Supplementary Discussion

Contamination Assessment

We evaluated the amount of human contamination in our viral DNA preparations by identifying sequences that matched human proteins using BLASTX. We found that human contamination was 0.07% of our sequencing dataset. These reads were omitted from downstream analysis.

The cesium chloride density gradient ultracentrifugation procedure used in this work is particularly designed for the purification of phage, however other viruses can be found at the density purified (1.5 g ml^{-1})[13]. We observed that 99.2% of viral sequences were of phage origin while the remaining 0.8% of viral sequences originated from eukaryotic viruses, confirming enrichment for phage populations during purification.

While cesium chloride purification is routinely used to remove contaminating bacterial cells and free DNA[32], it was important to evaluate the amount of chromosomal bacterial DNA which may have co-purified with our viral DNA preparations. Since the 16S ribosomal RNA gene is universal to bacteria but is not found in phage, we measured the amount of 16S rRNA in our viral samples to estimate the level of chromosomal bacterial contamination present in our sequencing dataset. To do so, we performed quantitative PCR for the 16S rRNA gene, using the universal primers 8F and 338R which span the variable region 2 (V2) and produce a 330 bp amplicon[9]. We compared the amount of 16S rRNA present in our viral samples to the amount of 16S rRNA present in equivalent concentrations of genomic DNA from *E. coli* MG1655 (purified in stationary phase) with a relative quantification analysis. This difference was calculated by $E^{(Ct\ bacterial\ DNA\ -\ Ct\ viral\ DNA)}$ (where E is efficiency, calculated using a standard curve), and we

assumed that this fold-difference was representative of the fraction of bacterial contamination. These analyses indicated that contamination was less than 0.1% in each viral DNA sample (Fig. S2). To account for this level of contamination, we assumed that contamination would be uniformly distributed and therefore we randomly discarded reads during statistical testing according to the amount of contamination detected in each viral sample. See Methods for further details.

Additional functions enriched in the phage metagenome under drug treatment

Our analysis revealed a subset of KEGG pathways that were commonly enriched in both ciprofloxacin and ampicillin treatment conditions (Fig. 2, orange-shaded nodes). Two-component systems (part of the broader signal transduction process) are important for environment sensing and response, including adaptation to antibiotic insult[33], and phage encoding of these systems may contribute to the development of resistance-conferring sensor-relay components[34]. Additionally, encapsulation of translation-related functions by phage belies their reliance on microbial hosts for protein synthesis. As increases in copies of translational machinery correlate to faster bacterial growth, and since the gut environment is a positive selection on this trait, particularly during colonization[35], it is interesting to consider that phage storage of these translational genes may be beneficial for the recovery of the gut microbiota following antibiotic treatment.

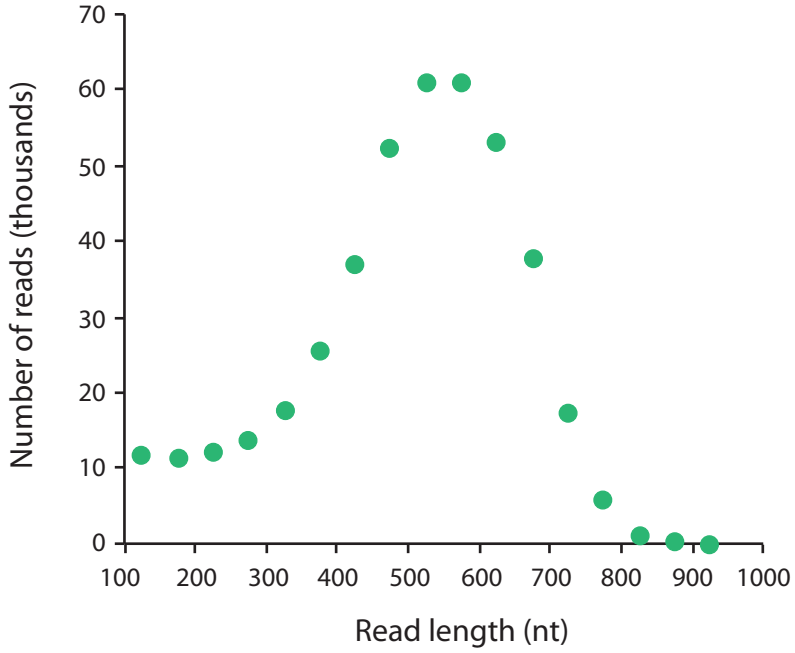
Supplementary References

32 Breitbart, M. *et al.* Genomic analysis of uncultured marine viral communities. *Proc Natl Acad Sci U S A* **99**, 14250-14255 (2002).

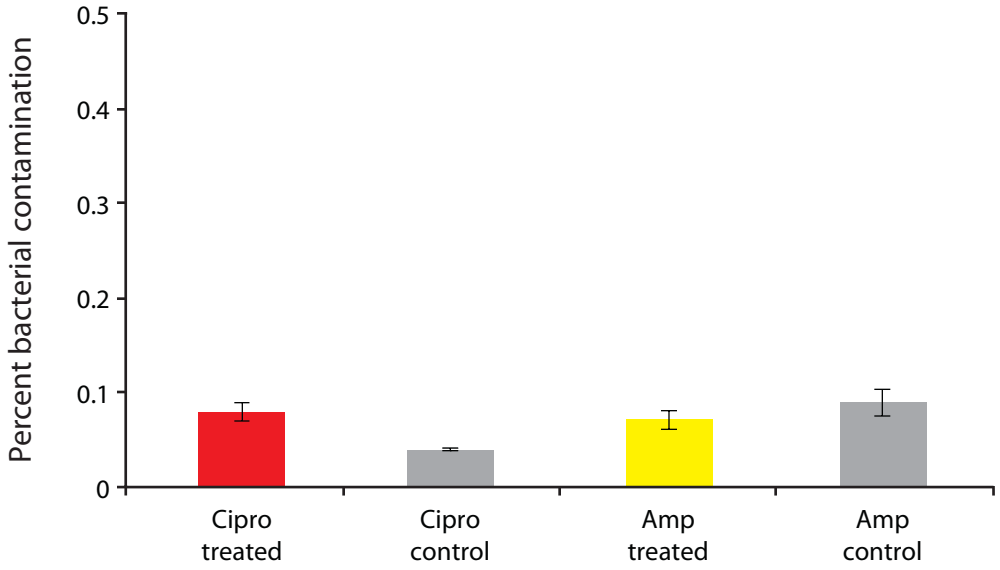
33 Miller, C. *et al.* SOS response induction by beta-lactams and bacterial defense against antibiotic lethality. *Science* **305**, 1629-1631 (2004).

34 Siryaporn, A., Perchuk, B. S., Laub, M. T. & Goulian, M. Evolving a robust signal transduction pathway from weak cross-talk. *Mol Syst Biol* **6**, 452 (2010).
35 Vieira-Silva, S. & Rocha, E. P. The systemic imprint of growth and its uses in ecological (meta)genomics. *PLoS Genet* **6**, e1000808 (2010).

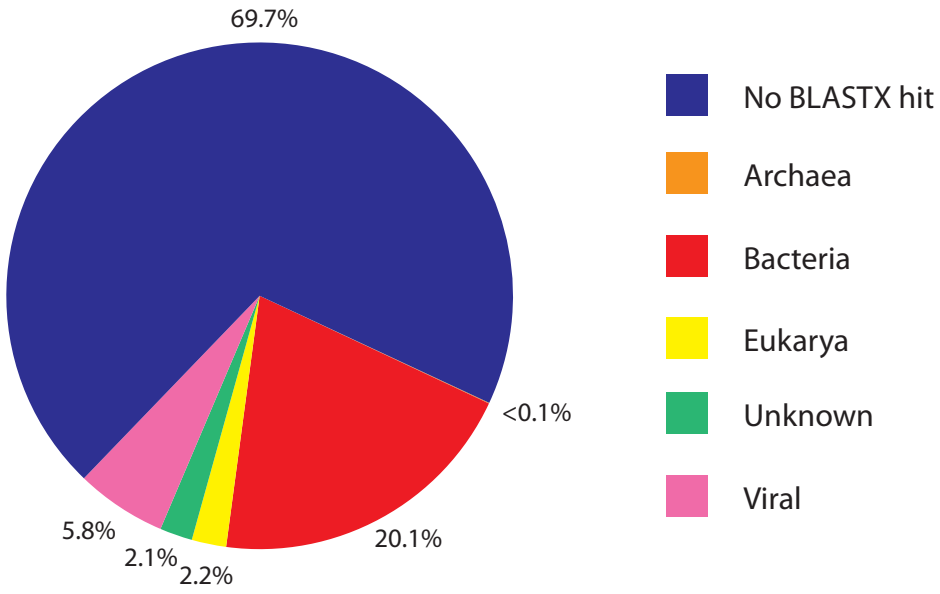
Supplementary Figures



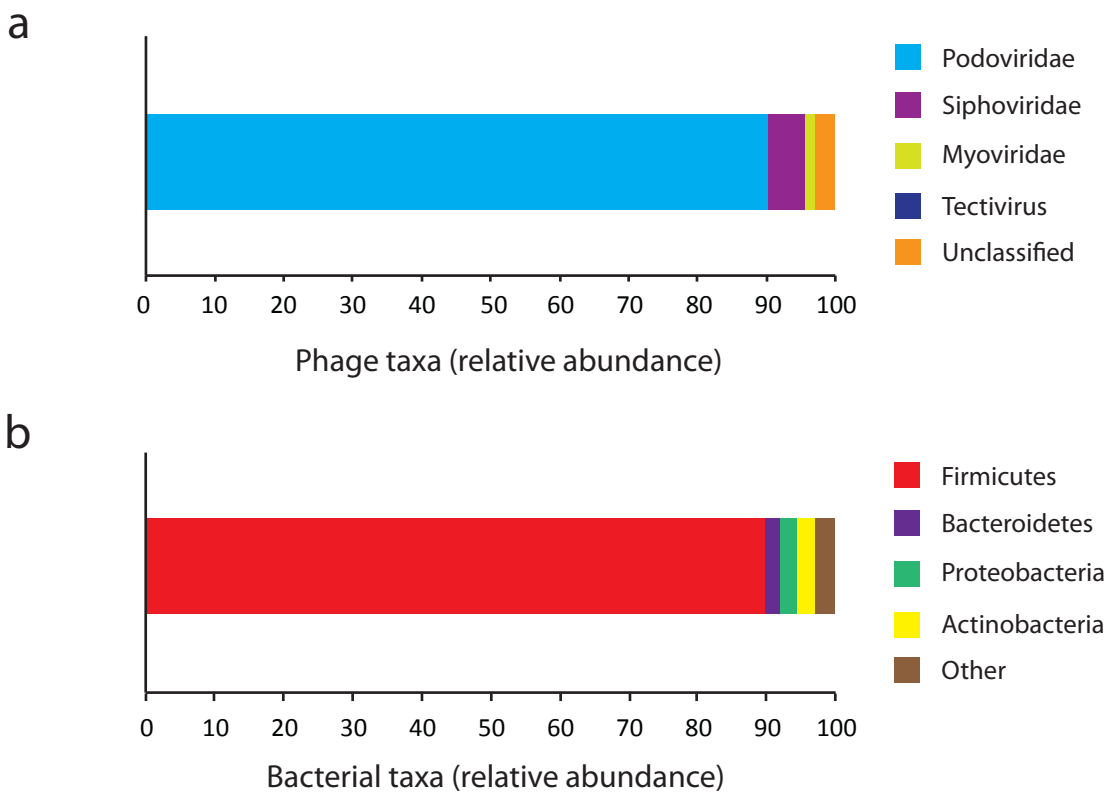
Supplementary Figure 1. Distribution of read lengths from 454 GS FLX+ pyrosequencing.



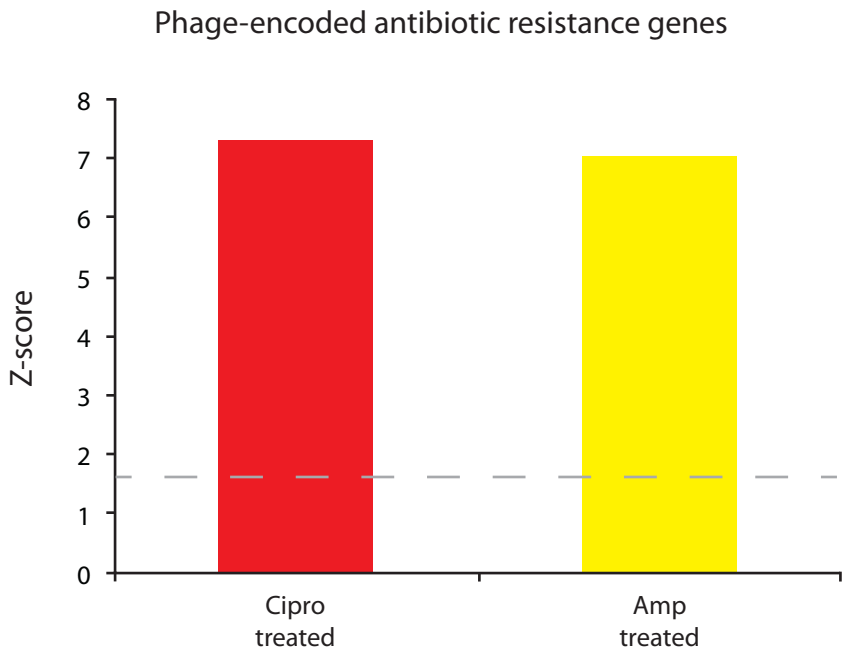
Supplementary Figure 2. Percent of bacterial contamination in each viral preparation as determined by quantitative PCR for the 16S rRNA gene ($n = 6$).



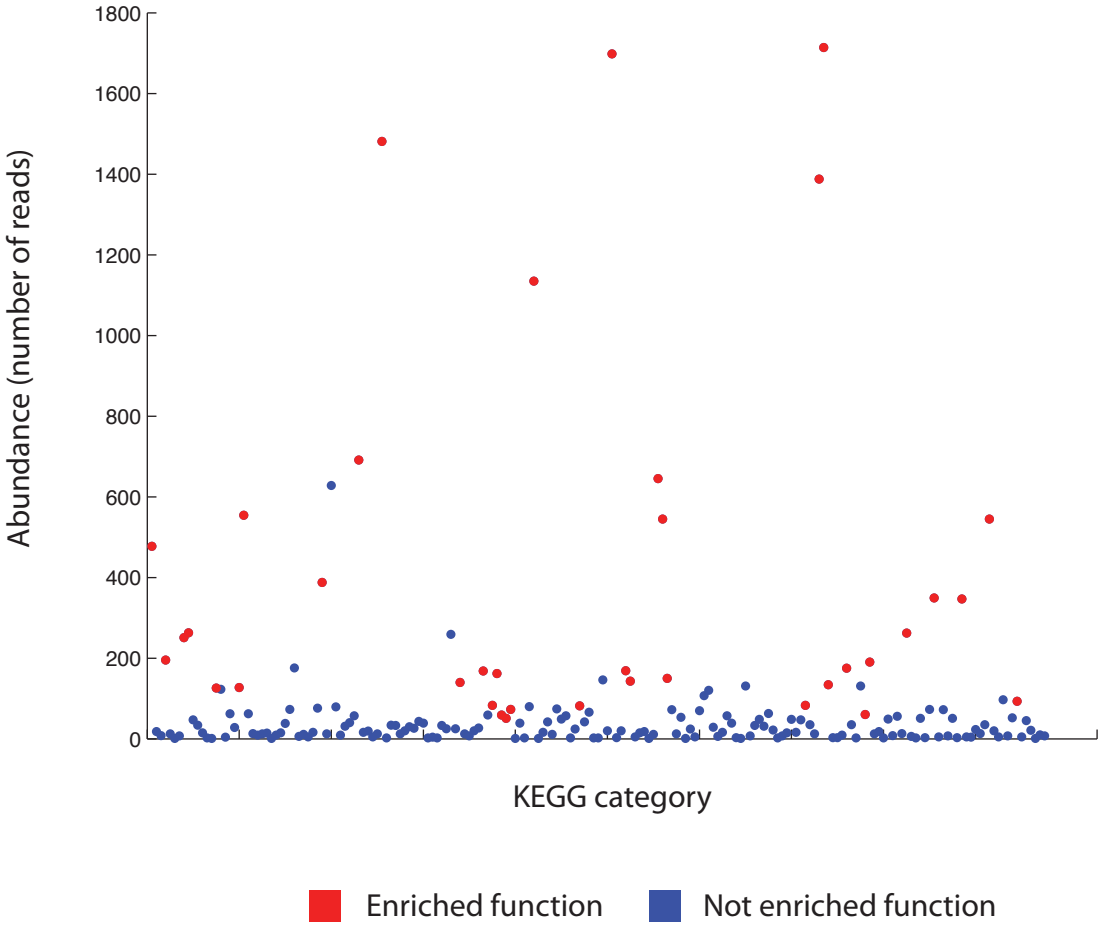
Supplementary Figure 3. Distribution of BLASTX hits among the major classes of biological entities.



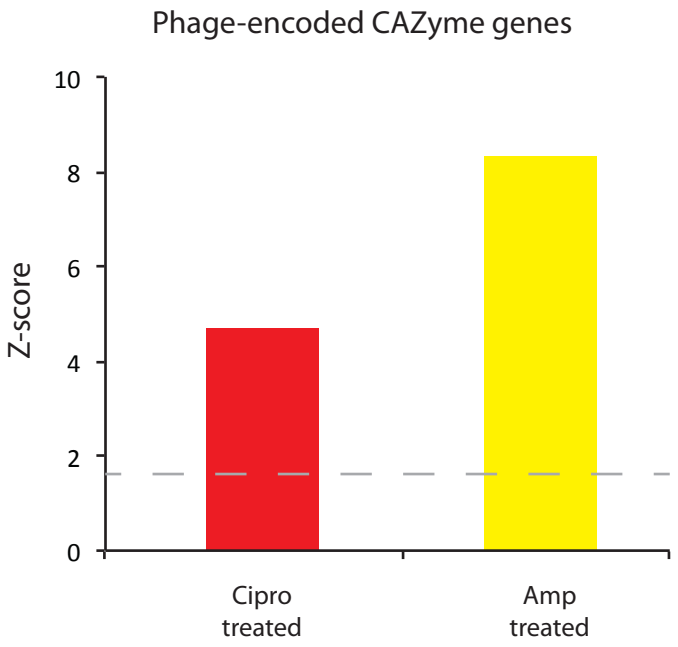
Supplementary Figure 4. a,b, Taxonomic distribution of phage (**a**) and bacteria (**b**) represented in phage metagenomes.



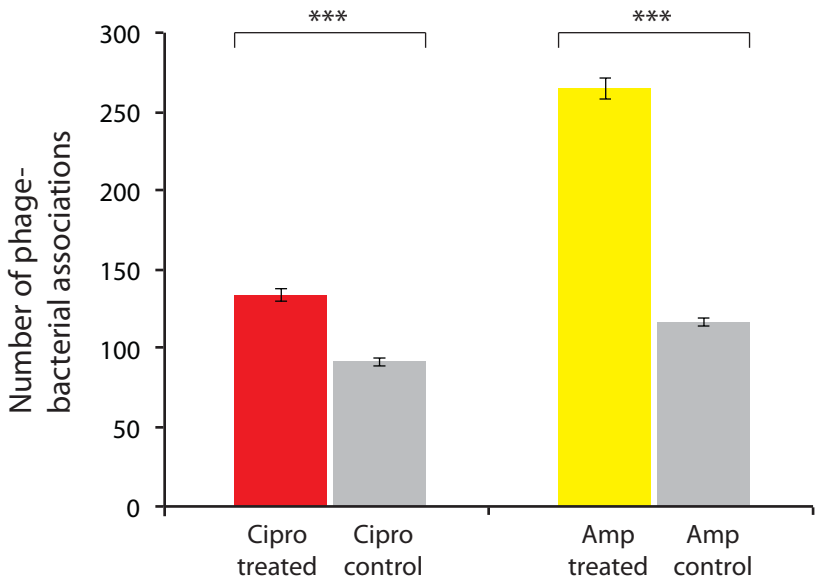
Supplementary Figure 5. Antibiotic resistance is enriched in fecal phage metagenomes following drug perturbation in mice. Z-scores are shown for sequences annotated as antibiotic resistance genes in phage from ciprofloxacin-treated mice (red) and phage from ampicillin-treated mice (yellow) in comparison to respective control mice. Dashed line corresponds to a Z-score of 1.65 ($P = 0.05$).



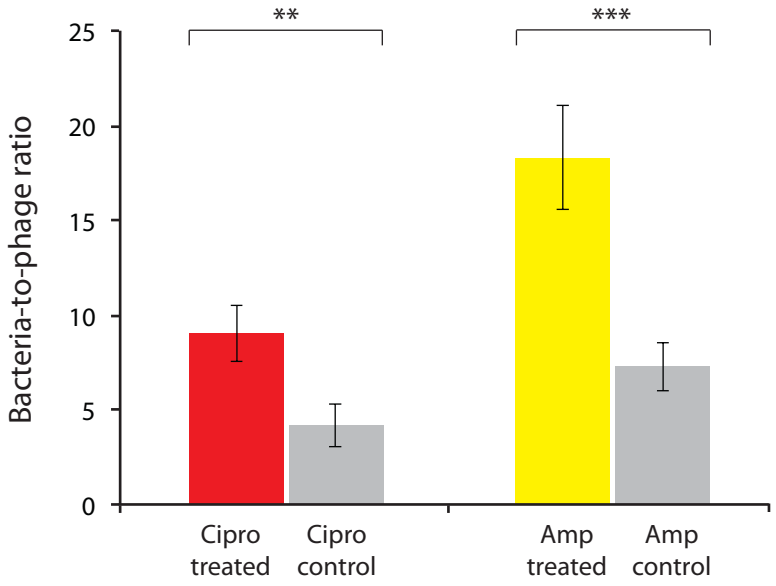
Supplementary Figure 6. Abundances (number of reads) associated with each KEGG category annotated in the fecal phage metagenome.



Supplementary Figure 7. CAZymes are enriched in fecal phage metagenomes following drug perturbation in mice. Z-scores are shown for sequences annotated as CAZyme genes in phage from ciprofloxacin-treated mice (red) and phage from ampicillin-treated mice (yellow) in comparison to respective control mice. Dashed line corresponds to a Z-score of 1.65 ($P = 0.05$).



Supplementary Figure 8. Number of phage-bacterial associations determined from the number of non-redundant associations found in each assembly trial (n = 50). P-value from Mann-Whitney U-test. Mean ± s.e.m. ***$P < 0.0005$.



Supplementary Figure 9. Bacteria-to-phage ratio from the number of bacterial species associated with a given phage for all phages using the union of associations from 50 assemblies shown in Fig. 4. *P*-value from Mann-Whitney *U*-test. Mean ± s.e.m. $^{**}P < 0.005$, $^{***}P < 0.0005$.

Supplementary Tables

Supplementary Table 1 (available as a separate Excel file). Reads annotated as encoding antibiotic resistance proteins in phage metagenomes. The first worksheet of this file summarizes all antibiotic resistance proteins identified and the corresponding number of reads in each phage metagenome. Subsequent worksheets provide detailed information on antibiotic resistance proteins found in each phage metagenome, including the corresponding drug class to which the protein confers resistance, the mechanism of resistance represented by the protein, protein similarity information determined by BLASTX using BLOSUM62, the strain of origin determined by BLASTX to the non-redundant NCBI databases, and protein family information by Clusters of Orthologous Groups (COG) annotation.

Supplementary Table 2. KEGG pathways enriched in fecal phage metagenomes following drug perturbation in mice. **a**, Pathways enriched in phage metageomes from ciprofloxacin-treated mice compared to those from control mice (24 out of 188 third-level KEGG pathways; Bonferroni-corrected $Z \geq 3.46$). **b**, Pathways enriched in phage metagenomes from ampicillin-treated mice compared to those from control mice (18 out of 178 third-level KEGG pathways; Bonferroni-corrected $Z \geq 3.43$).

a

KEGG pathway class 1	KEGG pathway class 2	KEGG pathway class 3	Z-score
Genetic Information Processing	Replication and Repair	Mismatch repair	12.29
Genetic Information Processing	Replication and Repair	DNA replication	12.21
Metabolism	Nucleotide Metabolism	Pyrimidine metabolism	11.71
Genetic Information Processing	Replication and Repair	Homologous recombination	11.17
Metabolism	Nucleotide Metabolism	Purine metabolism	10.78
Genetic Information Processing	Replication and Repair	Nucleotide excision repair	7.40
Environmental Information Processing	Membrane Transport	ABC transporters	7.39
Metabolism	Metabolism of Cofactors and Vitamins	Thiamine metabolism	6.17
Genetic Information Processing	Folding, Sorting and Degradation	Sulfur relay system	6.12
Genetic Information Processing	Replication and Repair	Base excision repair	6.07
Environmental Information Processing	Signal Transduction	Two-component system	6.05
Genetic Information Processing	Translation	Ribosome	4.88
Genetic Information Processing	Translation	Aminoacyl-tRNA biosynthesis	4.73
Metabolism	Metabolism of Cofactors and Vitamins	One carbon pool by folate	4.61
Genetic Information Processing	Translation	RNA transport	4.27
Metabolism	Metabolism of Cofactors and Vitamins	Nicotinate and nicotinamide metabolism	4.23
Metabolism	Amino Acid Metabolism	Cysteine and methionine metabolism	4.01
Metabolism	Energy Metabolism	Nitrogen metabolism	3.92
Environmental Information Processing	Membrane Transport	Bacterial secretion system	3.89
Genetic Information Processing	Folding, Sorting and Degradation	Protein export	3.87
Metabolism	Carbohydrate Metabolism	Glycolysis / Gluconeogenesis	3.63
Metabolism	Energy Metabolism	Carbon fixation pathways in prokaryotes	3.59
Metabolism	Amino Acid Metabolism	Lysine biosynthesis	3.59
Metabolism	Amino Acid Metabolism	Valine, leucine and isoleucine degradation	3.48

Supplementary Table 2 (continued).

b

KEGG pathway class 1	KEGG pathway class 2	KEGG pathway class 3	Z-score
Genetic Information Processing	Translation	Aminoacyl-tRNA biosynthesis	7.32
Environmental Information Processing	Membrane Transport	ABC transporters	6.18
Metabolism	Carbohydrate Metabolism	Starch and sucrose metabolism	5.93
Metabolism	Glycan Biosynthesis and Metabolism	Other glycan degradation	5.76
Metabolism	Carbohydrate Metabolism	Amino sugar and nucleotide sugar metabolism	5.61
Metabolism	Amino Acid Metabolism	Cysteine and methionine metabolism	5.44
Metabolism	Lipid Metabolism	Glycerolipid metabolism	5.34
Genetic Information Processing	Translation	Ribosome	5.20
Metabolism	Amino Acid Metabolism	Arginine and proline metabolism	4.35
Metabolism	Carbohydrate Metabolism	Pyruvate metabolism	4.31
Metabolism	Carbohydrate Metabolism	Galactose metabolism	4.22
Metabolism	Glycan Biosynthesis and Metabolism	Glycosphingolipid biosynthesis - globo series	4.11
Genetic Information Processing	Transcription	RNA polymerase	3.80
Metabolism	Glycan Biosynthesis and Metabolism	Glycosaminoglycan degradation	3.75
Metabolism	Amino Acid Metabolism	Alanine, aspartate and glutamate metabolism	3.68
Metabolism	Amino Acid Metabolism	Glycine, serine and threonine metabolism	3.53
Metabolism	Glycan Biosynthesis and Metabolism	Glycosphingolipid biosynthesis - ganglio series	3.49
Environmental Information Processing	Signal Transduction	Two-component system	3.43

Supplementary Table 3. CAZymes encoded in fecal phage metagenomes. **a**, Glycoside hydrolase (GH) families present in phage metagenomes. **b**, Glycosyltransferase (GT) families present in phage metagenomes. CAZyme families shaded in pink are unique to phage metagenomes from drug-treated mice compared to phage metagenomes from respective control mice.

a

Glycoside Hydrolase (GH) families

Cipro treated	Cipro control	Amp treated	Amp control
GH1	GH1	GH1	GH1
GH2	GH3	GH2	GH3
GH3	GH4	GH3	GH4
GH4	GH18	GH4	GH18
GH13	GH92	GH13	GH32
GH18		GH18	
GH29		GH28	
GH32		GH29	
GH53		GH32	
GH78		GH43	
GH84		GH53	
GH92		GH78	
		GH92	

b

Glycosyltransferase (GT) families

Cipro treated	Cipro control	Amp treated	Amp control
GT1	GT5	GT1	GT5
GT2	GT19	GT4	GT28
GT4	GT35	GT5	GT51
GT5	GT51	GT28	
GT6		GT35	
GT19		GT51	
GT28			
GT35			
GT51			