E, *et al*. Associations between nasal torquetenovirus load and spirometric indices in children with asthma. *J Infect Dis* 2005; 192: 1141–1148.
185 Pifferi M, Maggi F, Caramella D, *et al*. High torquetenovirus loads are correlated with bronchiectasis and peripheral airflow limitation in children. *Pediatr Infect Dis J* 2006; 25: 804–808.
186 Xie Y, Xue Q, Jiao W, *et al*. Associations between sputum torque teno virus load and lung function and disease severity in patients with chronic obstructive pulmonary disease. *Front Med (Lausanne)* 2021; 8: 618757.
187 Merenstein C, Liang G, Whiteside SA, *et al*. Signatures of COVID-19 severity and immune response in the respiratory tract microbiome. *mBio* 2021; 12: 1777–1798.
188 Freer G, Maggi F, Pifferi M, *et al*. The virome and its major component, anellovirus, a convoluted system molding human immune defenses and possibly affecting the development of asthma and respiratory diseases in childhood. *Front Microbiol* 2018; 8: 686.
189 Rocchi J, Ricci V, Albani M, *et al*. Torquetenovirus DNA drives proinflammatory cytokines production and secretion by immune cells *via* Toll-like receptor 9. *Virology* 2009; 394: 235–242.
190 van Rijn AL, Roos R, Dekker FW, *et al*. Torque teno virus load as marker of rejection and infection in solid organ transplantation – a systematic review and meta-analysis. *Rev Med Virol* 2023; 33: e2393.
191 Lim YW, Schmieder R, Haynes M, *et al*. Metagenomics and metatranscriptomics: windows on CF-associated viral and microbial communities. *J Cyst Fibros* 2013; 12: 154–164.
192 Zoll J, Rahamat-Langendoen J, Ahout I, *et al*. Direct multiplexed whole genome sequencing of respiratory tract samples reveals full viral genomic information. *J Clin Virol* 2015; 66: 6–11.

193 Ferravante C, Arslan-Gatz BS, Dell'Annunziata F, *et al*. Dynamics of nasopharyngeal tract phageome and association with disease severity and age of patients during three waves of COVID-19. *J Med Virol* 2022; 94: 5567–5573.

194 Garcia-Nuñez M, Gallego M, Monton C, *et al*. The respiratory virome in chronic obstructive pulmonary disease. *Future Virol* 2018; 13: 457–466.

195 Mao Q, Sun G, Qian Y, *et al*. Viral metagenomics of pharyngeal secretions from children with acute respiratory diseases with unknown etiology revealed diverse viruses. *Virus Res* 2022; 321: 198912.

196 Elbehery AHA, Feichtmayer J, Singh D, *et al*. The human virome protein cluster database (HVPC): a human viral metagenomic database for diversity and function annotation. *Front Microbiol* 2018; 9: 1110.

197 Van Rijn AL, Van Boheemen S, Sidorov I, *et al*. The respiratory virome and exacerbations in patients with chronic obstructive pulmonary disease. *PLoS One* 2019; 14: e0223952.

198 Liu B, Shao N, Wang J, *et al*. An optimized metagenomic approach for virome detection of clinical pharyngeal samples with respiratory infection. *Front Microbiol* 2020; 11: 1552.

199 Losada PM, Chouvarine P, Dorda M, *et al*. The cystic fibrosis lower airways microbial metagenome. *ERJ Open Res* 2016; 2: 00096-2015.

200 Lu D, Li C, Zhong Z, *et al*. Changes in the airway microbiome in patients with bronchiectasis. *Medicine (Baltimore)* 2023; 102: e36519.

201 Madi N, Al-Nakib W, Mustafa AS, *et al*. Metagenomic analysis of viral diversity in respiratory samples from patients with respiratory tract infections in Kuwait. *J Med Virol* 2018; 90: 412–420.

202 Ramesh A, Nakielny S, Hsu J, *et al*. Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PLoS One* 2019; 14: e0218318.

203 Hoque MN, Rahman MS, Ahmed R, *et al*. Diversity and genomic determinants of the microbiomes associated with COVID-19 and non-COVID respiratory diseases. *Gene Rep* 2021; 23: 101200.

204 Kim KW, Deveson IW, Nam C, *et al*. Respiratory viral co-infections among SARS-CoV-2 cases confirmed by virome capture sequencing. *Sci Rep* 2021; 11: 3934.

205 Iša P, Taboada B, García-López R, *et al*. Metagenomic analysis reveals differences in the co-occurrence and abundance of viral species in SARS-CoV-2 patients with different severity of disease. *BMC Infect Dis* 2022; 22: 792.

206 Li CX, Li W, Zhou J, *et al*. High resolution metagenomic characterization of complex infectomes in paediatric acute respiratory infection. *Sci Rep* 2020; 10: 3963.

207 Zhang C, Liu T, Wang Y, *et al*. Metagenomic next-generation sequencing of bronchoalveolar lavage fluid from children with severe pneumonia in pediatric intensive care unit. *Front Cell Infect Microbiol* 2023; 13: 1082925.

208 Hyblova M, Hadzega D, Babisova K, *et al*. Metatranscriptome analysis of nasopharyngeal swabs across the varying severity of COVID-19 disease demonstrated unprecedented species diversity. *Microorganisms* 2023; 11: 1804.

209 Willner D, Furlan M. Deciphering the role of phage in the cystic fibrosis airway. *Virulence* 2010; 1: 309–313.

210 Lim YW, Evangelista JS, Schmieder R, *et al*. Clinical insights from metagenomic analysis of sputum samples from patients with cystic fibrosis. *J Clin Microbiol* 2014; 52: 425–437.

211 Rovira Rubió J, Megremis S, Pasioti M, *et al*. Respiratory virome profiles reflect antiviral immune responses. *Allergy* 2023; 78: 1258–1268.

212 Goolam Mahomed T, Peters RPH, Allam M, *et al*. Lung microbiome of stable and exacerbated COPD patients in Tshwane, South Africa. *Sci Rep* 2021; 11: 19758.

213 Schlaberg R, Queen K, Simmon K, *et al*. Viral pathogen detection by metagenomics and pan-viral group polymerase chain reaction in children with pneumonia lacking identifiable etiology. *J Infect Dis* 2017; 215: 1407–1415.

214 Kustin T, Ling G, Sharabi S, *et al*. A method to identify respiratory virus infections in clinical samples using next-generation sequencing. *Sci Rep* 2019; 9: 2606.

215 Wu X, Li Y, Zhang M, *et al*. Etiology of severe community-acquired pneumonia in adults based on metagenomic next-generation sequencing: a prospective multicenter study. *Infect Dis Ther* 2020; 9: 1003–1015.

216 Fischer N, Indenbirken D, Meyer T, *et al*. Evaluation of unbiased next-generation sequencing of RNA (RNA-seq) as a diagnostic method in influenza virus-positive respiratory samples. *J Clin Microbiol* 2015; 53: 2238–2250.

217 Zhang P, Chen Y, Li S, *et al*. Metagenomic next-generation sequencing for the clinical diagnosis and prognosis of acute respiratory distress syndrome caused by severe pneumonia: a retrospective study. *PeerJ* 2020; 8: e9623.

218 Xie F, Duan Z, Zeng W, *et al*. Clinical metagenomics assessments improve diagnosis and outcomes in community-acquired pneumonia. *BMC Infect Dis* 2021; 21: 352.

219 Zhan Y, Xu T, He F, *et al*. Clinical evaluation of a metagenomics-based assay for pneumonia management. *Front Microbiol* 2021; 12: 751073.

220 Miao Q, Liang T, Pei N, *et al*. Evaluation of respiratory samples in etiology diagnosis and microbiome characterization by metagenomic sequencing. *Respir Res* 2022; 23: 345.

221 Shen H, Liu T, Shen M, *et al*. Utilizing metagenomic next-generation sequencing for diagnosis and lung microbiome probing of pediatric pneumonia through bronchoalveolar lavage fluid in pediatric intensive care unit: results from a large real-world cohort. *Front Cell Infect Microbiol* 2023; 13: 1200806.

222 Chang C, Wang H, Zhang L, *et al*. Clinical efficiency of metagenomic next-generation sequencing in sputum for pathogen detection of patients with pneumonia according to disease severity and host immune status. *Infect Drug Resist* 2023; 16: 5869–5885.

223 Billard L, Le Berre R, Pilorgé L, *et al*. Viruses in cystic fibrosis patients' airways. *Crit Rev Microbiol* 2017; 43: 690–708.

224 Mitchell AB, Mourad B, Buddle L, *et al*. Viruses in bronchiectasis: a pilot study to explore the presence of community acquired respiratory viruses in stable patients and during acute exacerbations. *BMC Pulm Med* 2018; 18: 84.

225 Ling KM, Stick SM, Kicic A. Pulmonary bacteriophage and cystic fibrosis airway mucus: friends or foes? *Front Med (Lausanne)* 2023; 10: 1088494.

226 Dedrick RM, Freeman KG, Nguyen JA, *et al*. Potent antibody-mediated neutralization limits bacteriophage treatment of a pulmonary *Mycobacterium abscessus* infection. *Nat Med* 2021; 27: 1357–1361.

227 van Beveren GJ, Said H, van Houten MA, *et al*. The respiratory microbiome in childhood asthma. *J Allergy Clin Immunol* 2023; 152: 1352–1367.

228 de Steenhuijsen Piters WAA, Watson RL, de Koff EM, *et al*. Early-life viral infections are associated with disadvantageous immune and microbiota profiles and recurrent respiratory infections. *Nat Microbiol* 2022; 7: 224–237.

229 Teo SM, Mok D, Pham K, *et al*. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* 2015; 17: 704–715.

230 Xepapadaki P, Megremis S, Rovina N, *et al*. Exploring the impact of airway microbiome on asthma morbidity: a focus on the "constructing a 'eubiosis reinstatement therapy' for asthma–CURE" project. *Pulm Ther* 2024; 10: 171–182.

231 Dransfield MT, Kunisaki KM, Strand MJ, *et al*. Acute exacerbations and lung function loss in smokers with and without chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2017; 195: 324–330.

232 Seemungal T, Harper-Owen R, Bhowmik A, *et al*. Respiratory viruses, symptoms, and inflammatory markers in acute exacerbations and stable chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001; 164: 1618–1623.

233 Papi A, Bellettato CM, Braccioni F, *et al*. Infections and airway inflammation in chronic obstructive pulmonary disease severe exacerbations. *Am J Respir Crit Care Med* 2006; 173: 1114–1121.

234 Tiew PY, Meldrum OW, Chotirmall SH. Applying next-generation sequencing and multi-omics in chronic obstructive pulmonary disease. *Int J Mol Sci* 2023; 24: 2955.

235 Gregory AC, Sullivan MB, Segal LN, *et al*. Smoking is associated with quantifiable differences in the human lung DNA virome and metabolome. *Respir Res* 2018; 19: 174.

236 Liu Y, Huang Q, Zhuang Z, *et al*. Gut virome alterations in patients with chronic obstructive pulmonary disease. *Microbiol Spectr* 2024; 12: e0428723.

237 Madapoosi SS, Cruickshank-Quin C, Opron K, *et al*. Lung microbiota and metabolites collectively associate with clinical outcomes in milder stage chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2022; 206: 427–439.