Title: Phageome transfer from gut to circulation and its regulation by human immunity

Aleksander Szymczak[1,8*], Katarzyna Gembara[1,3], Stanisław Ferenc[3], Joanna Majewska[1], Paulina Miernikiewicz[1], Marek Harhala[7], Izabela Rybicka[1], Dominik Strapagiel[2], Marcin Słomka[2], Jakub Lach[2], Jan Gnus[3,4], Paweł Stączek[5], Wojciech Witkiewicz[3], Matlock Jeffries[8], Krystyna Dąbrowska[1,3,6*]

1. Laboratory of Phage Molecular Biology, Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wroclaw, Poland
2. Biobank Lab, Department of Oncobiology and Epigenetics, University of Lodz, Lodz, Poland
3. Research and Development Center, Regional Specialist Hospital, Wroclaw, Poland
4. Wroclaw Medical University, Wroclaw, Poland
5. Department of Molecular Microbiology, Faculty of Biology and Environmental Protection, University of Lodz, Lodz, Poland
6. Faculty of Medicine, Wroclaw University of Science and Technology, Wroclaw, Poland
7. Institute of Human Biology and Evolution, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland
8. Arthritis & Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, USA

* correspondence to aleksander.szymczak@hirszfeld.pl and krystyna.dabrowska@hirszfeld.pl

Abstract

Whether phage signals in blood primarily reflect spillover from the gut reservoir or strong compartmental filtering remains unresolved at the scale of an entire phageome. We performed shotgun metagenomics on matched colon mucosa and serum from 37 adults, defined translocated vOTUs (viral Operational Taxonomy Units) as those detected in both compartments within the same individual, and contrasted them with intestine-only signals. Total phage abundance was markedly lower in serum than mucosa (98% loss; 3.43×10^8 vs 7.19×10^6 TPM). Most translocated vOTUs lacked UHGV(Unified Human Gut Virome Catalog)-based assignment (1,205/1,295; 93.1%), yet unassigned vOTUs formed reproducible co-enrichment links to assigned “anchor” families in a serum association network. Translocated vOTUs differed from intestine-only vOTUs in length and GC content, and compartment overlap among annotated proteins was highest for virion structural modules (923/2,861 CDS records; 32%). IgG-reactive protein groups frequently lacked metagenomic detection and IgG readouts showed weak associations with an intestine-to-serum ratio score. An oral T4 tracer experiment in mice (n=8) revealed stepwise attrition from gut content to blood (<10 PFU/mL). These results support a model in which systemic detectability is constrained by sequential barriers and loss across compartments.

Introduction

Bacteriophages are prominent components of the human gut virome, yet their fate beyond the intestinal mucosa remains poorly understood. Phages have been detected in blood and other extra-intestinal sites, with proposed routes including transcytosis across epithelial barriers and vesicular transport (Smith et al. 2023; Rybicka and Kaźmierczak 2025; Zhang et al. 2025). Once in circulation, phages encounter humoral immunity; phage-specific antibodies have been documented, but whether immune recognition influences which phages reach or persist systemically is unclear (Abed, Hasan, and Albayati 2025; Bichet et al. 2023; Nguyen et al. 2017).

A central limitation is the absence of paired intestinal–systemic sampling within individuals. Studies have characterized gut phageomes or circulating viromes separately, but without matched compartments it is difficult to determine which phages cross epithelial barriers, what proportion survives transit, or how antibody responses relate to systemic detection. The predominance of unannotated sequences in gut viromes further complicates efforts to link translocation to specific taxonomic or functional features.

Here we analyse phage translocation using matched mucosal biopsy–serum pairs from 37 individuals, combining metagenomic profiling with serological detection of phage-specific immunoglobulin G (IgG). Detectable phage sequence abundance in serum was reduced by approximately 98% relative to intestinal mucosa under the compartment-specific workflows used (VLP-enriched mucosa versus total serum DNA), and translocated populations were dominated by unannotated sequences. Structural virion proteins were enriched among translocated cargo. IgG reactivity targeted surface-exposed components, yet most IgG-reactive protein groups lacked metagenomic detection in either compartment—consistent with durable immune memory or epitope cross-reactivity. An in vivo tracer model confirmed stepwise attrition of orally administered phage across anatomical barriers, with blood titres near detection limits.

These findings indicate that phage translocation reflects stringent anatomical filtering rather than passive spillover from intestinal abundance and suggest that immune recognition may contribute to shaping which phages are detectable systemically.

Materials and Methods

Study design and participants

Matched colon mucosal biopsies and peripheral blood were collected from adult patients

undergoing diagnostic colonoscopy at the Endoscopy Laboratory of the Regional Specialist Hospital in Wrocław, Poland (Fig. 1). Participants were recruited based on gastrointestinal symptoms prompting colonoscopy. Individuals reporting antibiotic treatment within the preceding 6 months or diabetes treated with metformin were excluded. Intestinal biopsies were collected from regions showing no abnormalities on endoscopic examination. The cohort comprised 37 participants (21 women, 16 men; mean age 61 years; Supplementary Figure S1). The study was conducted in accordance with the Declaration of Helsinki and approved by the local Bioethical Commission of the Regional Specialist Hospital in Wrocław (KB/nr8/rok2017). Written informed consent was obtained from each participant.

VLP enrichment and DNA preparation

Biopsies were incubated in PBS for 2 h and filtered through 0.22 μm membrane filters. Viral particles were enriched by CsCl density-gradient ultracentrifugation (step gradient: 1.7, 1.5, 1.35, and 1.15 g ml□¹) performed overnight at 62,000 × g at 4 °C in a swinging-bucket rotor. A 0.5 ml fraction between 1.35 and 1.5 g ml□¹ was collected by syringe for DNAse I treatment (A&A Biotechnology). Phage DNA isolation (Sherlock AX kit, A&A Biotechnology), was followed by whole-genome amplification (GenomiPhi V2, Cytiva).
Whole blood was centrifuged at 2,000 RCF for 10 min at 4 °C. The serum layer was collected for total DNA isolation (Sherlock AX kit, A&A Biotechnology), followed by whole-genome amplification (GenomiPhi V2, Cytiva) before sequencing library preparation.

Library preparation and sequencing

Shotgun libraries were prepared using a Vazyme DNA library preparation kit and sequenced as paired-end 2×150 bp on Illumina NextSeq 500/550 using the v2.5 Kit (Medium) at two facilities. Thirty-seven matched intestine–serum pairs were analyzed by shotgun sequencing. QC details after sequencing for samples available in Supplementary file – Sequences Quality Control.

Read processing and host read removal

FASTQ files were quality-filtered and adapter-trimmed using fastp v0.24.2. Host reads were removed by alignment to human GRCh38.1 (UCSC) using Bowtie2 v2.4.5 with `--very-sensitive`; unmapped reads were retained.

Assembly, vOTU definition, and abundance estimation

Non-host reads were assembled using metaSPAdes v4.20. Contigs <1 kb were excluded. Viral contigs were identified using VirSorter2 v2.2.4. vOTUs were defined by clustering contigs using CheckV v1.0.3 average nucleotide identity (ANI)-based clustering at 95% with 85% (AF) alignment fraction; representative sequences were selected per cluster prioritizing higher CheckV completeness, with ties broken by longer length. Reads were mapped to vOTU representatives with Bowtie2 (--very-sensitive). Alignments were filtered at mapping quality ≥30. Per-vOTU counts were converted to TPM: TPM□ = ((count□/length□)/Σ□(count□/length□))×10□. A vOTU was considered detected in a compartment only if breadth of genome coverage was ≥10%, and this detection filter was applied before defining within-subject status. For subject-level comparisons, vOTU abundance was summed separately across all intestine-derived samples and all serum-derived samples per subject. vOTUs were labeled "translocated" when summed intestine abundance >0 and summed serum abundance >0 within the same subject, and "only gut presence" when summed intestine abundance >0 and summed serum abundance =0.

To assess whether serum detection primarily reflected matched-pair biology rather than sporadic contamination, we quantified serum–intestine concordance and tested it against a pairing null. Of 2,488 vOTUs detected in serum across the cohort, 2,436 (97.9%) were also detected in at least one matched intestine sample, whereas 52 (2.1%) were detected only in serum. Moreover, within-subject serum–intestine overlap exceeded a permuted pairing null (mean overlap proportion 0.464 observed versus 0.405 expected; permutation $P < 0.001$; Supplementary Table S5), supporting that serum detections are non-random with respect to subject pairing.

Taxonomy, host context, and lifestyle prediction

Taxonomic assignment was performed using against UHGV which contains collections from multiple databases (Metagenomic Gut Virus Compedium, Gut Phage Database, Metagenomic Mobile Genetic Elements Database, IMG Virus Resource v4, Hadza Hunter Gatherer Phage Catalog , Cenote Human Virome Database, Human Virome Database, Gut Virome Database, Atlas of Infant Gut DNA Virus Diversity, Circular Gut Phages from NCBI, Danish Enteric Virome Catalogue, Stability of the human gut virome and effect of gluten-free diet) with assignments requiring 95% ANI and 85% AF. Predicted lifestyle was obtained using BACPHLIP v0.96 with integrase annotations from PHROG and prophage predictions from geNomad v1.11.0.

Gene prediction and functional annotation

Protein-coding genes were predicted using Prodigal v2.6.3. Functional annotation was performed with eggNOG-mapper v2.11.2 (Cantalapiedra et al., 2021) and, in parallel, phage-focused

annotation using Pharokka v1.7.4 with PHROGs. KO/pathway summaries used mean values across CDS mapped to the same KO/pathway. Defense system calls used DefenseFinder v2.0.1.

IgG-reactivity library and readout

A phage-derived peptide library was designed from GenBank phage ORFs (accessed June 1, 2018) using inclusion keywords (‘structural’, ‘tail’, ‘capsid’, ‘baseplate’, ‘holin’, ‘lysin’, among others) and excluding ORFs with keywords ‘initiator’, ‘chaperone’, ‘maturation protease’, and ‘morphogenesis’. ORFs were tiled into 56-aa peptides starting every 28 aa (50% overlap), deduplicated, reverse-translated using E. coli–optimized codons, and synthesized (Agilent SurePrint), yielding 277,051 oligopeptides. The library was cloned into a T7 phage display system (T7Select 415-1, Merck Millipore) using Liga5 (A&A Biotechnology) in place of T4 ligase with EcoRI and HindIII sites. Immunoprecipitation and sequencing were performed following a modified protocol based on Xu et al., 2015 (details in Supplementary Methods). Sequencing reads were annotated using DIAMOND with BLASTx against a database built from the oligonucleotide FASTA. IgG-reactivity was quantified as raw reads per peptide; a protein was considered IgG-reactive if any peptide mapping to that protein had non-zero reads, and higher read counts indicated higher reactivity. Scheme presented in Figure 2, detailed description available in Supplementary Methods section. Library coverage relative to the metagenomic dataset was summarized post hoc by overlap between PhageScan-positive calls and metagenome-detected protein groups and host annotations (Supplementary Table S6).

Translocation score

Translocation score was defined as an **intestine-to-serum ratio score**, computed as **log2((intestine+1)/(serum+1))**. Under this definition, **lower values indicate comparatively higher relative serum abundance**, whereas higher values indicate higher relative intestinal abundance. Associations between IgG reactivity (raw reads) and the translocation score were evaluated using Spearman correlation and a linear mixed-effects model with translocation score as the outcome and IgG reads as a fixed effect, including a random intercept for subject; the implementation additionally included a category variance component (0 + C(category)) when enabled in the analysis script. Model fitting used restricted maximum likelihood (REML) with the L-BFGS optimizer.

Statistical analysis

All tests were two-sided unless otherwise stated. For comparisons of intestinal abundance between vOTUs classified as Transcytosed versus Only gut presence within taxonomic or host

categories, we used Wilcoxon rank-sum tests on log10(TPM+1) values and controlled multiple testing using Benjamini–Hochberg false discovery rate (BH-FDR) within each analysis family (ICTV family set; host-class set). For protein-group analyses comparing Translocated versus Only gut presence within subjects, we computed per-subject medians and applied paired Wilcoxon signed-rank tests, with BH-FDR correction across protein groups. Effect sizes are reported as median differences (Δmedian) and the log2 ratio of medians (log2[median_trans/median_not]) alongside P values and FDR (Supplementary Table S3) With 37 paired subjects, the study has sensitivity primarily for moderate-to-large paired effects; smaller effects may remain undetected, particularly after multiple-testing correction.

Animal model of phage translocation

C57BL/6J male mice (n=8) were maintained under SPF conditions. Experiments followed EU directive 2010/63/EU, were approved by the 1st Local Committee for Experiments with the Use of Laboratory Animals (Wrocław, Poland; project no. 76/2011) and followed ARRIVE guidelines. T4 phage (ATCC) was propagated on E. coli B, purified, and administered in drinking water (5×10□ pfu ml□¹). After 26 h, gut content, intestinal mucus, mesenteric lymph nodes, and blood were collected and phage titers quantified by double-layer plaque assays. PBS-treated controls (n=3) were processed in parallel.

Data and code availability

Sequencing Data is available in Sequencing Read Archive - PRJNA874475. Code for figures, and statistical analysis in available on github (Link). IgG-reactivity library available upon request.

Results

Intestinal mucosal viromes are dominated by Microviridae and Inoviridae

Intestinal viral operational taxonomic unit (vOTU) abundance, normalized as transcripts per million (TPM), was highly skewed (0.66–7.76×10^7 TPM; median 40.2 TPM; 90% range 3.45–3.63×10^3; Figure 3A–D). Within phage family–assigned intestinal TPM, Microviridae (48.9%) and Inoviridae (42.4%) dominated, followed by Straboviridae (4.3%) and Herelleviridae (3.3%), with all remaining families contributing <1%. Host-linked composition separated these signals: Bacteroidia-assigned TPM was almost entirely Microviridae (99.9%), whereas Inoviridae-assigned TPM partitioned between Clostridia (67.7%) and Bacilli (29.3%) (Figure 3B–D). At the order level, assigned intestinal TPM mapped mainly to Petitvirales (53.2%) and Tubulavirales (46.1%), with Crassvirales contributing 0.75% (Figure 3C). The abundance of bacteriophages in

the intestine was not a determining factor in transcytosis. Family-wise abundance differences were dominated by a small subset of groups after BH-FDR correction (for example, Straboviridae and Herelleviridae; Supplementary Table S3), whereas other families showed nominal but non-FDR-significant shifts (for example, Salasmaviridae).

Across 37 matched intestine–serum pairs, 29,733 proteins showed gut-only detection (intestine present, serum absent), spanning 301 protein groups; these counts reflected CDS calls pooled across subjects (Figure 4). Most gut-only proteins were annotated as unknown function (86.1%), leaving 4,126 annotated proteins (14%). Among annotated gut-only proteins, DNA/RNA/nucleotide metabolism (33.5%) and moron/auxiliary metabolic gene and host takeover (18.5%) were most common, followed by head/packaging (7.9%) and integration/excision (7.0%). Gut-only abundance summaries (log10[TPM]) highlighted frequently detected replication- and toxin-associated annotations, including replication initiation protein (36/37 subjects; median 0.952) and a RelE-like toxin (35/37; 0.850), alongside structural categories such as major head protein (36/37; 0.628) (Figure 4).

Taxonomic patterns among translocated vOTUs

Among phage family-assigned translocated abundance, Microviridae and Inoviridae dominated (8.69×10^6 TPM combined; Figure 3A,D). Because phi29 whole-genome amplification (WGA) can distort virome composition, we performed a computational assessment of length-dependent bias. Across subjects, vOTU contig length was not inversely associated with intestinal abundance; instead, within-subject Spearman correlations between log10(TPM+1) and log10(length) were weakly positive (median $\rho = 0.156$; IQR 0.117–0.215; 33/37 subjects positive; Supplementary Table S4). In addition, Microviridae vOTUs were not enriched among the shortest sequences (median length 5,786 bp versus 3,436 bp for other annotated families; only 10.9% of Microviridae fell in the shortest-length quartile; Supplementary Table S4). Together, these patterns are not consistent with a dominant global length-driven amplification artifact explaining Microviridae prominence. In family-wise comparisons of intestinal abundance, Straboviridae and Herelleviridae had significantly higher intestinal abundance among translocated versus intestine-only observations (medians 1,296–1,435 vs 60–87 TPM; FDR<0.01; Figure 3A). At order rank, Petitvirales and Tubulavirales dominated classified translocated abundance (Figure 3C).Host-associated patterns differed by group: unknown-host vOTUs had 21-fold higher intestinal abundance when translocated versus gut-only (FDR=2.03×10^−17; Figure 3B), and among host-assigned translocated abundance, Bacteroidia- and Clostridia-targeting phages were most prominent (Figure 3B–C). Inoviridae translocated predominantly via Bacilli hosts (94.8%; Figure 3D).

Functional composition of translocated proteins and contigs

Among module-assigned structural proteins, 32% of CDS records (923/2,861) were detected in both intestine and serum (Figure 4A). Translocation frequency ranged from 7.8% (Lysis) to 60.2% (Baseplate), with intermediate frequencies for Tail (24.9%), Capsid (40.1%), and other Structural groups (36.3%; Figure 4A). No protein group reached false discovery rate (FDR) significance (minimum q=0.521; Figure 4A), and the most frequently observed translocated protein groups were dominated by structural annotations (e.g., structural protein, tail fiber protein, capsid/scaffold, and baseplate) (Supplementary Table S1).

Translocated contigs with EggNOG-derived annotations mapped to 29 Kyoto Encyclopedia of Genes and Genomes (KEGG) orthology (KO) pathway identifiers across 10 phage families (102 pathway–family entries) in the KO-assigned subset (Figure 4B). Across this subset, the highest-abundance pathway identifiers were led by biosynthesis of antibiotics (ko01130), beta-lactam resistance (ko01501), and penicillin and cephalosporin biosynthesis (ko00311), followed by metabolic pathways (ko01100) and biosynthesis of secondary metabolites (ko01110) (Figure 4B). Straboviridae contributed the largest share of KO-assigned abundance (2,053.6 TPM; 36% of the KO-assigned total), and Clusters of Orthologous Groups (COG) categories were dominated by COG E (amino acid transport and metabolism) and COG V (defense mechanisms), together accounting for 58% of pathway–family entries (Figure 4B). DefenseFinder annotations comprised 142 total hits (104 defense; 38 anti-defense), with restriction–modification (RM; 50 hits) and NADP-associated types (22 hits) most frequent; putative AMR subclass annotations totaled 280 hits, dominated by beta-lactam- (169; 60%) and aminoglycoside-associated labels (Kanamycin; 61; 22%) (Figure 4C)(Supplementary Table S2)

IgG-reactivity versus compartment detection and translocation score

Immunoglobulin G (IgG) reactivity showed weak associations with the intestine-to-serum ratio score (log2((intestine+1)/(serum+1))), where lower values indicate comparatively higher relative serum abundance (phage host: ρ=-0.26, p=0.093; lifestyle: ρ=-0.09, p=0.57; protein function: ρ=0.14, p=0.20; Figure 5D). IgG-reactive protein groups overlapped with compartment-detected groups (32/163 intestine-detected; 26/95 serum-detected), but most IgG-reactive groups (102/134; 76.1%) lacked metagenomic detection in either compartment (Figure 5A). Among the 26 groups detected in intestine, serum and IgG, the highest IgG reactivity (mean raw reads across peptides 1,877–7,207) mapped to structural and tail annotations (e.g., virion/putative structural proteins and major capsid; Figure 5B) and spanned a broad range of mean TPM-normalized abundances (intestine 0.38–12.11; serum 0.31–10.81; Figure 5B). IgG-only groups also included strongly reactive tail-fibre categories (e.g., putative short tail fibre, tailfiber like protein) (Figure 5A–B). When stratified by predicted lifestyle, minor tail/capsid and major tail functions showed higher aggregated IgG in temperate predictions, whereas collar, tail sheath, major capsid, head–tail and baseplate-associated labels were higher in virulent predictions (Figure 5C). To quantify how well the our peptide library recapitulated signals observed in the current metagenomic vOTU set, we summarized overlap ("sensitivity") at multiple annotation levels. Among

metagenomically detected protein groups, 32/163 (19.6%) were also detected by PhageScan; at the predicted host genus level, 4/132 (3.0%) overlapped; and at normalized host labels, 14/109 (12.8%) overlapped (Supplementary Table S6). These estimates indicate that PhageScan captures a measurable subset of contemporaneous metagenomic signals while remaining incomplete, consistent with database growth since 2018 and with substantial viral dark matter in paired viromes.

Sequencing dark matter among translocated vOTUs and sequence-level signatures

Translocated vOTUs were dominated by sequencing dark matter: 1,205/1,295 (93.1%) lacked a confident assignment against the Gut Phage Database (GPD; 142,809 non-redundant gut phage genomes from 28,060 metagenomes), leaving 90 vOTUs as taxonomically "known" anchors (Figure 6A, D). Unassigned vOTUs nonetheless formed reproducible co-enrichment links to phage-assigned vOTUs in the serum association network, with connectivity concentrated around Herelleviridae, Straboviridae, Peduoviridae, and Inoviridae (each connected to 90–124 unique dark-matter partners; Figure 6A, D).

CpG residuals (observed/expected ratio after GC and length adjustment) varied widely, with 26 outliers clipped for display (Figure 6B). Sequence properties differed between translocated and intestine-only vOTUs: translocated vOTUs were longer (median 5,508 vs 3,265 bp; $p=5.2\times10^{-}$), had higher GC content (45.0% vs 41.9%; $p=1.1\times10^{-2}$), and higher mono-nucleotide Shannon entropy ($p=2.3\times10^{-}$), while open reading frame (ORF) density did not differ (p=0.85) (Figure 6C). Protein annotation remained incomplete, with 63.2% of translocated proteins labeled "hypothetical," and the most frequent specific labels mapping to virion structural components (e.g., tail fiber, baseplate, head proteins).

In vivo phage retention across compartments

As an in vivo test, T4 phage (Tequatrovirus; $5\times10^{-}$ PFU/mL) (Plaque Forming Units/milliliter) was administered continuously in drinking water to mice (n=8), and phage titers were quantified after 26 hours across four compartments (Figure 7). Mean titers decreased from gut content ($2.3\times10^{-}$ pfu/g) to intestinal mucus ($1.8\times10^{-}$ pfu/g), mesenteric lymph nodes (9.2×10^{2} pfu/g), and blood (<10 PFU/mL) (Figure 7). This corresponded to stepwise reductions of ~10-fold (gut content to mucus), ~10^5-fold (gut content to lymph nodes), and ~10^7-fold (gut content to blood). Blood titers were near the stated detection limit (<10 PFU/mL), while lymph node titers remained >10^5-fold lower than gut content (Figure 7, Supplementary Figure S2).

Discussion

Profiling of gut phage translocation across matched intestinal–serum samples shows that systemic detection reflects strong anatomical filtering rather than simple spillover from

intestines. Across the cohort, phage abundance decreased by ~98% from intestinal mucosa to serum (Figure 3A–B), and translocated vOTUs were dominated by unannotated sequences (93.1%; Figure 6A, D). In parallel, an in vivo tracer model showed stepwise attrition of orally administered T4 across compartments with blood titers near the detection limit after continuous exposure (Figure 7). Together, these data are consistent with low-level entry coupled to substantial loss across downstream compartments, limiting sustained detectability in blood (Nguyen et al. 2017; Berkson et al. 2024).

The magnitude of phage attrition is compatible with epithelial passage representing a major quantitative bottleneck. The ~98% reduction in detectable phage sequence abundance from mucosa to serum under the applied workflows (Figure 3A–B) is consistent with experimental systems in which only a small fraction of applied phages traverse epithelial layers over hours (Nguyen et al. 2017; Douadi et al. 2025). The in vivo titer gradient (Figure 7) further indicates that translocation is shaped not only at the epithelium but also by downstream checkpoints, including lymphatic filtration and systemic clearance mechanisms that can rapidly remove circulating phage (Berkson et al. 2024; Roozendaal, Mebius, and Kraal 2008). These constraints provide a parsimonious explanation for why serum viromes can be reproducibly detectable yet remain orders of magnitude lower than gut reservoirs.

Taxonomic and functional determinants of translocation

At the intestinal mucosa, the virome was dominated by Microviridae (48.9%) and Inoviridae (42.4%) (Figure 3A–D), families that are well represented in gut phage catalogues and are frequently detected across cohorts (Camarillo-Guerrero et al. 2021; Gregory et al. 2020). Filamentous Inoviridae proliferate without lysing host cells and can affect host fitness (Mai-Prochnow et al. 2015) , aligning with the pronounced partitioning of Inoviridae between Firmicutes classes reported herein (Figure 3B–D). By contrast, Crassvirales contributed 0.75% (Figure 3C), despite their reported abundance in many gut studies, suggesting that sampling context (for example, mucosal versus luminal) and/or cohort features may shift the balance among dominant phage groups (Smith et al. 2023; Zuo et al. 2019).

Functional annotations differed between intestine-only proteins and the subset showing compartment overlap. The large number of gut-only proteins is consistent (29,733 proteins; 301 groups; Figure 4) indicates that many prophage-associated functions remain localized to the gut environment, while 32% of structural protein-coding sequences were detected in both intestine and serum (Figure 4A), consistent with virion-linked passage across epithelial layers (Nguyen et al. 2017) Variation across structural modules (baseplate 60.2%, capsid 40.1%, tail 24.9%, lysis 7.8%; Figure 4A) is compatible with differences in detectability, stability, and/or annotation distributions across contigs meeting serum detection criteria. (Wang et al. 2019).

Functional enrichment signals among translocated contigs (Figure 4B–C) should be interpreted conservatively. In particular, KEGG terms linked to antibiotic resistance and the putative presence of AMR-like subclass hits (Figure 4C) do not, on their own, establish phage-encoded resistance, given evidence that virome pipelines frequently overcall ARGs and that many apparent hits reflect false positives or non-resistance homologs (Enault et al. 2017). These observations instead motivate targeted validation to distinguish bona fide phage functions from bacterial DNA fragments, prophage-adjacent regions, or annotation artifacts.

Dark matter dominance and sequence signatures of translocated vOTUs

An important part of the paired-compartment analysis is that most translocated vOTUs lacked confident database assignment (93.1%; Figure 6A,D), mirroring the wider "viral dark matter" problem in gut viromics (Fitzgerald et al. 2021; Santiago-Rodriguez and Hollister 2022). Network structure nonetheless provided reproducible anchors: taxonomically assigned families formed connectivity hubs linked to many dark-matter partners (Figure 6A,D), suggesting that co-enrichment relationships can give background for unassigned sequences when taxonomic resolution is limited among functionally related groups even when precise taxonomy is absent (Fitzgerald et al. 2021; Santiago-Rodriguez and Hollister 2022). Sequence features further separated translocated from intestine-only vOTUs (Figure 6B–C), including shifts in length and GC content (Figure 6C) and broad heterogeneity in CpG residuals (Figure 6B). Because CpG content and context have been linked to CpG-mediated immunogenicity in recent large-scale analyses, these patterns highlight nucleotide composition as a candidate correlate of systemic detectability, without establishing mechanism (Kharrat et al. 2025).

Immune recognition and clearance dynamics

Serology indicated IgG reactivity to phage-derived peptides, often in the absence of shotgun sequencing observation. Most IgG-reactive protein groups lacked metagenomic detection in either compartment (Figure 5A), consistent with durable humoral memory and/or cross-reactive epitopes.[2] (Kaźmierczak et al. 2021; Brzozowska et al. 2023). IgG reactivity was enriched for structural and tail-associated proteins (Figure 5B), aligning with immunodominance of abundant, surface-exposed virion components (Kaźmierczak et al. 2021; Le et al. 2025). The weak negative association between IgG reactivity and translocation score across host categories (Figure 5D) is compatible with partial neutralization and accelerated clearance of opsonized phage material before it accumulates in serum (Berkson et al. 2024). These dynamics have practical relevance: phage uptake and phage-specific immunity can act as sinks that reduce systemic persistence and may contribute to variability in outcomes for orally delivered or gut-targeted phage interventions (Bichet et al. 2021). The observation that epithelial transit can enhance immune activation in model systems further emphasizes that "how" phages cross may influence "what" the immune system sees (Carroll-Portillo et al. 2025).

Limitations and outlook

This study is constrained by the same factors that limit mechanistic inference across metagenomic viromics: high fractions of unclassified vOTUs and hypothetical proteins (Figure 6; Figure 4) restrict identification of the molecular determinants of epithelial uptake and immune recognition, and short-read data cannot always separate integrated prophage DNA from encapsidated genomes in ways that are directly actionable for lifecycle inference. In serum, differences in recovery of DNA versus proteins and rapid immune clearance can further contribute to non-detection. Future work should incorporate extraction blanks and synthetic spike-in standards processed alongside samples to quantify cross-sample carryover and establish sensitivity limits for low-biomass serum viromes. Finally, the in vivo tracer model focuses on a single phage and defined exposure window; extending the framework to broader phage diversity will be important for generalization.

Initial priorities are derived straight from the results. First, spatially resolved sampling (paired luminal–mucosal) combined with long-read sequencing would clarify how niche structure shapes the pool available for barrier passage. Second, mechanistic dissection using isolate panels and epithelial/colonoid systems can test whether particle architecture, surface properties, and receptor-binding modules predict transcytosis efficiency (Nguyen et al. 2017; Le et al. 2024). Third, focused functional validation—particularly of dark-matter proteins and ambiguous metabolic/AMR-like annotations—will be essential to connect sequence features to translocation competence and immune outcomes (Enault et al. 2017; Santiago-Rodriguez and Hollister 2022).

Conclusions

This study demonstrates that phage signals identified in blood are primarily indicative of selective passage through the intestinal barrier, followed by subsequent systemic elimination, rather than direct leakage from gut levels. By coupling matched colon mucosa–serum metagenomics with IgG profiling and an oral T4 tracer model, we outline a stepwise filtration framework in which only a small, compositionally biased fraction of gut phage material reaches—and persists in—circulation. As summarized in Figure 8., we further propose that low-frequency epithelial passage (Step 1) is followed by lymphatic redirection to draining lymph nodes (Step 2), a checkpoint that may promote mucosal tolerance and limit sustained serum detectability, with systemic persistence further constrained when a subset of particles elicits IgG responses (Step 3) that accelerate neutralization and clearance (Step 4). Enrichment of compartment overlap among virion structural modules, together with distinct nucleotide signatures in translocated sequences, is consistent with particle-linked features contributing to barrier passage, while predominant IgG reactivity to structural and tail proteins supports immune recognition as an additional sink shaping systemic detectability. Although extensive viral dark matter and short-read constraints limit assignment of specific uptake factors and separation of encapsulated genomes from integrated prophage DNA, the paired-compartment design narrows

candidate features for mechanistic follow-up. Next, tracking phage biodistribution from mucosa to lymph nodes and testing defined phage isolate panels in epithelial or colonoid systems—paired with IgG specificity/neutralization assays—can determine whether virion architectures or receptor-binding modules predict transcytosis efficiency and subsequent clearance.

Acknowledgments

This work was supported by National Science Center (Poland) grants: Preludium16 UMO-2018/31/N/NZ6/02584, Opus18 UMO-2019/35/B/NZ7/01824, and by the subsidy for scientific activity provided to the Department of Molecular Microbiology by the University of Lodz (B2111000000038.01).

Conflicts of interest

Authors declare no conflicts of interest.

Figures and Tables

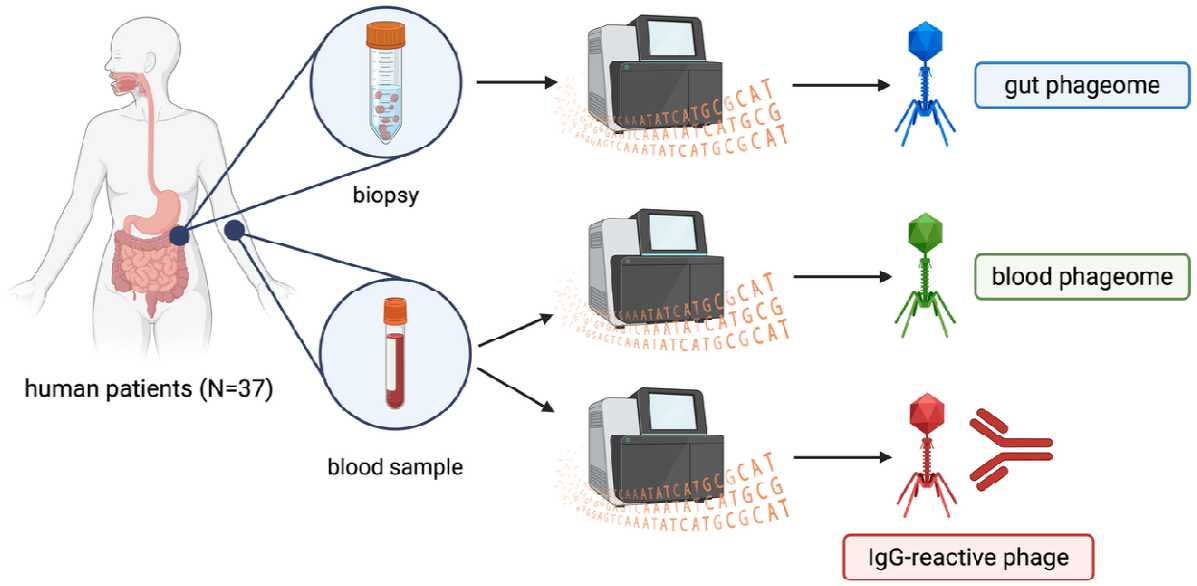


Figure 1. Human samples used in this study to investigate composition of gut and blood phageomes, and IgG antibody responses to the phageomes (BioRender).

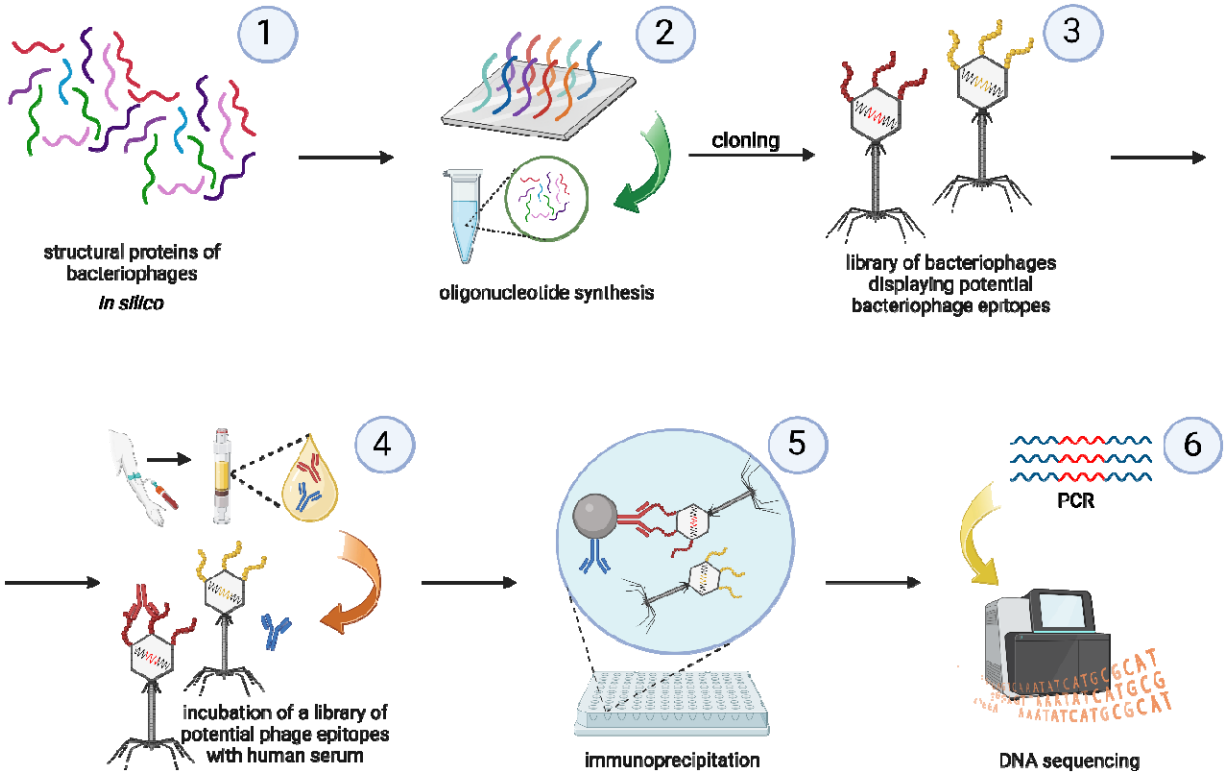


Figure 2. PhageScan technology (adapted from VirScan, Xu et al. (2015) (BioRender). Step 1: Structural protein sequences from bacteriophage genomes were computationally fragmented into overlapping 56-amino acid oligopeptides (28 aa offset). Step 2: Oligonucleotide sequences encoding 277,051 unique peptides were synthesized using SurePrint technology (Agilent). Step 3: Oligonucleotides were cloned into T7Select 415-1 vectors to generate a phage display library presenting potential bacteriophage epitopes. Step 4: The amplified phage library was incubated with human serum samples to allow antibody-epitope binding. Step 5: Antibody-bound phages were captured using Protein A/G magnetic beads and isolated by immunoprecipitation. Step 6: Enriched phage inserts were PCR-amplified and identified by Illumina sequencing with subsequent DIAMOND annotation against the reference oligopeptide database.

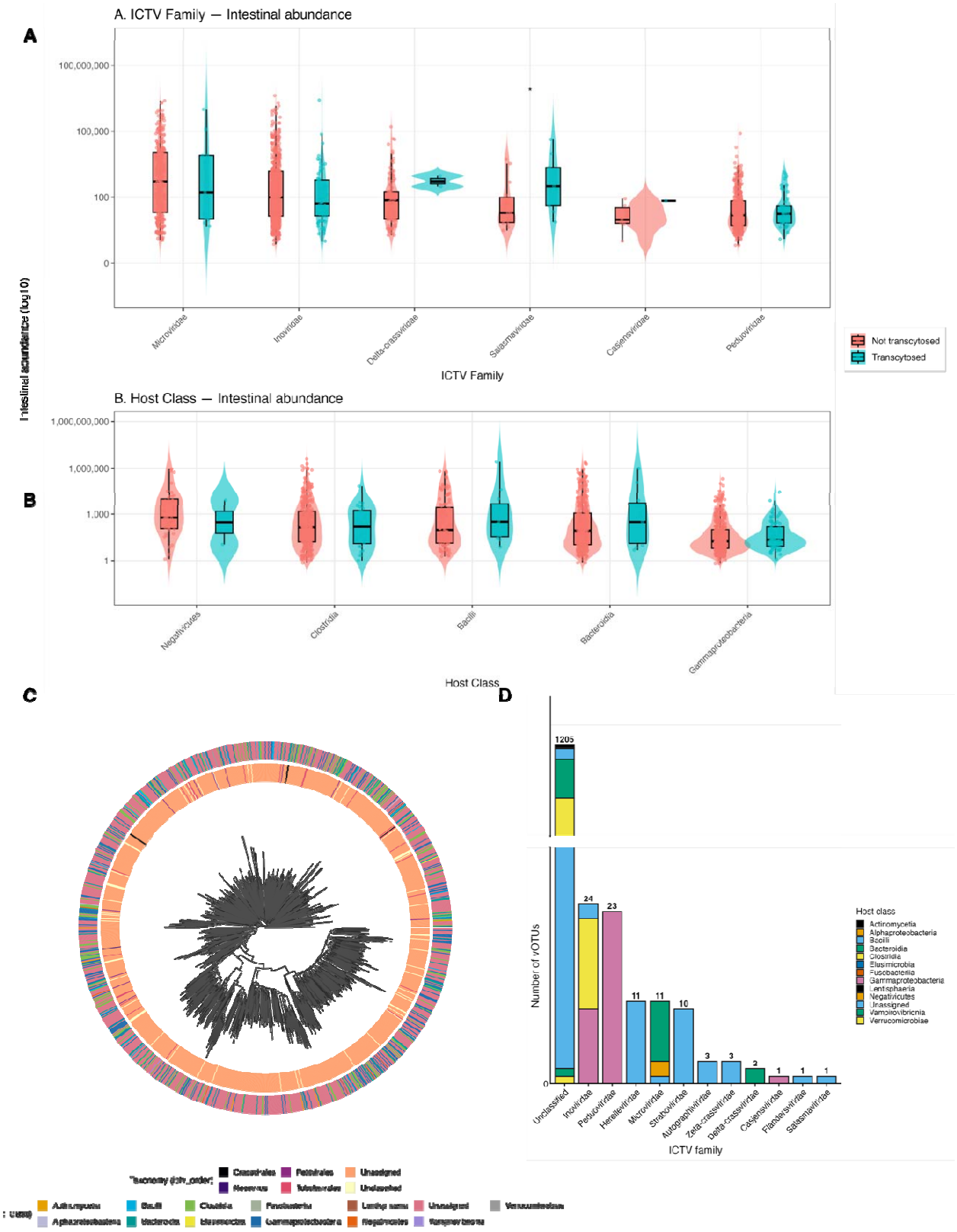


Figure 3. Taxonomic and host context of vOTUs detected in both intestine and serum. Intestinal abundance (log10[TPM+1]) for vOTUs grouped by ICTV family (A) or predicted host

class (B), stratified by detection status within subjects: **Transcytosed** (abundance >0 in both intestine and serum) versus **Only gut presence** (intestine >0, serum =0). Points represent individual vOTUs; boxplots show the median (center line), interquartile range (box), and 1.5×IQR whiskers. Two-sided Wilcoxon rank-sum tests compare Transcytosed versus Only gut presence within each category, with Benjamini–Hochberg false discovery rate (BH-FDR) correction across categories; n values per group are shown on the plot and summarized in Supplementary Table S3. (C) Circular phylogeny of the filtered vOTU set annotated by ICTV order and host class. (D) Stacked bar plot of vOTU counts per family (axis break), filled by host class; totals are labeled. vOTU, viral operational taxonomic unit.

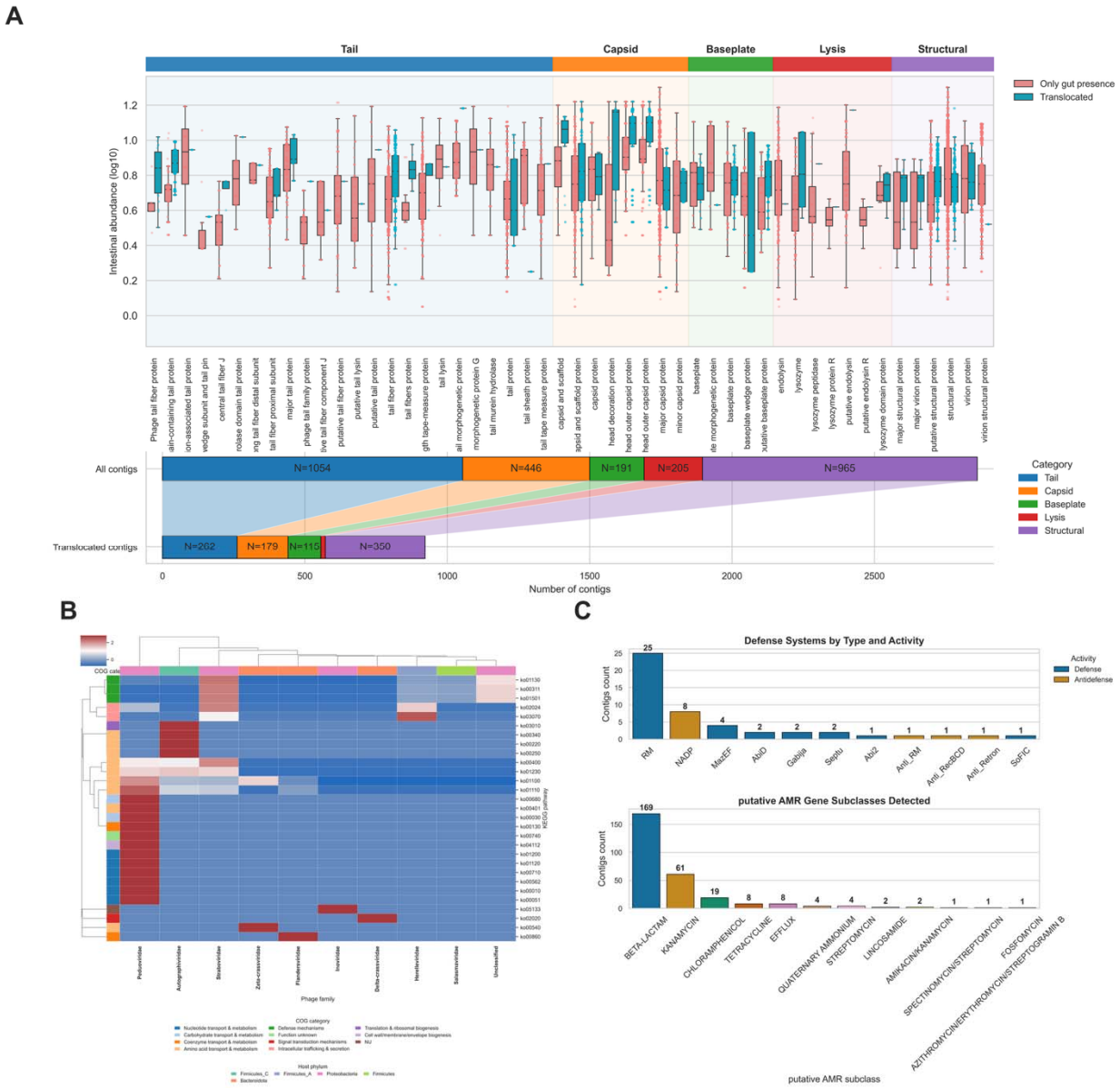


Figure 4 Functional and defense signatures among transcytosed phage contigs. (A) Intestinal abundance distributions (log10-transformed) of protein groups detected in the intestine, stratified by whether the corresponding contigs were also detected in serum (Translocated) or were restricted to gut detection (Only gut presence). Protein groups are ordered and visually grouped by predicted structural module (Tail, Capsid, Baseplate, Lysis, Structural; colored header bar), with shaded backgrounds indicating module blocks. Points represent per-sample values and boxplots summarize the distribution across samples; significance markers (asterisks) indicate multiple-testing–corrected paired differences between Translocated and Only gut presence within each protein group. The lower subpanel summarizes contig counts per structural module for all contigs versus translocated contigs, highlighting shifts in module representation between the full set and the translocated subset. (B) Heatmap of KEGG pathway enrichment patterns across Phage families, computed as row-wise Z-scores of total abundance aggregated per KEGG pathway × phage family. Rows (KEGG pathways) are annotated by the dominant COG functional category and columns (phage families) by the dominant predicted host phylum. Hierarchical clustering (dendrograms) organizes pathways and families by similarity in their Z-score profiles. Legends report the COG-category and host-phylum color mappings.(C) Summary of accessory functional potential detected among transcytosed contigs. The upper bar chart shows count of identified defense/anti-defense system types stratified by activity class, and the lower bar chart shows counts of antimicrobial resistance (AMR) gene subclasses detected. Bar labels indicate observed contig counts per category.

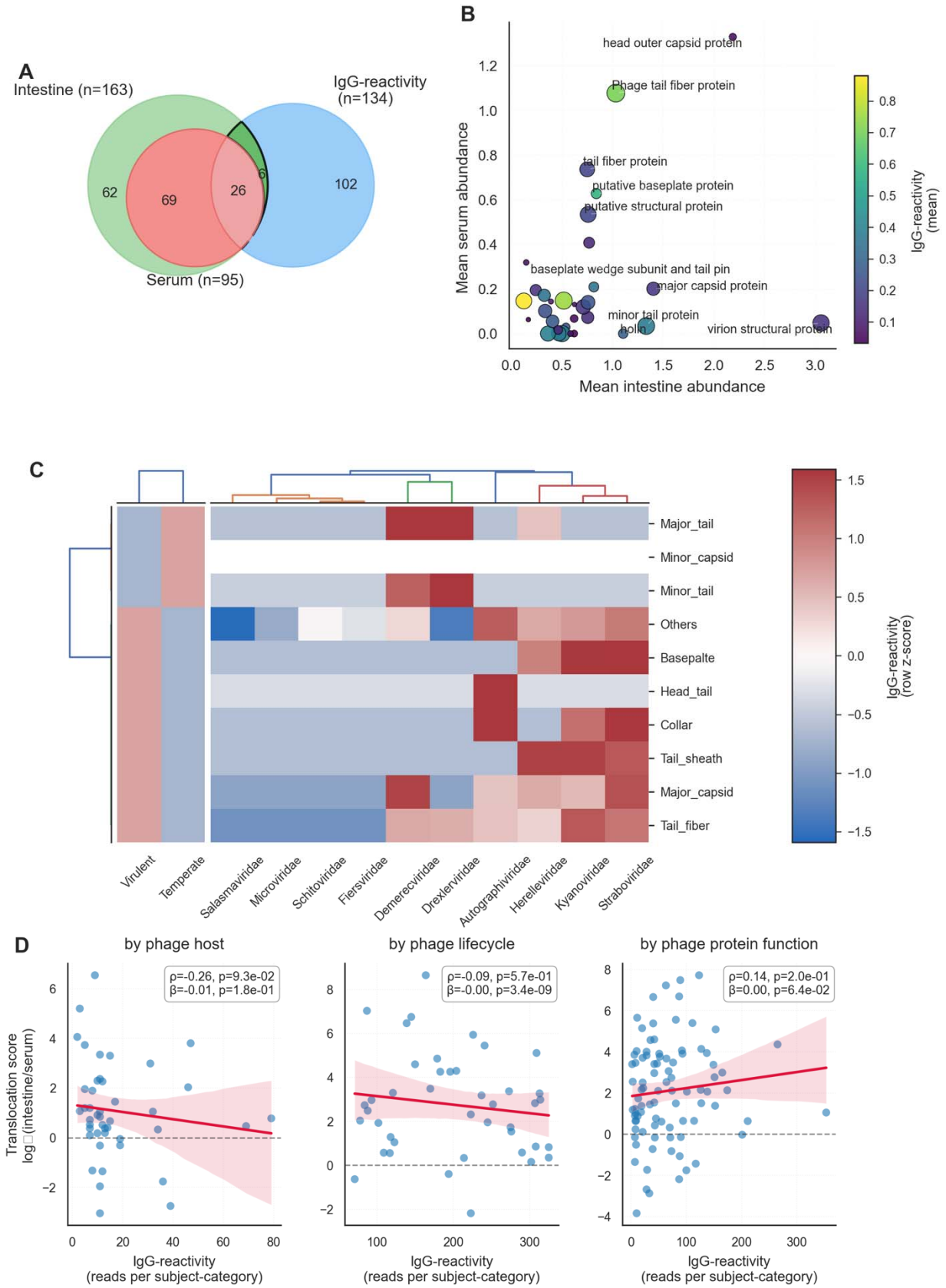

A
Intestine (n=163)
IgG-reactivity (n=134)
Serum (n=95)
62
69
26
6
102
B
head outer capsid protein
Phage tail fiber protein
tail fiber protein
putative baseplate protein
putative structural protein
baseplate wedge subunit and tail pin
major capsid protein
minor tail protein
holin
virion structural protein
Mean serum abundance
Mean intestine abundance
IgG-reactivity (mean)
C
Major_tail
Minor_capsid
Minor_tail
Others
Basepalte
Head_tail
Collar
Tail_sheath
Major_capsid
Tail_fiber
Virulent
Temperate
Salasmaviridae
Microviridae
Schitoviridae
Fiersviridae
Demerecviridae
Drexlerviridae
Autographiviridae
Herelleviridae
Kyanoviridae
Straboviridae
IgG-reactivity (row z-score)
D
by phage host
by phage lifecycle
by phage protein function
ρ=-0.26, p=9.3e-02
β=-0.01, p=1.8e-01
ρ=-0.09, p=5.7e-01
β=-0.00, p=3.4e-09
ρ=0.14, p=2.0e-01
β=0.00, p=6.4e-02
Translocation score
log□(intestine/serum)
IgG-reactivity
(reads per subject-category)

Figure 5. IgG-reactivity across gut–serum phage proteins and association with translocation. (A) Venn diagram showing overlap of protein groups detected in intestine, serum, and IgG-reactivity datasets (presence defined as non-zero signal) (B) Protein-group mean intestine abundance (x-axis) versus mean serum abundance (y-axis); point colour indicates mean IgG-reactivity (raw counts) and point size indicates total IgG-reactivity (log10-scaled sum), with labels shown for the highest IgG-reactive groups. (C) Dual clustered heatmaps of transformed IgG-reactivity reads (log1p, row z-score) summarised by phage cycle (left) and phage family (right), sharing a row dendrogram. (D) Intestine-to-serum ratio score log2((intestine+1)/(serum+1)) versus IgG-reactivity (raw reads per subject–category), stratified by phage hosts, predicted lifecycle, and protein function; lower scores indicate comparatively higher relative serum abundance. Inset statistics report Spearman's ρ and mixed-effects model slope (β).

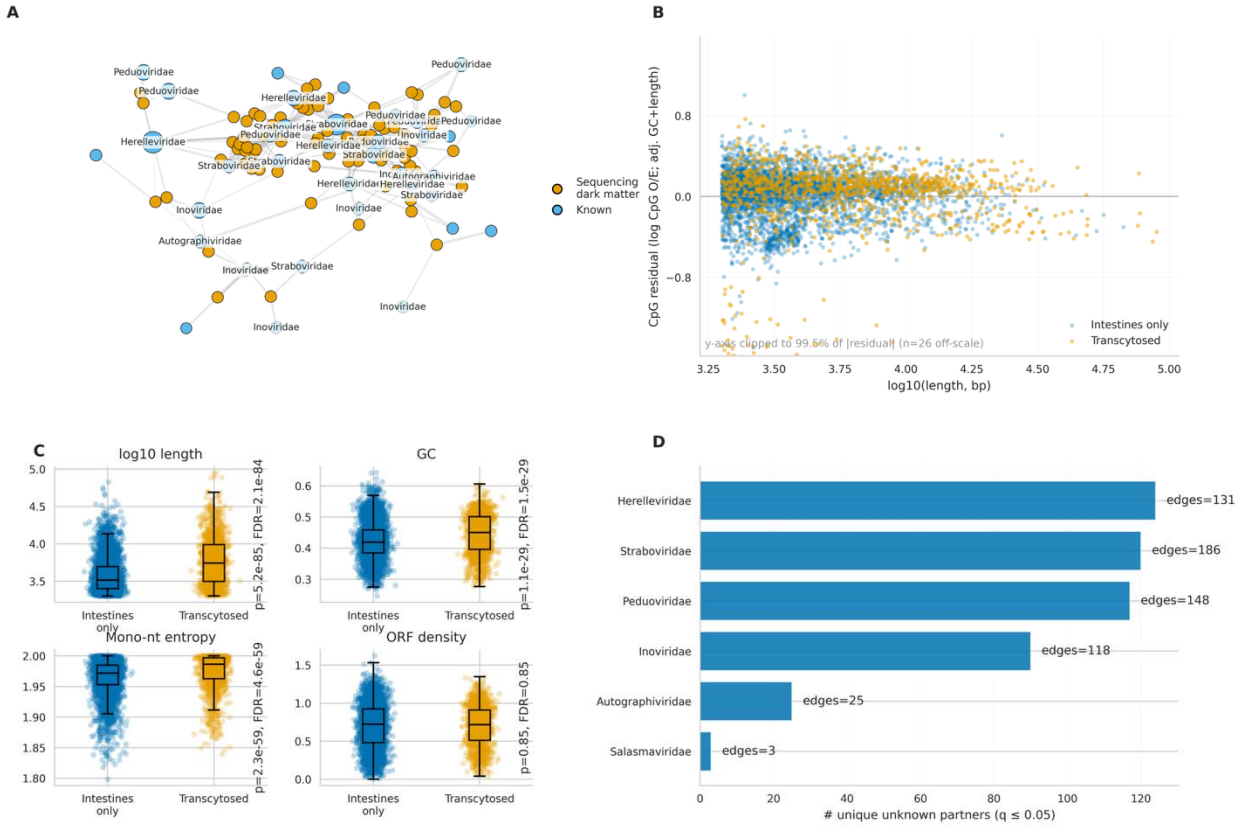


Figure 6. Co-transcytosis association network and sequence signatures of transcytosed viral contigs. (A) Unknown–known co-enrichment network of serum-enriched viral contigs across subjects (n=37). Nodes represent vOTUs and are colored by annotation class: Sequencing dark matter (contigs lacking a confident assignment) and Known (confidence ≥70). Serum enrichment is defined per subject as serum abundance ≥1.0 and log2((serum+0.5)/(intestine+0.5)) ≥1.0. Edges connect unknown–known pairs that co-occur above expectation; significance is assessed by one-sided hypergeometric tests and corrected by BH-FDR (q ≤ 0.05), with additional

thresholds for co-occurrence (≥3) and association strength (phi ≥ 0.25). Node size scales with degree and edge width scales with phi. (B) CpG residual versus vOTU length (log10 bp), colored by annotation class; outliers are clipped for display as indicated. (C) vOTU length comparison for Transcytosed versus Only gut presence groups. (D) Comparisons of GC content, mono-nucleotide Shannon entropy, and ORF density between Transcytosed and Only gut presence vOTUs. For panels C–D, two-sided Mann–Whitney U tests are used with BH-FDR correction across tested features; P and FDR values are shown on the plots. Boxplots show median, interquartile range, and 1.5×IQR whiskers.

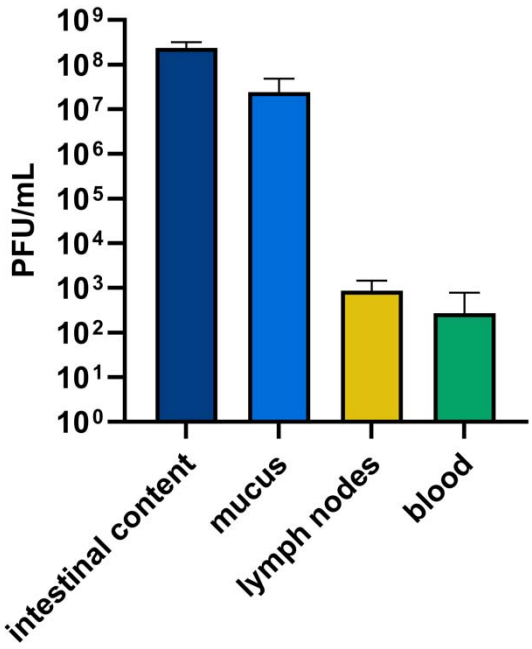


Figure 7. Mouse model of phage translocation and dissemination from gut. C57BL/6J normal male mice (N=8) were treated with Sepharose-purified T4 phage (ATCC) added to drinking water (5x109 PFU/mL) continuously for 26 hours. Murine blood, mesenteric lymph nodes, gut content, and scraped intestinal mucus were collected at 26 hours post-starting phage administration. Phage concentration was assessed by double-layer plate method or routine test dilution series. One exemplary experiment out of two conducted ones is presented.

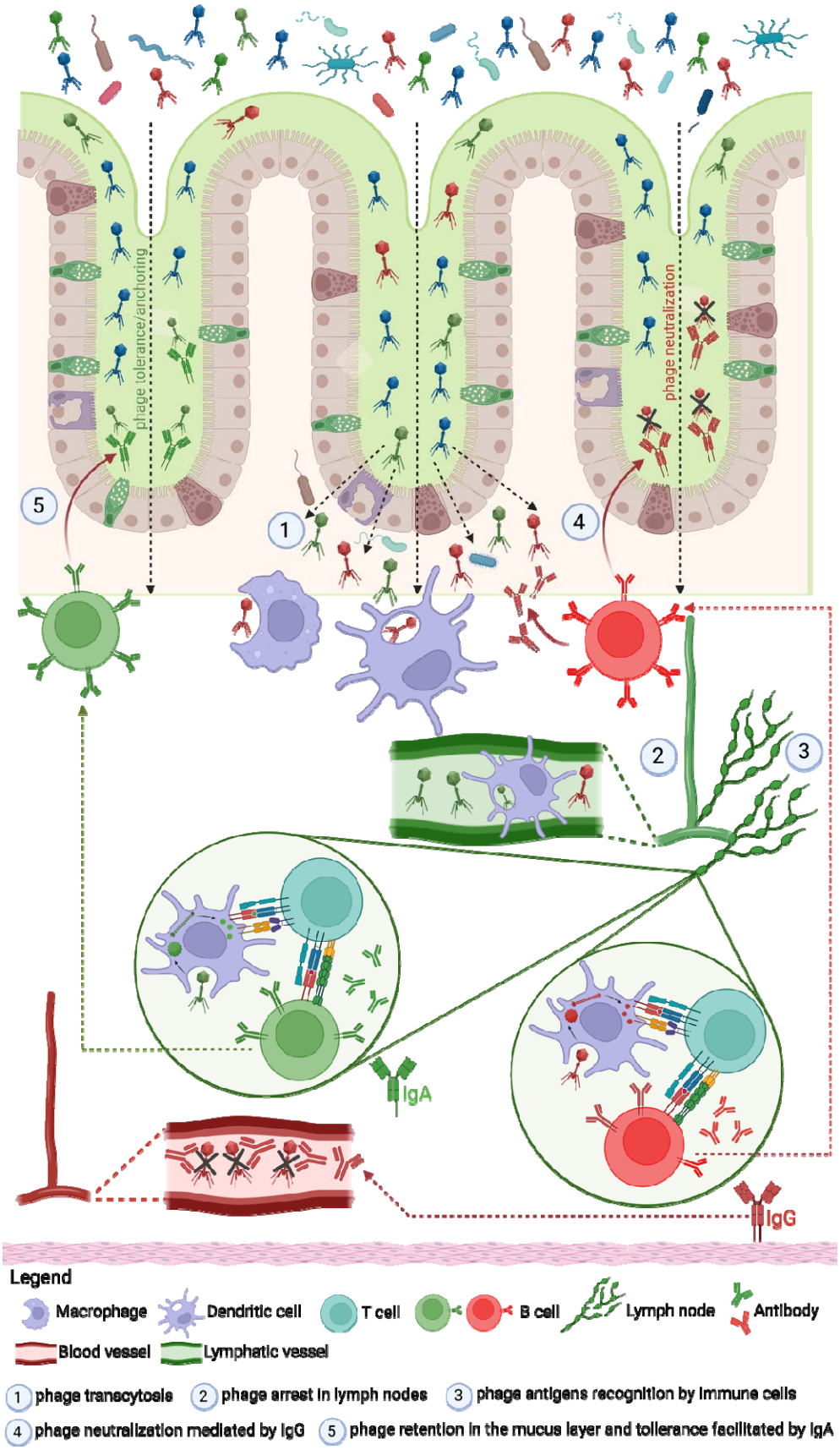

phage tolerance/anchoring
phage neutralization
5
1
4
2
3
IgA
IgG
Legend
Macrophage
Dendritic cell
T cell
B cell
Lymph node
Antibody
Blood vessel
Lymphatic vessel
1 phage transcytosis
2 phage arrest in lymph nodes
3 phage antigens recognition by immune cells
4 phage neutralization mediated by IgG
5 phage retention in the mucus layer and tolerance facilitated by IgA

Figure 8. Proposed immunological mechanism regulating phage translocation. 1. phage translocation across the epithelial barrier; 2. redirection of translocated phages to lymph nodes with the primarily effect of mucosal tolerance; 3. phages elicit systemic IgG responses; 4. phages targeted by IgGs are neutralized.

References

Abed, Wurood Kadhim, Thualfakar Hayder Hasan, and Maha Mashrq Albayati. 2025. "The Mutual Interaction between Microbial and Human Viruses in the Gut: The Role of the Virome in Modulating Immunity and Therapeutic Response." *Iraqi Journal of Natural Sciences and Nanotechnology* 6 (2): 145–57.

Berkson, Julia D., Claire E. Wate, Garrison B. Allen, Alyxandria M. Schubert, Kristin E. Dunbar, Michael P. Coryell, Rosa L. Sava, et al. 2024. "Phage-Specific Immunity Impairs Efficacy of Bacteriophage Targeting Vancomycin Resistant Enterococcus in a Murine Model." *Nature Communications* 15 (1): 2993.

Bichet, Marion C., Jack Adderley, Laura Avellaneda-Franco, Isabelle Magnin-Bougma, Natasha Torriero-Smith, Linden J. Gearing, Celine Deffrasnes, et al. 2023. "Mammalian Cells Internalize Bacteriophages and Use Them as a Resource to Enhance Cellular Growth and Survival." *PLoS Biology* 21 (10): e3002341.

Bichet, Marion C., Wai Hoe Chin, William Richards, Yu-Wei Lin, Laura Avellaneda-Franco, Catherine A. Hernandez, Arianna Oddo, et al. 2021. "Bacteriophage Uptake by Mammalian Cell Layers Represents a Potential Sink That May Impact Phage Therapy." *IScience* 24 (4): 102287.

Brzozowska, Ewa, Tomasz Lipiński, Agnieszka Korzeniowska-Kowal, Karolina Filik, Andrzej Górski, and Andrzej Gamian. 2023. "Detection and Level Evaluation of Antibodies Specific to Environmental Bacteriophage I11mO19 and Related Coliphages in Non-Immunized Human Sera." *Antibiotics (Basel, Switzerland)* 12 (3): 586.

Camarillo-Guerrero, Luis F., Alexandre Almeida, Guillermo Rangel-Pineros, Robert D. Finn, and Trevor D. Lawley. 2021. "Massive Expansion of Human Gut Bacteriophage Diversity." *Cell* 184 (4): 1098-1109.e9.

Carroll-Portillo, Amanda, October Barnes, Cristina N. Coffman, Cody A. Braun, Sudha B. Singh, and Henry C. Lin. 2025. "Transcytosis of T4 Bacteriophage through Intestinal Cells Enhances Its Immune Activation." *Viruses* 17 (1): 134.

Douadi, Clara, Quentin Lamy-Besnier, Ilias Theodorou, Olivier Schiettekatte, Yanis Sbardella, Loïc Brot, Paolo Emidio Costantini, et al. 2025. "Differential Translocation of Bacteriophages across the Intestinal Barrier in Health and Crohn's Disease." *Cell Reports* 45 (1): 116726.

Enault, François, Arnaud Briet, Léa Bouteille, Simon Roux, Matthew B. Sullivan, and Marie-Agnès Petit. 2017. "Phages Rarely Encode Antibiotic Resistance Genes: A Cautionary Tale for Virome Analyses." *The ISME Journal* 11 (1): 237–47.

Fitzgerald, C. Brian, Andrey N. Shkoporov, Aditya Upadrasta, Ekaterina V. Khokhlova, R. Paul Ross, and Colin Hill. 2021. "Probing the 'Dark Matter' of the Human Gut Phageome: Culture Assisted Metagenomics Enables Rapid Discovery and Host-Linking for Novel Bacteriophages." *Frontiers in Cellular and Infection Microbiology* 11 (March): 616918.

Gregory, Ann C., Olivier Zablocki, Ahmed A. Zayed, Allison Howell, Benjamin Bolduc, and Matthew B. Sullivan. 2020. “The Gut Virome Database Reveals Age-Dependent Patterns of Virome Diversity in the Human Gut.” *Cell Host & Microbe* 28 (5): 724-740.e8.

Kaźmierczak, Zuzanna, Joanna Majewska, Paulina Miernikiewicz, Ryszard Międzybrodzki, Sylwia Nowak, Marek Harhala, Dorota Lecion, et al. 2021. “Immune Response to Therapeutic Staphylococcal Bacteriophages in Mammals: Kinetics of Induction, Immunogenic Structural Proteins, Natural and Induced Antibodies.” *Frontiers in Immunology* 12 (June): 639570.

Kharrat, Lukeman, Camilo A. Garcia-Botero, Wade Ingersoll, Tiffany Luong, Alejandro Reyes, and Dwayne R. Roach. 2025. “Large-Scale Genomic Analysis of CpG-Mediated Immunogenicity in Bacteriophages and a Novel Predictive Risk Index.” *BioRxivorg*. https://doi.org/10.1101/2025.05.15.652987.

Le, Huu Thanh, Alicia Fajardo Lubian, Bethany Bowring, David van der Poorten, Jonathan Iredell, Jacob George, Carola Venturini, Golo Ahlenstiel, and Scott Read. 2024. “Using a Human Colonoid-Derived Monolayer to Study Bacteriophage Translocation.” *Gut Microbes* 16 (1): 2331520.

Le, Huu Thanh, Carola Venturini, Alicia Fajardo Lubian, Bethany Bowring, Jonathan Iredell, Jacob George, Golo Ahlenstiel, and Scott A. Read. 2025. “Differences in Phage Recognition and Immunogenicity Contribute to Divergent Human Immune Responses to Escherichia Coli and Klebsiella Pneumoniae Phages.” *European Journal of Immunology* 55 (3): e202451543.

Mai-Prochnow, Anne, Janice Gee Kay Hui, Staffan Kjelleberg, Jasna Rakonjac, Diane McDougald, and Scott A. Rice. 2015. “Big Things in Small Packages: The Genetics of Filamentous Phage and Effects on Fitness of Their Host.” *FEMS Microbiology Reviews* 39 (4): 465–87.

Nguyen, Sophie, Kristi Baker, Benjamin S. Padman, Ruzeen Patwa, Rhys A. Dunstan, Thomas A. Weston, Kyle Schlosser, et al. 2017. “Bacteriophage Transcytosis Provides a Mechanism to Cross Epithelial Cell Layers.” *MBio* 8 (6). https://doi.org/10.1128/mBio.01874-17.

Roozendaal, Ramon, Reina E. Mebius, and Georg Kraal. 2008. “The Conduit System of the Lymph Node.” *International Immunology* 20 (12): 1483–87.

Rybicka, Izabela, and Zuzanna Kaźmierczak. 2025. “The Human Phageome: Niche-Specific Distribution of Bacteriophages and Their Clinical Implications.” *Applied and Environmental Microbiology* 91 (6): e0178824.

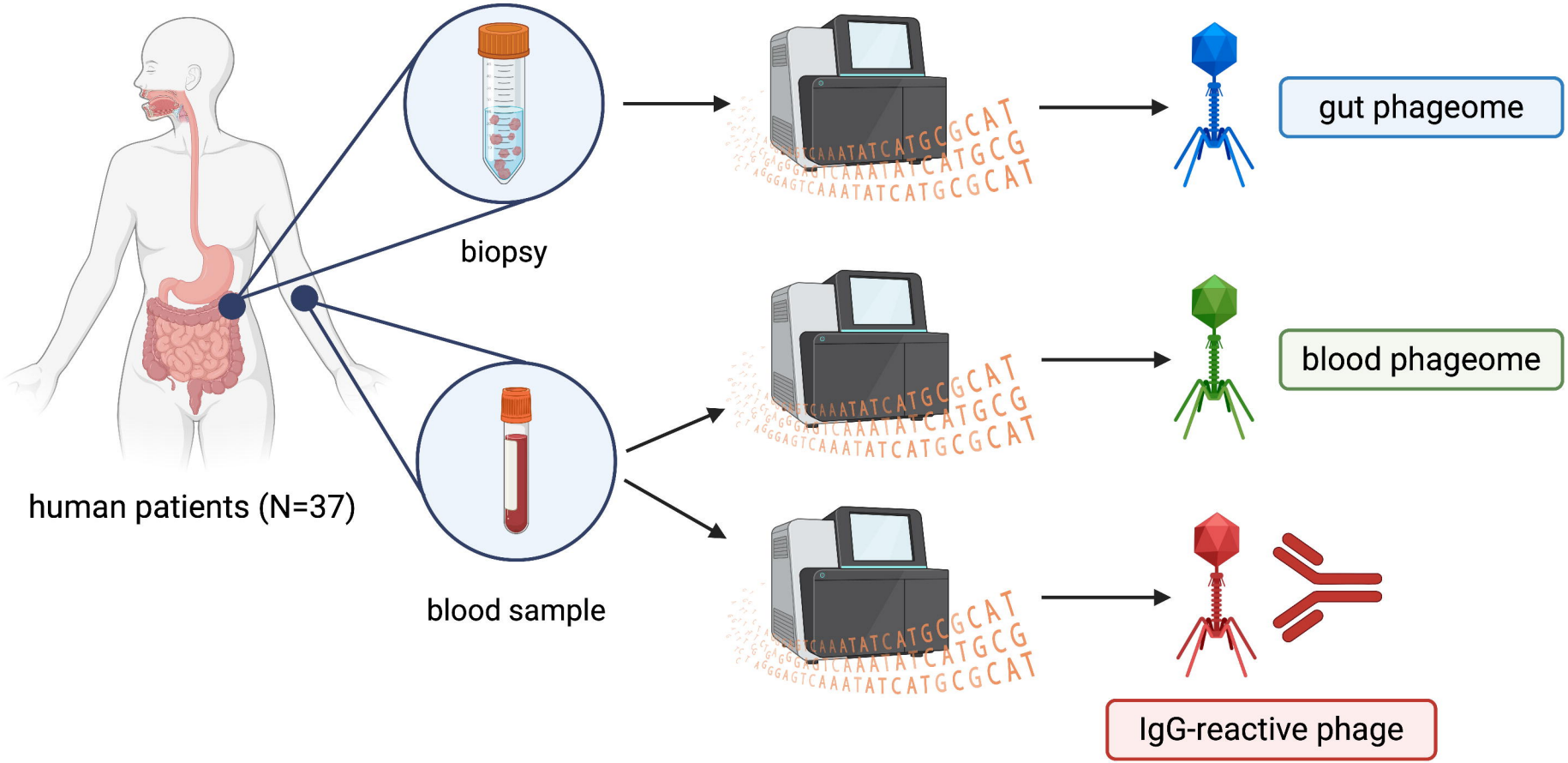
Santiago-Rodriguez, Tasha M., and Emily B. Hollister. 2022. “Unraveling the Viral Dark Matter through Viral Metagenomics.” *Frontiers in Immunology* 13 (September): 1005107.

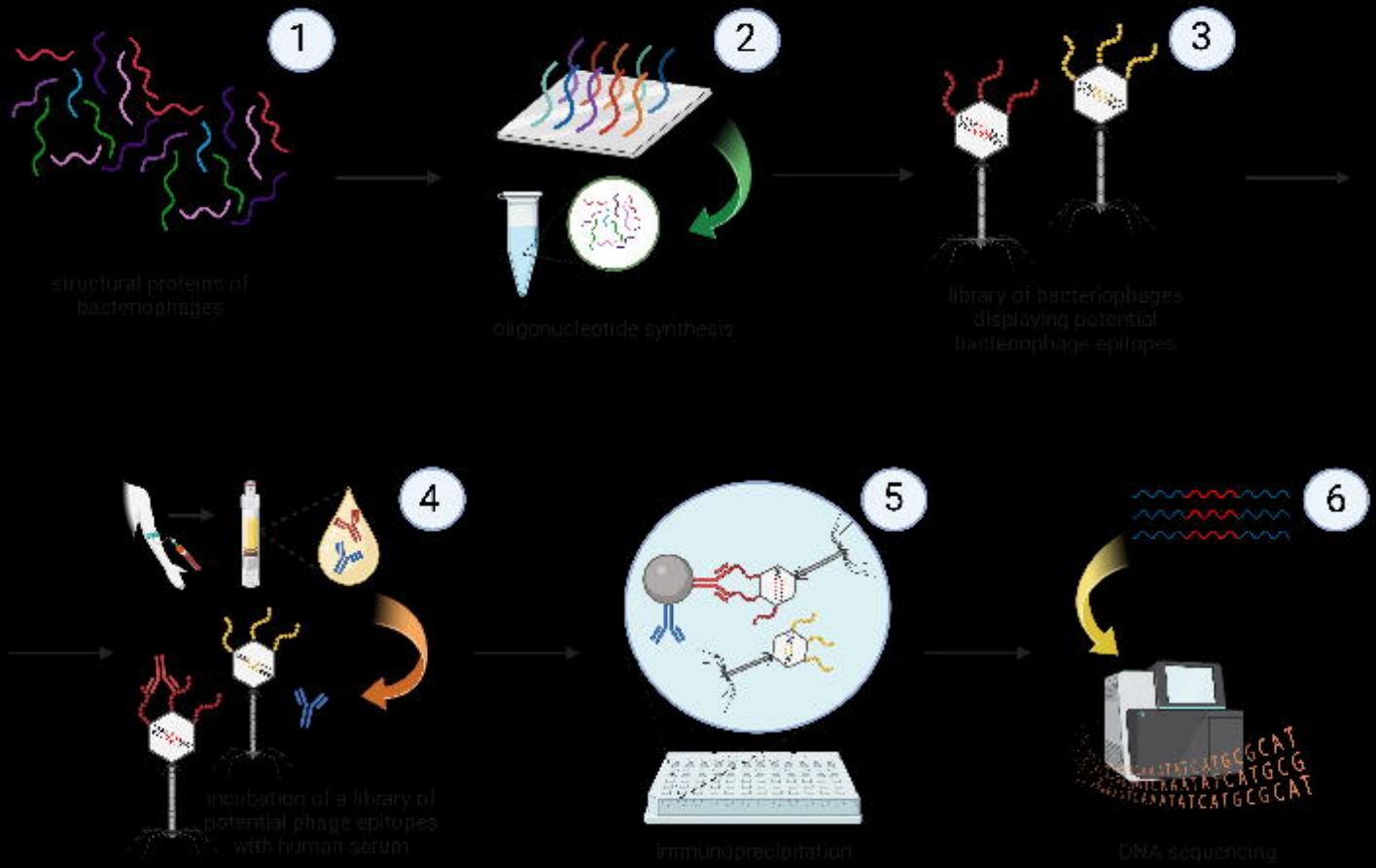
Smith, Linda, Ekaterina Goldobina, Bianca Govi, and Andrey N. Shkoporov. 2023. “Bacteriophages of the Order Crassvirales: What Do We Currently Know about This Keystone Component of the Human Gut Virome?” *Biomolecules* 13 (4): 584.

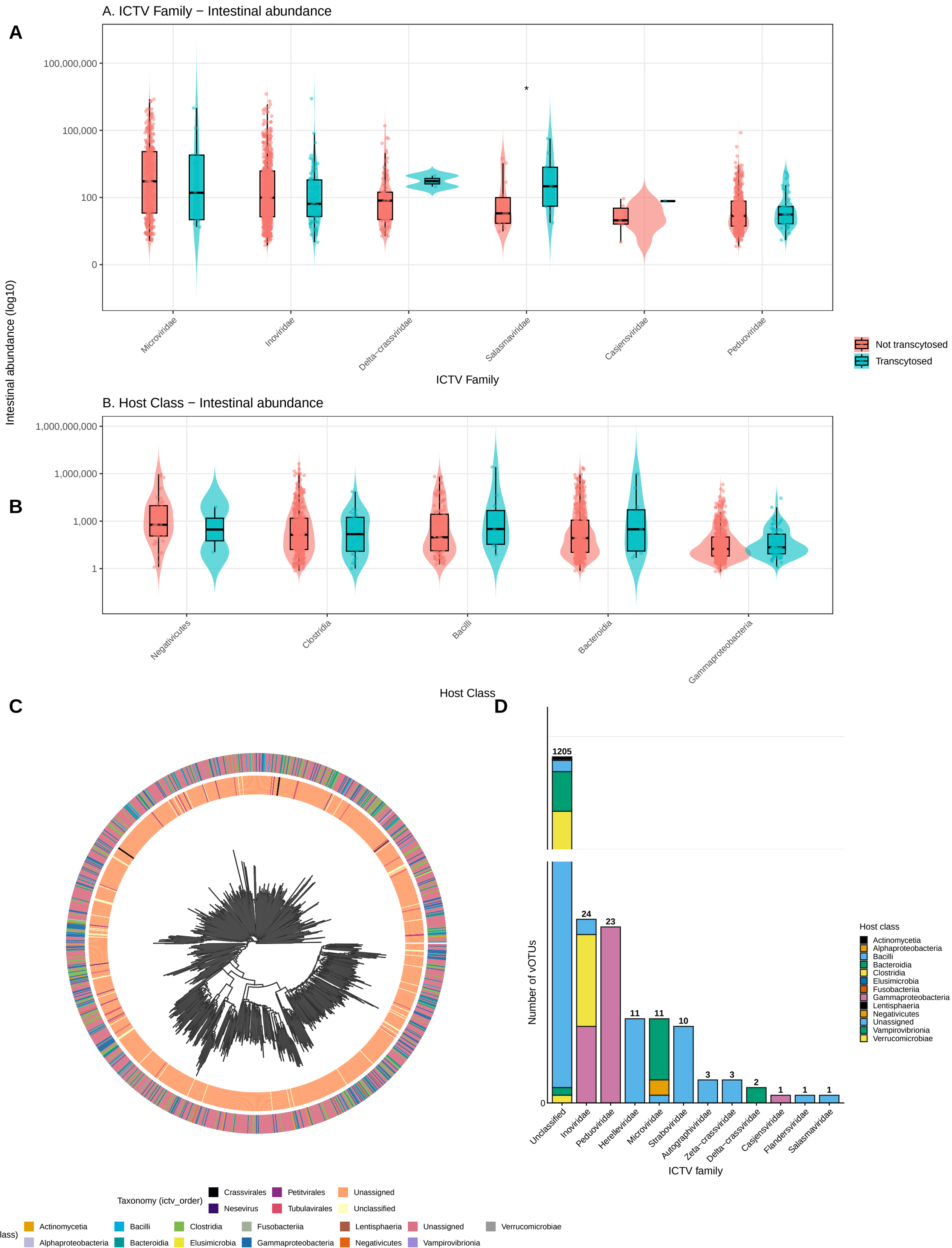
Wang, Chunyan, Jiagang Tu, Jun Liu, and Ian J. Molineux. 2019. “Structural Dynamics of Bacteriophage P22 Infection Initiation Revealed by Cryo-Electron Tomography.” *Nature Microbiology* 4 (6): 1049–56.

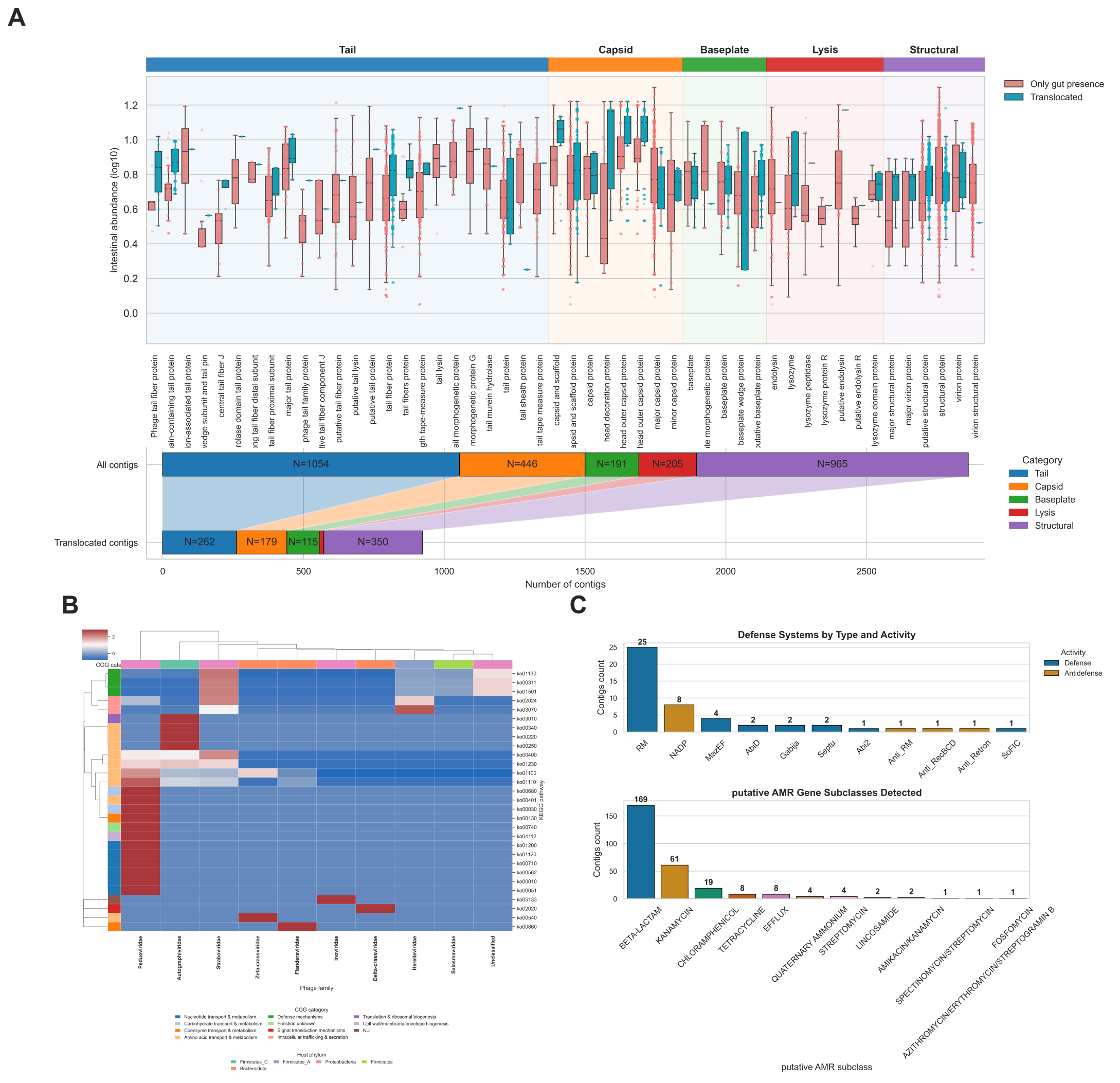
Zhang, Yidi, Ting Mei, Minhui Wang, Chunxia Huang, Hanzhi Yi, Yu Zhan, Sen Yang, et al. 2025. “Analysis of the Overall Development Trends and Hotspots in the Research Field of the Human Gut Virome.” *Virology Journal* 22 (1): 224.

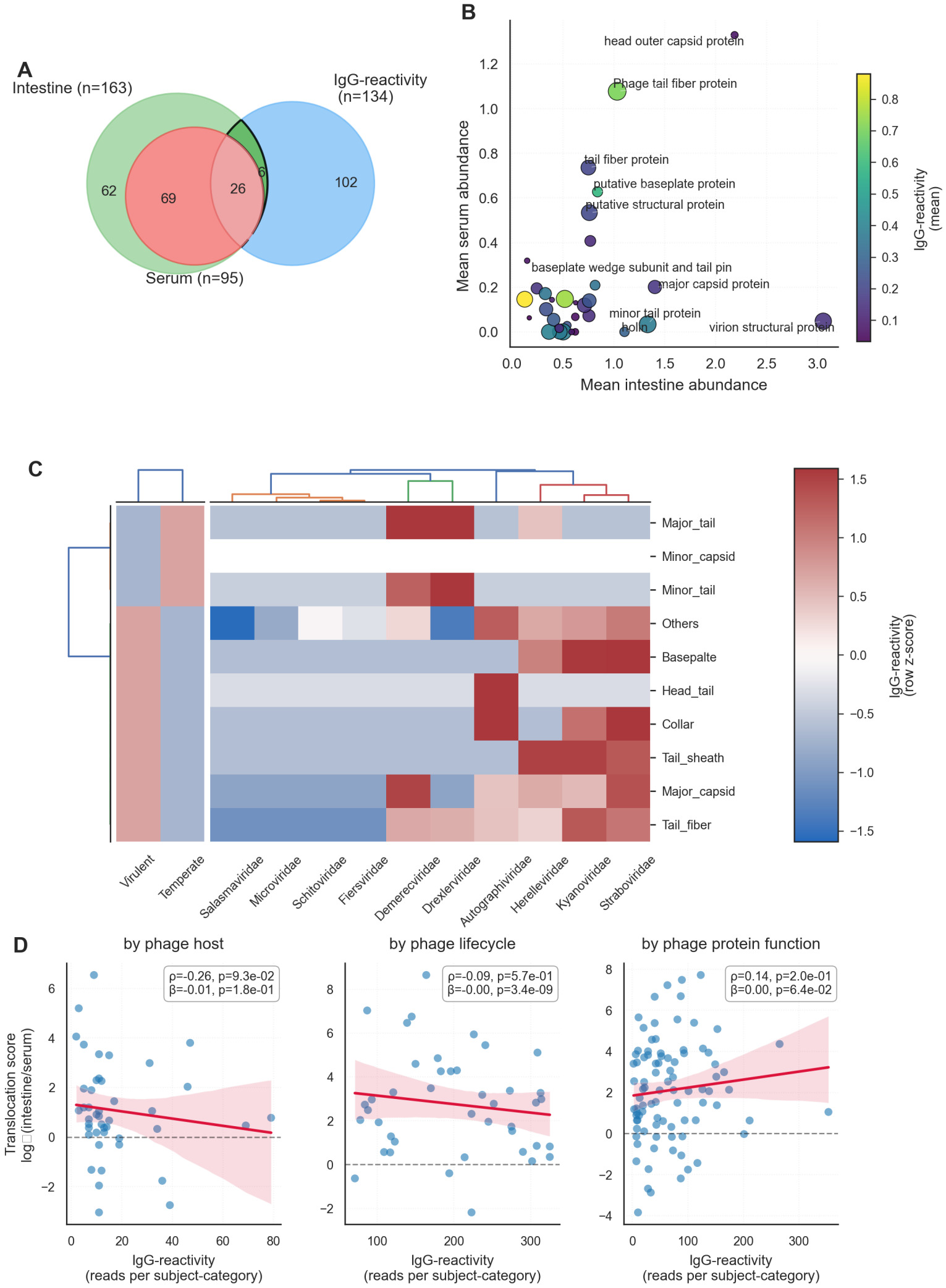
Zuo, Tao, Xiao-Juan Lu, Yu Zhang, Chun Pan Cheung, Siu Lam, Fen Zhang, Whitney Tang, et al. 2019. “Gut Mucosal Virome Alterations in Ulcerative Colitis.” *Gut* 68 (7): 1169–79.

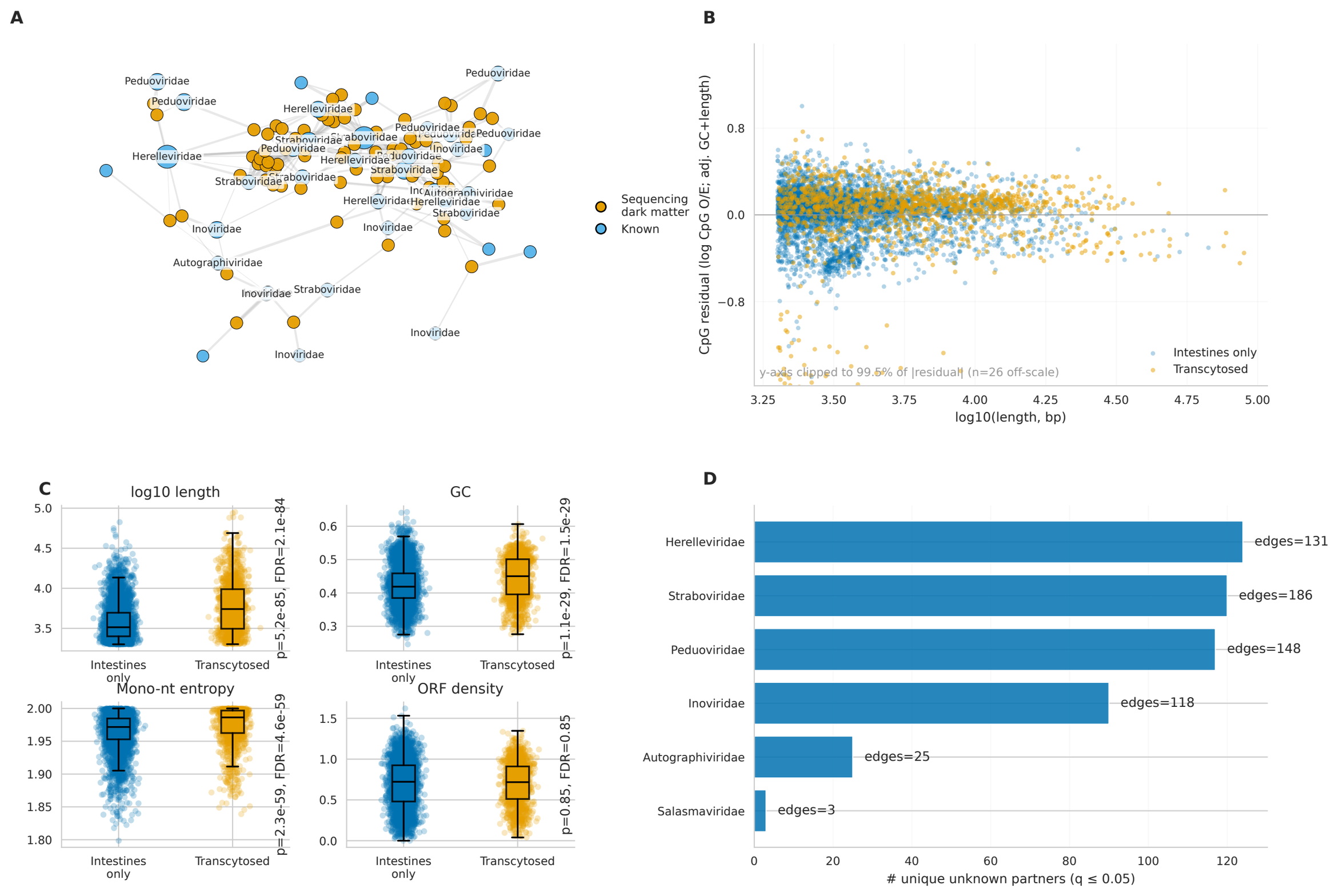

human patients (N=37)
biopsy
blood sample
gut phageome
blood phageome
IgG-reactive phage

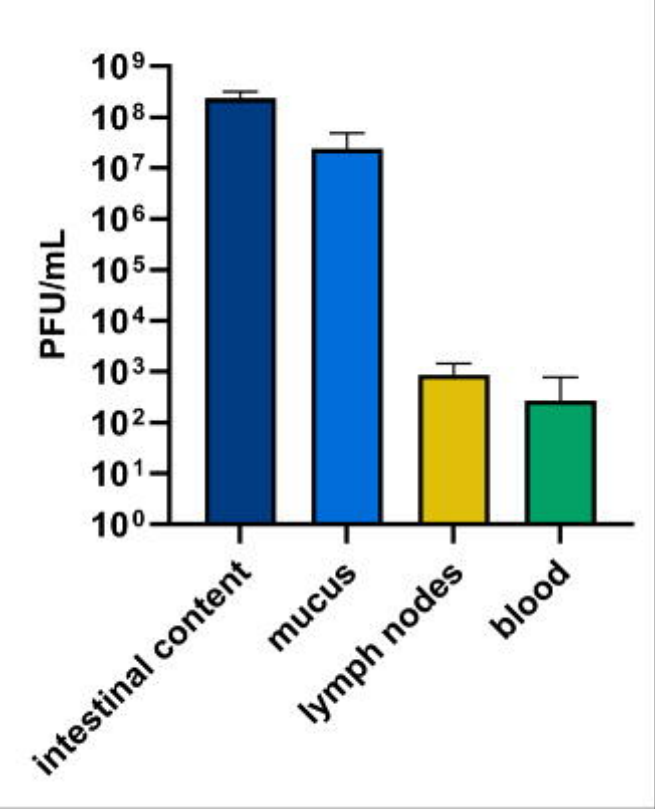

1
structural proteins of bacteriophages
2
oligonucleotide synthesis
3
library of bacteriophages displaying potential bacteriophage epitopes
4
incubation of a library of potential phage epitopes with human serum
5
immunoprecipitation
6
DNA sequencing

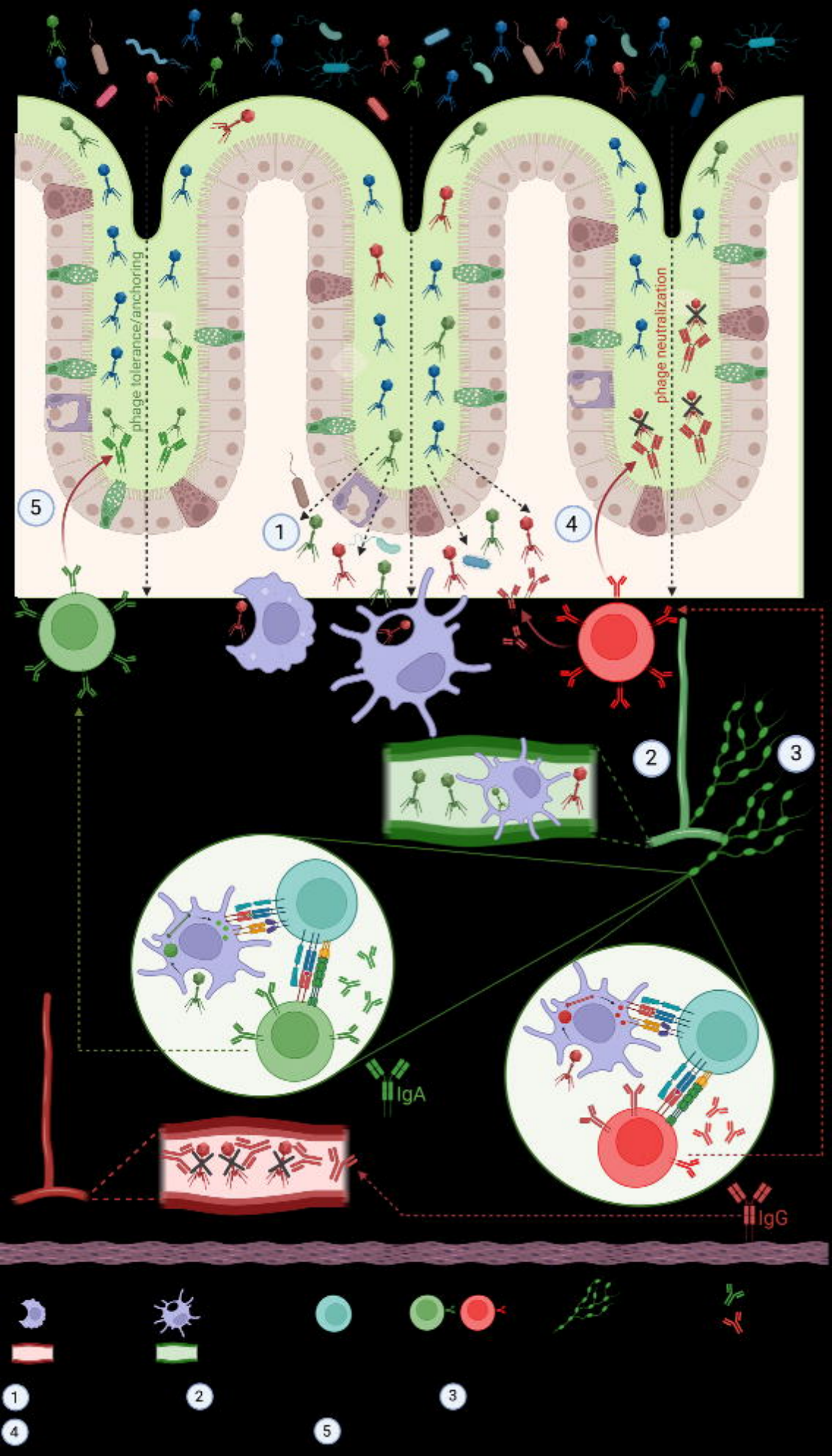

A
A. ICTV Family − Intestinal abundance
100,000,000
100,000
100
0
*
Microviridae
Inoviridae
Delta−crassviridae
Salasmaviridae
Casjensviridae
Peduoviridae
ICTV Family
Not transcytosed
Transcytosed
Intestinal abundance (log10)
B
B. Host Class − Intestinal abundance
1,000,000,000
1,000,000
1,000
1
Negativicutes
Clostridia
Bacilli
Bacteroidia
Gammaproteobacteria
Host Class
C
D
1205
24
23
11
11
10
3
3
2
1
1
1
0
Number of vOTUs
Unclassified
Inoviridae
Peduoviridae
Herelleviridae
Microviridae
Straboviridae
Autographiviridae
Zeta−crassviridae
Delta−crassviridae
Casjensviridae
Flandersviridae
Salasmaviridae
ICTV family
Host class
Actinomycetia
Alphaproteobacteria
Bacilli
Bacteroidia
Clostridia
Elusimicrobia
Fusobacteriia
Gammaproteobacteria
Lentisphaeria
Negativicutes
Unassigned
Vampirovibrionia
Verrucomicrobiae
Taxonomy (ictv_order)
Crassvirales
Petitvirales
Unassigned
Nesevirus
Tubulavirales
Unclassified
Actinomycetia
Bacilli
Clostridia
Fusobacteriia
Lentisphaeria
Unassigned
Verrucomicrobiae
Alphaproteobacteria
Bacteroidia
Elusimicrobia
Gammaproteobacteria
Negativicutes
Vampirovibrionia


A
Tail
Capsid
Baseplate
Lysis
Structural
Only gut presence
Translocated
Intestinal abundance (log10)
Phage tail fiber protein
central tail fiber J
tail fiber proximal subunit
major tail protein
phage tail family protein
putative tail fiber protein
putative tail lysin
putative tail protein
tail fiber protein
tail fibers protein
tail lysin
morphogenetic protein G
tail murein hydrolase
tail protein
tail sheath protein
tail tape measure protein
capsid and scaffold
capsid protein
head decoration protein
head outer capsid protein
major capsid protein
minor capsid protein
baseplate
baseplate protein
baseplate wedge protein
endolysin
lysozyme
lysozyme peptidase
lysozyme protein R
putative endolysin
putative endolysin R
lysozyme domain protein
major structural protein
major virion protein
putative structural protein
structural protein
virion protein
virion structural protein
All contigs
N=1054
N=446
N=191
N=205
N=965
Translocated contigs
N=262
N=179
N=115
N=350
Number of contigs
Category
Tail
Capsid
Baseplate
Lysis
Structural
B
COG category
Nucleotide transport & metabolism
Carbohydrate transport & metabolism
Coenzyme transport & metabolism
Amino acid transport & metabolism
Defense mechanisms
Function unknown
Signal transduction mechanisms
Intracellular trafficking & secretion
Translation & ribosomal biogenesis
Cell wall/membrane/envelope biogenesis
NU
KEGG pathway
Pedoviridae
Autographiviridae
Straboviridae
Zeta-crassviridae
Flandersviridae
Inoviridae
Delta-crassviridae
Herelleviridae
Salasmaviridae
Unclassified
Phage family
Host phylum
Firmicutes_C
Firmicutes_A
Proteobacteria
Firmicutes
Bacteroidota
C
Defense Systems by Type and Activity
Contigs count
Activity
Defense
Antidefense
RM
NADP
MazEF
AbiD
Gabija
Septu
Abi2
Anti_RM
Anti_RecBCD
Anti_Retron
SoFIC
putative AMR Gene Subclasses Detected
BETA-LACTAM
KANAMYCIN
CHLORAMPHENICOL
TETRACYCLINE
EFFLUX
QUATERNARY AMMONIUM
STREPTOMYCIN
LINCOSAMIDE
AMIKACIN/KANAMYCIN
SPECTINOMYCIN/STREPTOMYCIN
FOSFOMYCIN
AZITHROMYCIN/ERYTHROMYCIN/STREPTOGRAMIN B
putative AMR subclass


A
Intestine (n=163)
IgG-reactivity (n=134)
Serum (n=95)
62
69
26
6
102
B
head outer capsid protein
Phage tail fiber protein
tail fiber protein
putative baseplate protein
putative structural protein
baseplate wedge subunit and tail pin
major capsid protein
minor tail protein
holin
virion structural protein
Mean serum abundance
Mean intestine abundance
IgG-reactivity (mean)
C
Major_tail
Minor_capsid
Minor_tail
Others
Basepalte
Head_tail
Collar
Tail_sheath
Major_capsid
Tail_fiber
Virulent
Temperate
Salasmaviridae
Microviridae
Schitoviridae
Fiersviridae
Demerecviridae
Drexlerviridae
Autographiviridae
Herelleviridae
Kyanoviridae
Straboviridae
IgG-reactivity (row z-score)
D
by phage host
by phage lifecycle
by phage protein function
ρ=-0.26, p=9.3e-02
β=-0.01, p=1.8e-01
ρ=-0.09, p=5.7e-01
β=-0.00, p=3.4e-09
ρ=0.14, p=2.0e-01
β=0.00, p=6.4e-02
Translocation score
log□(intestine/serum)
IgG-reactivity
(reads per subject-category)


A
Peduoviridae
Herelleviridae
Straboviridae
Inoviridae
Autographiviridae
Sequencing dark matter
Known
B
CpG residual (log CpG O/E; adj. GC+length)
log10(length, bp)
y-axis clipped to 99.5% of |residual| (n=26 off-scale)
Intestines only
Transcytosed
C
log10 length
p=5.2e-85, FDR=2.1e-84
GC
p=1.1e-29, FDR=1.5e-29
Mono-nt entropy
p=2.3e-59, FDR=4.6e-59
ORF density
p=0.85, FDR=0.85
Intestines only
Transcytosed
D
Herelleviridae
Straboviridae
Peduoviridae
Inoviridae
Autographiviridae
Salasmaviridae
edges=131
edges=186
edges=148
edges=118
edges=25
edges=3
# unique unknown partners (q ≤ 0.05)


PFU/mL
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
intestinal content
mucus
lymph nodes
blood


phage tolerance/anchoring
phage neutralization
5
1
4
2
3
IgA
IgG
1
2
3
4
5