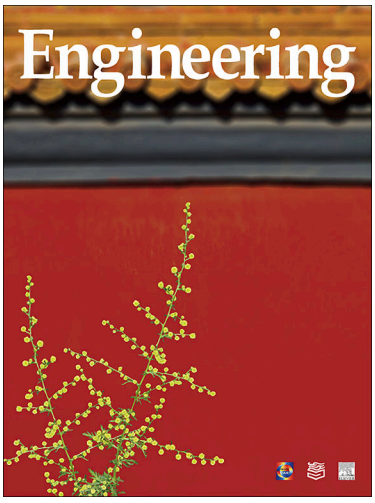
Journal Pre-proofs



Views & Comments

The Infant Gut Virome: Knowns, Unknowns, and Avenues for Future Studies

Nataliia Kuzub, Asier Fernández-Pato, Sanzhima Garmaeva, Alexandra Zhernakova

PII: S2095-8099(26)00187-6
DOI: https://doi.org/10.1016/j.eng.2026.03.018
Reference: ENG 2313

To appear in: *Engineering*

Received Date: 18 April 2025
Revised Date: 10 December 2025
Accepted Date: 6 March 2026

Please cite this article as: N. Kuzub, A. Fernández-Pato, S. Garmaeva, A. Zhernakova, The Infant Gut Virome: Knowns, Unknowns, and Avenues for Future Studies, *Engineering* (2026), doi: https://doi.org/10.1016/j.eng.2026.03.018

The Infant Gut Virome: Knowns, Unknowns, and Avenues for Future Studies

Nataliia Kuzub[#], Asier Fernández-Pato[#], Sanzhima Garmaeva, Alexandra Zhernakova*

Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, The Netherlands

[#] These authors contributed equally to this work.
* Corresponding author.
E-mail address: a.zhernakova@umcg.nl (A. Zhernakova)

The role of the gut microbiome in both infant and adult health is widely recognized. Until recently, most studies have focused on bacterial communities, often neglecting other components of the gut ecosystem. However, the virome—a highly abundant and diverse community of viruses (predominantly bacteriophages)—plays a crucial role in shaping the gut environment. In this review, we describe the composition of the gut virome, summarize laboratory and bioinformatic challenges associated with virome analysis, and provide an overview of studies examining the gut virome in early life. This work has highlighted dynamic changes during infancy and associations between viruses and infant diseases. While research on the gut virome is still in its early stages, the ability of viruses to influence the ecosystem means the gut virome holds significant potential for improving future microbiome-based treatments.

. General overview of viruses in the human gut

The human gut harbors a vast and diverse community of microorganisms that includes bacteria [1], archaea [2], fungi [3], protozoa [4], and viruses [5], which is often collectively referred to as the "microbiome." Until recently, the use of the term "microbiome" in relation to the human gut was almost exclusively reserved for bacterial communities due to the limitations of amplicon-based sequencing, and this focus largely persisted even after the introduction of whole-genome sequencing [6]. Extensive work has now linked gut bacteria to various human traits and diseases, including metabolic [7–9], cardiovascular [10,11], immune diseases [12], mental health [13,14], cancer [15], and other conditions [16]. In parallel, numerous studies have characterized the rapid assembly and development of the bacteriome in early life and its links to infant and long-term health outcomes, including immune [17] and neurocognitive development [18–20], metabolic function and obesity [21], susceptibility to allergies [22,23], and other conditions [24].

In contrast, the composition and significance of the other members of the microbial community, such as viruses, archaea, and fungi, particularly in the developing gut, remain poorly understood. While archaea and fungi comprise a modest proportion of the human gut microbiome [2,25,26], viruses are highly abundant, and their number is similar to the number of gut bacteria [27,28]. Viruses residing in the gastrointestinal tract, which collectively form the gut virome, number up to ~10^{11} particles [28]. They are predominantly viruses of bacteria, called bacteriophages, which comprise over 95% of the gut viral community [29,30]. The remaining fraction comprises diverse viruses that infect other organisms, including viruses of other gut prokaryotes, archaea [31–33], and viruses of eukaryotes [34,35], including human cells [36]. Most eukaryotic viruses in the human gut, however, are usually viruses of plants derived from dietary or environmental sources [34,35].

The adult human gut virome shows high inter-individual diversity and relative compositional stability over short-term [37] and long-term periods [38]. Its diversity and stability can be attributed to the most numerous and abundant members of the double-stranded DNA (dsDNA) gut phageome, class Caudoviricetes [38] and order Crassvirales [39]. Members of the order Crassvirales, including the well-studied crAssphage group, often dominate the phage population in the human gut, with crAssphages alone accounting for up to 90% of viral sequences in some individuals, while being absent in others [37,39–41]. Notably, crAss-like phages are rarely detected in infants at birth but become increasingly prevalent in the microbiome after one month of life, likely driven by the acquisition of host bacterial taxa [29,35,42]. Generally, the virome expands rapidly during the first month of life. Studies estimate that there are approximately 10^8–10^9 viral particles per gram of infant feces by a few weeks to one month of age [35,36,43]. This highlights the early establishment of the virome and its potential influence on infant health and development.

Viruses in the infant gut exist in close association with bacterial populations, shaping each other's ecology through a dynamic bidirectional relationship. Bacterial communities strongly influence phage populations through host availability [37], antiviral defense systems [44], modulation of prophage induction [45,46], and the structuring of mucus and biofilm niches that limit viral activity [47,48]. In turn, phages can modulate their bacterial hosts and affect early gut colonization, with pioneering bacteriophages potentially shaping the temporal development of both bacterial and viral communities [49]. Moreover, the gut phageome influences evolution through horizontal gene transfer [50], together with other mobile genetic elements such as plasmids or integrative and conjugative elements. For instance, phages facilitate the exchange of genes like those encoding metabolic enzymes or virulence factors, as well as the dissemination of antibiotic resistance genes within gut bacterial populations [51–55]. The bacterial community can also be altered indirectly during infections by eukaryotic enteric viruses, particularly those associated with gastroenteritis, as viral infections and diarrheal episodes are consistently accompanied by shifts in gut microbial composition [56]. Notably, several viruses associated with gastroenteritis, including

members of Anelloviridae and certain Parvoviridae, can remain in the gut without causing disease [57–59]. The significance of th[illegible]obial community in a context-dependent manner [57–61].

Beyond ecological interactions, both phages and eukaryotic viruses can also influence early immune and intestinal development. Eukaryotic viruses do so by directly infecting host cells, and their effects can vary from pathogenic to beneficial. For example, studies in germ-free mice have shown that certain enteric viruses can partially compensate for the absence of the gut microbiota by supporting intestinal maturation and immune development [62]. At the same time, gut phages may contribute to infant immune system development, either directly through interaction with immune cells (after crossing the epithelial cell layer) [63,64] or indirectly by shaping bacterial communities via processes such as phage-mediated bacterial lysis (releasing DNA and inflammatory molecules) [65–68], prophage-driven increases in pathogenicity (via virulence factors) [50], or phage-binding to immune mediators [69].

2. Challenges of virome research

Despite its significance, large-scale research on the infant gut virome faces substantial challenges across multiple stages, from study design to laboratory methods to data analysis and interpretation (Fig. 1). As the infant gut virome shows high inter-individual variability and dynamic temporal behavior [49,70–72], large longitudinal population studies with frequent sampling and careful control for confounding factors (e.g., diet, delivery mode, and antibiotic exposure) are needed to unravel its role in infant health and the factors shaping its development.

The reproducibility of findings in infant gut virome research is also highly dependent on the laboratory methods and protocols used. Gut viruses show high variability in genome composition, consisting of single- or double-stranded DNA (ssDNA, dsDNA) or RNA (ssRNA, dsRNA), and in genome size, which can range from approximately 3.5 to 540 kb (phages) [73]. However, most early-life human gut studies to date have focused solely on dsDNA viruses, often omitting second-strand synthesis or reverse transcription steps during nucleic acid processing. As a result, ssDNA viruses (e.g., Microviridae, Anelloviridae) and RNA viruses (e.g., Picornaviridae-like enteroviruses, Caliciviridae-like noroviruses) may be overlooked, despite their potential relevance to infant health and early immune modulation [74–77]. Furthermore, the viral fraction that is recovered can be significantly influenced by methodological choices, including the nucleic acid extraction protocols and viral enrichment procedures used [78,79]. Viral enrichment, often obtained through the filtration of particles above a certain size, enhances overall detection sensitivity (enabling the identification of rare viruses), improves the ability to capture actively replicating viruses, and increases confidence in the viral origin of sequences. However, this approach may hinder the detection of larger viruses, such as megaphages, as well as temperate or persistent viral infections [80]. Ideally, combining viral enrichment methods with total metagenome sequencing would provide a more comprehensive view of gut viral communities. However, few studies adopt this dual approach, often resulting in incomplete virome profiles. This issue is particularly pronounced in infant studies, where the low biomass in fecal samples further complicates the detection of microbial populations and often leads to high contamination [81,82]. In addition, sequencing depth plays a critical role, as insufficient depth can limit detection of low-abundance viral populations, reduce assembly quality, and bias the estimation of viral diversity, making comprehensive virome characterization more difficult. Simulation studies in microbial communities have shown that increasing sequencing depth improves genome recovery and community resolution, highlighting the importance of high-throughput sequencing for accurate metagenomic analyses [83].

In addition to challenges in virus isolation and recovery, accurately identifying viral genomes from sequencing data represents an additional hurdle. Unlike bacteria, which can be studied using conserved genetic markers like 16S rRNA, viruses lack universal markers, making their identification, classification, and analysis more challenging [84]. An overview of the computational tools developed to address these challenges and enable comprehensive characterization of viral genomes is provided in Table 1 [85–113]. Several tools have been developed in recent years to identify viral genomes within complex metagenomes. These fall into three main categories: alignment-free or deep-learning-based methods (e.g., DeepVirFinder [85], Seeker [86]), which use machine learning to detect viral signatures based on sequence features; gene-based methods that rely on comparison or alignment to reference databases to identify viral protein markers (e.g., VirSorter2 [87], VIBRANT [88]); and hybrid methods that combine these approaches for improved accuracy (e.g., geNomad) [89]. Further, increasing efforts have been made to facilitate viral profiling from metagenomes by integrating it with bacterial profiling methods to enable more comprehensive microbiome analyses [90,114] and develop uniformly processed public gut viral databases that compile viral genomes from multiple metagenomic studies and public repositories [5,115–117]. Despite these efforts, metagenomic studies have revealed that a high proportion of the viral sequences in the infant gut show no similarity to known viruses [60,71,72,118,119], highlighting the vast unexplored diversity of the infant virome and its underrepresentation in existing viral databases.

The rapid discovery of novel viruses through metagenomics, along with the need for a standardized, evolution-based viral classification approach, has led the International Committee on Taxonomy of Viruses (ICTV) to adopt a genome-similarity-based framework for viral taxonomy, integrating metagenomically discovered viruses into the official classification scheme [120,121]. Newly developed computational methods now allow ICTV taxonomy to be assigned directly from genomic sequences [91,92]. However, unlike bacteria, the rapid evolution of viruses and frequent horizontal gene transfer challenge traditional taxonomic boundaries, complicating systematic analysis of the gut virome [122]. Ecology- or function-centered approaches offer an alternative to address this by grouping viruses based on shared traits, such as virus−host interactions or viral lifestyles (virulent or temperate) [93,94]. For instance, viral host prediction can significantly improve our understanding of phage−bacteria dynamics and the ecological roles viruses play within the developing gut ecosystem [49,71]. Additionally,

determining the functional potential of gut viruses is essential for understanding their contributions to infant health. Although sequ[illegible]es in alignment-free approaches offer promising solutions. In particular, methods based on protein embeddings derived from protein language models have been shown to improve viral protein annotation [95,96]. Complementing these approaches, recent progress in protein structure prediction has enabled the creation of extensive protein repositories, such as the AlphaFold Protein Structure Database [123] and the ESM Atlas [124], as well as virus-specific databases like the Big Fantastic Virus Database [125], Viro3D [126], or the Viral AlphaFold Database [127].

Despite these advances, a substantial gap remains in our understanding of the actual functional activity of viruses in the infant gut. Phage activity and phage–host interactions are highly dynamic, context-dependent, transient, and often spatially heterogeneous, making them difficult to capture or study using metagenomic approaches. This gap is increasingly being addressed using methods such as single-cell RNA sequencing (RNA-seq), which captures phage and host transcriptional activity at the level of individual cells, revealing which viral genes are actively expressed *in situ* [128]. Techniques that link phages to their bacterial hosts, such as viral tagging and high-throughput/resolution chromosome conformation capture (Hi-C), provide complementary insights into host specificity and ecological interactions [129,130]. In addition, live imaging techniques offer spatial and temporal visualization of phage replication and spread within living hosts, adding a dynamic perspective on viral activity [131].

Taken together, progress in virome research will depend on continued development and standardization of laboratory and computational frameworks. Current protocols for DNA isolation and sequencing in viromics introduce distinct recovery biases that strongly shape downstream results [78,79]. Because these protocols also differ in cost and processing time, full standardization may not be immediately feasible, but explicit reporting and awareness of methodological biases are essential. Similar issues arise in bioinformatic pipelines, where diverse viral identification and analysis strategies limit comparability across studies. In the coming years, the field is likely to benefit from unified, high quality viral reference databases that support hybrid approaches that combine assembly-based and read-based detection, thereby improving the recovery and profiling of rare and divergent viruses. In parallel, complementing virus-centric analyses with function-focused perspectives, such as grouping viruses by shared accessory genes or predicted functional potential, may provide a more robust framework for linking the virome to ecosystem dynamics. Finally, moving from association to mechanism will require the expansion and broader use of experimental toolkits, including single-cell RNA-seq, Hi-C, live imaging techniques, and systematic culturing of phages, ideally coupled with organoid and animal models to reveal how virome variation shapes gut ecology and function.

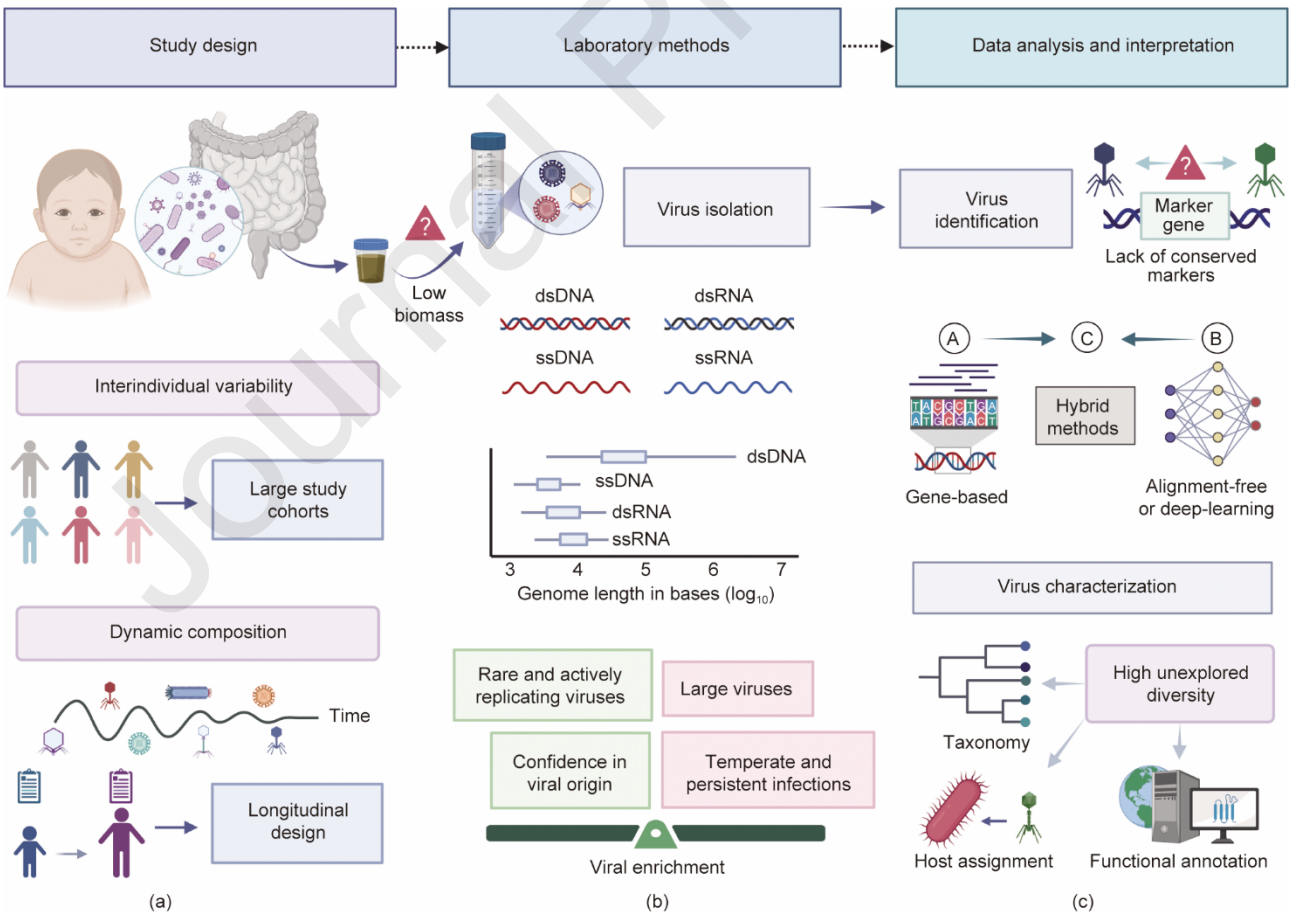


Fig. 1. Challenges in studying the infant gut virome. This diagram illustrates the multi-stage challenges in studying the infant gut virome that

hinder consistent and reproducible findings critical for understanding the virome's role in early life. (a) Challenges in study design. The infant gut viron[illegible] and standardized protocols. (b) Challenges in laboratory methods. The low biomass in infant fecal samples and high viral genomic diversity complicate virus isolation. Most studies currently focus on dsDNA viruses, potentially overlooking ssDNA and RNA viruses that are relevant to infant health. Viral enrichment methods enhance detection sensitivity but may miss larger viruses and temperate/persistent infections. (c) Challenges in data analysis and interpretation. The absence of universal viral markers, the diversity of viral genome identification methods, and the vast unexplored viral diversity all make systematic analysis of the infant gut virome challenging. Integrating viral taxonomy with ecological and functional characterization will be essential to provide a better understanding of the virome's role in early-life development and its potential impact on infant health.

3. Early-life gut virome: dynamics, sources, and health implications

Despite the challenges in the field, a number of studies on the early-life human gut virome have provided insights into its dynamics and composition (Table 2 [29,35,43,49,55,60,66,67,70,72,118,119,132–154]).

Table 1. Bioinformatic tools and resources for virome analysis.

Analysis category	Tool	Last update	Input data	Algorithm/ Computational approach	Description and source
Virus identification	Seeker [86]	v1.0.3 (2020)	DNA sequences (FASTA), typically contigs from metagenomic data	Long short-term memory recurrent neural network, alignment-free method	Trained to distinguish phage vs bacterial DNA without needing similarity to known genomes, gives a "phage score" per fragment and averages over fragments
	DeepVirFinder [85]	v1.0 (2020)	DNA sequences (FASTA), typically contigs from metagenomic data	Deep learning, convolutional neural networks (CNN), k-mer features, motif extraction	Alignment-free viral contig classifier using CNNs trained on k-mer motif signatures, outputs score + p-value per sequence
	VIBRANT [88]	v1.2.1 (2020)	(Meta)genomic contigs/assemblies (FASTA) (DNA or RNA)	Neural network (multi-layer perceptron) on protein annotation signatures, similarity-based (hidden markov model (HMM)) features, v-score metric	Automated recovery and annotation of viral genomes determines genome completeness, excises integrated proviruses, and identifies auxiliary metabolic genes (AMGs) for functional characterization
	VirSorter2 [87]	v2.2.4 (2023)	(Meta)genomic contigs/assemblies (FASTA) (DNA or RNA)	HMM-based feature extraction, random forest classifiers trained per viral group	Modular pipeline that detects diverse viral sequences (dsDNA, ssDNA, RNA, NCLDV, virophages), trims host regions, and outputs viral scores per sequence
	VIRify [97]	v3.1.0 (2025)	Metagenomic or metatranscriptomic contigs	HMMs (profile HMMs) for virus-specific protein domains,	Pipeline to detect, annotate, and

			FASTQ files can be provided for assembly	taxon-specific HMM hits	classify viral contigs (including prophages), using curated viral protein models (ViPhOGs)
	Phanta [90]	v1.1.1 (2025)	Metagenomic reads (FASTQ)	Read classification using Kraken2/Bracken with a curated viral database, abundance estimation, viral–bacterial integration detection	Reference-based pipeline for identifying and quantifying viruses from metagenomic reads supports abundance profiling and detection of integrated prophages
	geNomad [89]	2025 (v1.11.2)	Whole genomes, metagenomes, metatranscriptomes (FASTA)	Hybrid classification framework: IGLOO-based neural network classifier + marker-based classifier using informative protein markers	Identifies viral and plasmid genomes, detects integrated proviruses, annotates sequences via marker proteins, and assigns up-to-date ICTV taxonomy
Virus binning/contig extension	vRhyme [98]	v1.1.0 (2022)	Viral scaffolds/ contigs (FASTA), coverage profiles across samples	Supervised machine learning (neural networks + decision trees), read coverage variance, network clustering, protein-redundancy scoring	Constructs viral metagenome-assembled genomes (vMAGs) by grouping contigs with similar coverage patterns and nucleotide features, and refines bins using viral-specific features such as low redundancy of proteins
	PHAMB [99]	v1.0.1 (2022)	Viral contigs from metagenomes/metaviromes (FASTA)	Deep learning metagenomic binning algorithm (VAMB), paired metagenome/metavirome features, random forest classifier for viral bin identification	Bins viral genomes directly from bulk metagenomic data, and groups them into taxonomic viral populations, which also helps infer viral-host interactions
	COBRA [100]	v1.3.0 (2025)	Assembled contigs/scaffolds (FASTA), coverage table, read mapping (BAM/SAM)	De Bruijn graph-based reassembly, contig overlap detection, paired-read linkage, coverage comparison	Extends and joins fragmented viral contigs to improve genome completeness and contiguity, potentially forming circular genomes, and solves assembly breakpoints by analyzing end overlap and read-pair information

Qua assessment	CheckV	v1.0.3 (2024)	Viral contigs or complete metagenomes or isolates	Reference-based comparison to identify viral genes, and genome termini analysis	Detects and removes host contamination, estimates viral genome completeness, classifying sequences as: complete, high quality, medium quality, or low quality
Taxonomic assignment	Vcontact2 [102]	v0.11.3 (2022)	Viral protein sequences (FASTA), gene-to-genome mapping file	Gene-sharing network, protein clustering, similarity-based clustering, network propagation	Classifies viral genomes into viral clusters based on shared gene content, and infers taxonomic relationships across viruses
	PhaGCN2 [103,104]	v2.3.0 (2024)	Viral genome sequences (FASTA)	Graph convolutional networks (GCN), protein clustering, sequence feature embeddings, and supervised learning	Predicts viral taxonomy using graph-based deep learning, uses protein similarity networks, and sequence features to assign viral genomes to taxonomic groups
	VITAP [91]	v1.7.1 (2025)	Viral genome sequences or contigs (FASTA)	Alignment-based protein scoring, weighted cumulative taxonomic scoring, graph-based taxonomic path selection, and automated ICTV database updates	Classifies DNA and RNA viral sequences up to genus level, providing confidence levels for each assignment, and automatically updates ICTV-based reference database
	taxmyPhage [92]	v0.3.4 (2025)	dsDNA bacteriophage genomes (FASTA)	MASH-based similarity search against ICTV-classified genomes, BLASTn intergenomic similarity calculation	Classifies dsDNA phage genomes at genus and species levels using updated ICTV taxonomy
Host prediction	CHERRY [105]	2023	Viral genomes (FASTA)	Graph-based model using protein-organization similarity, sequence similarity, clustered regularly interspaced short palindromic repeats (CRISPR) matches, and k-mer features, GCN encoder + link-prediction decoder	Predicts prokaryotic hosts for viruses (phylum to species). Integrates multiple evidence types in a unified graph model to improve host-prediction accuracy
	iPHOP [93]	v1.4.1 (2025)	Viral contigs or genomes (viruses of archaea or bacteria) (FASTA)	Sequence similarity, CRISPR spacer matching, k-mer frequency correlation, probabilistic scoring	Assigns potential hosts to viruses of archaea and bacteria using multiple evidence sources, including CRISPR matches and sequence composition, and outputs confidence scores for predictions

Phage Lifestyle prediction	BACPHLIP [94]	v0.9.6 (2020)	Bacteriophage genomes (FASTA)	Random forest classifier, protein domain presence/absence.	Predicts bacteriophage lifestyle (temperate vs virulent) based on conserved protein domains
	PhaTYP [106]	v0.3.0 (2023)	Bacteriophage genomes (FASTA)	Contextualized protein embeddings using BERT, a deep learning classifier	Predicts bacteriophage lifestyle (temperate vs virulent) based on BERT-based contextual embeddings for protein composition and associations
Functional annotation	Pharokka [107]	v1.8.2 (2025)	Viral contigs or complete genomes (FASTA)	PHANOTATE or Prodigal for CDS prediction, mmseqs2 and PyHMMER protein searches, PHROGs-, CARD-, and VFDB-based functional assignment, tRNAs predicted with tRNAscan-SE 2.0, tmRNAs with Aragorn, CRISPRs with CRISPR recognition tool (CRT)	Predicts CDS, tRNAs, tmRNAs, and CRISPRs, assigns protein functions using the PHROGs database, and identifies virulence and antimicrobial resistance genes via VFDB and CARD
	VPF-PLM [95]	v1.0 (2023)	Viral protein sequences (FASTA)	Protein language models (PLMs), embedding-based classification, multi-class predictive model	Uses pretrained protein language models to embed viral proteins and classify their function. Captures remote functional homology beyond sequence similarity
	PHOLD [108]	v1.1.0 (2025)	Phage genomes (FASTA, GenBank)	Protein structural alignment, ProtT5 embeddings, Foldseek, large phage protein structure database	Annotates phage proteins using structural information and embeddings
	Empathi [96]	v1.0.6 (2025)	Viral protein sequences (FASTA). Comma-separated values (CSV) file of protein embeddings	Protein language model embeddings (ProtT5), support vector machine (SVM)-based hierarchical binary classification, and PHROG-derived functional ontology	Predicts phage protein functions using ProtT5 protein embeddings, It assigns proteins to newly defined functional groups better suited for ML-based predictions

Inte pipelines	ViroProfiler	v0.2.5 (2023)	Metagenomic reads	Modular pipeline: read quality (metaSPAdes), viral contig identification (VirSorter2, VIBRANT, DeepVirFinder, CheckV), binning (vRhyme or phamb), gene function annotation (DRAM-v, EggNOG, abricate), viral lifestyle prediction (BACPHLIP, Replidec), taxonomy assignment (vConTACT2, MMseqs2), host prediction (iPHoP), visualization	End-to-end modular iral metagenomics: performs assembly, quality control (QC), viral discovery, annotation (AMGs, gene function), lifestyle and host prediction, taxonomic assignment, and result visualization
	ViWRAP [110]	v1.3.1(2024)	Metagenomic assemblies and reads (FASTA/FASTQ)	Modular integration of multiple tools: virus identification and annotation (VIBRANT, VirSorter2, DeepVirFinder), virus binning (vRhyme), clustering (vConTACT2, dRep), taxonomy classification (NCBI RefSeq, VOG HMMs, IMG/VR v4), genome quality assessment (CheckV), host prediction (iPHoP)	Integrated pipeline for viral metagenome analysis, providing identification, binning, clustering, taxonomy, genome quality, host prediction, and summary visualizations
	MVP [111]	v.1.1.4 (2024)	Sequencing reads (FASTQ) and metagenomic contigs/assemblies (FASTA)	Integrated pipeline: viral genome recovery (geNomad), quality assessment (CheckV), viral operational taxonomic unit (vOTU) clustering (ANI-based), read mapping (Bowtie2/minimap2), functional annotation (PHROGs, Pfam, DRAM-v), viral binning (vRhyme), host prediction (iPHoP), abundance and summary generation	End-to-end modular viromics pipeline: identifies viruses and proviruses, assesses genome quality, clusters viral contigs, annotates genes and AMGs, predicts viral lifestyle, generates coverage tables, and supports public database submission
	PhaBOX [112]	v2.1.13(2025)	Viral genome sequences (FASTA) with/without metagenomic reads	Integrates multiple specialized tools: PhaMer for phage identification, PhaGCN for taxonomy classification, CHERRY for host prediction, PhaTYP for lifestyle prediction, and PhaVIP for viral protein annotation	Comprehensive viral genome analysis pipeline for detection, classification, host prediction, lifestyle annotation, functional annotation, and phylogenetics
	Hecatomb [113]	v.1.3.4 (2025)	Short- and long-read metagenomic data (FASTQ)	Four-module workflow: (1) read QC, contaminant/host removal (fastp/BBTools, Minimap2) and clustering, (2) MMseqs2 tiered annotation, (3) per-sample and merged assembly (MEGAHIT/Canu + Flye), (4) contig-based annotation and abundance estimation (MMseqs2, BBMap)	Pipeline performing read filtering and annotation, metagenomic assembly, contig taxonomic annotation, and abundance profiling.

Table 2. Overview of sequencing-based gut virome studies. In the fourth column, "N" denotes the number of subjects and "S" indicates the number of samples. Only sequenced samples are included in this count. *Breitbart et al. [43] used PCR and microarray techniques to longitudinally analyze the infant gut virome. ** Yinda et al. [143], only samples collected from children aged 0–3 years were counted. VLP: viral-like particles (sequencing), MGS: metagenomic sequencing.

Study	VLP/	Country (participants)	N	Inf (S	Term/	Timepoints	Design	Main findings on the infant gut

	MGS		Inf); N Mo)	preterm	(infants		virome
Breitbart et al. 2008 [43]	VLP	USA	1 (1); —	Full-term	1 week	Longitudinal*	• Infant gut virome dominated by bacteriophages • Most prevalent fecal viruses not derived from breast milk or formula, suggesting a non-dietary origin • Infant gut virome changes dramatically during the first 2 weeks
Kramna et al. 2015 [132]	VLP	Finland	38 (114); —	Not specified	3, 6, and 9 months	Longitudinal	• No association between virome composition and islet autoimmunity
Reyes et al. 2015 [133]	VLP	Malawi	40 (222); 9 (9)	Not specified	0–30 months	Longitudinal	• Gut virome development disrupted in children with severe acute malnutrition and not corrected by therapeutic interventions • *Anelloviridae* and *Circoviridae* distinguish malnourished children from healthy ones
Lim et al. 2015 [134]	VLP	USA	8 (40); —	Not specified	0, 3, 6, 12, 18, and 24 months	Longitudinal	• Gut virome more similar between co-twins than between unrelated individuals • Eukaryotic virome expands while phage virome contracts and shifts in composition during the first 2 years
Zhao et al. 2017 [135]	VLP	Finland/Estonia	22 (220); —	Not specified	0–36 months	Longitudinal	• Children with autoantibodies associated with type 1 diabetes (T1D) have less diverse gut viromes than children without autoimmunity. • Higher prevalence and abundance of *Circoviridae* in controls compared to cases • Disease-associated phage correlated with abundance of *Bacteroides* and *Bifidobacterium* contigs
Asnicar et al. 2017 [136]	MGS	Italy	5 (8); 5 (8)	Not specified	3, 5, and 10 months	Longitudinal/ cross-sectional	• *Enterobacter* and *Shigella* phages most prevalent • Pepper mild mottle virus present and very abundant in 2 of the 5 infant pairs
Olm et al. 2017 [137]	MGS	USA	2 (37); —	Preterm, gestational age: < 30 weeks	0–1 months	Longitudinal	• Phage blooms shift microbial composition in the infant gut
Duranti et al. 2017 [138]	MGS	Italy	25 (48); 21 (21)	Not specified	7 days, 1 month	Longitudinal	• Bifidophages vertically transmitted from mother to infant
McCann et al. 2018 [139]	VLP	Ireland	20 (20); —	Full-term	1 year	Cross-sectional	• Infants born via Cesarean section have lower alpha diversity and a distinct virome compared to infants born via spontaneous vaginal delivery
Pannaraj et al. 2018 [140]	VLP	USA	10 (10); 10 (10)	Full-term	0 month	Cross-sectional	• The majority of viruses in the infant gut virome are bacteriophages • Different virome composition in infant stool and maternal milk • A significant portion of viruses

							shared with maternal milk
Aiemjoy et al. 2019 [141]	VLP	Ethiopia	269 (269); —	Not specified	0 months–5 years	Cross-sectional	• Stool consistency correlates with virome composition • Children with watery stools had higher viral diversity • Norovirus GII, Aichivirus A, and adeno-associated virus 2 more abundant in children with loose or watery stools
Maqsood et al. 2019 [142]	VLP	USA	50 (50); 27 (27)	Full-term	0 month	Cross-sectional	• 15% of infant virome shared with maternal virome • Viromes of infant twin pairs more similar to each other than to unrelated infants • Mother–infant virome sharedness not impacted by delivery route
Yinda et al. 2019 [143]	VLP	Cameroon	80 (80); —**	Not specified	0–89 years	Cross-sectional	• *Picobirnaviruses* less abundant in infants and children • *Anelloviridae* more prevalent in children and the elderly
Vehik et al. 2019 [144]	MGS	USA, Germany, Sweden, Finland	495 (12034); —	Not specified	3 months–4 years	Longitudinal	• Prolonged shedding of enterovirus EV-B serotype correlates with islet autoimmunity • Early-life infection with HAdV-C associated with a decreased risk of islet autoimmunity and T1D
Gregory et al. 2020 [29]	VLP	USA, Ireland	27 (68); —	Not specified	0 month–2 years	Longitudinal	• Infant gut virome diversity highest at birth and declines over the first 2 years of life • Most gut viruses are temperate bacteriophages
Liang et al. 2020 [35]	VLP	USA	166 (206); —	Full-term	0, 1, 3, 4 months; 2–5 years	Longitudinal/ cross-sectional	• Prophages, induced from the bacteria and vertically transmitted from mother to infant, provide the predominant population of VLPs • Virome composition shaped by feeding mode (breastfeeding vs. formula feeding) • Human viruses more abundant in exclusively formula-fed infants
Lindfors et al. 2020 [145]	MGS	USA, Germany, Sweden, Finland	166 (1507); —	Not specified	3 months–2 years	Longitudinal	• Frequent exposure to enterovirus between ages 1 and 2 is associated with celiac disease
Taboada et al. 2021 [146]	MGS	Mexico	3 (35); —	Full-term	2 weeks–1 year	Longitudinal	• Eukaryotic virome abundance and richness significantly increase from 6 to 12 months of life • Most detected viruses were from six bacteriophage families, including five from the Caudovirales order
Beller et al. 2022 [60]	VLP	Belgium	8 (8); —	Full-term	0–12 months	Longitudinal	• Infant gut virome is highly individual-specific (70.5% of phages infant-specific + only 7.8% present in > 50% of infants) • Infant gut virome shows rapid increase in richness and decreasing proportion of temperate phages during the first year of life • Phages show temporal correlations with their hosts

							• Infection rate with ses dramatically increased upon entering daycare, but rarely caused clinical manifestations • Infant gut virome harbors plant viruses and phages of Lactococcus, which is used in dairy production
Kaelin et al. 2022 [147]	VLP	USA	23 (138); —	Preterm, gestational age: < 27 weeks	0–11 weeks	Longitudinal	• Infant gut virome is individual-specific, showing lower within-individual distances than between-individual distances • Virome beta-diversity in infants who developed NEC reduced 10 days prior to NEC onset • No change in richness and Shannon diversity over time in preterm infants • 137 viral sequences associated with NEC 10 days prior to NEC onset
Li et al. 2023 [148]	VLP	China	90 (90); —	Not specified	0 month, 5 years	Cross-sectional	• Infants have lower diversity and a distinct composition compared to 5-year-old children
Shah et al. 2023 [72]	VLP	Denmark	647 (647); —	Not specified	1 year	Cross-sectional	• Discovered 232 previously unknown viral family-level clades belonging to Caudoviricetes class • Temperate phages dominate gut virome of infants • Virulent phages more abundant in one-year-old infants, temperate phages more prevalent and diverse • Skunaviridae phages infecting Lactococcus (used in dairy products) were the most abundant family in infants
Walters et al. 2023 [149]	VLP	USA	53 (454); 53 (233)	Not specified	0 months–3 years	Longitudinal	• Human host virome prevalent in infants but not mothers • The infant phageome, but not the eukaryotic virome, resembled adult-like state by age 3 • Infant gut virome richness and abundance increase over time • Early life virome development is shaped mainly by environmental rather than maternal exposure
Leal Rodríguez et al. 2024 [67]	VLP	Denmark	647 (647); —	Not specified	1 year	Cross-sectional	• 19 viral families, mainly infecting *Faecalibacterium* and *Ruminococcus*, had decreased abundance in children who subsequently developed asthma • Asthma risk score calculated based on virome composition independently predicted asthma occurrence in children • Virome-asthma risk score was associated with the transient asthma phenotype and the infant's TLR9 rs187084 variant
Garmaeva et al. 2024 [70]	VLP	Netherlands	32 (89); 30 (116)	Full-term	1, 2, 3, 6, and 12 months	Longitudinal	• Phages are vertically co-transmitted from mothers to infants together with their bacterial hosts • Infant gut viromes have a higher

							abundance of actively replicating adult gut viromes, and this decreases over the first year of life • Feeding mode (breastfeeding vs. formula feeding) and place of delivery influence infant gut virome composition • Formula-fed infants have a higher abundance of actively replicating temperate phages
Zheng et al. 2024 [118]	MGS & VLP	Belgium, USA, Finland, Ireland, Denmark, Cameroon, Estonia, Italy, Sweden, Bangladesh, New Zealand, Singapore, UK, Russia, and Luxembourg	2969 (8130); —	Full-term/preterm	0–3 years	Longitudinal/ cross-sectional	• Created an early-life human gut virome catalog spanning 82 141 species-level vOTUs • During the first 3 years, infant gut virome richness gradually increased, while the proportion of temperate phages gradually decreased • Delivery mode, gestational age, and feeding mode influence infant gut virome composition • Infant gut virome and bacteriome show a correlation in abundance and diversity
Byrne et al. 2024 [150]	MGS	South Africa	69 (69); —	Full-term	4 weeks	Cross-sectional	• Uninfected but exposed to HIV infants (nHEU) had an altered gut virome compared to unexposed infants • Virome features in nHEU were associated with inflammation
Maqsood et al. 2024 [151]	VLP	Kenya	72 (147); 63 (159)	Not specified	1, 2, 3, 6, 12, 15, 18, and 21 months	Longitudinal	• Infant gut virome changes are driven by age and not impacted by SARS-CoV-2 infection
Lou et al. 2024 [119]	MGS	USA	52 (735); 42 (84)	Full-term/preterm, gestational age: < 34 weeks	0–3 years	Longitudinal	• Infant gut virome diversity resembles adult by age 3 • Approximately 9% of phages persist over time, most transmitted from the mother • Phages with stop-codon reassignment showed increased persistence • Phages with diversity-generating retroelements increased with age
Redgwell et al. 2025 [66]	VLP	Denmark	647 (647); —	Not specified	1 year	Cross-sectional	• Infant gut virome is individual-specific and does not have a typical core, with the most prevalent vOTU detected in ~70% of samples • The Hanky P00' phage-related cluster, which infects Bacteroides and carries diversity-generating retroelements and genes involved in capsular polysaccharide synthesis, was the most dominant in the cohort • Most prophages identified were induced, with no environmental exposures influencing them
Tisza et al. 2025 [49]	MGS	USA, Germany, Sweden, Finland	887 (12,262); —	Not specified	0–4 years	Longitudinal	• Type 1 diabetes linked to a slower shift in bacterial and viral community composition in children between age one and two • Phage populations unique to each individual, but their

							ecological succession patterns • Phage data incorporation enhanced machine learning models in the ability to distinguish samples by country
Fernández-Pato et al. 2025 [55]	MGS	Netherlands	28 (172); 23 (23)	Full-term	1–6 weeks; 6, 12 months	Longitudinal	• Infant gut virome is individual-specific and shows temporal stability during the first 6 weeks • Formula-fed infants show higher richness and distinct virome composition compared with breastfed infants • Mother–infant sharing of specific gut viral strains is limited in the first months and increases only after 6 months
Zhang et al. 2025 [152]	VLP	Denmark	647 (647); —	Not specified	1 year	Cross-sectional	• Virome composition differs by early-life environment, mainly with older siblings and urban vs rural residence • Dietary factors (maternal fish oil supplementation and egg introduction) associated with distinct virome patterns • Phage abundances co-varied with their bacterial hosts, and some environmentally associated phages carried metabolic genes
Ren at al. 2025 [153]	VLP	China	107 (107); —	Preterm, gestational age: 26.9–36.9 weeks	1 week	Cross-sectional	• In preterm infants, higher virome (but not bacterial) diversity and distinct virome profiles are linked to higher risk of early onset anemia • Enrichment of Circoviridae, an uncultured Caudoviridae phage, and depletion of certain CRESS viruses are associated with higher risk of neurodevelopmental delay • *Geobacillus* virus Tp84 linked to lower risk of neuro-developmental delay but higher risk of early onset anemia
Subramanian et al. 2025 [154]	MGS	USA	55 (272)	Not specified	1–2 days, 2, 6, 12, and 24 months	Longitudinal	• Compared to Cesarean-delivered infants, vaginally delivered infants have higher gut viral diversity at 2 months and a distinct virome up to 12 months, with different predicted bacterial hosts and more AMGs

Researchers increasingly agree that the maternal womb is a sterile environment, with initial seeding of the infant gut mici with viral-like particles (VLPs) detectable in the first days of life [35,43]. Although some studies report that virome diversity peaks at birth [29,134,142] and declines during the first years of life [29,134], most studies agree that the early-life virome is characterized by low diversity, which is followed by a steady increase with age [35,60,70,148,149]. Similarly to the adult gut virome, the expanding infant gut virome is dominated by bacteriophages [35,37,70]. Studies have shown that the phage component of the gut virome reaches an adult-like state by the age of three [96,149]. However, even though eukaryotic virus diversity declines with infant age, it remains higher in three-year-olds compared to their mothers, suggesting that virome maturation is still ongoing at this age [149].

Another key feature of early-life gut virome dynamics is a higher fraction of actively replicating temperate phages when compared to maternal viromes [29,70]. Most bacteriophages exist in either a lytic cycle, where they infect and lyse their bacterial hosts, producing VLPs, or a lysogenic cycle, where they integrate into bacterial DNA and replicate with their hosts. Phages capable of switching between these cycles tend to favor lysogeny in the adult gut environment [70,157–159]. In contrast, the infant gut virome has a higher proportion of temperate phages that peaks a few weeks after birth and then declines, but remains elevated at one year [70,72]. This abundance of temperate phages likely reflects unique bacteria−phage dynamics in the newly colonized infant gut, possibly driven by phages responding to bacterial stress [160] and mediated by various factors, including diet [70,161].

The source of the initial viral seeding remains a topic of debate. Some studies claim that the infant gut virome is primarily shaped by the environment [149], while others report mother−infant virus sharing starting right after birth [142] and continuing throughout the first years of infant life [70,96]. Temperate phages have been shown to be pioneering colonizers of the infant gut [35] and to be transmitted from the maternal gut in the form of prophages integrated into the genomes of host bacteria [70], with subsequent induction upon bacterial seeding [35,70]. Additionally, both lytic and temperate viral strains have been shown to co-transmit with their bacterial hosts from mother to infant, further highlighting the role of the maternal microbiome as a key source of the infant gut virome [70].

A shared environment with other individuals also plays a crucial role in virome development. For example, Beller et al. [60] reported a rapid expansion of eukaryotic gastroenteritis-associated viruses in infants after entering daycare, suggesting active viral transmission in communal settings. Another confirmation of the importance of shared environment for infant gut virome development comes from twin studies showing that twins' viromes are similar in infancy [133,134] but not adulthood [162]. Diet can also serve as a source for the infant gut virome, although it is likely to be a transient source that does not result in long-term colonization. Studies have documented the acquisition of plant viruses [60,149] and phages from industrial bacterial strains used in dairy products [60,72], both of which appear in the infant gut following solid food introduction, suggesting a dietary origin.

Beyond the sources of the early-life virome, factors influencing its development have also been identified. A study by McCan et al. [139] found lower viral Shannon diversity and distinct composition in one-year-old infants born via Cesarean section compared to those born via vaginal delivery. While one might argue that these changes reflect differences in bacterial microbiome composition, research shows that such shifts occur in the infant gut virome independently of bacterial dynamics, suggesting distinct virome signatures [139]. Similarly, several studies have shown that the location of delivery (home vs hospital) [70] and feeding mode (breastfeeding vs formula feeding) [35,70] also affect gut virome dynamics in early life. Several studies have indicated that formula-fed infants exhibit higher viral diversity than breastfed infants, with elevated prophage induction levels [70] and an increased number of eukaryotic viruses [35]. The latter difference is thought to stem from the protective role of breast milk and its immune-modulating components, such as maternal antibodies [35]. Additionally, it was hypothesized that formula feeding components might trigger induction of prophages [70] in the same way that other dietary components were shown to induce prophages [161], although additional evidence is needed to validate both hypotheses.

While being influenced by various factors and exposures, the gut virome has also been shown to affect its human host in multiple ways. These include direct infection by eukaryotic viruses, modulation of bacterial populations by bacteriophages, and direct interaction of bacteriophages with the host immune system [64,65,163,164]. These interactions are hypothesized to play a role in various diseases, as growing evidence suggests associations between the composition of the infant gut virome and several conditions, including necrotizing enterocolitis (NEC) [147], asthma [67], and autoimmunity [49,144].

NEC is a leading cause of death among preterm infants, with an incidence of 2%–7% and a mortality rate of around 30% in infants weighing less than 1500 grams at birth [147,165]. While numerous studies have explored microbiome associations with NEC onset and progression [166–169], recent evidence suggests a role for the gut virome in its development [147]. A study by Kaelin et al. [147] that included longitudinal sampling of nine infants who developed NEC and 14 matched controls found that the viromes of the affected infants converged toward higher similarity approximately 10 days before onset. Kaelin et al. [147] also identified 137 viral contigs associated with NEC onset within this timeframe. To elucidate the functional importance of the gut virome in NEC development, Offersen et al. [170] used a preterm piglet NEC model in which VLPs were isolated from donor fecal filtrate and transferred to susceptible recipients. Results showed that fecal virome transfer (FVT) alone was sufficient to lessen NEC-like intestinal injury and lowered the relative abundance of NEC-associated bacteria, including *Klebsiella pneumoniae* and *Clostridium perfringens*.

Other important links have been identified between the gut virome and predisposition to asthma. Asthma is the most common chronic lung disease, affecting millions of children worldwide, and its onset is influenced by environmental, genetic, and immune factors [171–173]. Recent research suggests there is an association between the virome component of the gut microbiome and asthma development [174,175]. A study of 647 one-year-old found that specific virome features were linked to the later onset of asthma [67]. Notably, the abundances of 19 asthma-associated viral families, primarily infecting

Faecalibacterium and *Ruminococcus*, were decreased in children who developed asthma later in life. Additionally, the study estal[Journal Pre-proofs]or 9 (TLR9) rs187084 variant, suggesting direct interactions between phages and the child's immune system. Phages have also been reported to directly interact with TLR9, involved in antiviral response in inflammatory bowel disease in adults [176], thereby activating both phage-specific and non-specific immune responses.

The infant gut virome has also been associated with the development of autoimmune diseases, including type 1 diabetes (T1D) and celiac disease. A study by Tisza et al. [49] showed that children who eventually developed T1D exhibited a reduced rate of change in their gut viral and bacterial communities between 400 and 700 days of life, compared to those who did not develop the disease. This period of decreased microbial community dynamics preceded the average age of T1D diagnosis, which was approximately 2.4 years [49]. Several studies have found links between specific eukaryotic viruses and these non-communicable diseases [144,145]. Notably, exposure to enterovirus between ages 1 and 2 was associated with celiac disease, while prolonged shedding of the Enterovirus-B serotype correlated with islet autoimmunity [144,145]. Conversely, early-life infection with Human Adenovirus-C was linked to a decreased risk of islet autoimmunity and T1D [144]. These findings highlight the significance of eukaryotic viruses in non-communicable disease development, although the exact mechanisms remain unclear.

Together, these findings show that the infant gut virome is a dynamic community shaped by multiple factors, and early-life viral changes may contribute to diseases such as NEC, asthma, T1D, and celiac disease. The results of current studies may also help identify at-risk infants and guide future preventive and therapeutic virome-based interventions. Although these studies provide important insights into human gut virome development, they are largely descriptive and require mechanistic validation in experimental systems, especially to establish causal relationships between specific viruses, their bacterial hosts, and infant health. The interaction of bacteriophages with the human host represents a particularly important direction for future studies, since phages may influence host cells, not only by modulating bacterial populations and metabolism, but also by crossing epithelial barriers and interacting directly with gut and immune cells [177–179]. However, to clarify the role of phages in shaping infant health and immunity, the current association-based metagenomic approaches need to be complemented with additional methods.

Existing studies focused on translational virome research might offer useful conceptual and technical guidance on approaches to study the infant gut virome, including FVT and phage-focused interventions that show functional effects on metabolism, pathogen colonization resistance, and immune responses in animal models and early clinical settings [170,180,181]. Moving forward, integrating infant-derived viral communities with intestinal organoids, gnotobiotic animal models, and synthetic communities, together with targeted in vitro studies of disease-associated viruses, will help the field move beyond associations toward defining viral functions that cannot be inferred from sequence data alone.

4. Conclusions

Compared with studies of the bacteriome, far fewer studies to date have examined the gut virome, even though viral communities are similarly abundant and, at the same time, considerably more variable and individual-specific. Current virome research offers only limited insight into its functional role and sometimes lacks reproducibility, largely due to technical challenges and the need for large longitudinal cohorts for association studies, which are not always feasible. Although pipelines for virome analysis have advanced rapidly in recent years, with multiple tools and databases now available, substantial methodological challenges remain. Further development, eventual standardization, and integration with complementary approaches will be essential to move the field beyond associations toward mechanistic understanding.

Existing studies already provide valuable insights into early-life virome composition and dynamics, the origin of infant gut viruses, and the environmental factors shaping virome development. They also highlight that viral composition often mirrors the bacterial profile, underscoring the tight interactions between viruses and bacteria in the infant gut. This interdependence, however, adds complexity when disentangling bacterial versus viral contributions to health outcomes. On the other hand, evidence on the direct impact of viruses on infant health is emerging. In addition to the capacity of human viruses to modulate immunity through host-cell infection, the interaction of bacteriophages with innate immunity through Toll-like receptors and their associations with immune-mediated diseases like T1D raise an important question: To what extent do viruses in the infant gut influence and shape infant immune development?

While the gut virome field is in its infancy, it holds large potential for health practices. Bacteriophages have the ability to mediate microbiome composition and might be used as alternatives to prebiotics and antibiotics, offering insights into the role of gut viruses and viral−bacterial dynamics in gut ecosystem development and health improvement.

Acknowledgments

We thank Kate Mc Intyre for editing the manuscript. This research is supported by the Netherlands Organization for Scientific Research (NWO) VICI (VI.C.232.074), NWO Gravitation (ExposomeNL 024.004.017), NWO KIC (KICH1.LWV04.21.01), ZonMW ME/CFS (10091012110015), and European Union Horizon Europe Program (INITIALISE 101094099 and DarkMatter 101136582).

References

[1] Nayfach S, Shi ZJ, Seshadri R, Pollard KS, Kyrpides NC. New insights from uncultivated genomes of the global human gut microbiome. Nature 2019;568(7753):505–10.

[2] Chibani CM, Mahnert A, Borrel G, Almeida A, Werner A, Brugère JF, et al. A catalogue of 1,167 genomes from the human gut archaeome. Nat Microbiol 2021;7(1):48–61.

[3] Yan Q, Li S, Yan Q, Huo X, Wang C, Wang X, et al. A genomic compendium of cultivated human gut fungi characterizes the gut mycobiome and its relevance to common diseases. Cell 2024;187(12):2969–89.e24.

[4] Piperni E, Nguyen LH, Manghi P, Kim H, Pasolli E, Andreu-Sánchez S, et al. Intestinal Blastocystis is linked to healthier diets and more favorable cardiometabolic outcomes in 56,989 individuals from 32 countries. Cell 2024;187(17):4554–70.

[5] Nayfach S, Páez-Espino D, Call L, Low SJ, Sberro H, Ivanova NN, et al. Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. Nat Microbiol 2021;6(7):960–70.

[6] Jansen D, Matthijnssens J. The emerging role of the gut virome in health and inflammatory bowel disease: challenges, covariates and a viral imbalance. Viruses 2023;15(1):173.

[7] Liu R, Hong J, Xu X, Feng Q, Zhang D, Gu Y, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. Nat Med 2017;23(7):859–68.

[8] Chen L, van den Munckhof ICL, Schraa K, Ter Horst R, Koehorst M, van Faassen M, et al. Genetic and microbial associations to plasma and fecal bile acids in obesity relate to plasma lipids and liver fat content. Cell Rep 2020;33(1):108212.

[9] Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Võsa U, et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. Nat Genet 2019;51(4):600–5.

[10] Vieira-Silva S, Falony G, Belda E, Nielsen T, Aron-Wisnewsky J, Chakaroun R, et al. Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. Nature 2020;581(7808):310–5.

[11] Kurilshikov A, van den Munckhof ICL, Chen L, Bonder MJ, Schraa K, Rutten JHW, et al. Gut microbial associations to plasma metabolites linked to cardiovascular phenotypes and risk. Circ Res 2019;124(12):1808–20.

[12] Vatanen T, Franzosa EA, Schwager R, Tripathi S, Arthur TD, Vehik K, et al. The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. Nature 2018;562(7728):589–94.

[13] Valles-Colomer M, Falony G, Darzi Y, Tigchelaar EF, Wang J, Tito RY, et al. The neuroactive potential of the human gut microbiota in quality of life and depression. Nat Microbiol 2019;4(4):623–32.

[14] Fan Y, Støving RK, Berreira Ibraim S, Hyötyläinen T, Thirion F, Arora T, et al. The gut microbiota contributes to the pathogenesis of anorexia nervosa in humans and mice. Nat Microbiol 2023;8(5):787–802.

[15] Bossuet-Greif N, Vignard J, Taieb F, Mirey G, Dubois D, Petit C, et al. The colibactin genotoxin generates dna interstrand cross-links in infected cells. MBio 2018;9(2):e02393–17.

[16] Kunath BJ, De Rudder C, Laczny CC, Letellier E, Wilmes P. The oral-gut microbiome axis in health and disease. Nat Rev Microbiol 2024;22(12):791–805.

[17] Yao Y, Cai X, Ye Y, Wang F, Chen F, Zheng C. The role of microbiota in infant health: from early life to adulthood. Front Immunol 2021;12:708472.

[18] Vuong HE, Pronovost GN, Williams DW, Coley EJL, Siegler EL, Qiu A, et al. The maternal microbiome modulates fetal neurodevelopment in mice. Nature 2020;586(7828):281–6.

[19] Kim E, Paik D, Ramirez RN, Biggs DG, Park Y, Kwon HK, et al. Maternal gut bacteria drive intestinal inflammation in offspring with neurodevelopmental disorders by altering the chromatin landscape of CD4+ T cells. Immunity 2022;55(1):145–58.

[20] Cerdó T, Ruiz-Rodríguez A, Acuña I, Torres-Espínola FJ, Menchén-Márquez S, Gámiz F, et al. Infant gut microbiota contributes to cognitive performance in mice. Cell Host Microbe 2023;31(12):1974–88.

[21] Soderborg TK, Clark SE, Mulligan CE, Janssen RC, Babcock L, Ir D, et al. The gut microbiota in infants of obese mothers increases inflammation and susceptibility to NAFLD. Nat Commun 2018;9(1):4462.

[22] Hoskinson C, Dai DLY, Del Bel KL, Becker AB, Moraes TJ, Mandhane PJ, et al. Delayed gut microbiota maturation in the first year of life is a hallmark of pediatric allergic disease. Nat Commun 2023;14(1):4785.

[23] Sinha T, Brushett S, Fernández-Pato A, Garmaeva S, Andreu-Sánchez S, Spreckels J, et al. Pregnancy and early life gut microbiome: influencing factors and health implications. 2024. Preprint at: https://doi.org/10.21203/rs.3.rs-5334252/v1.

[24] Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. Cell Res 2020;30(6):492–506.

[25] Sender R, Fuchs S, Milo R. Revised estimates for the number of human and bacteria cells in the body. PLoS Biol 2016;14(8):e1002533.

[26] Nash AK, Auchtung TA, Wong MC, Smith DP, Gesell JR, Ross MC, et al. The gut mycobiome of the Human Microbiome Project healthy cohort. Microbiome 2017;5(1):153.

[27] López-García P, Gutiérrez-Preciado A, Krupovic M, Ciobanu M, Deschamps P, Jardillier L, et al. Metagenome-derived virus-microbe ratios across ecosystems. ISME J 2023;17(10):1552–63.

[28] Shkoporov AN, Hill C. Bacteriophages of the human gut: the “known unknown” of the Microbiome. Cell Host Microbe 2019;25(2):195–209.

[29] Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. The gut virome database reveals age-dependent patterns of virome diversity in the human gut. Cell Host Microbe 2020;28(5):724–40.

[30] Garmaeva S, Gulyaeva A, Sinha T, Shkoporov AN, Clooney AG, Stockdale SR, et al. Stability of the human gut virome and effect of gluten-free diet. Cell Rep 2021;35(7):109132.

[31] Medvedeva S, Borrel G, Krupovic M, Gribaldo S. A compendium of viruses from methanogenic archaea reveals their diversity and adaptations to the gut environment. Nat Microbiol 2023;8(11):2170–82.

[32] Baquero DP, Medvedeva S, Martin-Gallausiaux C, Pende N, Sartori-Rupp A, Tachon S, et al. Stable coexistence between an archaeal virus and the dominant methanogen of the human gut. Nat Commun 2024;15(1):7702.

[33] Li R, Wang Y, Hu H, Tan Y, Ma Y. Metagenomic analysis reveals unexplored diversity of archaeal virome in the human gut. Nat Commun 2022;13(1):7978.

[34] Aguado-García Y, Taboada B, Morán P, Rivera-Gutiérrez X, Serrano-Vázquez A, Iša P, et al. Tobamoviruses can be frequently present in the oropharynx and gut of infants during their first year of life. Sci Rep 2020;10(1):13595.

[35] Liang G, Zhao C, Zhang H, Mattei L, Sherrill-Mix S, Bittinger K, et al. The stepwise assembly of the neonatal virome is modulated by breastfeeding. Nature 2020;581(7809):470–4.

[36] Liang G, Bushman FD. The human virome: assembly, composition and host interactions. Nat Rev Microbiol 2021;19(8):514–27.

[37] Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, Nolan JA, et al. The human gut virome is highly diverse, stable, and individual specific. Cell Host Microbe 2019;26(4):527–41.

[38] Gulyaeva A, Garmaeva S, Kurilshikov A, Vich Vila A, Riksen NP, Netea MG, et al. Diversity and ecology of Caudoviricetes phages with genome terminal repeats in fecal metagenomes from four Dutch cohorts. Viruses 2022;14(10):2305.

[39] Gulyaeva A, Garmaeva S, Ruigrok RAAA, Wang D, Riksen NP, Netea MG, et al. Discovery, diversity, and functional associations of crAss-like phages in human gut metagenomes from four Dutch cohorts. Cell Rep 2022;38(2):110204.

[40] Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GGZ, Boling L, et al. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat Commun 2014;5(1):4498.

[41] Yutin N, Makarova KS, Gussow AB, Krupovic M, Segall A, Edwards RA, et al. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. Nat Microbiol 2017;3(1):38–46.

[42] Siranosian BA, Tamburini FB, Sherlock G, Bhatt AS. Acquisition, transmission and strain diversity of human gut-colonizing crAss-like phages. Nat Commun 2020;11(1):280.

[43] Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, et al. Viral diversity and dynamics in an infant gut. Res Microbiol 2008;159(5):367–73.

[44] Georjon H, Bernheim A. The highly diverse antiphage defence systems of bacteria. Nat Rev Microbiol 2023;21(10):686–700.

[45] Guler P, Bendori SO, Borenstein T, Aframian N, Kessel A, Eldar A. Arbitrium communication controls phage lysogeny through non-lethal modulation of a host toxin-antitoxin defence system. Nat Microbiol 2024;9(1):150–60.

[46] Henrot C, Petit MA. Signals triggering prophage induction in the gut microbiota. Mol Microbiol 2022;118(5):494–502.

[47] Lourenço M, Chaffringeon L, Lamy-Besnier Q, Pédron T, Campagne P, Eberl C, et al. The spatial heterogeneity of the gut limits predation and fosters coexistence of bacteria and bacteriophages. Cell Host Microbe 2020;28(3):390–401.

[48] Mortensen AS, Nielsen DS, Røder HL, Sølbeck Rasmussen T. Mucofilm: a nexus for phage-microbiome interactions in gut ecology. Appl Environ Microbiol 2025;91(11):e0126925.

[49] Tisza MJ, Lloyd RE, Hoffman K, Smith DP, Rewers M, Javornik Cregeen SJ, et al. Longitudinal phage-bacteria dynamics in the early life gut microbiome. Nat Microbiol 2025;10(2):420–30.

[50] Diard M, Bakkeren E, Cornuault JK, Moor K, Hausmann A, Sellin ME, et al. Inflammation boosts bacteriophage transfer between Salmonella spp. Science 2017;355(6330):1211–15.

[51] Kent AG, Vill AC, Shi Q, Satlin MJ, Brito IL. Widespread transfer of mobile antibiotic resistance genes within individual gut microbiomes revealed throu

[52] Diebold PJ, New FN, Hovan M, Satlin MJ, Brito IL. Linking plasmid-based beta-lactamases to their bacterial hosts using single-cell fusion PCR. eLife 2021;10:e66834.

[53] Durrant MG, Li MM, Siranosian BA, Montgomery SB, Bhatt AS. A bioinformatic analysis of integrative mobile genetic elements highlights their role in bacterial adaptation. Cell Host Microbe 2020;27(1):140–53.

[54] Borodovich T, Shkoporov AN, Ross RP, Hill C. Phage-mediated horizontal gene transfer and its implications for the human gut microbiome. Gastroenterol Rep 2022;10:goac012.

[55] Fernández-Pato A, Sinha T, Garmaeva S, Gulyaeva A, Kuzub N, Roux S, et al. Early-life development of the gut virome and plasmidome: a longitudinal study in cesarean-born infants. Cell Rep 2025;44(6):115731.

[56] Borrego-Ruiz A, Borrego JJ. The gut microbiome in enteric viral infections: underlying mechanisms and therapeutic approaches. Microorganisms 2025;13(10):2247.

[57] Sabbaghian M, Gheitasi H, Shekarchi AA, Tavakoli A, Poortahmasebi V. The mysterious anelloviruses: investigating its role in human diseases. BMC Microbiol 2024;24(1):40.

[58] Lolomadze EA, Rebrikov DV. Constant companion: clinical and developmental aspects of torque teno virus infections. Arch Virol 2020;165(12):2749–57.

[59] Ong DSY, Schuurman R, Heikens E. Human bocavirus in stool: a true pathogen or an innocent bystander? J Clin Virol 2016;74:45–49.

[60] Beller L, Deboutte W, Vieira-Silva S, Falony G, Tito RY, Rymenans L, et al. The virota and its transkingdom interactions in the healthy infant gut. Proc Natl Acad Sci USA 2022;119(13):e2114619119.

[61] Feng Z, Burgermeister E, Philips A, Zuo T, Wen W. The gut virome in association with the bacteriome in gastrointestinal diseases and beyond: roles, mechanisms, and clinical applications. Precis Clin Med 2025;8(2):pbaf010.

[62] Kernbauer E, Ding Y, Cadwell K. An enteric virus can replace the beneficial function of commensal bacteria. Nature 2014;516(7529):94–8.

[63] Nguyen S, Baker K, Padman BS, Patwa R, Dunstan RA, Weston TA, et al. Bacteriophage transcytosis provides a mechanism to cross epithelial cell layers. MBio 2017;8(6):e01874–17.

[64] Bichet MC, Adderley J, Avellaneda-Franco L, Magnin-Bougma I, Torriero-Smith N, Gearing LJ, et al. Mammalian cells internalize bacteriophages and use them as a resource to enhance cellular growth and survival. PLoS Biol 2023;21(10):e3002341.

[65] Barr JJ, Auro R, Furlan M, Whiteson KL, Erb ML, Pogliano J, et al. Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc Natl Acad Sci USA 2013;110(26):10771–6.

[66] Redgwell TA, Thorsen J, Petit MA, Deng L, Vestergaard G, Russel J, et al. Prophages in the infant gut are pervasively induced and may modulate the functionality of their hosts. NPJ Biofilms Microbiomes 2025;11(1):46.

[67] Leal Rodríguez C, Shah SA, Rasmussen MA, Thorsen J, Boulund U, Pedersen CT, et al. The infant gut virome is associated with preschool asthma risk independently of bacteria. Nat Med 2024;30(1):138–48.

[68] Zuppi M, Hendrickson HL, O'Sullivan JM, Vatanen T. Phages in the gut ecosystem. Front Cell Infect Microbiol 2022;11:822562.

[69] Miernikiewicz P, Kłopot A, Soluch R, Szkuta P, Kęska W, Hodyra-Stefaniak K, et al. T4 phage tail adhesin gp12 counteracts LPS-induced inflammation in vivo. Front Microbiol 2016;7:1112.

[70] Garmaeva S, Sinha T, Gulyaeva A, Kuzub N, Spreckels JE, Andreu-Sánchez S, et al. Transmission and dynamics of mother-infant gut viruses during pregnancy and early life. Nat Commun 2024;15(1):1945.

[71] Peng Y, Zhu J, Wang S, Liu Y, Liu X, DeLeon O, et al. A metagenome-assembled genome inventory for children reveals early-life gut bacteriome and virome dynamics. Cell Host Microbe 2024;32(12):2212–30.

[72] Shah SA, Deng L, Thorsen J, Pedersen AG, Dion MB, Castro-Mejía JL, et al. Expanding known viral diversity in the healthy infant gut. Nat Microbiol 2023;8(5):986–98.

[73] Camargo AP, Nayfach S, Chen IA, Palaniappan K, Ratner A, Chu K, et al. IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. Nucleic Acids Res 2023;51(D1):D733–43.

[74] Wolsk HM, Følsgaard NV, Birch S, Brix S, Hansel TT, Johnston SL, et al. Picornavirus-induced airway mucosa immune profile in asymptomatic neonates. J Infect Dis 2016;213(8):1262–70.

[75] Freer G, Maggi F, Pifferi M, Di Cicco ME, Peroni DG, Pistello M. 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;9:686.

[76] Fernandes MA, Verstraete SG, Phan TG, Deng X, Stekol E, LaMere B, et al. Enteric virome and bacterial microbiota in children with ulcerative colitis

[77] Hassan E, Baldridge MT. Norovirus encounters in the gut: multifaceted interactions and disease outcomes. Mucosal Immunol 2019;12(6):1259–67.

[78] Kleiner M, Hooper LV, Duerkop BA. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. BMC Genomics 2015;16(1):7.

[79] Kosmopoulos JC, Klier KM, Langwig MV, Tran PQ, Anantharaman K. Viromes vs. mixed community metagenomes: choice of method dictates interpretation of viral community ecology. 2023. Preprint at: https://doi.org/10.1101/2023.10.15.562385.

[80] Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, Krupovic M, et al. Minimum information about an uncultivated virus genome (MIUViG). Nat Biotechnol 2019;37(1):29–37.

[81] Kennedy KM, de Goffau MC, Perez-Muñoz ME, Arrieta MC, Bäckhed F, Bork P, et al. Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. Nature 2023;613(7945):639–49.

[82] Kuzub N, Kurilshikov A, Zhernakova A, Garmaeva S. Negativeome characterization and decontamination in early-life virome studies. Nat Commun 2025;16(1):6190.

[83] Rocha U, Kasmanas JC, Toscan R, Sanches DS, Magnusdottir S, Saraiva JP. Simulation of 69 microbial communities indicates sequencing depth and false positives are major drivers of bias in prokaryotic metagenome-assembled genome recovery. PLOS Comput Biol 2024;20(10):e1012530.

[84] Sullivan MB. Viromes, not gene markers, for studying double-stranded DNA virus communities. J Virol 2015;89(5):2459–61.

[85] Ren J, Song K, Deng C, Ahlgren NA, Fuhrman JA, Li Y, et al. Identifying viruses from metagenomic data using deep learning. Quant Biol 2020;8(1):64–77.

[86] Auslander N, Gussow AB, Benler S, Wolf YI, Koonin EV. Seeker: alignment-free identification of bacteriophage genomes by deep learning. Nucleic Acids Res 2020;48(21):e121.

[87] Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome 2021;9(1):37.

[88] Kieft K, Zhou Z, Anantharaman K. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome 2020;8(1):90.

[89] Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, et al. Identification of mobile genetic elements with geNomad. Nat Biotechnol 2024;42(8):1303–12.

[90] Pinto Y, Chakraborty M, Jain N, Bhatt AS. Phage-inclusive profiling of human gut microbiomes with Phanta. Nat Biotechnol 2024;42(4):651–62.

[91] Zheng K, Sun J, Liang Y, Kong L, Paez-Espino D, Mcminn A, et al. VITAP: a high precision tool for DNA and RNA viral classification based on meta-omic data. Nat Commun 2025;16(1):2226.

[92] Millard A, Denise R, Lestido M, Thomas MT, Webster D, Turner D, et al. taxMyPhage: automated Taxonomy of dsDNA Phage Genomes at the Genus and Species Level. Phage 2025;6(1):5–11.

[93] Roux S, Camargo AP, Coutinho FH, Dabdoub SM, Dutilh BE, Nayfach S, et al. iPHoP: an integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. PLoS Biol 2023;21(4):e3002083.

[94] Hockenberry AJ, Wilke CO. BACPHLIP: predicting bacteriophage lifestyle from conserved protein domains. PeerJ 2021;9:e11396.

[95] Flamholz ZN, Biller SJ, Kelly L. Large language models improve annotation of prokaryotic viral proteins. Nat Microbiol 2024;9(2):537–49.

[96] Boulay A, Leprince A, Enault F, Rousseau E, Galiez C. Empathi: embedding-based phage protein annotation tool by hierarchical assignment. Nat Commun 2025;16(1):9114.

[97] Rangel-Pineros G, Almeida A, Beracochea M, Sakharova E, Marz M, Reyes Muñoz A, et al. VIRify: an integrated detection, annotation and taxonomic classification pipeline using virus-specific protein profile hidden Markov models. PLOS Comput Biol 2023;19(8):e1011422.

[98] Kieft K, Adams A, Salamzade R, Kalan L, Anantharaman K. vRhyme enables binning of viral genomes from metagenomes. Nucleic Acids Res 2022;50(14):e83.

[99] Johansen J, Plichta DR, Nissen JN, Jespersen ML, Shah SA, Deng L, et al. Genome binning of viral entities from bulk metagenomics data. Nat Commun 2022;13(1):965.

[100] Chen L, Banfield JF. COBRA improves the completeness and contiguity of viral genomes assembled from metagenomes. Nat Microbiol 2024;9(3):737–50.

[101] Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol 2021;39(5):578–85.

[102] Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabl

[103] Shang J, Jiang J, Sun Y. Bacteriophage classification for assembled contigs using graph convolutional network. Bioinformatics 2021;37(Suppl 1):i25–33.

[104] Jiang JZ, Yuan WG, Shang J, Shi YH, Yang LL, Liu M, et al. Virus classification for viral genomic fragments using PhaGCN2. Brief Bioinform 2023;24(1):bbac505.

[105] Shang J, Sun Y. CHERRY: a computational method for accurate prediction of virus-prokaryotic interactions using a graph encoder-decoder model. Brief Bioinform 2022;23(5):bbac182.

[106] Shang J, Tang X, Sun Y. PhaTYP: predicting the lifestyle for bacteriophages using BERT. Brief Bioinform 2023;24(1):bbac487.

[107] Bouras G, Nepal R, Houtak G, Psaltis AJ, Wormald PJ, Vreugde S. Pharokka: a fast scalable bacteriophage annotation tool. Bioinformatics 2023;39(1):btac776.

[108] Bouras G, Grigson SR, Mirdita M, Heinzinger M, Papudeshi B, Mallawaarachchi V, et al. Protein structure informed bacteriophage genome annotation with Phold 2026;54(1):gkaf1448.

[109] Ru J, Khan Mirzaei M, Xue J, Peng X, Deng L. ViroProfiler: a containerized bioinformatics pipeline for viral metagenomic data analysis. Gut Microbes 2023;15(1):2192522.

[110] Zhou Z, Martin C, Kosmopoulos JC, Anantharaman K. ViWrap: a modular pipeline to identify, bin, classify, and predict viral-host relationships for viruses from metagenomes. iMeta 2023;2(3):e118.

[111] Coclet C, Camargo AP, Roux S. MVP: a modular viromics pipeline to identify, filter, cluster, annotate, and bin viruses from metagenomes. mSystems 2024;9(10):e0088824–e24.

[112] Shang J, Peng C, Liao H, Tang X, Sun Y. PhaBOX: a web server for identifying and characterizing phage contigs in metagenomic data. Bioinform Adv 2023;3(1):vbad101.

[113] Roach MJ, Beecroft SJ, Mihindukulasuriya KA, Wang L, Paredes A, Cárdenas LAC, et al. Hecatomb: an integrated software platform for viral metagenomics. Gigascience 2024;13:giae020.

[114] Zolfo M, Silverj A, Blanco-Miguez A, Manghi P, Rota-Stabelli O, Heidrich V, et al. Discovering and exploring the hidden diversity of human gut viruses using highly enriched virome samples. 2024. Preprint at: https://doi.org/10.1101/2024.02.19.580813

[115] Camarillo-Guerrero LF, Almeida A, Rangel-Pineros G, Finn RD, Lawley TD. Massive expansion of human gut bacteriophage diversity. Cell 2021;184(4):1098–1109.e9.

[116] Camargo AP, Baltoumas FA, Ndela EO, Fiamenghi MB, Merrill BD, Carter MM, et al. A genomic atlas of the human gut virome elucidates genetic factors shaping host interactions. 2025. Preprint at: https://doi.org/10.1101/2025.11.01.686033.

[117] Nayfach S, Camargo A. Unified human gut virome (UHGV) [Internet]. Genève: Zenodo; 2025 Oct 24 [cited 2026 Mar 3]. Available from: https://zenodo.org/records/17402089.

[118] Zeng S, Almeida A, Li S, Ying J, Wang H, Qu Y, et al. A metagenomic catalog of the early-life human gut virome. Nat Commun 2024;15(1):1864.

[119] Lou YC, Chen L, Borges AL, West-Roberts J, Firek BA, Morowitz MJ, et al. Infant gut DNA bacteriophage strain persistence during the first 3 years of life. Cell Host Microbe 2024;32(1):35–47.e6.

[120] Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, et al. Virus taxonomy in the age of metagenomics. Nat Rev Microbiol 2017;15(3):161–8.

[121] Gorbalenya AE, Krupovic M, Mushegian A, Kropinski AM, Siddell SG, Varsani A, et al. The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks. Nat Microbiol 2020;5(5):668–74.

[122] Caetano-Anollés G, Claverie JM, Nasir A. A critical analysis of the current state of virus taxonomy. Front Microbiol 2023;14:1240993.

[123] Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, Radhakrishnan M, et al. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. Nucleic Acids Res 2024;52(D1):D368–75.

[124] Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science 2023;379(6637):1123–30.

[125] Kim RS, Levy Karin E, Mirdita M, Chikhi R, Steinegger M. BFVD-a large repository of predicted viral protein structures. Nucleic Acids Res 2025;53(D1):D340–7.

[126] Litvin U, Lytras S, Jack A, Robertson DL, Hughes J, Grove J. Viro3D: a comprehensive database of virus protein structure predictions. Mol Syst Biol 2025;21(11):1599–617.

[127] Odai R, Leemann M, Al-Murad T, Abdullah M, Shyrokova L, Tenson T, et al. The viral alphafold database of monomers and homodimers reveals conse

[128] Shen Y, Qian Q, Ding L, Qu W, Zhang T, Song M, et al. High-throughput single-microbe RNA sequencing reveals adaptive state heterogeneity and host-phage activity associations in human gut microbiome. Protein Cell 2025;16(3):211–26.

[129] Džunková M, Low SJ, Daly JN, Deng L, Rinke C, Hugenholtz P. Defining the human gut host-phage network through single-cell viral tagging. Nat Microbiol 2019;4(12):2192–203.

[130] Wu R, Davison MR, Nelson WC, Smith ML, Lipton MS, Jansson JK, et al. Hi-C metagenome sequencing reveals soil phage-host interactions. Nat Commun 2023;14(1):7666.

[131] Ortiz de Ora L, Wiles ET, Zünd M, Bañuelos MS, Haro-Ramirez N, Suder DS, et al. Phollow reveals in situ phage transmission dynamics in the zebrafish gut microbiome at single-virion resolution. Nat Microbiol 2025;10(5):1067–83.

[132] Kramná L, Kolářová K, Oikarinen S, Pursiheimo JP, Ilonen J, Simell O, et al. Gut virome sequencing in children with early islet autoimmunity. Diabetes Care 2015;38(5):930–3.

[133] Reyes A, Blanton LV, Cao S, Zhao G, Manary M, Trehan I, et al. Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. Proc Natl Acad Sci USA 2015;112(38):11941–6.

[134] Lim ES, Zhou Y, Zhao G, Bauer IK, Droit L, Ndao IM, et al. Early life dynamics of the human gut virome and bacterial microbiome in infants. Nat Med 2015;21(10):1228–34.

[135] Zhao G, Vatanen T, Droit L, Park A, Kostic AD, Poon TW, et al. Intestinal virome changes precede autoimmunity in type I diabetes-susceptible children. Proc Natl Acad Sci USA 2017;114(30):E6166–75.

[136] Asnicar F, Manara S, Zolfo M, Truong DT, Scholz M, Armanini F, et al. Studying vertical microbiome transmission from mothers to infants by strain-level metagenomic profiling. mSystems 2017;2(1):e00164–16.

[137] Olm MR, Brown CT, Brooks B, Firek B, Baker R, Burstein D, et al. Identical bacterial populations colonize premature infant gut, skin, and oral microbiomes and exhibit different in situ growth rates. Genome Res 2017;27(4):601–12.

[138] Duranti S, Lugli GA, Mancabelli L, Armanini F, Turroni F, James K, et al. Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. Microbiome 2017;5(1):66.

[139] McCann A, Ryan FJ, Stockdale SR, Dalmasso M, Blake T, Ryan CA, et al. Viromes of one year old infants reveal the impact of birth mode on microbiome diversity. PeerJ 2018;6:e4694.

[140] Pannaraj PS, Ly M, Cerini C, Saavedra M, Aldrovandi GM, Saboory AA, et al. Shared and distinct features of human milk and infant stool viromes. Front Microbiol 2018;9:1162.

[141] Aiemjoy K, Altan E, Aragie S, Fry DM, Phan TG, Deng X, et al. Viral species richness and composition in young children with loose or watery stool in Ethiopia. BMC Infect Dis 2019;19(1):53.

[142] Maqsood R, Rodgers R, Rodriguez C, Handley SA, Ndao IM, Tarr PI, et al. Discordant transmission of bacteria and viruses from mothers to babies at birth. Microbiome 2019;7(1):156.

[143] Yinda CK, Vanhulle E, Conceição-Neto N, Beller L, Deboutte W, Shi C, et al. Gut virome analysis of cameroonians reveals high diversity of enteric viruses, including potential interspecies transmitted viruses. mSphere 2019;4(1):e00585-18

[144] Vehik K, Lynch KF, Wong MC, Tian X, Ross MC, Gibbs RA, et al. Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. Nat Med 2019;25(12):1865–72.

[145] Lindfors K, Lin J, Lee HS, Hyöty H, Nykter M, Kurppa K, et al. Metagenomics of the faecal virome indicate a cumulative effect of enterovirus and gluten amount on the risk of coeliac disease autoimmunity in genetically at risk children: the TEDDY study. Gut 2020;69(8):1416–22.

[146] Taboada B, Morán P, Serrano-Vázquez A, Iša P, Rojas-Velázquez L, Pérez-Juárez H, et al. The gut virome of healthy children during the first year of life is diverse and dynamic. PLoS One 2021;16(4):e0240958.

[147] Kaelin EA, Rodriguez C, Hall-Moore C, Hoffmann JA, Linneman LA, Ndao IM, et al. Longitudinal gut virome analysis identifies specific viral signatures that precede necrotizing enterocolitis onset in preterm infants. Nat Microbiol 2022;7(5):653–62.

[148] Li H, Wang H, Ju H, Lv J, Yang S, Zhang W, et al. Comparison of gut viral communities in children under 5 years old and newborns. Virol J 2023;20(1):52.

[149] Walters WA, Granados AC, Ley C, Federman S, Stryke D, Santos Y, et al. Longitudinal comparison of the developing gut virome in infants and their mothers. Cell Host Microbe 2023;31(2):187–98.e3.

[150] Byrne A, Diener C, Brown BP, Maust BS, Feng C, Alinde BL, et al. Neonates exposed to HIV but uninfected exhibit an altered gut microbiota and inflammation associated with impaired breast milk antibody function. Microbiome 2024;12(1):261.

[151] Maqsood R, Holland LA, Wu LI, Begnel ER, Adhiambo J, Owiti P, et al. Gut virome and microbiome dynamics before and after SARS-CoV-2 infec

[152] Zhang Y, Castro-Mejía JL, Deng L, Shah SA, Thorsen J, Leal Rodríguez C, et al. The influence of early life exposures on the infant gut virome. Gut Microbes 2025;17(1):2501194.

[153] Ren S, Zhang D, Shi X, Li T, Hu Q, Feng Y, et al. Gut virome characteristics associated with early onset of anemia and neurodevelopmental delay in preterm infants. iScience 2025;28(10):113578.

[154] Subramanian P, Romero-Soto HN, Stern DB, Maxwell GL, Levy S, Hourigan SK. Delivery mode impacts gut bacteriophage colonization during infancy. Gut Microbes Rep 2025;2(1):2464631.

[155] Kennedy KM, Gerlach MJ, Adam T, Heimesaat MM, Rossi L, Surette MG, et al. Fetal meconium does not have a detectable microbiota before birth. Nat Microbiol 2021;6(7):865–73.

[156] Panzer JJ, Romero R, Greenberg JM, Winters AD, Galaz J, Gomez-Lopez N, et al. Is there a placental microbiota? A critical review and re-analysis of published placental microbiota datasets. BMC Microbiol 2023;23(1):76.

[157] Sutcliffe SG, Reyes A, Maurice CF. Bacteriophages playing nice: lysogenic bacteriophage replication stable in the human gut microbiota. iScience 2023;26(2):106007.

[158] Silveira CB, Rohwer FL. Piggyback-the-winner in host-associated microbial communities. NPJ Biofilms Microbiomes 2016;2(1):16010.

[159] Knowles B, Silveira CB, Bailey BA, Barott K, Cantu VA, Cobián-Güemes AG, et al. Lytic to temperate switching of viral communities. Nature 2016;531(7595):466–70.

[160] Nanda AM, Thormann K, Frunzke J. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. J Bacteriol 2015;197(3):410–9.

[161] Boling L, Cuevas DA, Grasis JA, Kang HS, Knowles B, Levi K, et al. Dietary prophage inducers and antimicrobials: toward landscaping the human gut microbiome. Gut Microbes 2020;11(4):721–34.

[162] Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, et al. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 2010;466(7304):334–8.

[163] Andreu-Sánchez S, Bourgonje AR, Vogl T, Kurilshikov A, Leviatan S, Ruiz-Moreno AJ, et al. Phage display sequencing reveals that genetic, environmental, and intrinsic factors influence variation of human antibody epitope repertoire. Immunity 2023;56(6):1376–92.e8.

[164] Berkson JD, Wate CE, Allen GB, Schubert AM, Dunbar KE, Coryell MP, et al. Phage-specific immunity impairs efficacy of bacteriophage targeting vancomycin resistant enterococcus in a murine model. Nat Commun 2024;15(1):2993.

[165] Alsaied A, Islam N, Thalib L. Global incidence of necrotizing enterocolitis: a systematic review and meta-analysis. BMC Pediatr 2020;20(1):344.

[166] Warner BB, Deych E, Zhou Y, Hall-Moore C, Weinstock GM, Sodergren E, et al. Gut bacteria dysbiosis and necrotising enterocolitis in very low birthweight infants: a prospective case-control study. Lancet 2016;387(10031):1928–36.

[167] Olm MR, Bhattacharya N, Crits-Christoph A, Firek BA, Baker R, Song YS, et al. Necrotizing enterocolitis is preceded by increased gut bacterial replication, Klebsiella, and fimbriae-encoding bacteria. Sci Adv 2019;5(12):eaax5727.

[168] Chae H, Kim SY, Kang HM, Im SA, Youn YA. Dysbiosis of the initial stool microbiota increases the risk of developing necrotizing enterocolitis or feeding intolerance in newborns. Sci Rep 2024;14(1):24416.

[169] Neumann CJ, Mahnert A, Kumpitsch C, Kiu R, Dalby MJ, Kujawska M, et al. Clinical NEC prevention practices drive different microbiome profiles and functional responses in the preterm intestine. Nat Commun 2023;14(1):1349.

[170] Offersen SM, Mao X, Spiegelhauer MR, Larsen F, Li VR, Sandris Nielsen D, et al. Fecal virus-like particles are sufficient to reduce necrotizing enterocolitis. Gut Microbes 2024;16(1):2392876.

[171] Miligkos M, Oh J, Kwon R, Konstantinou GN, Kim S, Yon DK, et al. Epidemiology of asthma across the ages. Ann Allergy Asthma Immunol 2025;134(4):376–84.e13.

[172] Wang Y, Liu L. Immunological factors, important players in the development of asthma. BMC Immunol 2024;25(1):50.

[173] Miteva D, Lazova S, Velikova T. Genetic and epigenetic factors in risk and susceptibility for childhood asthma. Allergies 2023;3(2):115–33.

[174] Stokholm J, Blaser MJ, Thorsen J, Rasmussen MA, Waage J, Vinding RK, et al. Maturation of the gut microbiome and risk of asthma in childhood. Nat Commun 2018;9(1):141.

[175] Depner M, Taft DH, Kirjavainen PV, Kalanetra KM, Karvonen AM, Peschel S, et al. Maturation of the gut microbiome during the first year of life contributes to the protective farm effect on childhood asthma. Nat Med 2020;26(11):1766–75.

[176] Gogokhia L, Buhrke K, Bell R, Hoffman B, Brown DG, Hanke-Gogokhia C, et al. Expansion of bacteriophages is linked to aggravated intestinal inflammation and colitis. Cell Host Microbe 2019;25(2):285–99.e8.

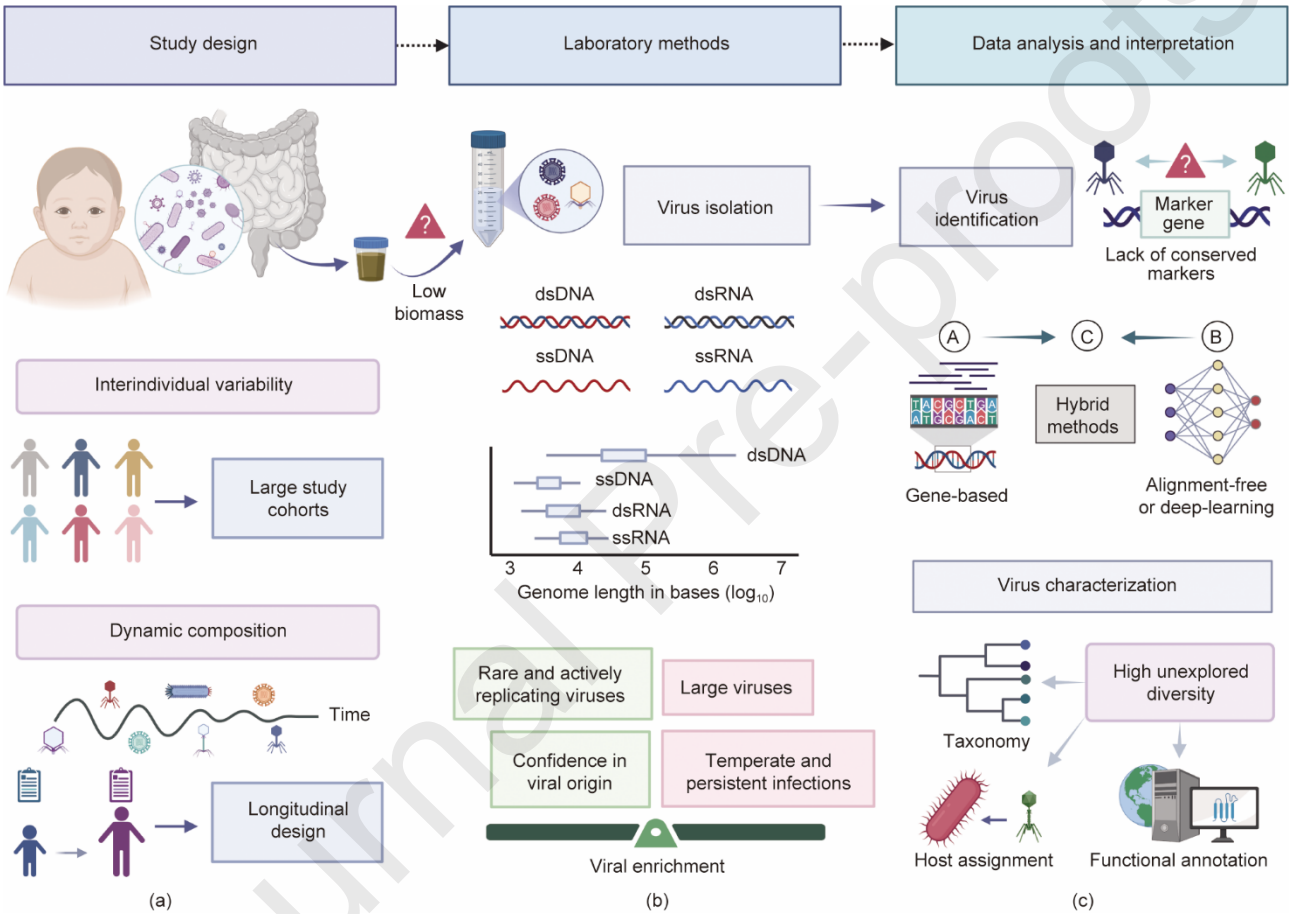
[177] Cao Z, Sugimura N, Burgermeister E, Ebert MP, Zuo T, Lan P. The gut virome: a new microbiome component in health and disease. EBio

[178] Ritz NL, Draper LA, Bastiaanssen TFS, Turkington CJR, Peterson VL, van de Wouw M, et al. The gut virome is associated with stress-induced changes in behaviour and immune responses in mice. Nat Microbiol 2024;9(2):359–76.

[179] Bichet MC, Chin WH, Richards W, Lin YW, Avellaneda-Franco L, Hernandez CA, et al. Bacteriophage uptake by mammalian cell layers represents a potential sink that may impact phage therapy. iScience 2021;24(4):102287.

[180] Wortelboer K, de Jonge PA, Scheithauer TPM, Attaye I, Kemper EM, Nieuwdorp M, et al. Phage-microbe dynamics after sterile faecal filtrate transplantation in individuals with metabolic syndrome: a double-blind, randomised, placebo-controlled clinical trial assessing efficacy and safety. Nat Commun 2023;14(1):5600.

[181] Mao X, Larsen SB, Zachariassen LSF, Brunse A, Adamberg S, Mejia JLC, et al. Transfer of modified gut viromes improves symptoms associated with metabolic syndrome in obese male mice. Nat Commun 2024;15(1):4704.



Declaration of Interest Statement

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for this journal and was not involved in the editorial review or the decision to publish this article.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: