Title

Third-generation cephalosporin antibiotics induce phage bursts in the human gut microbiome.

Author list

Eugen Pfeifer*[1,2] Camille d'Humières[3,4,5], Quentin Lamy-Besnier[1], Florian Plaza-Onate[2], Remi Denise[5], Sara Dion[3], Bénédicte Condamine[3], Marie Touchon[5], Laurence Ma[6], Charles Burdet[3], France Mentre[3], Erick Denamur[7#]**, Eduardo Rocha[5]***[#] and the PrediRes study group.

Author affiliation and footnotes

Lead contact:

eugen.pfeifer@inrae.fr

*-*** corresponding authors

eugen.pfeifer@inrae.fr

erick.denamur@inserm.fr

eduardo.rocha@pasteur.fr

[#]co-last authors

1 Université Paris-Saclay, INRAE, AgroParisTech, MICALIS, 78350 Jouy-en-Josas, France

2 Université Paris-Saclay, INRAE, MetaGenoPolis, 78350 Jouy-en-Josas, France

3 Université Paris Cité, IAME, INSERM, 75018 Paris, France

4 AP-HP, Hôpital Bichat, Laboratoire de Bactériologie, 75018 Paris, France

5 Institut Pasteur, Université Paris Cité, CNRS, UMR3525, Microbial Evolutionary Genomics, 75015 Paris, France

6 Genomics Platform, BIOMICS, Institut Pasteur, 25-28 rue Dr Roux, Paris, 75015, France

7 AP-HP, Hôpital Bichat, Laboratoire de Génétique Moléculaire, 75018 Paris, France

Running title: Antibiotics perturb gut phage dynamics

Summary (150 words)

The use of antibiotics disrupts the gut microbiota, potentially leading to long-term health issues and the spread of resistance. To investigate the impact of antibiotics on phage populations, we followed 22 healthy individuals two weeks before and up to six months after a three-day course of 3rd-generation cephalosporins. The populations of phages encoded very rarely antibiotic resistance genes and were mostly temperate including many phage-plasmids. Gut phages remained individual-specific even after microbiome perturbation by antibiotics. Yet, we found a 20% decline in phage diversity and abundance the day after treatment, caused by the burst of a few (mostly virulent) phages. We suggest they contribute to the recovery of gut bacterial diversity, since several targeted *Parabacteroides distasonis*, a bacterium thriving after cephalosporin treatment that only proliferated in the absence of these phages. Our findings point out that phages play a crucial role in the gut microbiota's response to antibiotics by restoring microbial balance and diversity.

Key words

Antibiotics, human gut microbiota, gut perturbation, gut virome, metagenomics

Introduction

Changes in the gut microbiome's composition frequently act as risk factors for the emergence of other diseases, such as the proliferation of opportunistic microbes that thrive when the microbiome is destabilized[1]. It plays a crucial role in the emergence of antibiotic resistance genes (ARGs), as horizontal gene transfer within the gut environment facilitates their spread[2,3]. This process is further accelerated by the use of antibiotics favoring the growth of ARG-carrying bacteria. Gut commensals like *Candidatus Borkfalkia ceftriaxoniphila* or *Parabacterium distasonis* are known to bloom after treatment with cephalosporins[4]. Some species also release enzymes protecting vulnerable groups, as shown for gut commensal *Bacteroides spp*, who produce cephalosporinases that enhance the survival of susceptible community members from lethal doses of cefotaxime[5]. These dynamics lead to changes in the resistome (sum of ARGs within a microbiome) and can cause an unbalanced gut microbiota (microbial dysbiosis), which is commonly linked to gut-related diseases[6,7]. Antibiotics may be present in the gut when they are taken orally or when they are excreted through the digestive tract[8]. This is the case of ceftriaxone and cefotaxime, both 3rd generation cephalosporins, which upon injection in patients are eliminated partly through the intestinal route and partly by the urinary tract[9,10]. The fraction of the antibiotics that is excreted by the digestive tract is subject to strong inter-individual variability, which is directly related with the impact of the antibiotic treatment on gut bacteria[11]. It is therefore essential to detail the impact of antibiotics on the human gut microbiome to better understand their effects on human health and on the spread of antibiotic resistance. Yet, most past studies have focused primarily on the bacterial fraction of the gut microbiota.

Bacterial viruses (phages) are highly abundant in the human gut microbiome and exhibit significant inter-individual variation within the intestinal tract[12]. It is suggested that they drive bacterial population diversity and dynamics [13,14]. Some phages are virulent, they infect, replicate and lyse their hosts. Temperate phages can either behave similarly or become prophages, where they are kept in bacterial genomes (lysogen). Most prophage genes are silent, but some may provide adaptive traits to the lysogen. Some temperate phages propagate as phage-plasmids (P-Ps), i.e. they transfer vertically in the cell lineage as plasmids and horizontally as phages. We recently showed that P-Ps are widespread in bacteria[15]. Within the gut phageome (collective of phages), temperate phages were reported to be highly prevalent, with their numbers exceeding those of virulent phages[16,17]. In addition, it is estimated that 70-80% of gut bacteria carry prophages within their genomes[18,19], suggesting that temperate phages have key roles in the gut environment. The impact of phages may be affected by antibiotics in multiple ways. First, antibiotics may induce temperate phages and result in lysis of bacterial populations. This has been reported for DNA damaging drugs such as fluoroquinolones[20,21]. Second, the change in bacterial populations following antibiotic therapy has almost necessarily an impact on their phages (which prey on them). Finally, while ARGs are typically spread by conjugative plasmids requiring cell-to-cell contact, they are being increasingly identified in P-Ps that can transfer them and induce resistance by lysogenic conversion[22], raising the possibility that P-P lysogens may protect bacteria from antibiotics. A few works focused on the impact of antibiotics on phages in the human gut. In particular, significant alterations of phage populations in the gut have been observed during antibiotic treatment for *Helicobacter pylori*, involving a proton pump inhibitor, amoxicillin, and clarithromycin[23]. Furthermore, a pioneering study including ten healthy volunteers, who received four antibiotics (quinolones, β-lactams, tetracyclines, and macrolides), suggested an expansion of ARGs in phages, although phages with ARGs

remained at low frequency[24,25]. A general large-scale analysis of the impact of clinically-relevant antibiotics on the phage populations and on how this can feedback on the bacterial fraction is missing.

To improve our understanding of antibiotic-induced perturbation effects on the gut microbiome, we conducted the CEREMI clinical trial[26,27], where ceftriaxone and cefotaxime were given intravenously over three days (clinical standards) to 22 healthy volunteers. Samples were taken 14 days before (day -14) and up to 180 days (day 180) following treatment, including before (day -1) and after (day 4) treatment (see Figure S1 for detailed sampling scheme). Stool samples were analyzed using measures of antibiotics concentrations, metabolomics and also metagenomics of bacteria, fungi and phages[26]. The samples before antibiotic treatment were used as controls since the strong inter-individual variability of phage composition[12] means that using other individuals provides a weak control. Since we found no significant differences between the two antibiotic treatments in our previous studies[26,27], we analyzed these samples together in this work. In our previous and global study, we observed disturbances in the population structures of bacteria and interestingly also of phages following treatment[26]. Moreover, maximal fungal load was found to be inversely correlated with total bacterial counts, indicating that fungi thrive when antibiotics suppress bacterial populations. In all but two cases, the antibiotics were not detectable in the stool samples, suggesting β-lactam degradation by gut anaerobes, as often previously described[28]. We had observed a marked increase in the abundance of genes encoding β-lactamases in the bacterial fraction[26]. These findings have raised several questions: i) How strong and targeted is the treatment's impact on the gut phageome? ii) Do gut phages carry β-lactam resistance genes, potentially contributing to their dissemination? iii) If so, are P-Ps involved and how common are they? iv) Finally, how does the treatment alter phage population dynamics? To take on these questions, we conducted a comprehensive, in-depth analysis of the phageome dataset[26], applying state-of-the-art assembly techniques and metagenomic tools to expand and annotate the repertoire of gut phages. We searched for P-Ps by comparing the updated phageome data with P-P genomes[29], employed a unified gut phageome database, and annotated ARGs using the NCBI pipeline[30]. This enabled us to track changes in the populations of phages in terms of richness, relative abundances, and gene context throughout the surveillance period.

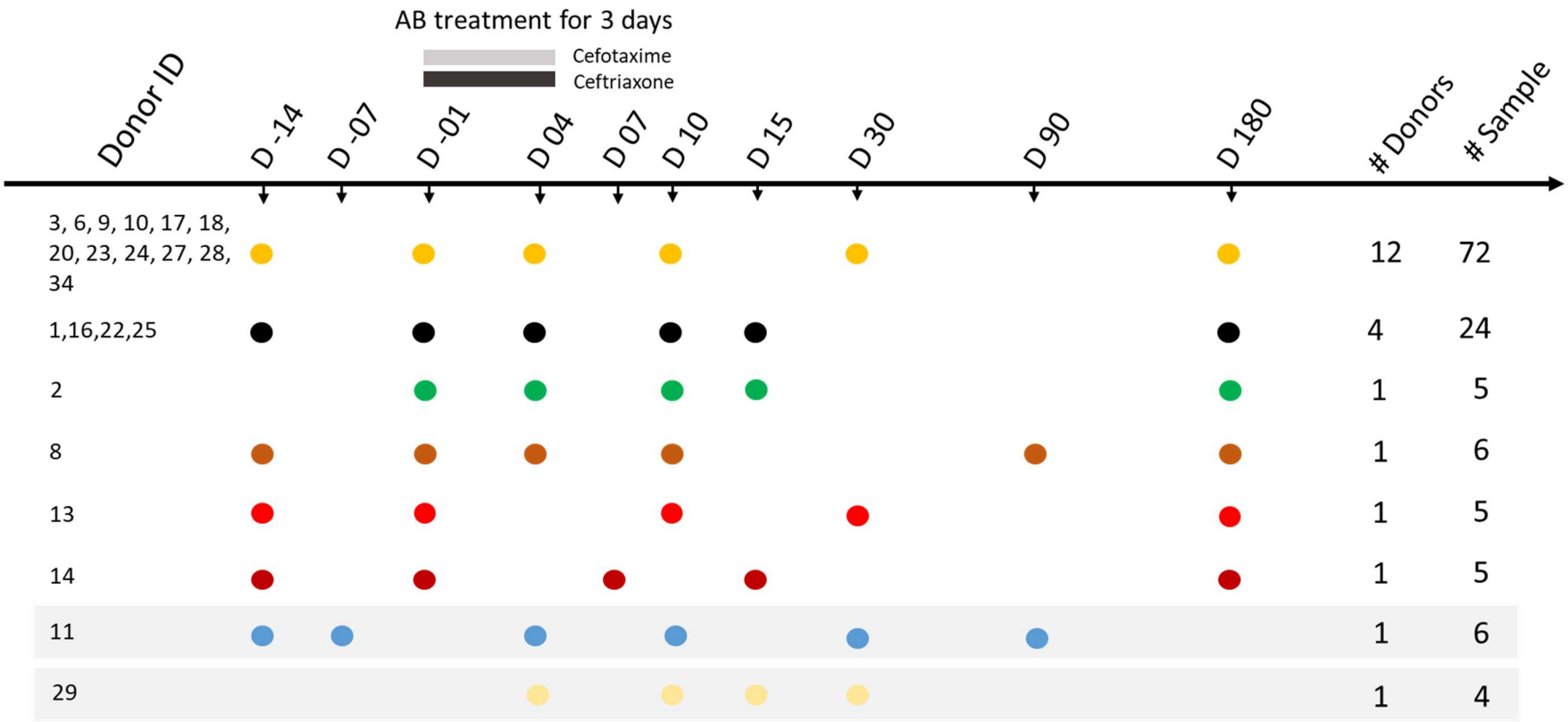


Figure S1. Detailed overview of sampling scheme and timeline of the clinical trial[26,27] [REF]. Samples from two donors (marked with grey bars) were partly excluded. Briefly, samples of donor 29 were excluded from ecological analyses (incl. diversity and abundance analysis) due to absence of any samples prior to the treatment. However, reads were included in the assembly and processing of viral contigs (see Methods). Donor 11 was excluded, if samples from the day -1 (before treatment) were required (specified in the analysis).

Results

Many novel and high-quality genomes detected in the CEREMI dataset

Given the great advances in assembling and annotating phage metagenomics data in the last few years, we completely re-analysed the CEREMI phageome data set using two different assembly strategies (see Methods). We employed the most commonly used and recent programs (geNomad[31], viralVerify[32] and VIBRANT[33]) to fetch contigs classed as viral by at least one method (Figure 1A). We then further recovered phage genomes with lower coverage by mapping all the reads on the large catalog of human gut virome datasets, the Unified Human Gut Virome (UHGV) database. We pooled all viral sequences to define dereplicated, viral operational taxonomic units (vOTUs). Relative to our previous study[26], this resulted in a 2.6-fold increase in viral sequences to 6,467 unique vOTUs (Table S1), most of which originated from UHGV (Figure 1B). We found only 31 species of viruses infecting eukaryotes which were present at low relative frequency. Hence, we will refer to vOTUs as phage species interchangeably throughout this study. The majority of sequences (84.8%) exceeded 10 kb in size (Figure 1C, Table S1). The sequence size distribution and the quality metrics of our dataset indicate a dataset with quality similar to that of UHGV (Figure S2AB), even if our dataset encompassed a broader size range (≥3 kb, for UHGV ≥10 kb). Among the vOTUs, 796 were classed as complete by CheckV[34] and 1150 as high-quality, accounting for 30.1% of the total (Figure 1D). The comparison of the vOTUs with all sequences from UHGV suggests that 1,951 are novel species (Figure S2CD). Although sequences of novel species were more likely to fail the CheckV[34] quality control (15.6% versus 5.8% for the entire dataset, Table S1, Figure 1C, Figure S2C), this may in part stem from the difficulty of the tool in classifying novel sequences. A viral taxonomy was assigned for the majority of the vOTUs (98.1%), of which 98.5% are tailed dsDNA phages (Caudoviricetes). Only 73 vOTUs are ssDNA phages (Ino- and Microviridae) (Table S1).

To exhaustively allocate the reads to the genomes, we sequentially mapped them against: the vOTUs, assemblies classified as non-viral, predicted prophages (identified by geNomad, VIBRANT, and CheckV), the UHGV, and the Unified Human Gastrointestinal Genome (UHGG) database (see Methods). Notably, we used the UHGG database to identify read sequences originating from bacteria, as it comprises over 200,000 non-redundant genomes from more than 4,644 gut prokaryotic species[35]. On average, 57.8% of the reads mapped to the vOTUs (Figure S3, Table S2), 18.9% aligned to sequences that could not be classified as viral, likely representing bacterial or highly divergent viral sequences, and only 0.32% mapped to prophages suggesting that their contribution to the viral signal is minor. Around 4.58% of the reads mapped to viral contigs from UHGV that we did not include in our analysis due to very low sequence coverage. Finally, 11.8% of the reads mapped to sequences from UHGG, potentially representing bacterial contamination or highly diverse prophages not yet identified or represented in the databases used in this study. Interestingly, we identified six vOTUs with exceptionally large genome sizes (Table S1). These included five phages (>200 kb[36]) and a novel member of *Megaviricetes* (vOTU_5310), a class of eukaryotic giant viruses.

We then used iPHoP to predict the bacterial hosts, which was possible for 4,206 vOTUs (Table S1). The hosts were found among the taxonomic classes of bacteria detected in the bacterial fractions[26] with a preponderance of Clostridia, Bacteroidia, and Bacilli (Figure S4AB). However, relative to the bacterial fractions, we found a much higher prevalence of vOTUs infecting Bacteroidia and a lower prevalence of

those infecting Clostridia. The comparison and detection of phage-host pairs are crucial to analyze the underlying dynamic interactions, which we present and discuss in the following sections.

Temperate phages including phage-plasmids are prevalent in the gut

We applied BACPHLIP[37] to predict the lifecycle of phages having complete and high-quality genomes (predicted by CheckV[34], n=1946). The majority was classed as temperate (64.1%, Figure 1E). We then searched specifically for P-Ps, which are also temperate phages but typically lack integrases and are often misclassified by tools predicting phage lifecycles. We grouped our database of 1416 P-Ps[29] with all vOTUs and used protein-sharing networks for the clustering (see Methods). We could thus identify 399 vOTUs that grouped with P-Ps representing 6.2% of all vOTUs (Figure 1F). Nearly all (87.2%) of these vOTUs clustered with putative P-Ps infecting Firmicutes and Actinobacteria, including *Bacillus thuringiensis, Clostridium beijerinckii*, *Corynebacterium atypicum*, *Clostridium perfringens* and *Enterococcus faecium* (Table S3, supplementary file 1). In particular, 198 clustered with a putative P-P from *B. thuringiensis* CTC, an episomal element described to be a 25-kb linear plasmid with similarities to a phage of *Lactococcus lactis*[38]. 61 vOTUs are similar to P-Ps of *C. beijerinckii* and/or *E. faecium,* that are described as linear elements with sizes of 16 and 18 kb, respectively. Among these, ϕ6423 (of *C. beijerinckii*) is suggested to be a phage[39] whereas no information on a lytic cycle is available for p63-3 of *E. faecium*. Many vOTUs (n=44) grouped with a putative 48-kb P-P from *C. atypicum,* for which a phage lifecycle has been proposed[40], but confirmation is lacking. 32 vOTUs clustered with vB_CpeS-CP51, a temperate phage infecting *Clostridium perfringens*[41], for which no experimental evidence of a plasmid-life cycle is reported. With *Carjivirus communis*, the prototypic crAssphage, we grouped 11 vOTUs. *C. communis*, was recently reported to propagate as a P-P in its host[42].

Notably, many vOTUs cluster with P-Ps that lack direct empirical evidence. *C. communis* is one example, whose lysogenic lifecycle remains under debate. It lacks genes typically associated with functions of integrases and the genome has not been observed in an integrated state. Furthermore, plaque formation for *C. communis* has not yet been demonstrated[42]. Our previous work identified plasmid-like replication genes in *C. communis*[15], suggesting a P-P lifecycle. This is further supported by a recent study that experimentally demonstrated the propagation of *C. communis* as a P-P and reported its primary plasmid-like lifestyle[42]. Similarly, our search for plasmid genes identified vB_CpeS-CP51 (from *C. perfringens*) as a potential P-P. It encodes a replication gene similar to those found in *C. communis* and to ϕSM101, who has been classified as a P-P in *C. perfringens*[43]. While the presence of plasmid genes strongly suggests a P-P lifecycle, this classification remains tentative due to the limited availability of experimental evidence. The 25-kb episomal element reported as a plasmid in *B. thuringiensis* CTC shares homology with a virulent phage of *L. lactis*[38], but lacks experimental evidence to propagate as a plasmid. Details on 40 other vOTUs that clustered with a range of putative and confirmed P-Ps are described in Table S3.

In summary, most phages are temperate and a sizeable fraction group together with known and putative P-Ps.

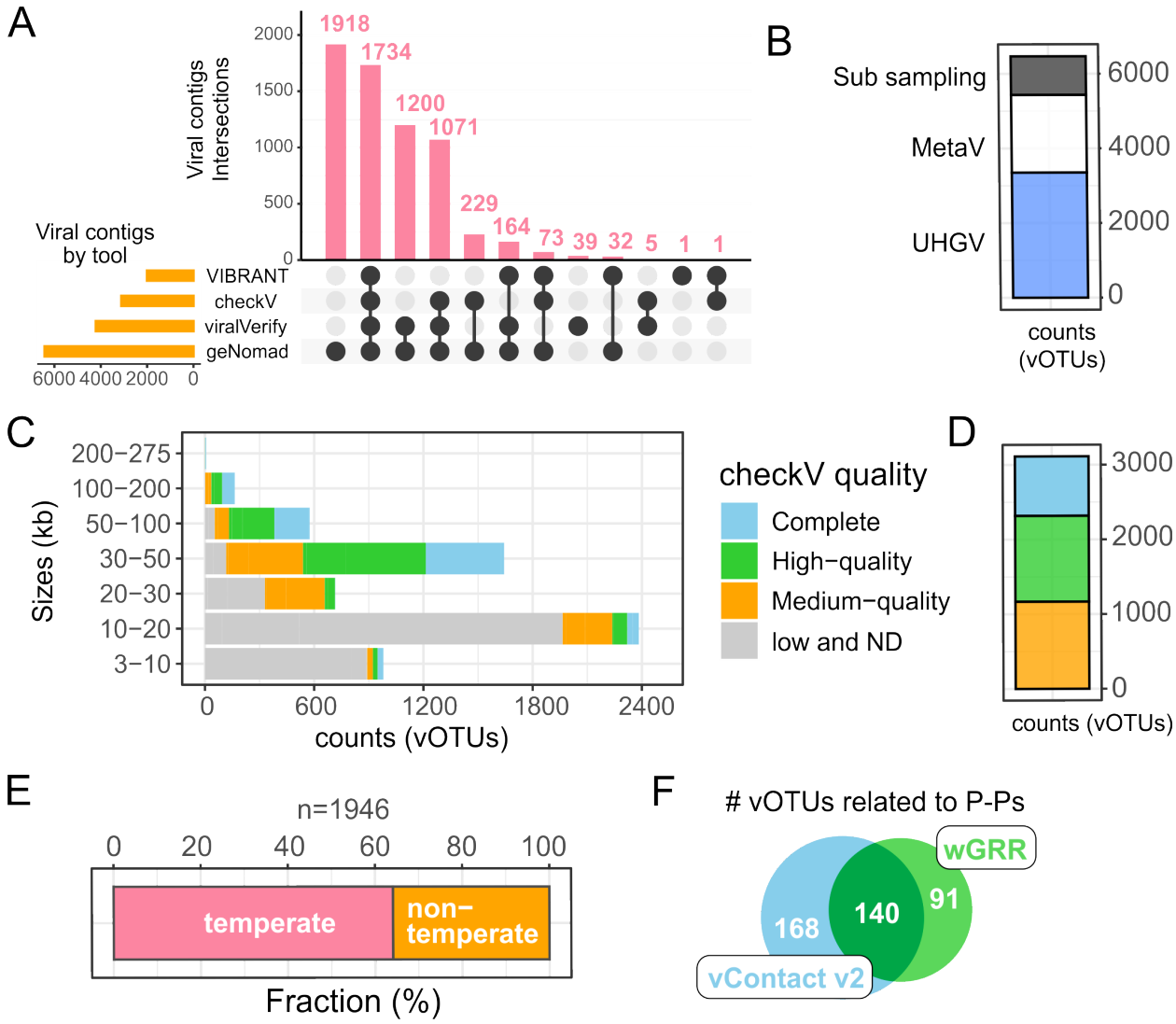


Figure 1: Overview of the gut phageome retrieved from 22 healthy donors.
A. Number of viral contigs identified by different classification tools. B. In total 6,467 viral contigs were detected by mapping reads to the UHGV and following two assembly strategies (see Methods). D. Histogram displaying the sizes of all viral contigs, ranging from 3 kb to 275 kb, along with quality information (from CheckV[34]). D. Quality assessment of viral contigs was performed using CheckV[34] [REF], with counts of medium, high-quality, and complete sequences shown (n=3,113). E. Lifecycles of 1,946 vOTUs (complete and high quality, assessed by CheckV[34]) were predicted using BACPHLIP[37]. F. Number of vOTUs related to P-Ps from a database of our previous work[29]. We used two methods, vConTACT v2[44] and gene repertoire relatedness (wGRR), that employ shared protein networks to compute genome-to-genome relatedness (see Methods).

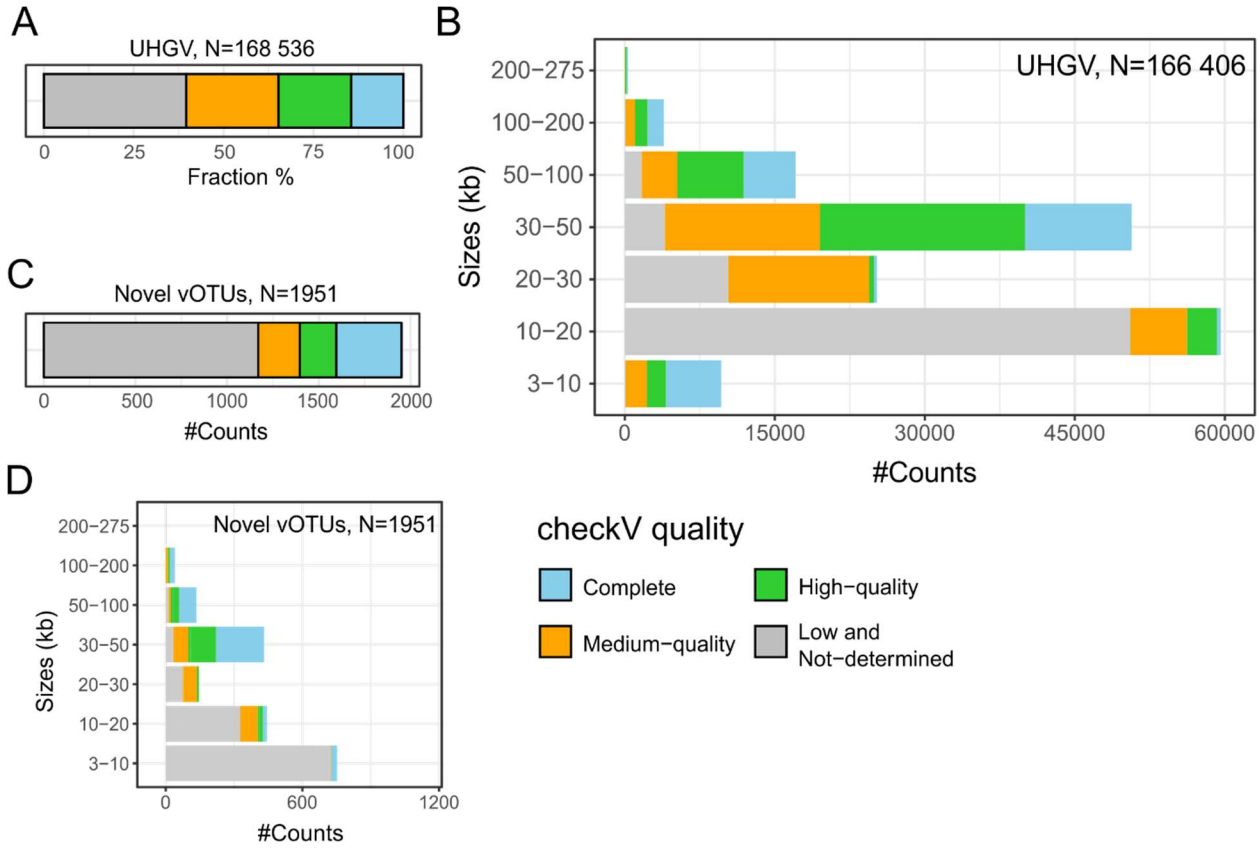


Figure S2. Completeness, quality, and sizes of sequences from UHGV and newly identified vOTUs. A and C: Quality assessment of sequences from UHGV (A) and of novel vOTUs (n=1,951) (C). B and D: Histogram showing size distribution in relation to the quality of sequences from UHGV (B) and new vOTUs (D).

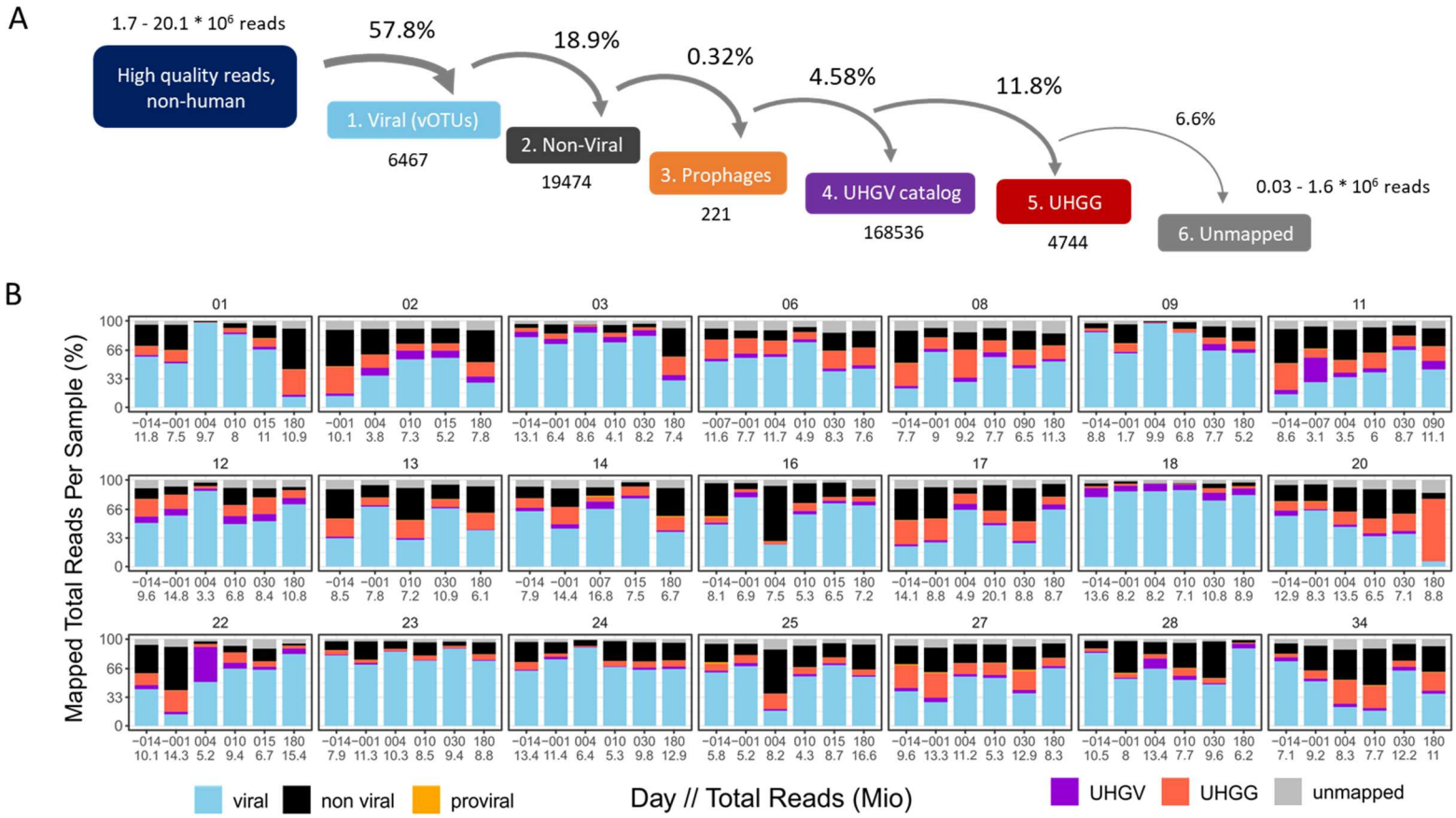


Figure S3: Allocation of the read signal to different gut-microbiome-related datasets. Reads from all samples were step-wise mapped on all vOTUs, non-viral assemblies, proviruses (predicted by geNomad[31], CheckV[34] and VIBRANT[33]), on the UHGV catalog and lastly on sequences of UHGG (see Methods). Fractions of the mapped signal are listed in Table S2. A. represents the average of all samples that are shown in bar plots B. In A, numbers below the boxes represent non-redundant sequences or representative genomes from the database. Numbers above the boxes represent the number of reads and fractions mapping on the datasets. In B, numbers below the bar plots indicate the sampling day and the number of reads in millions.

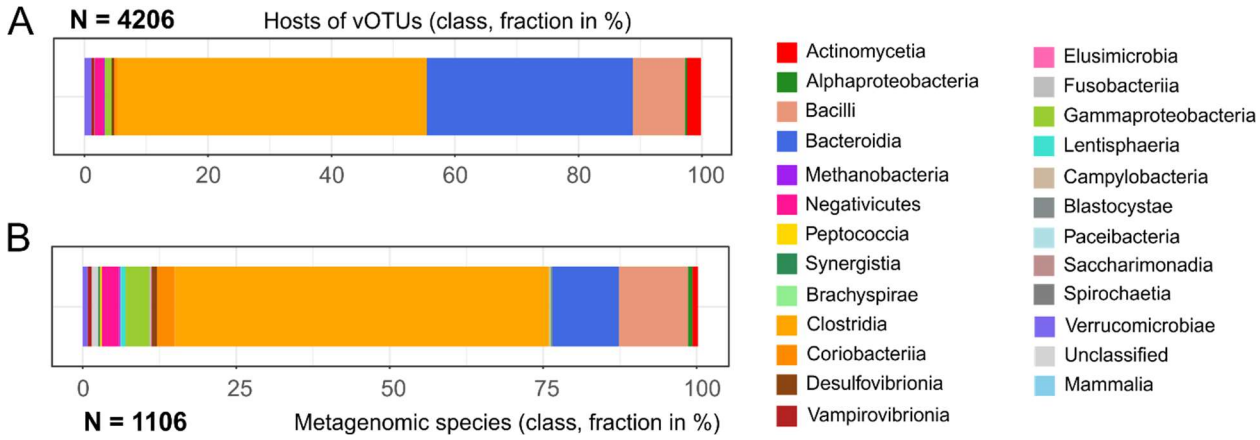


Figure S4. Taxa of hosts of vOTUs (of phage samples) and detected metagenomic species (of bacterial fractions). A. Hosts for 4206 vOTUs (on the genus level, prediction score >90) were predicted using iPHoP[45] (Table S1). Shown is the fraction of the assigned classes B. 1106 metagenomic species were detected by short reads from whole microbiome (mostly bacteria) samples obtained in our previous work[26] (see Methods).

β-lactam antibiotics induce a strong perturbation in the gut phageome with a distinct response for each individual

We assessed changes induced by the cephalosporin antibiotics along time by monitoring the total species richness (Figure 2A). On average, we detected 374 different vOTUs per sample, but the individual values were widely variable ranging from 38 to 771 (Figure S5A). Due to the high variability and specificity of phage species within individuals[12], their numbers strongly correlate with the number of samples analyzed. We therefore assessed changes in phage richness by using similar numbers of samples per time point (N=19 to 20). We excluded days with a low number of samples (two samples of day -7, and two samples of day 90), and pooled donor-specific samples of day 4 (N=19) with day 7 (N=1) and day 15 (N=6) with day 30 (N=14) (see sampling scheme in Figure S1). We observed a drop of 20.4% in total phage richness following the beginning of antibiotic treatment (days 4-7), which is comparable to the drop in bacterial species richness (19.3%, obtained from the bacterial fractions[26], Figure 2A). Since we do not expect cephalosporins to directly affect viral particles, the similar reduction in richness likely reflects the impact on the phage population of the substantial decrease in bacterial hosts following the treatment.

To assess the strength and specificity of the perturbation, we compared compositional changes per volunteer over the trial. Only donors who gave samples on the day before treatment (day -1) were considered, excluding subject 11 and 29. We computed a Shannon index normalized per donor with respect to day -1. We observed major fluctuations before and after treatment, regardless of the type of antibiotics used (Figure S5B), with a higher number of donors whose diversity decreased in the days following treatment (N=14) when compared to those where it increased (N=6, Figure 2B). In some cases, the phage populations experienced a substantial shift in diversity, while others showed a subtle change, in gaining and a decline in diversity. After 180 days, in 15 out of 20 donors, similar levels of diversity were observed in comparison to compositions before treatment. These changes parallelled those of the bacterial population, where in 18 out of 20 donors, we observed a decline in bacterial diversity (Figure S5C). Interestingly, in the two donors that showed increased bacterial diversity, we did not observe an increase in phage diversity, showing that bacterial and phage diversity are not always directly correlated.

To evaluate how the samples grouped by antibiotic type, time points, and donor, we computed the Bray-Curtis distance between all samples (Figure 2C). This revealed lower distances between samples from the same donors ($Mean_{diff}$ = 0.58) than when comparing samples across days ($Mean_{diff}$ = 0.85) or antibiotics ($Mean_{diff}$ = 0.93). This suggests that the overall individuality of each human individual[12] outweighs the effects of time and antibiotic type.

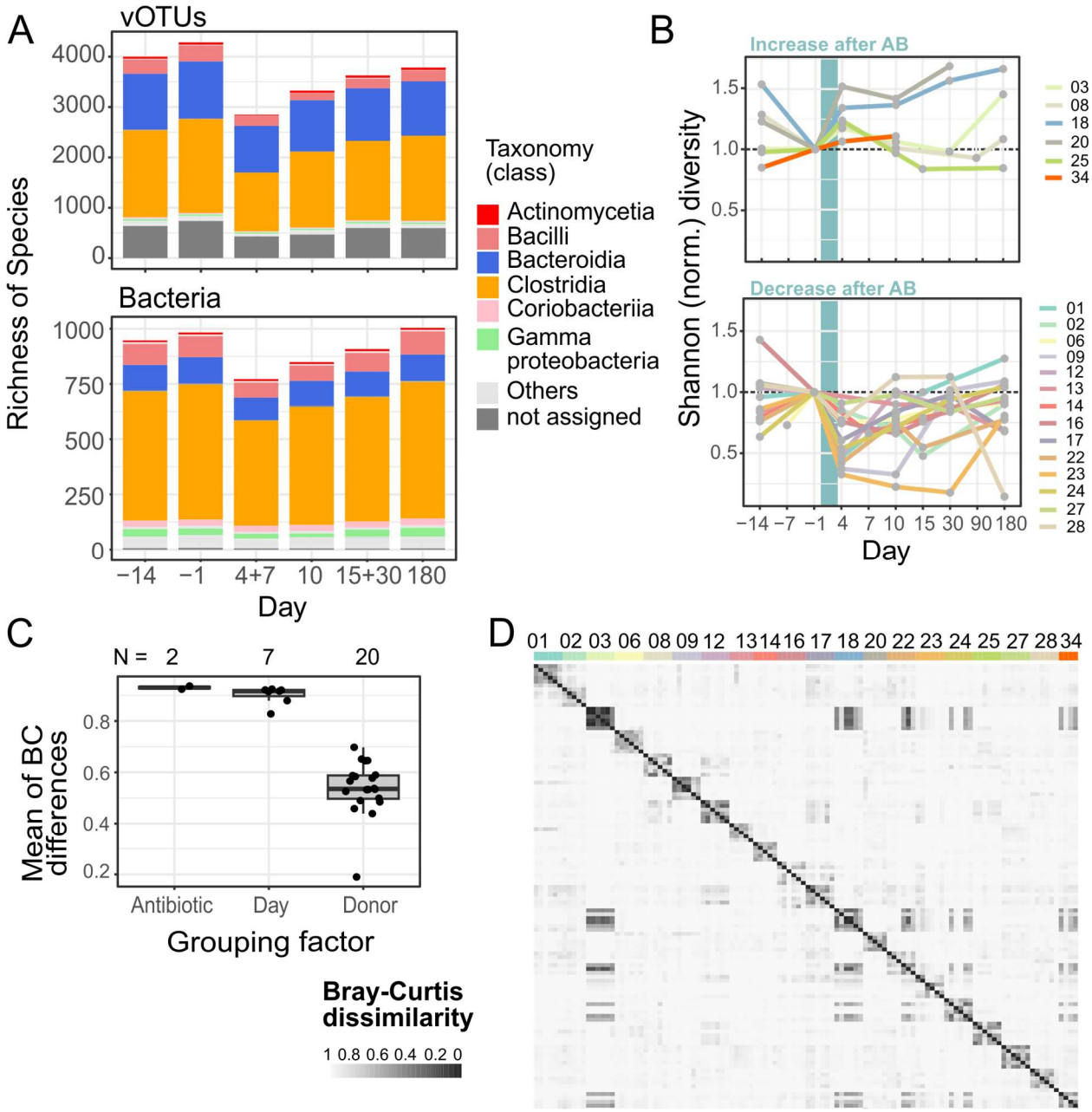


Figure 2: Perturbation effect assessed by total richness and samples'-diversities. A. Richness of vOTU and bacterial species per day of sampling. Not all volunteers provided samples on every collection day, which reduced species richness on certain days. To maintain a comparable sample size across time points, we pooled samples from day 4 with day 7, and from day 15 with day 30. For each of these pairs of days, volunteers provided a sample on only one of the two days. B. Shannon diversity index normalized to the day before treatment (day -1) for samples from all donors throughout the clinical trial (AB: antibiotics). The period of antibiotic treatment is indicated by a Cyan vertical bar. C. Mean variances between samples based on Bray-Curtis distances, grouped by antibiotic type, sampling day, and donor. D. Matrix showing Bray-Curtis distances between samples, ordered by donor and within each donor ordered by sampling day.

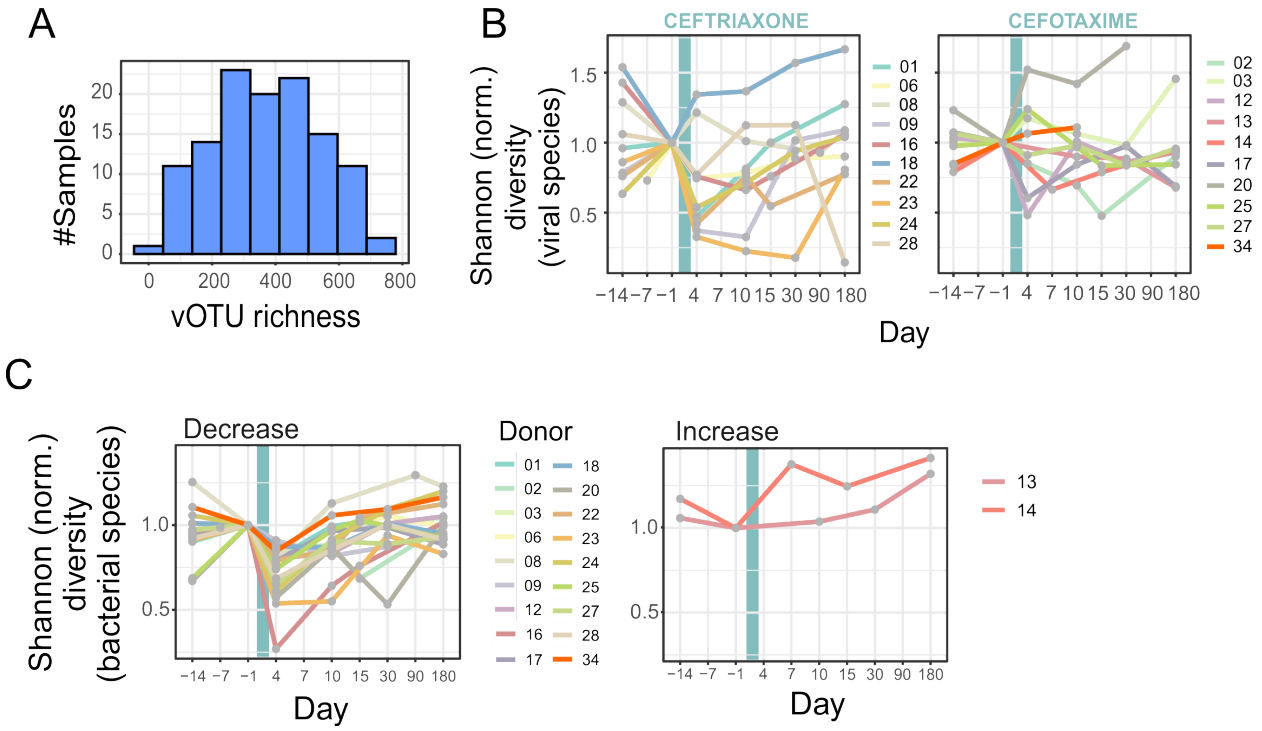


Figure S5. vOTUs and bacterial richness and diversity. A. Histogram displaying the range of richness per sample across a total of 123 samples. B. Shannon diversity of vOTUs for the samples separated by type of cephalosporin. The period of antibiotic treatment is indicated by a Cyan vertical bar. C. Normalized diversity of the bacterial species. Abundances of bacterial species were computed from full metagenome samples (mostly bacterial species) and used to calculate the Shannon diversity per sample. This index was then normalized (per donor) to the day before treatment. The individuals were split in terms of decreasing or increasing diversity after treatment.

Cephalosporin treatment does not promote phages with resistance genes

In our previous work, we observed that the frequency of genes encoding β -lactamases increased after administration of the cephalosporins[26]. We suspected that spread of resistance genes might be facilitated by gut phages encoding ARGs. To detect them, we annotated all vOTUs with AMRfinderPlus[30], and found nine phages with resistance genes. Among these phages, five were classed with high quality, four of the nine are temperate (none of which is a P-P). Four encode resistance against β-lactams, three against tetracycline, one against vancomycin, and one against streptogramin antibiotics (Figure S7). All the four β-lactamases are classed as *cfxA* that encodes a class A β-lactamase that is often detected in *Bacteroides* species[46]. This matches the corresponding genera of the vOTU hosts that were predicted to be *Bacteroides* and *Prevotella*. Notably, we did not detect any genes encoding extended-spectrum β-lactamases typical for Enterobacterales in the phage sequences, even though these genes had been found in our previous work (in bacterial fractions)[26,27].

We inspected the relative abundances of these vOTUs, focusing on those encoding β-lactam resistances and on their variation in frequency following administration of the antibiotics (Figure 3). The relative frequency remained quite constant with maximum levels being less than 0.6% in all samples. In 16 donors, no major changes were detected along the sampling period. vOTU_3934 was found in nearly all volunteers but usually at very low frequency (n=19, mean abundance 0.07% per donor and sample). However, in three donors (#34, #25 & #27) this phage showed increased levels after cephalosporin treatment. In donor #34 (orange in Figure 3), all four *cfxA* encoding vOTUs were detected, only on the day after treatment, and were not detectable anymore after day 10 (post treatment). 180 days following the administration of antibiotics, we observed an increased frequency of vOTU_3115 and vOTU_3934 in donor #1 (shown in teal in Figure 3). This change is too late in the sampling period to be related to the use of antibiotics.

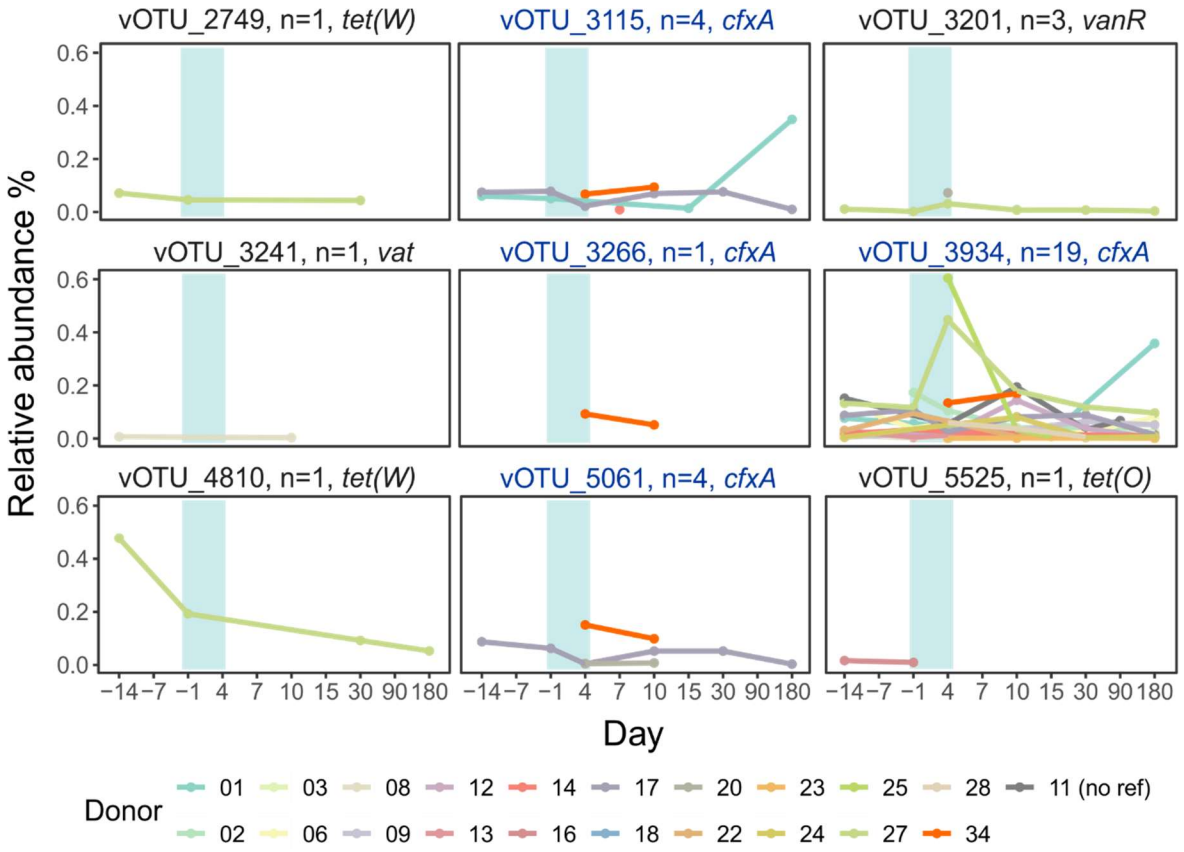


Figure 3: Relative abundance of vOTUs encoding ARGs. AMRfinderPlus detected nine vOTUs with genes encoding antibiotic resistance against: β-lactam (*cfxA*), tetracycline (*tet(O)* and *tet(W)*), vancomycin (*vanR*) and streptogramin (*vat*). N indicates the number of donors in which the vOTU was detected. The period of antibiotic treatment is indicated by the cyan bar. For donor 11, the reference sample from day -1 was not available.

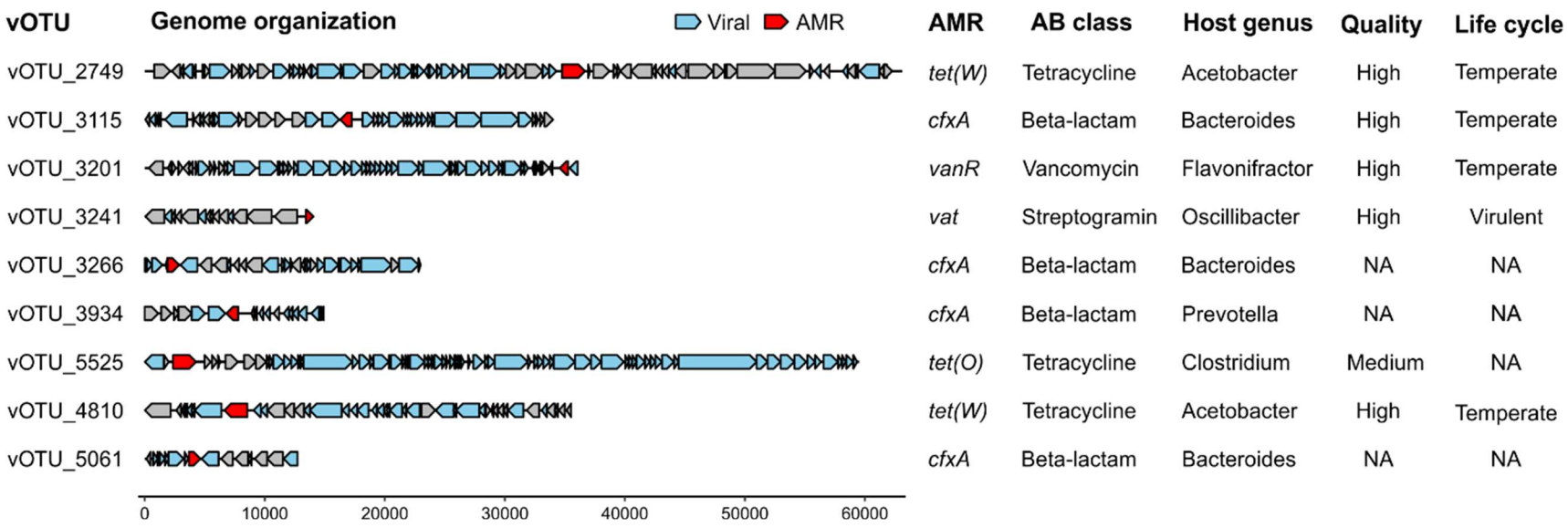


Figure S6. Genome organization of vOTUs encoding antibiotic resistance genes. Viral genes annotated by geNomad are shown in blue and ARGs by AMRFinderPlus in red.

Populations of phages undergo high fluctuations, with bursts after cephalosporin treatment

Our longitudinal study enables tracking the changes of phage species along time to assess the community dynamics. We quantified the relative abundances of all vOTUs and observed that phage populations fluctuate more widely than those of bacterial species as pointed out by higher coefficients of variation (CV) of relative abundances (Figure 4A, and Figure S7, S8A). On average, a vOTU had a relative abundance of 0.06%, and only a few species were highly abundant per sample (1% of all vOTUs with abundances >15% in at least one sample and 0.5% of vOTUs with abundances >25%) (Figure 4B). We defined phages with quantities higher than 25% as dominant in the focal sample. We found 30 such phages (Figure 4A). The dominant phages were detected in most donors, except in #08 and #34, albeit usually at much lower frequencies. Surprisingly, although most phages are predicted to be temperate, most dominant vOTUs were predicted to be virulent. In a few cases, the dominance was very high: four phages (vOTU_5377 infecting Spirochaetia, vOTU_6373 infecting *Prevotella*, vOTU_6374, and vOTU_6307) accounted each for more than 80% of the viral reads in the respective samples.

Dominant phages undergo strong fluctuations in population sizes and only five dominant phages were found to persist across multiple time points in a few individuals. We tested if these fluctuations (counts of phages per day for the 30 dominant phages) follow a Poisson regression model and found that these events follow the expected distribution on most days, with the notable except of the day after treatment (day 4). On this day, we detected a significant increase (p-value = 0.012, Figure S8C) of the number of dominant phages (Figure 4C). This conclusion was robust to changing the threshold values for the dominant vOTUs in the regression analysis (Figure S8B). Overall, this indicates that the cephalosporin treatment affects the gut microbiota in a manner that results in spikes of high frequency of certain (mostly virulent) phages.

We did a similar analysis to the bacterial fraction obtained in our previous work, using the same threshold of 25% to detect dominant bacterial species. We identified 13 different bacterial species that became highly abundant in at least one sample (Figure S7), 10 of the class Bacteroidia, two of Bacilli and one of Clostridia. Of note, in seven individuals (#06, #11, #12, #13, #17, #18, #34), we did not detect any dominant bacterial species (according to our definition). Six out of the 13 species were found to be dominant across samples indicating that dominant gut bacteria tend to fluctuate less widely than highly abundant phages and persist once strongly abundant. In line with another work[4], we observed that *P. distasonis* dominated the sample in donor #22 following cephalosporin treatment (Figure S7). In donor #16, however, we detected high abundance levels (63.5%) of *Candidatus Borkfalkia sp003343765*. Notably, another species of the same genus, *Candidatus B. ceftriaxoniphila*, has been reported to reach monodominance (abundances up to 92% after ceftriaxone administration)[4]. To explore the relationship between these bacteria and their phages, we examined their co-occurrence. While we could not assign any phages to *Candidatus Borkfalkia sp003343765,* we found a few phages associated with *P. distasonis.* Interestingly, these phages became dominant in donors #01, #17, and #24, where *P. distasonis* has been detected but it did not thrive (Figure 5). In contrast, in donor #22 (*P. distasonis* became dominant) none of the associated phages were detectable. These dynamics suggest that phages infecting *P. distasonis* may inhibit or substantially reduce its growth after cephalosporin treatment.

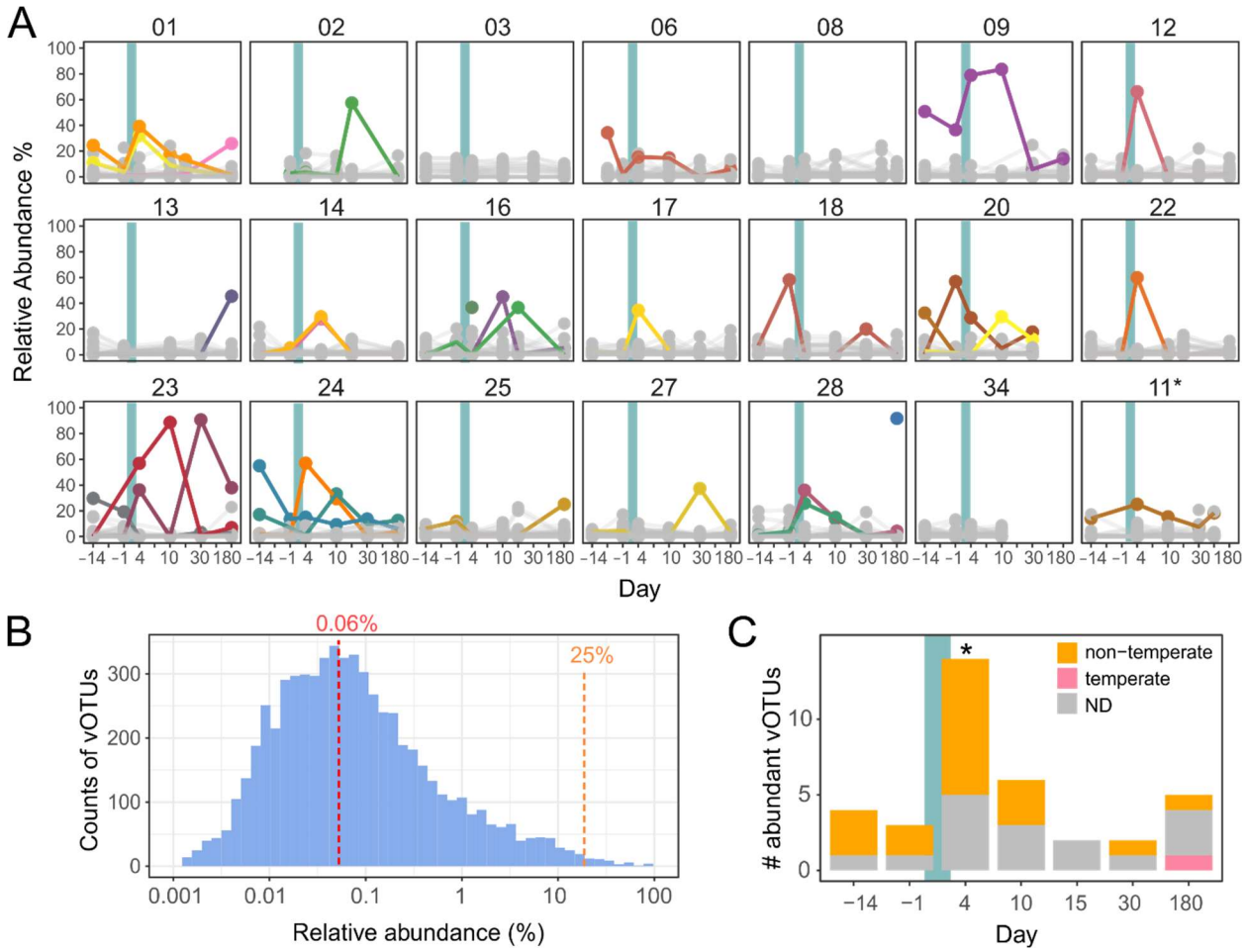


Figure 4: Dynamics of dominant gut phages. A. Shown are relative abundances of all vOTUs in all donor samples before treatment (N=21). The period of treatment with the cephalosporin antibiotics is indicated by cyan vertical bars. vOTUs were colored when their frequency exceeded 25%. B. Histogram showing maximum abundance levels of all vOTUs. If a vOTU was detected in several donors, the highest level was considered. Average level is 0.06% (in red). The orange dashed vertical line indicate the threshold of 25% used to define the dominant species. C. Daily counts of dominant vOTUs (those with at least 25%) over all donors. As in A, the period of cephalosporin treatment is indicated by a cyan vertical bar. To determine significance, daily counts were compared to a Poisson distribution (see Methods), using the day before treatment (day -1) as the baseline. For day 4, a p-value of 0.012 was obtained (Figure S8 BC), indicating counts for this day are significantly different from the Poisson model.

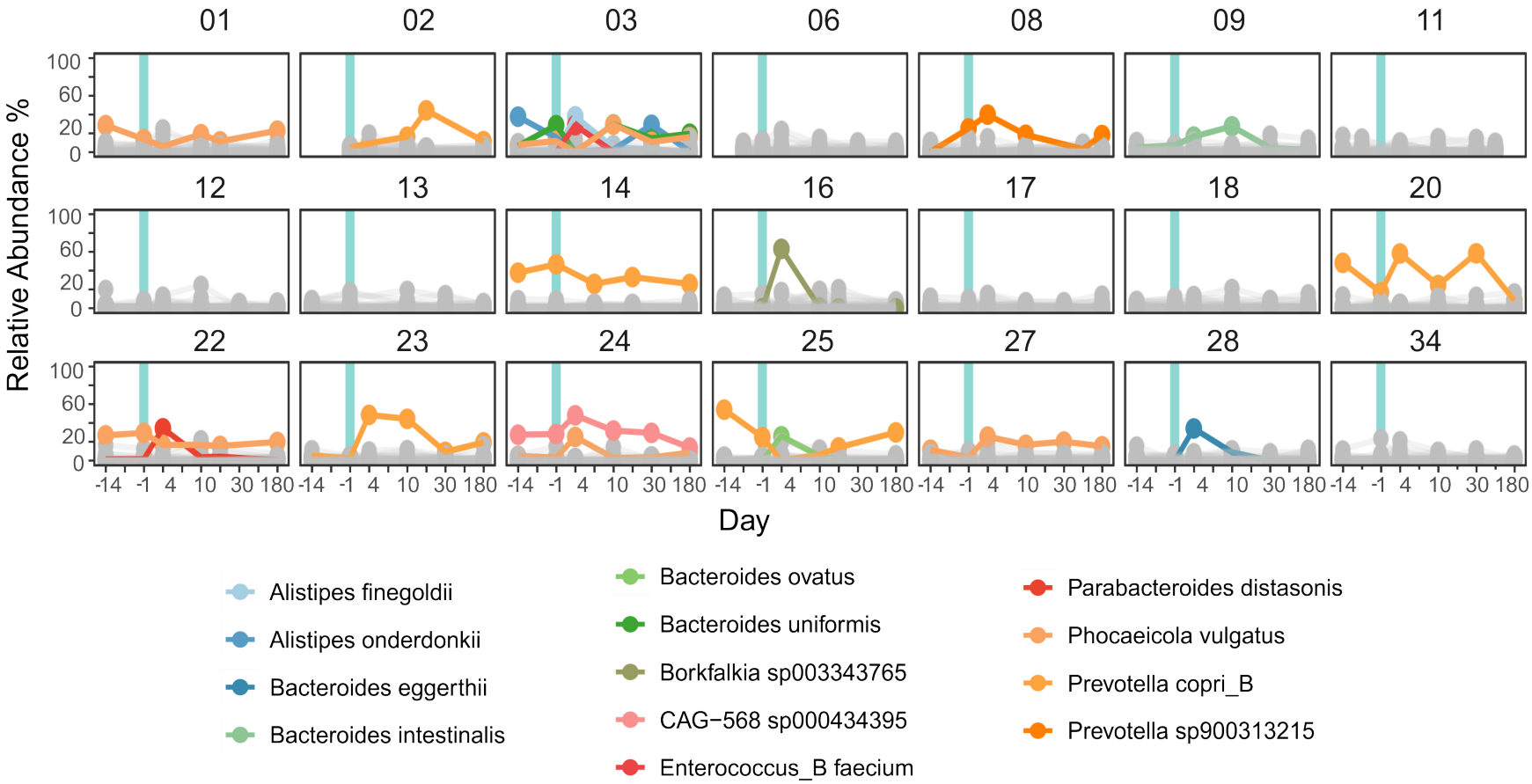
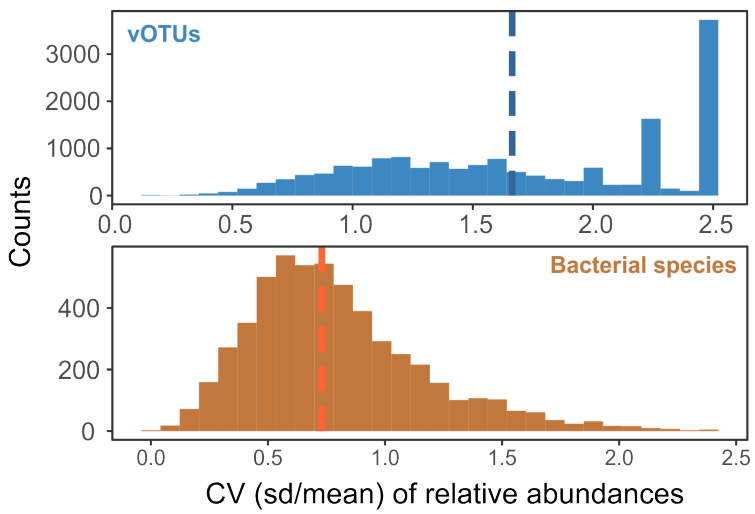
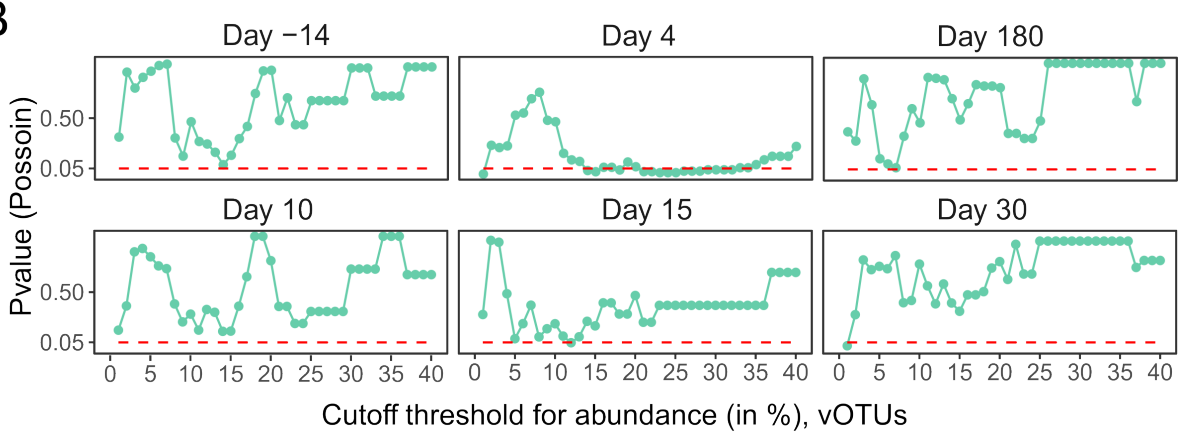


Figure S7: Relative abundances of metagenomics species (mostly bacteria) over the course of the clinical trial (as in Figure 4A for vOTUs).

A



B



C

	Day	-014	004	010	015	030	180
Reference Day-001	p-value	0.657	0.012	0.327	0.382	0.978	0.399
	Significance	ns	*	ns	ns	ns	ns

Figure S8: Analysis of gut phage dynamics. A. Relative standard deviation of abundance levels for phage and bacterial species. The coefficient of variation (CV) was calculated for all phage species (upper panel, blue) and bacterial species (lower panel, orange), with each CV based on a minimum of three values. Median values are indicated by dashed lines. Species absent in all samples from a donor were excluded from the analysis. For species detected at least once in a donor, absences were assigned a value of 0. B. The number of dominant phage species depends on the definition of dominance: in our study we used the threshold 25%. We also tested various other values in the range 0.1% to 40% for defining dominance. At each cutoff, we counted the number of dominant species and tested if these events follow a Poisson distribution. The day before treatment (day -001), was set as the reference value in the GLM (see Methods). A p-value of <0.05 (dashed, horizontal red line) indicates these events are significantly different. C. p-Values computed at a cutoff of 25% abundance (to define dominant species) are shown (exemplary).

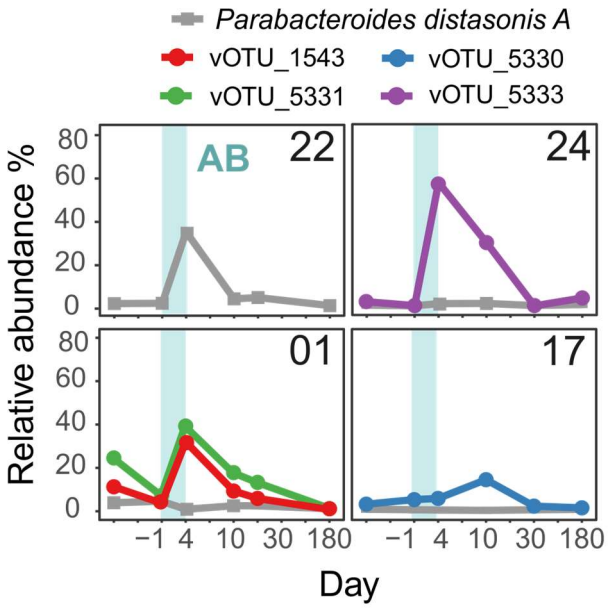


Figure 5. Abundance profile of *P. distasonis A* (grey) and its phages (purple, red, green, blue) in donor #01, #17, #22 and #24. Treatment by cephalosporin antibiotics are indicated by the cyan bar. All donors received ceftriaxone, except donor 17, who was treated with cefotaxime.

Discussion

We aimed to understand the impact of antibiotics on the gut phage diversity by analyzing the phage fraction of healthy individuals receiving a standard antibiotic treatment with 3rd generation cephalosporins. We used established protocols (fine-tuned in a past study[47]) and most recent computational resources, including the newest version of a harmonized human gut virome database, to recover a high quality dataset. Furthermore, we identified >1900 novel phage species, most of which were not closely related with any phage with a genome in GenBank. In consequence, most vOTUs (87.4%) could not be classified deeper than the class level (Caudoviricetes) and their bacterial host could only be assigned for roughly 65% of the vOTUs. These numbers are in line with those of other studies focusing on human gut phage diversity[17,48–50]. We used this vOTU characterization to track changes in bacterial and phage populations upon antibiotic treatment and found that phages exhibited stronger fluctuations in abundances compared to bacterial species. We propose that several factors are affecting these dynamics. First, the higher diversity (richness) of phages compared to bacteria suggests intense competition for bacterial hosts, leading to shifting dominance where different phages succeed over time. Second, phages replicate quickly with high burst sizes, causing a rapid increase in population numbers under favorable conditions. Finally, the high prevalence of lysogens (70–80%[18,19]) adds another layer of unpredictability. Since prophage induction varies in function of conditions and mechanisms, that are yet poorly understood, it may lead to the activation of different phages in different individuals.

We observed that the phage diversity is unique for the individual. This distinctiveness has been described before[12]. Here, we show that after antibiotic-induced perturbation, the phage communities remain specific (for the individuals). The clustering of the phage communities revealed that the differences caused by the antibiotics are of a minor impact, in confirmation with our previous results[26], and that the effect of the day of sampling is much smaller than the effect of the donors. Hence, antibiotics may disturb the phage composition, but the latter remain largely a function of the individual.

In our previous studies, we observed an increase in genes encoding β-lactamases following cephalosporin treatment and isolated *E. coli* strains harboring β-lactamase genes[26,27]. We hypothesized that gut phages might contribute to the dissemination of these resistance genes. However, our analysis reveals that only a few gut phages encode β-lactamase genes. These phages appear to play a minor role in the microbial perturbation induced by antibiotics since they are found at low abundances even just after the end of the treatment when antibiotic impact is highest. Also, several of these phages are not ranked as high quality and could be other types of mobile genetic elements. Conservatively, only one out of the four identified vOTUs (vOTU_3115) ranks high in the characteristics of a phage providing resistance to bacteria: it has a high-quality genome of more than 30 kb (typical of tailed phages) with a high number of phage genes and is predicted to be temperate thereby potentially providing resistance by lysogenic conversion. Overall, this suggests that temperate phages carrying antibiotic resistance genes are rare and remain rare upon antibiotic treatment, presumably having minor effects in these individuals.

We found that temperate phages outnumber of almost 2 to 1 the virulent ones. This is in line with past work, detecting 76% of temperate phages in a Belgian cohort[16] and 78% in A Danish cohort[17], consistent with recent findings from a quantitative study on their numbers[51]. Interestingly, we found many vOTUs (>6%) related to P-Ps, especially from *Bacillus*, *Clostridium, Corynebacterium* and *Enterococcus* species, which have rarely been described in previous studies. The identification of P-Ps poses specific challenges since these are temperate phages that usually lack an integrase and there has been a paucity of methods to identify them in metagenomes. Our analysis highlights the prevalence of P-Ps among gut phages, likely driven by the high abundance of temperate phages. These results suggest that P-Ps play a

significant role in ecological dynamics of gut phage communities, even as questions remain about their exact lifecycle mechanisms.

Even if temperate phages are usually the most abundant, the phage species that we identified as becoming dominant following cephalosporin treatment are almost all predicted to be virulent. Some studies have reported an increase in virulent gut phages following perturbations of the gut microbiota. In a 1983 study, fecal samples were analyzed for phage content, revealing that lower titers from healthy individuals were predominantly composed of temperate phages, while higher titers from patients with intestinal diseases were mainly composed of virulent phages[52]. Similar findings were reported in a 2004 study, where stool samples from 140 children with severe gastroenteritis were analyzed for coliphages, and roughly 90% were identified as virulent T-even phages[53]. Nevertheless, the increase in virulent phages observed in or study was somewhat unexpected, given that β-lactam antibiotics are known to induce the bacterial SOS response in species like *E. coli*[54], which can trigger the activation of prophages. However, the dominant phages are not predicted to be hosted by enterobacteria and we propose that the observed rise in virulent phages was driven through other circumstances. Specifically, we hypothesize that some of these phages prey on bacterial populations that temporarily became dominant after being stimulated by the antibiotic-induced perturbation. This hypothesis is supported by the here-observed dynamics of *P. distasonis* and its phages in multiple donors: high phage levels were correlated with low abundance of *P. distasonis*, while low phage levels coincided with a high abundance of the host (Figure 5). Species within the *Bacteroides* and *Parabacteroides* genera have increasingly developed resistance to certain antibiotics,[55], particularly *P. distasonis* isolates against β-lactams[56]. This increasing resistance contributes to their ability to thrive. We suggest that their phages could help to restore the diversity in the gut environment. By infecting and killing bacteria that are rising in frequency, they control the number of their hosts and regulate the bacterial populations that would otherwise have become, and potentially stay, dominant. By this, they create niches for other species to colonize the gut.

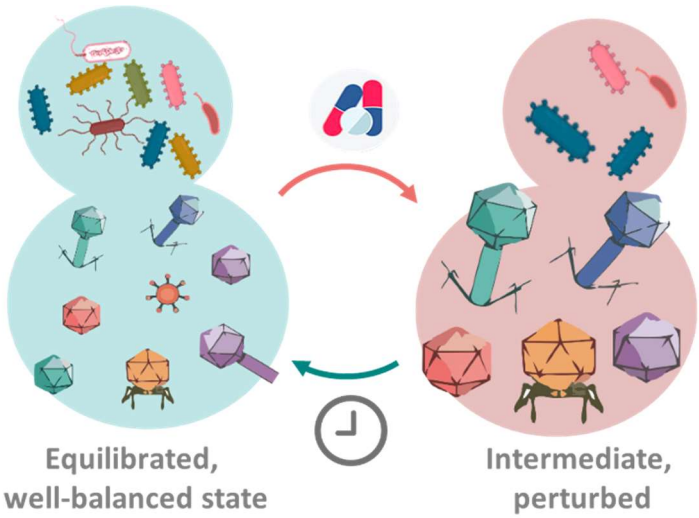
.

Limitations of the study

Our study design is limited in that it provides only a snapshot analysis, allowing us to analyze phage frequencies only in relative terms and preventing us from capturing full dynamics over time. This could have implications in certain results, notably in the study of the dominant virulent phages as it could be argued that it is not these phages that become abundant, but all the others that became extremely rare. We find this hypothesis less parsimonious, because we previously showed that many bacterial clades were little affected by the antibiotics, providing little reasons of why the frequency of their phages would be affected. Additionally, given the high phage diversity, it is more reasonable to conclude that one phage became highly frequent rather than to suggest that nearly all other phages disappeared from the samples. Future studies need to focus on these specific interactions. In particular, in vitro experiments conducted in controlled environments, such as bioreactors, could provide valuable insights. By enabling precise cultivation and sampling under antibiotic exposure[57], these experiments may significantly enhance our understanding of these dynamic interactions.

Our results suggest that over 6% of the gut phage vOTUs were P-Ps. Detecting P-Ps is particularly challenging in metagenomic datasets. We clustered the detected vOTUs with a previously reported P-P database[29]. This approach fails to detect novel P-Ps because it is limited to the diversity of the known ones, and is prone to their classification or misclassification. While P-Ps are temperate phages, they do not integrate into the bacterial chromosome and typically lack integrases. Many tools designed to type viral life cycles, including BACPHLIP[37], are trained on complete genomes and rely on typical phage genes involved in the lysogenic cycle as hallmarks (e.g. integrases and repressors). However, metagenome assemblies are often incomplete and may lack key functional modules, particularly those related to integration and lysogeny, leading to the misclassification of some phages, including P-Ps, as non-temperate.

Graphical abstract



References

1. Palleja, A., Mikkelsen, K.H., Forslund, S.K., Kashani, A., Allin, K.H., Nielsen, T., Hansen, T.H., Liang, S., Feng, Q., Zhang, C., et al. (2018). Recovery of gut microbiota of healthy adults following antibiotic exposure. Nat. Microbiol. *3*, 1255–1265. https://doi.org/10.1038/s41564-018-0257-9.

2. Crits-Christoph, A., Hallowell, H.A., Koutouvalis, K., and Suez, J. Good microbes, bad genes? The dissemination of antimicrobial resistance in the human microbiome. Gut Microbes *14*, 2055944. https://doi.org/10.1080/19490976.2022.2055944.

3. Anthony, W.E., Burnham, C.-A.D., Dantas, G., and Kwon, J.H. (2020). The Gut Microbiome as a Reservoir for Antimicrobial Resistance. J. Infect. Dis. *223*, S209–S213. https://doi.org/10.1093/infdis/jiaa497.

4. Hildebrand, F., Moitinho-Silva, L., Blasche, S., Jahn, M.T., Gossmann, T.I., Huerta-Cepas, J., Hercog, R., Luetge, M., Bahram, M., Pryszlak, A., et al. (2019). Antibiotics-induced monodominance of a novel gut bacterial order. Gut *68*, 1781–1790. https://doi.org/10.1136/gutjnl-2018-317715.

5. Stentz, R., Horn, N., Cross, K., Salt, L., Brearley, C., Livermore, D.M., and Carding, S.R. (2015). Cephalosporinases associated with outer membrane vesicles released by *Bacteroides spp.* protect gut pathogens and commensals against β-lactam antibiotics. J. Antimicrob. Chemother. *70*, 701–709. https://doi.org/10.1093/jac/dku466.

6. DeJong, E.N., Surette, M.G., and Bowdish, D.M.E. (2020). The Gut Microbiota and Unhealthy Aging: Disentangling Cause from Consequence. Cell Host Microbe *28*, 180–189. https://doi.org/10.1016/j.chom.2020.07.013.

7. Fan, Y., and Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. Nat. Rev. Microbiol. *19*, 55–71. https://doi.org/10.1038/s41579-020-0433-9.

8. Sommer, F., and Bäckhed, F. (2013). The gut microbiota--masters of host development and physiology. Nat. Rev. Microbiol. *11*, 227–238. https://doi.org/10.1038/nrmicro2974.

9. Jehl, F., Peter, J.D., Picard, A., Dupeyron, J.P., Marescaux, J., Sibilly, A., and Monteil, H. (1987). [Biliary excretion of cefotaxime and desacetylcefotaxime]. Rev. Med. Interne *8*, 487–492. https://doi.org/10.1016/s0248-8663(87)80198-4.

10. Bergan, T. (1985). [Pharmacokinetics and tissue penetration of ceftriaxone]. Rev. Med. Interne *6*, 163–177. https://doi.org/10.1016/s0248-8663(85)80061-8.

11. Burdet, C., Nguyen, T.T., Duval, X., Ferreira, S., Andremont, A., Guedj, J., and Mentré, F. (2019). Impact of Antibiotic Gut Exposure on the Temporal Changes in Microbiome Diversity. Antimicrob. Agents Chemother. *63*, e00820-19. https://doi.org/10.1128/AAC.00820-19.

12. Shkoporov, A.N., Clooney, A.G., Sutton, T.D.S., Ryan, F.J., Daly, K.M., Nolan, J.A., McDonnell, S.A., Khokhlova, E.V., Draper, L.A., Forde, A., et al. (2019). The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. Cell Host Microbe *26*, 527-541.e5. https://doi.org/10.1016/j.chom.2019.09.009.

13. Breitbart, M., Bonnain, C., Malki, K., and Sawaya, N.A. (2018). Phage puppet masters of the marine microbial realm. Nat. Microbiol. *3*, 754–766. https://doi.org/10.1038/s41564-018-0166-y.

14. Castledine, M., and Buckling, A. (2024). Critically evaluating the relative importance of phage in shaping microbial community composition. Trends Microbiol. *32*, 957–969. https://doi.org/10.1016/j.tim.2024.02.014.

15. Pfeifer, E., Moura de Sousa, J.A., Touchon, M., and Rocha, E.P.C. (2021). Bacteria have numerous distinctive groups of phage-plasmids with conserved phage and variable plasmid gene repertoires. Nucleic Acids Res. *49*, 2655–2673. https://doi.org/10.1093/nar/gkab064.

16. Beller, L., Deboutte, W., Vieira-Silva, S., Falony, G., Tito, R.Y., Rymenans, L., Yinda, C.K., Vanmechelen, B., Van Espen, L., Jansen, D., et al. (2022). The virota and its transkingdom interactions in the healthy infant gut. Proc. Natl. Acad. Sci. U. S. A. *119*, e2114619119. https://doi.org/10.1073/pnas.2114619119.

17. Shah, S.A., Deng, L., Thorsen, J., Pedersen, A.G., Dion, M.B., Castro-Mejía, J.L., Silins, R., Romme, F.O., Sausset, R., Jessen, L.E., et al. (2023). Expanding known viral diversity in the healthy infant gut. Nat. Microbiol. *8*, 986–998. https://doi.org/10.1038/s41564-023-01345-7.

18. Kim, M.-S., and Bae, J.-W. (2018). Lysogeny is prevalent and widely distributed in the murine gut microbiota. ISME J. *12*, 1127–1141. https://doi.org/10.1038/s41396-018-0061-9.

19. Sutcliffe, S.G., Reyes, A., and Maurice, C.F. (2023). Bacteriophages playing nice: Lysogenic bacteriophage replication stable in the human gut microbiota. iScience *26*, 106007. https://doi.org/10.1016/j.isci.2023.106007.

20. Ubeda, C., Maiques, E., Knecht, E., Lasa, I., Novick, R.P., and Penadés, J.R. (2005). Antibiotic-induced SOS response promotes horizontal dissemination of pathogenicity island-encoded virulence factors in staphylococci. Mol. Microbiol. *56*, 836–844. https://doi.org/10.1111/j.1365-2958.2005.04584.x.

21. Misson, P., Bruder, E., Cornuault, J.K., De Paepe, M., Nicolas, P., Demarre, G., Lakisic, G., Petit, M.-A., Espeli, O., and Lecointe, F. (2023). Phage production is blocked in the adherent-invasive *Escherichia coli* LF82 upon macrophage infection. PLoS Pathog. *19*, e1011127. https://doi.org/10.1371/journal.ppat.1011127.

22. Pfeifer, E., Bonnin, R.A., and Rocha, E.P.C. (2022). Phage-Plasmids Spread Antibiotic Resistance Genes through Infection and Lysogenic Conversion. mBio *13*, e0185122. https://doi.org/10.1128/mbio.01851-22.

23. Wang, L., Yao, H., Morgan, D.C., Lau, K.S., Leung, S.Y., Ho, J.W.K., and Leung, W.K. (2023). Altered human gut virome in patients undergoing antibiotics therapy for *Helicobacter pylori*. Nat. Commun. *14*, 2196. https://doi.org/10.1038/s41467-023-37975-y.

24. Enault, F., Briet, A., Bouteille, L., Roux, S., Sullivan, M.B., and Petit, M.-A. (2017). Phages rarely encode antibiotic resistance genes: a cautionary tale for virome analyses. ISME J. *11*, 237–247. https://doi.org/10.1038/ismej.2016.90.

25. Kang, K., Imamovic, L., Misiakou, M.-A., Bornakke Sørensen, M., Heshiki, Y., Ni, Y., Zheng, T., Li, J., Ellabaan, M.M.H., Colomer-Lluch, M., et al. (2021). Expansion and persistence of antibiotic-specific resistance genes following antibiotic treatment. Gut Microbes *13*, 1–19. https://doi.org/10.1080/19490976.2021.1900995.

26. d'Humières, C., Delavy, M., Alla, L., Ichou, F., Gauliard, E., Ghozlane, A., Levenez, F., Galleron, N., Quinquis, B., Pons, N., et al. (2024). Perturbation and resilience of the gut microbiome up to 3 months after β-lactams exposure in healthy volunteers suggest an important role of microbial β-lactamases. Microbiome *12*, 50. https://doi.org/10.1186/s40168-023-01746-0.

27. Burdet, C., Grall, N., Linard, M., Bridier-Nahmias, A., Benhayoun, M., Bourabha, K., Magnan, M., Clermont, O., d'Humières, C., Tenaillon, O., et al. (2019). Ceftriaxone and Cefotaxime Have Similar Effects on the Intestinal Microbiota in Human Volunteers Treated by Standard-Dose Regimens. Antimicrob. Agents Chemother. *63*, e02244-18. https://doi.org/10.1128/AAC.02244-18.

28. Léonard, F., Andremont, A., Leclerq, B., Labia, R., and Tancrède, C. (1989). Use of β-Lactamase-Producing Anaerobes to Prevent Ceftriaxone from Degrading Intestinal Resistance to Colonization. J. Infect. Dis. *160*, 274–280. https://doi.org/10.1093/infdis/160.2.274.

29. Pfeifer, E., and Rocha, E.P.C. (2024). Phage-plasmids promote recombination and emergence of phages and plasmids. Nat. Commun. *15*, 1545. https://doi.org/10.1038/s41467-024-45757-3.

30. Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J.G., Haendiges, J., Haft, D.H., Hoffmann, M., Pettengill, J.B., Prasad, A.B., Tillman, G.E., et al. (2021). AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci. Rep. *11*, 12728. https://doi.org/10.1038/s41598-021-91456-0.

31. Camargo, A.P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B., Chain, P.S.G., Nayfach, S., and Kyrpides, N.C. (2024). Identification of mobile genetic elements with geNomad. Nat. Biotechnol. *42*, 1303–1312. https://doi.org/10.1038/s41587-023-01953-y.

32. Antipov, D., Raiko, M., Lapidus, A., and Pevzner, P.A. (2020). MetaviralSPAdes: assembly of viruses from metagenomic data. Bioinformatics *36*, 4126–4129. https://doi.org/10.1093/bioinformatics/btaa490.

33. Kieft, K., Zhou, Z., and Anantharaman, K. (2020). VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome *8*, 90. https://doi.org/10.1186/s40168-020-00867-0.

34. Nayfach, S., Camargo, A.P., Schulz, F., Eloe-Fadrosh, E., Roux, S., and Kyrpides, N.C. (2021). CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat. Biotechnol. *39*, 578–585. https://doi.org/10.1038/s41587-020-00774-7.

35. Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z.J., Pollard, K.S., Sakharova, E., Parks, D.H., Hugenholtz, P., et al. (2021). A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat. Biotechnol. *39*, 105–114. https://doi.org/10.1038/s41587-020-0603-3.

36. Harding, K.R., Kyte, N., and Fineran, P.C. (2023). Jumbo phages. Curr. Biol. CB *33*, R750–R751. https://doi.org/10.1016/j.cub.2023.05.056.

37. Hockenberry, A.J., and Wilke, C.O. (2021). BACPHLIP: predicting bacteriophage lifestyle from conserved protein domains. PeerJ *9*, e11396. https://doi.org/10.7717/peerj.11396.

38. Dong, Z., Li, J., Zheng, J., Geng, C., Peng, D., and Sun, M. (2016). Complete genome sequence of *Bacillus thuringiensis* CTC-A typical strain with high production of S-layer proteins. J. Biotechnol. *220*, 100–101. https://doi.org/10.1016/j.jbiotec.2015.12.027.

39. Máté de Gérando, H., Wasels, F., Bisson, A., Clement, B., Bidard, F., Jourdier, E., López-Contreras, A.M., and Lopes Ferreira, N. (2018). Genome and transcriptome of the natural isopropanol producer *Clostridium beijerinckii* DSM6423. BMC Genomics *19*, 242. https://doi.org/10.1186/s12864-018-4636-7.

40. Tippelt, A., Möllmann, S., Albersmeier, A., Jaenicke, S., Rückert, C., and Tauch, A. (2014). Mycolic Acid Biosynthesis Genes in the Genome Sequence of *Corynebacterium atypicum* DSM 44849. Genome Announc. *2*, e00845-14. https://doi.org/10.1128/genomeA.00845-14.

41. Gervasi, T., Curto, R.L., Narbad, A., and Mayer, M.J. (2013). Complete genome sequence of ΦCP51, a temperate bacteriophage of *Clostridium perfringens*. Arch. Virol. *158*, 2015–2017. https://doi.org/10.1007/s00705-013-1647-1.

42. Schmidtke, D.T., Hickey, A.S., Liachko, I., Sherlock, G., and Bhatt, A.S. (2024). Analysis and culturing of the prototypic crAssphage reveals a phage-plasmid lifestyle. Preprint at bioRxiv, https://doi.org/10.1101/2024.03.20.585998 https://doi.org/10.1101/2024.03.20.585998.

43. Tamai, E., Yoshida, H., Sekiya, H., Nariya, H., Miyata, S., Okabe, A., Kuwahara, T., Maki, J., and Kamitori, S. (2014). X-ray structure of a novel endolysin encoded by episomal phage phiSM101 of *Clostridium perfringens*. Mol. Microbiol. *92*, 326–337. https://doi.org/10.1111/mmi.12559.

44. Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J.H., Roux, S., Adriaenssens, E.M., Brister, J.R., Kropinski, A.M., Krupovic, M., Lavigne, R., et al. (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat. Biotechnol. *37*, 632–639. https://doi.org/10.1038/s41587-019-0100-8.

45. Roux, S., Camargo, A.P., Coutinho, F.H., Dabdoub, S.M., Dutilh, B.E., Nayfach, S., and Tritt, A. (2023). iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. PLOS Biol. *21*, e3002083. https://doi.org/10.1371/journal.pbio.3002083.

46. Smith, C.J. (2002). What Constitutes an Extended-Spectrum β-lactamase? Antimicrob. Agents Chemother. *46*, 600–601. https://doi.org/10.1128/AAC.46.2.600-601.2002.

47. d'Humières, C., Touchon, M., Dion, S., Cury, J., Ghozlane, A., Garcia-Garcera, M., Bouchier, C., Ma, L., Denamur, E., and P.C.Rocha, E. (2019). A simple, reproducible and cost-effective procedure to analyse gut phageome: from phage isolation to bioinformatic approach. Sci. Rep. *9*, 11331. https://doi.org/10.1038/s41598-019-47656-w.

48. Li, J., Yang, F., Xiao, M., and Li, A. (2022). Advances and challenges in cataloging the human gut virome. Cell Host Microbe *30*, 908–916. https://doi.org/10.1016/j.chom.2022.06.003.

49. Camarillo-Guerrero, L.F., Almeida, A., Rangel-Pineros, G., Finn, R.D., and Lawley, T.D. (2021). Massive expansion of human gut bacteriophage diversity. Cell *184*, 1098-1109.e9. https://doi.org/10.1016/j.cell.2021.01.029.

50. Zeng, S., Almeida, A., Li, S., Ying, J., Wang, H., Qu, Y., Paul Ross, R., Stanton, C., Zhou, Z., Niu, X., et al. (2024). A metagenomic catalog of the early-life human gut virome. Nat. Commun. *15*, 1864. https://doi.org/10.1038/s41467-024-45793-z.

51. Lopez, J.A., McKeithen-Mead, S., Shi, H., Nguyen, T.H., Huang, K.C., and Good, B.H. (2024). Abundance measurements reveal the balance between lysis and lysogeny in the human gut microbiome. BioRxiv Prepr. Serv. Biol., 2024.09.27.614587. https://doi.org/10.1101/2024.09.27.614587.

52. Furuse, K., Osawa, S., Kawashiro, J., Tanaka, R., Ozawa, A., Sawamura, S., Yanagawa, Y., Nagao, T., and Watanabe, I. (1983). Bacteriophage distribution in human faeces: continuous survey of healthy subjects and patients with internal and leukaemic diseases. J. Gen. Virol. *64 (Pt 9)*, 2039–2043. https://doi.org/10.1099/0022-1317-64-9-2039.

53. Chibani-Chennoufi, S., Sidoti, J., Bruttin, A., Dillmann, M.-L., Kutter, E., Qadri, F., Sarker, S.A., and Brüssow, H. (2004). Isolation of *Escherichia coli* Bacteriophages from the Stool of Pediatric Diarrhea Patients in Bangladesh. J. Bacteriol. *186*, 8287–8294. https://doi.org/10.1128/JB.186.24.8287-8294.2004.

54. Miller, C., Thomsen, L.E., Gaggero, C., Mosseri, R., Ingmer, H., and Cohen, S.N. (2004). SOS response induction by beta-lactams and bacterial defense against antibiotic lethality. Science *305*, 1629–1631. https://doi.org/10.1126/science.1101630.

55. Kierzkowska, M., Majewska, A., and Mlynarczyk, G. (2020). Trends and Impact in Antimicrobial Resistance Among Bacteroides and Parabacteroides Species in 2007-2012 Compared to 2013-2017. Microb. Drug Resist. Larchmt. N *26*, 1452–1457. https://doi.org/10.1089/mdr.2019.0462.

56. Ezeji, J.C., Sarikonda, D.K., Hopperton, A., Erkkila, H.L., Cohen, D.E., Martinez, S.P., Cominelli, F., Kuwahara, T., Dichosa, A.E.K., Good, C.E., et al. *Parabacteroides distasonis*: intriguing aerotolerant gut anaerobe with emerging antimicrobial resistance and pathogenic and probiotic roles in human health. Gut Microbes *13*, 1922241. https://doi.org/10.1080/19490976.2021.1922241.

57. Guzman-Rodriguez, M., McDonald, J.A.K., Hyde, R., Allen-Vercoe, E., Claud, E.C., Sheth, P.M., and Petrof, E.O. (2018). Using bioreactors to study the effects of drugs on the human microbiota. Methods San Diego Calif *149*, 31–41. https://doi.org/10.1016/j.ymeth.2018.08.003.

58. Didion, J.P., Martin, M., and Collins, F.S. (2017). Atropos: specific, sensitive, and speedy trimming of sequencing reads. PeerJ *5*, e3720. https://doi.org/10.7717/peerj.3720.

59. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

60. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: A New Genome Assembly Algorithm and Its

Applications to Single-Cell Sequencing. J. Comput. Biol. *19*, 455–477. https://doi.org/10.1089/cmb.2012.0021.

61. Vasimuddin, Md., Misra, S., Li, H., and Aluru, S. (2019). Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 314–324. https://doi.org/10.1109/IPDPS.2019.00041.

62. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

63. Marçais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., and Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. PLoS Comput. Biol. *14*, e1005944. https://doi.org/10.1371/journal.pcbi.1005944.

64. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421. https://doi.org/10.1186/1471-2105-10-421.

65. Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. Proc. Int. AAAI Conf. Web Soc. Media *3*, 361–362. https://doi.org/10.1609/icwsm.v3i1.13937.

66. Zeileis, A., Kleiber, C., and Jackman, S. (2008). Regression Models for Count Data in R. J. Stat. Softw. *27*, 1–25. https://doi.org/10.18637/jss.v027.i08.

Resource availability

Materials availability

CEREMI trial (ClinicalTrials.gov identifier NCT02659033)

Data and code availability

The metagenomic shotgun sequencing data of bacterial microbiome are available from the European Nucleotide Archive under accession number PRJEB58157 (www.ebi.ac.uk/ena/browser/view/PRJEB58157). The metagenomic shotgun sequencing data of phage microbiome are available from the European Nucleotide Archive under accession number PRJEB58815 (www.ebi.ac.uk/ena/browser/view/PRJEB58815).
Rscripts, data tables, sequence and meta-information to reproduce figures and anaylysis are available on GitHub and Zenodo repositories under following links: https://github.com/EpfeiferNutri/PrediRes.git and https://doi.org/10.5281/zenodo.14758483.

Consortia

Members of the PrediRes study group not otherwise listed as authors are listed here:
Xavier Duval (INSERM, Univ Paris Diderot, APHP-Bichat Hospital), Dusko Ehrlich (INRAE Metagenopolis), Laurie Alla (INRAE Metagenopolis), Emmanuelle Le Chatelier (INRAE Metagenopolis), Florence Levenez (INRAE Metagenopolis), Nathalie Galleron (INRAE Metagenopolis), Nicolas Pons (INRAE Metagenopolis), Benoît Quinquis (INRAE Metagenopolis), Khadija Bourabha (INSERM), Antoine Bridier Nahmias (INSERM, Univ Paris Diderot), Olivier Clermont (INSERM, Univ Paris Diderot), Mélanie Magnan (INSERM, Univ Paris Diderot), Dominique Rainteau (INSERM, Univ Pierre et Marie Curie, APHP–Saint Antoine Hospital), Antonin Lamazière (INSERM, Univ Pierre et Marie Curie, APHP–Saint Antoine Hospital), Emilie Gauliard (INSERM, Univ Pierre et Marie Curie, APHP–Saint Antoine Hospital), Farid Ichou (ICAN), Philippe Lesnik (ICAN), Marie Lhomme (ICAN). Jimmy Mullaert (INSERM, Univ Paris Diderot, APHP—Bichat Hospital), Thu Thuy Nguyen (INSERM).

Acknowledgements

This work was supported in part by the Laboratoire d'Excellence IBEID Integrative Biology of Emerging Infectious Diseases (grant ANR-10-LABX-62-IBEID), the PREDIRES project (grant ANR-16-CE15-0022), and the French junior professor chair ANR grant (ANR-22-CPJ1-0041-01).

High-throughput sequencing was performed on the Institut Pasteur Genomics Platform, member of "France Génomique" consortium (ANR10-INBS-09-08). This work used the computational and storage services (TARS cluster) provided by the IT department at Institut Pasteur, Paris.

Author contributions

Conceptualization: EP, CdH, EPCR, ED
Biological analysis of the samples: CdH, SD
Computational analyses: EP, CdH, QLB, BC, FPO
Contributed material, software or expertise: MT, CB, RD, FPO
Sequencing of the samples: LM
Writing—original draft and designed figures: EP
Writing—review & editing: EP, CdH, ED, EPCR;
All authors revised and contributed to the current draft of the manuscript.
Supervision and funding acquisition: CB, FM, ED, EPCR.

Declaration of interest

The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work the authors used DeepL (DeepL SE, (2017)) and ChatGPT (OpenAI, (2021)) in order to improve the language and readability. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Supplemental information titles and legends

Supplementary file 1: Homology between phage-plasmids and vOTUs.
File contains comparison of vOTU detection methods (slide 2), summaries of vOTU and phage-plasmid networks based on wGRR (slide 3) and vConTACT v2 (slide 4). Synteny plots (generated with gggenomes, slide 5-32) show detailed comparisons of examples (vOTU and P-Ps).

Table S1: Summary of vOTU characteristics.
Table S2: Read signal and its allocation on different gut microbiome related datasets.
Table S3: Summary of phage-plasmids to vOTU comparison.
Table S4: Abundance profiles of vOTUs.
Table S5: Summary of microbial species.
Table S6: Relative abundance profile of microbial species.

Methods

Collection and Background of Fecal Samples

This study is based on a previous work[26] of the CEREMI clinical trial. In particular, 22 healthy volunteers, both male and female, aged 18 to 65 years, with no exposure to antibiotics in the preceding 3 months or hospitalization in the last 6 months, participated in the study. In the clinical trial, donors received intravenous treatment with either 1 g ceftriaxone once daily (n=11) or 1 g cefotaxime three times a day (n=11) for three consecutive days. Fecal samples were collected starting from 15 days before to 180 days after treatment for most donors (detailed sampling scheme shown in Figure S1). For the majority (n=18 donors), six time points are available, for 3 donors five samples and for one donor with four samples, no time point before treatment was available. Only for donor 29, no samples prior to the treatment were available. For this reason, we used the read sets from donor 29 only used for assemblies and their characterization, but excluded these sample from all ecological and abundance analysis.

Preparation of virome DNA and sequencing

Viral particles were prepared as described in previous work[26,47] from samples collected on days ranging from -15, to 180 days after treatment with the two antibiotics (Figure S1). Phages were precipitated following a polyethylene glycol (PEG) protocol as evaluated in a previous work[47]. Briefly, 1 gram of fecal sample was homogenized in 40 mL of phosphate-buffered saline (PBS) (Sigma-Aldrich, Saint-Louis, MO, USA), agitated with a mechanical laboratory agitator for 1 hour at 4 °C, centrifuged at 17,000 g for 5 min, and filtered at 2 µm and 0.45 µm. Phages were precipitated using PEG by adding 1 M solid NaCl and 10% (v/v) PEG 8000 (Sigma-Aldrich, Saint-Louis, MO, USA), and incubated overnight at 4 °C. Particles were pelleted (at 5,250 g for 1 hour at 4 °C), and re-suspended in 500 µL of SM buffer (NaCl 100mM, $MgSO_4$.7xH2O 8mM, Tris-Cl 50mM). Bacterial DNA were digested by DNAse (10 U/mL) (Sigma-Aldrich, Saint-Louis, MO, USA) for 30 min at 37 °C and then heated to 65 °C for 10 min to stop the reaction. DNA extraction was done using the "Phage DNA extraction" kit (Norgen Biotek Corp, Thorold, ON, Canada), purified on a Sephadex column (Sigma-Aldrich, Saint-Louis, MO, USA), quantified using the Qubit dsDNA HS Assay kit (ThermoFisher Scientific, Waltham, MA, USA), and sequenced using the Illumina HiSeq2500 PE_250 method with the Nextera XT kit and an input of 1 ng DNA.

Read Processing, Assembly, and Curation of Multimeric Sequences

Illumina adapters and low-quality reads were removed using Atropos (v1.1.18)[58] (trim -m 100 –q 20,20 –trim-n), and subsequently the reads were mapped against the human genome (GRCh38) with Bowtie2[59] v.2.5.1 (--very-sensitive), retaining only high-quality non-human (HQ-nH) reads for downstream analysis. On average, 8.85 +/- 3.04 Mio. reads were available (see Figure S3, Table S2).

Contigs were assembled by following three approaches:

1. HQ-nH reads were pooled per donor and assembled using SPAdes[60] v.3.15.5 in metaviral mode (---metaviral)[32]. Assemblies larger 3kb, were validated by mapping reads back to them using bwa-mem2[61] v. 2.2.1 and keeping only alignments of good quality using msamtools v.1.1.3 (https://github.com/arumugamlab/msamtools) (filter -S -b -l 80 -p 95 -z 80).
2. Reads not matching were saved (via with samtools[62] 1.4 (samtools fastq -1 read.1.fq.gz -2 read.2.fq.gz -s /dev/null -o /dev/null -c 6 -N -) and mapped per sample to the unified human gut virome database (https://github.com/snayfach/UHGV) (UHGV) (https://portal.nersc.gov/UHGV/, last accessed on 19.01.2024). Viral operational taxonomic units (vOTUs) that were covered by at least 50% of sequence length with an average depth of 1.

3. In a few samples, more than 90% of reads did not align with any sequences obtained by the first two approaches and or to any further datasets (see Assessing Quality of Read Signal by Mapping). Of these low-covered samples, we randomly extracted reads (10% and 1%) using seqtk v.1.3 (https://github.com/lh3/seqtk). For 10%: sample -s100 read1.fq 0.1 > sub01.fq), and repeated the assembly with SPAdes[60] v.3.15.5 (in default mode. The assemblies, referred as sub sampling approach, resulted in a higher final coverage (fractions shown in Figure S3).

Contigs from the three approaches were pooled, and only sequences >3kb were used to predict viral contigs (N=8966).
In addition, we screened these sequences for concatemeric forms using MUMmer[63] 4.0.0 (exact-tandems contigs.fna 1000), and by reporting only sequences with tandems >1000 bp. Positive candidates were manually inspected and resolved to monomeric forms. In total, we found two of the total 8966 sequences as concatemeric.

Defining viral contigs

Contigs were classed as viral if geNomad[31] v.1.7.4 (end-to-end), VIBRANT[33] v. 1.2.1 (VIBRANT_run.py), or viralVerify v.1.1[32] (default) detected them as positive hits. Proviral sequences reported by geNomad and VIBRANT were added in a separate category ("proviral") and not classed as viral. Intersection of the viral classification tools are shown in Figure 1.

Clustering viral contigs into viral Operational Taxonomical Units

To define viral Operational Taxonomical Units (vOTUs) and curate the sequences from redundancy, we pooled all viral, assembled contigs with the ones detected in UHGV. We clustered sequences into vOTUs using a customized workflow. First, we computed the average nucleotide identity (ANI) between all sequences using blastn v.2.15.0[64] and applied the anicalc.py script by CheckV[34] 1.0.1 to compute the average nucleotide identity (ANI) between all contigs. Then, we constructed a network, where nodes are viral contigs and edges represent ANI values matching the viral species definitions (at least 95% ANI and 85% alignment fraction (AF) in reference to the shorter contig). Largest viral contigs were chosen as the representative sequences (vOTUs). Connected nodes were extracted as clusters. Singletons were directly considered as vOTUs. We observed that in some clusters some nodes were not directly connected to each other (belong to the same species), but by inter-connecting nodes which may bridge two related species level. To control for this, we performed an iterative curation of the clustered network: We first ranked contigs by size (largest first, since it is the representative) in ranges of 3 kb. Then, all sequences related to the one with the highest rank were removed (from all clusters) from the network since they represent redundant sequences and belong to the same species. The elimination results in a new, less complex and connected, network. We then defined again connected nodes as clusters, repeat the elimination and curate step-wise for inter-connected sequences (from the largest to the smallest contigs) until the network consist of only of singletons with defined vOTUs. Overall, this approach resulted in 6467 non-redundant vOTUs (see Table S1).

Taxonomical and functional annotation of viral contigs

We applied CheckV[34] v.1.01 (with database version 1.5) to assess the quality and completeness of viral contigs. Taxonomy was predicted with geNomad[31] and vConTACT v2[44]. geNomad gives directly a viral taxonomy in the output table. For vConTACT v2, we clustered 27 022 phage sequences (retrieved from GenBank (accessed in February 2024) and 168536 viral sequences from UHGV with all 6467 vOTUs. If taxonomical assignments were inconsistent, we reported predictions from the two tools. For the vOTUs retrieved from UHGV, we kept the assigned taxonomy. iPHoP[45] v.1.3.3 (database version September 2021) was applied to predict the bacterial hosts, assigning 4206 vOTUs. Notably, iPHoP is primarily

trained for bacterial and archaeal viruses and exhibits a high false-positive rate for eukaryotic viruses[45]. It predicted a bacterial host for 27 out of the 31 eukaryotic vOTUs, consequently, we have low confidence in these predictions. BACPHLIP[37] v.0.9.6 was applied on contigs of complete and high quality to predict viral life cycles. Antibiotic resistance genes were annotated by AMRfinderPlus[30] v.3.12.8 on genes called by geNomad[31].

Clustering of vOTUs with phage-plasmids

We used our database of 1416 P-Ps to search for homologous sequences among the vOTUs. To achieve this, we grouped them by two approaches: 1) vConTACT v2[44] (default) and 2) weighted gene repertoire relatedness (wGRR). wGRR was computed as described in previous works[15,29]. vOTUs with at least 10 genes and a relatedness of wGRR > 0.3 to a P-P were considered as potential P-Ps. Networks were visualized using Gephi[65] and genome relations were visualized using gggenomes (https://thackl.github.io/gggenomes/). A brief comparison of the two methods showed that they work in a nearly complementary manner. While vConTACT v2 detected distant relationships that do not meet the wGRR >0.3 threshold, it tends to miss smaller genomes or genomes with only a few edges (despite strong homology). These cases are instead captured by the gene repertoire approach.

Assessing Quality of Read Signal by Mapping

We evaluated assembly and read quality by mapping HQ-nH reads to different gut-related datasets (per sample) using bwa-mem2 with msamtools (as described above). Reads aligning to viral contigs were referred to as the viral signal. Non-mapping reads were sequentially mapped to non-viral contigs (assembled but not classed as viral), proviral sequences, UHGV, and UHGG[35]. The proportions of reads mapping to each dataset per sample and on average are shown in Figure S3 and Tablse S2.

Quantitative and statistical analysis of vOTUs

To assess viral species abundance, we removed highly contaminated samples or those with insufficient viral signal (<10% viral signal and <1 million viral reads), removing 2 of 127 samples. We used bwa-mem2 for mapping and retained only reads with good-quality alignments using msamtools (filter -S -b -l 80 -p 95 -z 80). vOTUs were considered as present in a sample if the sequence was at least half covered with an average read depth of 1. To assess the diversity of samples, we computed the abundances of vOTUs. For this, we normalized the counts of reads mapping on a contig to the contig size, and rarefied (down-sized) these counts according to the minimal library size. For within-donor analyses (alpha diversity, relative abundance), rarefaction matched the smallest read count per donor. For beta-diversity, the smallest read count of all considered samples was chosen for the rarefaction. Alpha diversity was indexed by the Shannon and beta diversity by Bray-Curtis distance (with the vegan package in R (https://vegandevs.github.io/vegan/)). Relative abundance was defined as the fraction of total abundance of all present contigs in a sample.

To assess phage population dynamics, we calculated the mean coefficient of variation (CV) for each vOTU within each donor. The CV, a measure of relative variability, was determined by dividing the standard deviation by the mean of relative abundance values. vOTUs absent from all samples of a donor were excluded from the analysis. For vOTUs present in at least one sample per donor, zero values were assigned to absent time points. The mean CV was calculated only for vOTUs with at least three values.

Abundance profiling of metagenomic species

The abundances of bacterial species were retrieved from previous work[26]. Briefly, our catalogue of gut species is based on 2,741 unique metagenomic species. Microbial gene count table was created using METEOR (https://github.com/metagenopolis/meteor). Relative abundances of the species were calculated as the mean of their 100 signature genes, which show the highest correlation with each other

(per species). If fewer than 10% of these genes were detected, the abundance was set to 0. Relative abundance was computed by normalizing to the sum of all species to estimate the proportion. Bacterial richness in each sample was defined as the number of unique species. An overview of the microbial species is listed in Table S5 and their relative abundances in Table S6. To assess fluctuations of species per donor, we computed the CV as described for the vOTUs.

Goodness-of fit test for the number of dominant phages on distinct days

To test for significant differences in the number of highly abundant vOTUs on distinct days, we used a Poisson regression model[66], by setting day before treatment (day -1) as the baseline, and considering the number of samples per day. We used the glm() function of the stats package in R to model the Poisson distribution with 'family = poisson(link = "log")' and 'offset = (log (n_samples_per_day)'. We considered only days with more than three samples. P-values for all days were computed for different thresholds of dominant vOTUs starting from 0.1% to 40.0% (Figure S8BC).