

Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome

Sheetal R. Modi¹, Henry H. Lee^{1†}, Catherine S. Spina^{1,2,3} & James J. Collins^{1,2,3}

The mammalian gut ecosystem has considerable influence on host physiology^{1–4}, but the mechanisms that sustain this complex environment in the face of different stresses remain obscure. Perturbations to the gut ecosystem, such as through antibiotic treatment or diet, are at present interpreted at the level of bacterial phylogeny^{5–7}. Less is known about the contributions of the abundant population of phages to this ecological network. Here we explore the phageome as a potential genetic reservoir for bacterial adaptation by sequencing murine faecal phage populations following antibiotic perturbation. We show that antibiotic treatment leads to the enrichment of phage-encoded genes that confer resistance via disparate mechanisms to the administered drug, as well as genes that confer resistance to antibiotics unrelated to the administered drug, and we demonstrate experimentally that phages from treated mice provide aerobically cultured naive microbiota with increased resistance. Systems-wide analyses uncovered post-treatment phage-encoded processes related to host colonization and growth adaptation, indicating that the phageome becomes broadly enriched for functionally beneficial genes under stress-related conditions. We also show that antibiotic treatment expands the interactions between phage and bacterial species, leading to a more highly connected phage–bacterial network for gene exchange. Our work implicates the phageome in the emergence of multidrug resistance, and indicates that the adaptive capacity of the phageome may represent a community-based mechanism for protecting the gut microflora, preserving its functional robustness during antibiotic stress.

Antibiotic treatment, an important and often necessary therapeutic intervention, can negatively affect the mammalian gut environment, potentially giving rise to immune² and metabolic deficiencies⁸. Studies on the disruption of intestinal homeostasis have focused on the resulting alterations in microbial composition^{6,7}. However, investigation of the gut ecosystem has uncovered a myriad of resident phages⁹, and it remains unclear how perturbation of the gut environment affects these symbionts. Phages can contribute genes that are advantageous to their microbial hosts^{10,11}, in turn promoting their own survival and propagation¹². This gene flow suggests that phages may have an important role in the adaptation of the microbiome to stressful environments. We used a comparative metagenomic approach to explore the effects of antibiotic perturbation on functions encoded in the phageome, as well as to examine how antibiotic treatment alters the phage–bacterial ecological network.

We treated groups of young adult mice ($n = 5$) orally with physiologically relevant concentrations of ciprofloxacin (a quinolone that inhibits DNA synthesis) or ampicillin (a β -lactam that inhibits cell-wall synthesis), each with a respective control. We obtained collective faecal samples from each group after 8 weeks of treatment and purified phages as previously described^{9,13}. DNA was extracted from phages and whole-genome amplified before performing shotgun 454 GS FLX+ pyrosequencing. We obtained a total of 440,792 quality reads, with a median read length of 477 nucleotides (210 megabases in total; Supplementary Fig. 1). Evaluation of contamination by quantification

of bacterial 16S ribosomal RNA genes indicated that contaminating bacterial sequences represented less than 0.1% of our data, which was subsequently accounted for in all statistical analyses (see Supplementary Discussion and Supplementary Fig. 2).

Phage DNA sequences were compared to the non-redundant National Center for Biotechnology Information (NCBI) protein and environmental protein databases (BLASTX; E value $< 10^{-5}$). Approximately 70% of reads were not assigned to previously sequenced genes (Supplementary Fig. 3), suggesting that the mouse phageome, like many other viral communities^{14,15}, harbours uncharacterized genetic material. We used the most significant BLAST alignment of a sequence, when available, to determine its phylogenetic origin. Most of the identifiable phages in our mouse phageomes (Supplementary Fig. 4a) were from the Caudovirales order, comprising the tailed phage families Siphoviridae, Podoviridae and Myoviridae, many of which are known to have a temperate life cycle. Because phage genomes incorporate bacterial genes, we also identified bacterial taxa; we found that 97% of phage-encoded bacterial genes were attributable to the four phyla known to dominate the gut (Firmicutes, Bacteroidetes, Proteobacteria and Actinobacteria; Supplementary Fig. 4b), consistent with the known hosts of the phages we detected.

We wondered whether antibiotic treatment leads to increases in phage-encoded genes for drug resistance. To investigate this, we compared DNA sequences in the phageome to an assembled database of antibiotic-resistance proteins (BLASTX; E value $< 10^{-3}$, see Methods). We found that reads annotated as antibiotic-resistance genes were highly enriched in phage metagenomes from mice treated with ciprofloxacin or ampicillin compared with those from control mice (Z score = 7.3 and $Z = 7.0$, respectively; Supplementary Fig. 5; read annotations enumerated in Supplementary Table 1). We catalogued the resistance reservoir by annotating phage-encoded genes based on the drug class to which they confer resistance (Fig. 1a). Our analysis revealed that resistance to the administered drug class was enriched in phage metagenomes from antibiotic-treated mice, such that resistance to DNA-synthesis inhibitors was enriched in ciprofloxacin treatment ($Z = 2.6$), and resistance to cell-wall-synthesis inhibitors was enriched in ampicillin treatment ($Z = 5.0$). Additionally, upon drug treatment, new resistance genes were found in the phageome. For example, phages from ciprofloxacin-treated mice carried genes encoding numerous quinolone efflux pumps (for example, *norM*, *mexD*, *mexF*), and phages from ampicillin-treated mice carried genes encoding sensor and response regulators of cell-wall-synthesis inhibitors (for example, *vanRS*).

Of note, resistance to other, orthogonal drug classes was also over-represented in the antibiotic-perturbed phageomes. Both treatments led to significant enrichment of resistance to antibiotics that target protein synthesis, and ciprofloxacin treatment also led to significant enrichment of resistance to cell-wall synthesis inhibitors (Fig. 1a). This cross-resistance was mediated by drug-specific inactivators (for example, chloramphenicol acetyltransferases), as well as multidrug-resistance exporters (for example, *mdtK*; Supplementary Table 1). Together, these

¹Howard Hughes Medical Institute, Department of Biomedical Engineering, and Center for BioDynamics, Boston University, Boston, Massachusetts 02215, USA. ²Boston University School of Medicine, 715 Albany Street, Boston, Massachusetts 02118, USA. ³Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02118, USA. [†]Present address: Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.

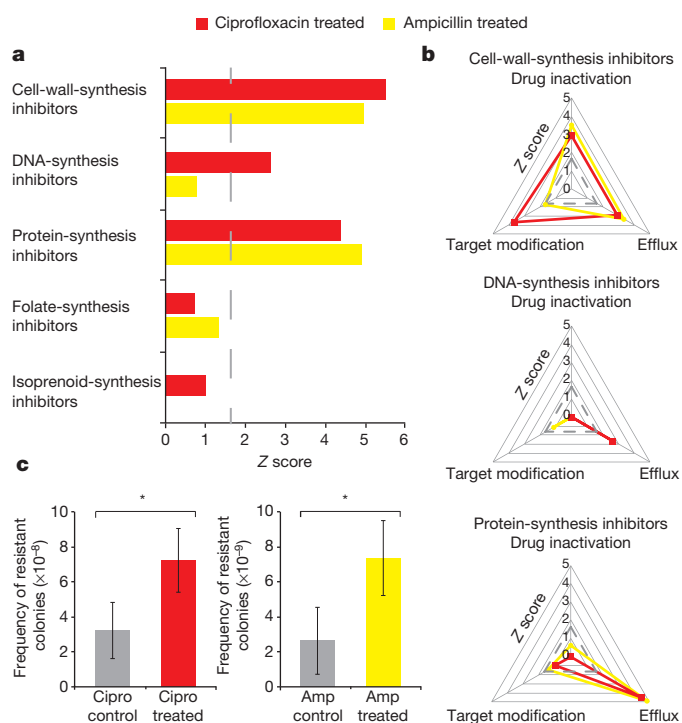


Figure 1 | Antibiotic resistance is enriched in phage metagenomes following drug perturbation in mice. **a**, **b**, Z scores are shown for sequencing reads annotated as antibiotic-resistance genes in phages from ciprofloxacin-treated (red) and ampicillin-treated (yellow) mice in comparison with respective control mice. Dashed lines correspond to a Z score of 1.65 ($P = 0.05$). Phage-encoded resistance genes were classified according to the drug class to which they confer resistance (**a**) and by their mechanism of resistance (**b**). **c**, Frequency of colonies resistant to ciprofloxacin (Cipro; $1 \mu\text{g ml}^{-1}$) upon infection of microbiota with phages from ciprofloxacin-treated mice or phages from control mice (left), and frequency of colonies resistant to ampicillin (Amp; $4 \mu\text{g ml}^{-1}$) upon infection of microbiota with phages from ampicillin-treated mice or phages from control mice (right). P values from Mann–Whitney U -test; $n > 12$. Data show mean \pm standard error of the mean (s.e.m.). * $P < 0.05$.

findings implicate the phage metagenome as a potential source of multidrug resistance during antibiotic treatment of the host.

We aimed to understand the specific mechanisms represented in phage-encoded genes conferring resistance to the most significantly enriched classes of antibiotics: inhibitors of cell-wall synthesis, DNA synthesis, and protein synthesis. We categorized resistance genes according to primary resistance mechanisms, which include modification or protection of the drug target (target modification), enzymatic inactivation of the drug (drug inactivation), and transport of the drug out of the cell (efflux)¹⁶. Using this framework to classify phage-encoded resistance genes, our analysis revealed that antibiotic treatment led to disparate resistance mechanism profiles for each drug class (Fig. 1b). Analysis of resistance to cell-wall-synthesis inhibitors showed that all types of resistance mechanisms were significantly enriched with both ciprofloxacin treatment and ampicillin treatment. By contrast, resistance to DNA-synthesis and protein-synthesis inhibitors occurred predominantly by efflux. Resistance to protein-synthesis inhibitors occurred by target modification and drug inactivation mechanisms at low levels and, in accordance with its rarity, resistance to DNA-synthesis inhibitors through drug inactivation was not detected. These data probably reflect resistance mechanisms that are both environmentally available in the gut ecosystem and impose sustainable *in vivo* fitness costs. As continued treatment with an antibiotic invariably leads to its own resistance, mechanisms that enable cross-resistance encoded by the phageome may be an important consideration when selecting subsequent therapeutics.

We next sought to demonstrate that phages from antibiotic-treated mice confer increased drug resistance to the host-associated bacterial

community. We assessed the frequency of resistant isolates from aerobically cultured naive microbiota that were infected *ex vivo* with phages from antibiotic-treated or control (untreated) mice. Our results show that this fraction of microbiota infected with phages from mice administered ciprofloxacin or ampicillin yielded two to three times more colonies resistant to the respective drug than the aerobically cultured fraction of microbiota infected with phages from control mice (Fig. 1c). These data indicate that phages from antibiotic-treated mice can contribute relevant functional advantages to their microbial hosts.

We next took a systems-level approach and classified other phage-encoded genes into functional pathways described by the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. We depicted enriched functional changes in phage metagenomes after antibiotic treatment as a network diagram (Fig. 2; abundances shown in Supplementary Fig. 6). Among the most significantly enriched pathways were functional properties related to the mode of action of the administered drug (Supplementary Table 2). Phageomes from ampicillin-treated mice were enriched for the amino sugar and nucleotide metabolism pathway (part of the broader carbohydrate metabolism process; $Z = 5.6$), indicating overrepresentation of genes related to synthesis of cell-wall constituents, and increases in these components have been shown to be requisite for drug resistance in clinical isolates¹⁷. Additionally, we found that phageomes from ciprofloxacin-treated mice were enriched for replication- and repair-related pathways, including base excision repair ($Z = 6.1$), nucleotide excision repair ($Z = 7.4$) and homologous recombination ($Z = 11.2$). Included in these pathways are members of the GO system for the repair of DNA oxidative lesions, which has been demonstrated to reduce cytotoxicity due to a range of antibiotic classes¹⁸. Also represented are members of the DNA-damage-inducible SOS system, known to provide protection against antibiotic-mediated cell death and induce the development of resistance-conferring mutations¹⁹. Furthermore, hyper-recombination has been shown to promote multi-drug resistance phenotypes²⁰. These results show that under drug treatment, the phageome encodes diverse mechanisms for modulating antibiotic susceptibility.

We also observed that phage metagenomes from antibiotic-treated mice were enriched for microbial functions that contribute to host metabolism (Fig. 2 and Supplementary Table 2). Phageomes from ciprofloxacin-treated mice were uniquely enriched for pathways relevant to the metabolism of cofactors and vitamins, including thiamine, an essential nutrient provided by the microbiome. Microbiota ferment polysaccharides indigestible by the host alone; metabolism of these sugars enables bacterial survival in and colonization of the gut environment²¹ and, as a beneficial consequence, provides energy to the host⁴. We found that polysaccharide-degradation genes, specifically related to metabolism of starch, cellulose, lactose and fructans (plant-derived fructose polymers), were enriched with antibiotic treatment (Fig. 3a). Genes coding for carbohydrate active enzymes (CAZymes), which enable bacteria to ferment a variety of dietary- and host-sourced glycans, were also enriched with antibiotic treatment and were represented by a range of glycoside hydrolase and glycosyltransferase families (Supplementary Fig. 7 and Supplementary Table 3). Because many gut microbes express only a specific array of carbohydrate-degrading enzymes²², bacteria that acquire these genes from the phage reservoir may gain additional foraging capacity and, consequently, a selective growth advantage. These results suggest that the phageome may be an adaptive repository for functions important for the host–commensal relationship that may otherwise be depleted by antibiotic perturbation. (See Supplementary Information for additional discussion.)

We next aimed to elucidate the phylogenetic basis of these phage-encoded bacterial functions. In Fig. 3b, we illustrate the taxonomic composition of all sequences of bacterial origin and sequences annotated with enriched functions following drug perturbation. Examining bacterial phylotypes that contribute antibiotic-resistance genes to the phage metagenome, we found a comparatively high representation of the Clostridia class and a low representation of the Bacilli class.

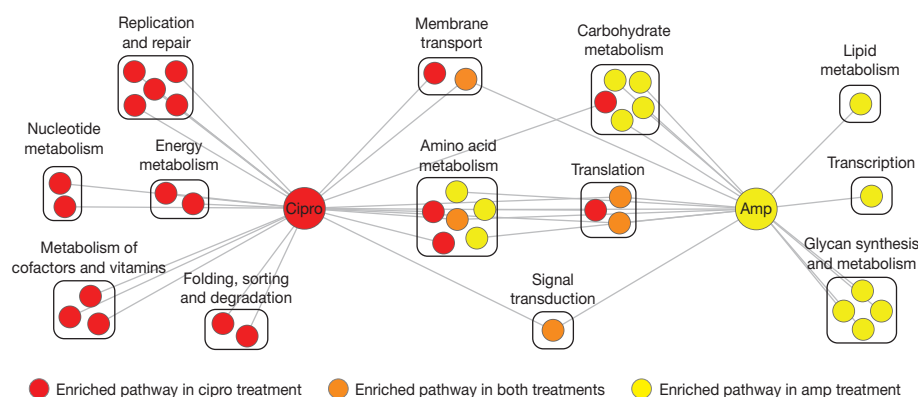


Figure 2 | Broad bacterial functions are enriched in phage metagenomes following drug perturbation in mice. Network depicts KEGG pathways significantly enriched under antibiotic treatment compared with controls. Treatments are represented by large nodes; enriched pathways are represented by small nodes, grouped by their higher-level processes and coloured by the

treatment condition (red, ciprofloxacin; yellow, ampicillin; orange, common to both treatments). In total, we identified 24 out of 188 pathways that were enriched with ciprofloxacin treatment ($Z \geq 3.46$, Bonferroni corrected), and 18 out of 178 pathways that were enriched with ampicillin treatment ($Z \geq 3.43$, Bonferroni corrected). Amp, ampicillin; cipro, ciprofloxacin.

Notably, a large fraction of CAZyme-annotated sequences originated from the Bacteroidetes class, which comprises members found to have diverse capabilities for carbohydrate metabolism²². Investigation of thiamine metabolism reveals that the Bacilli and Verrucomicrobia classes constitute a large proportion of these annotations. As the phageome reflects emergent properties of its environment¹⁵, phylogenetic analyses of phage-encoded elements may more broadly enable the identification of bacteria actively contributing to specific functions in the gut environment.

Our results show that the phageome harbours a diversity of potentially beneficial functional elements in the face of antibiotic perturbation. However, the extent to which this genetic reservoir is accessible to members of the microbiota remains unclear²³. To investigate this, we sought to elucidate the phage–bacterial ecological network and how

it changes under antibiotic treatment. We approximated the network of phage–microbe interactions with relationships identified through the reconstruction and analysis of individual viral genomes. *De novo* assembly was accomplished using stringent parameters (see Methods). Reconstructed viral genomes are composed of a mosaic of bacterial genes, and we used the phylogenetic origins of these sequences to determine putative phage–bacterial associations (Fig. 4). Our resulting network recapitulated known interactions, including the lysogenic relationships of foodborne pathogens, such as bacteriophage ϕ 3626 infection of *Clostridium perfringens*²⁴ and Siphoviridae Listeria phage A500 infection of *Listeria monocytogenes*²⁵. Importantly, antibiotic treatment leads to widespread restructuring of the phage–bacterial ecological network (Fig. 4). These data show that new links between phages and bacteria are formed with drug treatment, giving rise to significantly greater network connectivity (Supplementary Fig. 8). This increased connectivity is reflective of phages broadly, as more bacterial species are associated with a given phage (Supplementary Fig. 9). These results suggest that antibiotic treatment increases the frequency of phage integration and stimulates broad host range, which promotes a functional reservoir that is both genetically diverse and highly accessible to gut bacteria.

Although the phageome is a highly connected network for gene exchange, the functional consequence of acquiring genetic material from this reservoir depends on the molecular context of the host bacterium. Acquisition of a single gene may enhance an existing function by gene dosage or enable a novel phenotype. As some proteins rely on additional machinery, subsequent horizontal gene-transfer events may be required to produce a phenotype. Moreover, redundantly acquired genes may gain new functionality through paralogous evolution.

We demonstrate that antibiotic treatment enriches the phage metagenome for stress-specific and niche-specific functions, while mediating changes in the topology of the phage–bacterial ecological network to potentiate accessibility of these genetic elements. Functional resilience of the microbiome following environmental perturbation has been empirically documented and has engendered interest in the restorative forces that return the commensal flora to its pre-perturbed state^{6,7,26}. Our results implicate phage encapsulation of adaptive signatures as a community-based mechanism for functional robustness in the gut environment during stress. Of note, antibiotic treatment can also prime the gut environment for pathogen invasion², and our findings have potential implications for the emergence of drug resistance and evasion strategies in pathogenic populations. Cohabitation of phages and bacteria in the gut ecosystem is probably governed by complex and dynamic interactions, particularly during stress perturbation. Additional work is needed to discern the selective pressures imposed on each member of this community and the resulting mechanisms that influence the

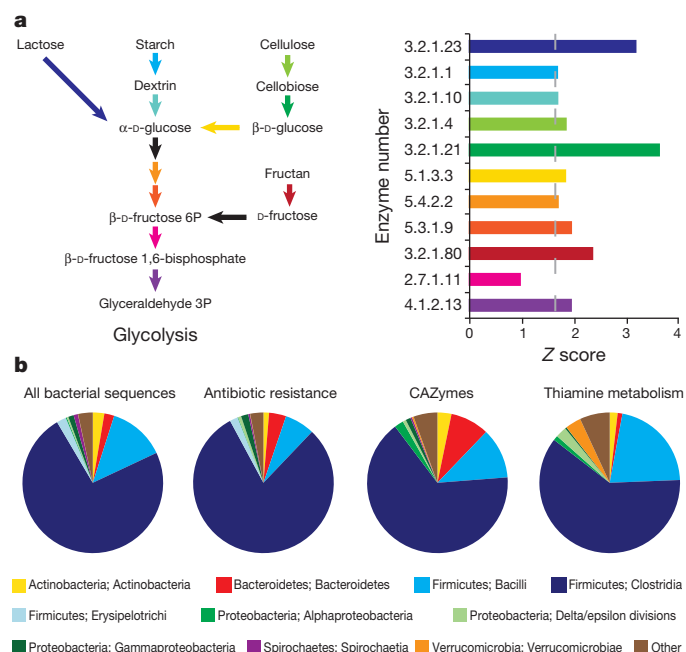


Figure 3 | Investigation of bacterial functions encoded in phages. **a**, Bacterial enzymes from sugar metabolism to glycolysis (left) with corresponding Z scores in phages from drug-treated mice in comparison with control mice (right). Dashed line corresponds to a Z score of 1.65 ($P = 0.05$). **b**, Class-level taxonomic distribution of all sequences of bacterial origin identified in phage sequencing data (far left) and sequences annotated with enriched functions following drug perturbation. ‘Other’ constitutes taxa that contributed less than 1% to all distributions.

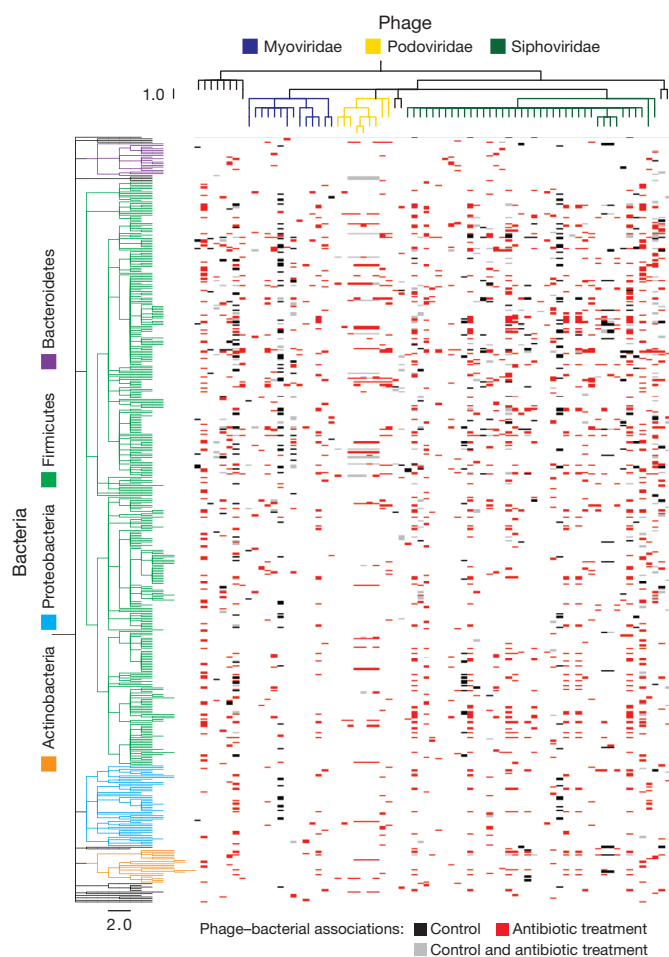


Figure 4 | The phage-bacterial ecological network. Dashes represent associations between virotypes and bacterial species identified from phylogenetic analysis of reconstructed phage genomes. Phage-bacterial associations only in control metagenomes (black), only in drug-treated metagenomes (red), and commonly identified in control and drug-treated metagenomes (grey). Data are the union of associations identified in 50 assemblies of randomly sampled reads from each treatment.

encoding and progressive enrichment of functionally beneficial genes in the phageome. Phage-mediated gene flow may be an important phenotypic buffer for bacterial communities, and further investigation of the adaptive reservoir of the phageome and the dynamic nature of the phage-bacterial ecological network may prove critical to understanding the influence of the gut ecosystem on host physiology.

METHODS SUMMARY

Groups of 6-week-old female FVB mice were treated with antibiotics in their drinking water to achieve doses of $28.5 \text{ mg kg}^{-1} \text{ day}^{-1}$ ampicillin or $12.5 \text{ mg kg}^{-1} \text{ day}^{-1}$ ciprofloxacin. Control groups were supplied with standard drinking water (ampicillin) or alkaline water (ciprofloxacin). After 8 weeks, we harvested collective faecal samples from each group. Viral purification from faecal samples was performed as previously described^{9,13}. DNA was extracted from each viral sample and whole-genome amplified in three separate reactions. Equimolar concentrations of multiplexed samples were pooled on a single plate for 454 GS FLX+ pyrosequencing. To analyse sequencing reads, an antibiotic resistance database was assembled using sequences from the Antibiotic Resistance Genes Database (ARDB)²⁷ and UniProt proteins were annotated with the Gene Ontology term “antibiotic breakdown”. Custom perl scripts were written to annotate sequences with KEGG (v.61.0). Enrichment between treatment and control was calculated by random sampling with replacement ($n = 10,000$), and Z scores were computed from the resulting normal distribution. Contigs were assembled using the Roche 454 GS *De novo* Assembler with default parameters, except for a minimum overlap of 100 bp and a minimum identity of 100%.

Full Methods and any associated references are available in the online version of the paper.

Received 3 July 2012; accepted 22 April 2013.

Published online 9 June 2013.

- Atarashi, K. *et al.* Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science* **331**, 337–341 (2011).
- Brandl, K. *et al.* Vancomycin-resistant enterococci exploit antibiotic-induced innate immune deficits. *Nature* **455**, 804–807 (2008).
- Smillie, C. S. *et al.* Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480**, 241–244 (2011).
- Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).
- Faith, J. J., McNulty, N. P., Rey, F. E. & Gordon, J. I. Predicting a human gut microbiota's response to diet in gnotobiotic mice. *Science* **333**, 101–104 (2011).
- Dethlefsen, L., Huse, S., Sogin, M. L. & Relman, D. A. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* **6**, e280 (2008).
- Dethlefsen, L. & Relman, D. A. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl Acad. Sci. USA* **108** (Suppl. 1), 4554–4561 (2011).
- Antunes, L. C. *et al.* Effect of antibiotic treatment on the intestinal metabolome. *Antimicrob. Agents Chemother.* **55**, 1494–1503 (2011).
- Reyes, A. *et al.* Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **466**, 334–338 (2010).
- Oliver, K. M., Degnan, P. H., Hunter, M. S. & Moran, N. A. Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* **325**, 992–994 (2009).
- Chen, J. & Novick, R. P. Phage-mediated intergeneric transfer of toxin genes. *Science* **323**, 139–141 (2009).
- Lindell, D., Jaffe, J. D., Johnson, Z. I., Church, G. M. & Chisholm, S. W. Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* **438**, 86–89 (2005).
- Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L. & Rohwer, F. Laboratory procedures to generate viral metagenomes. *Nature Protocols* **4**, 470–483 (2009).
- Breitbart, M. *et al.* Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* **185**, 6220–6223 (2003).
- Dinsdale, E. A. *et al.* Functional metagenomic profiling of nine biomes. *Nature* **452**, 629–632 (2008).
- Walsh, C. Molecular mechanisms that confer antibacterial drug resistance. *Nature* **406**, 775–781 (2000).
- Hanaki, H. *et al.* Activated cell-wall synthesis is associated with vancomycin resistance in methicillin-resistant *Staphylococcus aureus* clinical strains Mu3 and Mu50. *J. Antimicrob. Chemother.* **42**, 199–209 (1998).
- Foti, J. J., Devadoss, B., Winkler, J. A., Collins, J. J. & Walker, G. C. Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* **336**, 315–319 (2012).
- Kohanski, M. A., Dwyer, D. J. & Collins, J. J. How antibiotics kill bacteria: from targets to networks. *Nature Rev. Microbiol.* **8**, 423–435 (2010).
- Hanage, W. P., Fraser, C., Tang, J., Connor, T. R. & Corander, J. Hyper-recombination, diversity, and antibiotic resistance in pneumococcus. *Science* **324**, 1454–1457 (2009).
- Chang, D. E. *et al.* Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc. Natl Acad. Sci. USA* **101**, 7427–7432 (2004).
- Xu, J. & Gordon, J. I. Honor thy symbionts. *Proc. Natl Acad. Sci. USA* **100**, 10452–10459 (2003).
- Stern, A., Mick, E., Tirosh, I., Sagy, O. & Sorek, R. CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. *Genome Res.* **22**, 1984–1994 (2012).
- Zimmer, M., Scherer, S. & Loessner, M. J. Genomic analysis of *Clostridium perfringens* bacteriophage ϕ 3626, which integrates into *guaA* and possibly affects sporulation. *J. Bacteriol.* **184**, 4359–4368 (2002).
- Dorscht, J. *et al.* Comparative genome analysis of *Listeria* bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. *J. Bacteriol.* **191**, 7206–7215 (2009).
- Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).
- Liu, B. & Pop, M. ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Res.* **37**, D443–D447 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank K. Bodi and J. Schiemer at the Tufts Genomic Core for their sequencing assistance and A. Green and K. Pardee for ultracentrifugation help. This work was supported by the Howard Hughes Medical Institute and the National Institutes of Health Director's Pioneer Award Program.

Author Contributions All authors designed the study. C.S.S. oversaw the mouse work. S.R.M. and H.H.L. performed and analysed the experiments, with conceptual insight provided by J.J.C. S.R.M., H.H.L. and J.J.C. prepared the manuscript.

Author Information Virome data sets have been deposited in the NCBI Sequence Read Archive under accession SRP021521. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.J.C. (jjcollins@bu.edu).

METHODS

Mouse study. All experiments involving animals were pre-reviewed and approved by the Boston Children's Hospital Institutional Animal Care and Use Committee. Groups of separately housed 6-week-old female FVB/NJ mice ($n = 5$; Jackson Laboratory) were treated with ampicillin (142.5 mg l^{-1}) or ciprofloxacin (62.5 mg l^{-1}) in their drinking water. The corresponding dosage was $28.5 \text{ mg kg}^{-1} \text{ day}^{-1}$ ampicillin and $12.5 \text{ mg kg}^{-1} \text{ day}^{-1}$ ciprofloxacin, based on the average mouse weight of 20 g and an approximate intake of 4 ml per day. Control mouse groups were supplied with standard drinking water (ampicillin) or alkaline water (ciprofloxacin). Treatments were refreshed twice per week. Mice were housed in sterile conditions and received autoclaved chow during the course of this study. We harvested fresh collective faecal samples from each group to obtain ample material (3–4 g) for purification. Samples were stored at -80°C before use.

Viral purification and preparation of genomic DNA. Viral purification from faecal samples of mice after 8 weeks was performed as previously described^{9,13}. An aliquot of the viral preparation (1.5 g ml^{-1} layer from ultracentrifugation) was stained with SYBR gold and visualized with epifluorescence microscopy to verify the absence of bacterial contamination. Viral particles were concentrated and desalted using an Ultra-4 Centrifugal Filter Unit (Ultracel-30K MWCO; Millipore) to a volume of $\sim 200 \mu\text{l}$. Concentrated viral samples were treated with DNase (0.2 mg ml^{-1}) and samples were passed through a $0.22 \mu\text{m}$ filter to ensure no procedural contamination was introduced. Genomic DNA was extracted using the QiaAMP DNA mini kit (Qiagen) as per the protocol for viral DNA detailed in the manual. Genomic DNA was amplified using the Illustra Genomiphi v2 kit (GE) according to the manufacturer's instructions. For each sample, we pooled amplified DNA from three separate reactions to minimize bias.

Next-generation sequencing. Viral DNA was submitted to the Tufts Genomic Core for library preparation and shotgun sequencing. Equimolar concentrations of multiplexed (Rapid Library MID) samples were pooled on a single plate and pyrosequenced using the 454 GS FLX+ platform. Resulting sequences were filtered by removing duplicates using the tool available at <http://microbiomes.msu.edu/replicates/>²⁶ with the following parameters: sequence identity cut-off = 97%; length difference requirement = 0; number of beginning base pairs to check = 20. **Antibiotic-resistance annotations.** To facilitate annotation of antibiotic-resistance genes, we assembled a database consisting of the ARDB²⁷ and UniProt proteins annotated with the Gene Ontology (GO) term "antibiotic breakdown" (GO: 0017001) to achieve a total of 12,687 protein sequences. The functional annotations of the proteins in this database are supported by either experimental validation or high-quality computational prediction. Viral DNA sequences were compared with this database using BLASTX, and sequences with an E value $< 10^{-5}$ were deemed significant. This cut-off was selected to maintain a consistent stringency, by accounting for database size, with functional annotation to the NCBI databases.

Functional annotations. Phage DNA sequences were compared with the non-redundant NCBI protein (nr) and environmental protein (env_nr) databases (BLASTX; E value $< 10^{-5}$). Sequences were annotated with KEGG²⁹ (v.61.0) using custom Perl scripts interfaced with the KEGG API. The most significant BLAST hit that resulted in a KEGG annotation was used, and we included all annotations in the event that a hit had multiple annotations. KEGG orthologue annotations were compared with glycoside hydrolases and glycosyltransferases found on <http://www.cazy.org>³⁰ to identify CAZyme-encoding genes.

Statistical testing. To compare the functional annotations of two metagenomic data sets, A and B, we generated a distribution for each data set, A and B, reflecting the number of annotations from 10,000 trials of random sampling with replacement, sampled at the number of reads in the comparison data set. To account for the effects of contaminating bacterial DNA in our comparative metagenomic analyses, we assumed that contamination would be uniformly distributed and therefore we randomly discarded bacterial reads in each sampling trial according to the amount of contamination detected by quantitative PCR. We compared the number of annotations identified in a given sample to the comparison distribution and determined a Z score, calculated as $(x - \mu)/\sigma$, where, for example, x is the raw number of annotations in sample A, μ is the mean number of annotations in the distribution for sample B, and σ is the standard deviation of B's distribution. In

essence, this results in two Z scores, one comparing A to B and one comparing B to A, and consequently the minimum $|Z|$ was used. According to the central limit theorem, random sampling with replacement results in a normal distribution, so Z scores > 1.65 ($P < 0.05$) were considered enriched. In KEGG pathway analysis, Bonferroni was used to correct for multiple hypotheses, where the P -value cut-off $0.05/n$ (n was the number of total third-level pathways identified in our phage metagenomes) was converted to a Z score.

Phage infection of microbiota. Microbiota were isolated from faecal samples of naive mice as previously described³¹, except PBS plus 0.1% cysteine supernatants were plated on four separate Luria–Bertani (LB) agar plates. Colonies were grown aerobically for 24 h before plates were scraped with 2 ml LB, amassed, and cultured at 37°C and 300 r.p.m. for 2 h. This mixture was stored as $150 \mu\text{l}$ aliquots in 15% glycerol at -80°C . For each experiment, an entire aliquot was used as the inoculum to minimize growth biases. Phages were isolated as described above. Owing to limited sample availability, we harvested phages from mice that had been treated for 5 weeks for ciprofloxacin experiments and phages from mice that had been treated for 3 weeks for ampicillin experiments, along with phages from control mice, respectively. Phage preparations from drug-treated mice and control mice were diluted to the same volume, split equivalently, and incubated with 0.25 ml microbiota (cultured to exponential phase in LB plus 0.2% maltose) with 5 mM CaCl_2 and 10 mM MgSO_4 . Phage–microbiota mixtures were allowed to adsorb for one hour at 37°C (no shaking). Phage-infected microbiota were then pelleted, resuspended in fresh LB, and plated on LB agar plates with ciprofloxacin ($1 \mu\text{g ml}^{-1}$) or LB agar plates with ampicillin ($4 \mu\text{g ml}^{-1}$). A $10 \mu\text{l}$ aliquot was serially diluted and plated onto no-drug LB agar plates. Frequency was calculated as: number of colonies on drug plate divided by number of colonies on no-drug plate. Additionally, the basal frequency of resistant isolates from microbiota was measured as described in the previous sentence in the absence of phages, and this frequency was confirmed to be lower than that from microbiota infected with phages from either treated or untreated mice.

Quantification of 16S rRNA. Quantitative (q)PCR was used to measure levels of the 16S rRNA gene in our viral preparations. We used the universal primers 8F and 338R³, and qPCR was performed using the SYBR Green I Master kit and the LightCycler 480 (Roche) according to the manufacturer's instructions.

Contig assembly and identification of phage–bacterial associations. Contigs were assembled using the Roche 454 GS *De novo* Assembler with default parameters, except for a minimum overlap of 100 bp and a minimum identity of 100% to minimize erroneous alignments. Phage–bacterial associations were determined by computing the combination of phage phylogenetic annotations and bacterial phylogenetic annotations on a given contig, and non-redundant phage–bacterial associations were amassed for all contigs. To evenly compare the number of phage–bacterial associations across samples, we performed 50 assemblies using 60,000 randomly selected sequences with replacement from each sample. The data presented in Fig. 4 are representative of these 50 assemblies such that an association was illustrated if it was present in at least one assembly analysis. We computed the mean number of phage–bacterial associations for a given sample (Supplementary Fig. 8) from the number of non-redundant associations found in each assembly trial. Significance was determined by comparing sample values using the Mann–Whitney U -test. We computed the mean bacteria-to-phage ratio for a given sample (Supplementary Fig. 9) 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. Significance was determined by comparing sample values using the Mann–Whitney U -test.

28. Gomez-Alvarez, V., Teal, T. K. & Schmidt, T. M. Systematic artifacts in metagenomes from complex microbial communities. *ISME J.* **3**, 1314–1317 (2009).
29. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
30. Cantarel, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37**, D233–D238 (2009).
31. Goodman, A. L. *et al.* Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc. Natl Acad. Sci. USA* **108**, 6252–6257 (2011).