Bacteria deplete deoxynucleotides to defend against bacteriophage infection

Bacteria deplete deoxynucleotides to defend against bacteriophage infection

Nitzan Tal[1], Adi Millman[1], Avigail Stokar-Avihail[1], Taya Fedorenko[1], Azita Leavitt[1], Sarah Melamed[1], Erez Yirmiya[1], Carmel Avraham[1], Alexander Brandis[2], Tevie Mehlman[2], Gil Amitai[1], Rotem Sorek[1,#]

[1] Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 7610001, Israel

[2] Life Sciences Core Facilities, Weizmann Institute of Science, 7670001 Rehovot, Israel.

Corresponding author: rotem.sorek@weizmann.ac.il

Abstract

DNA viruses and retroviruses consume large quantities of deoxynucleotides (dNTPs) when replicating. The human antiviral factor SAMHD1 takes advantage of this vulnerability in the viral lifecycle, and inhibits viral replication by degrading dNTPs into their constituent deoxynucleosides and inorganic phosphate. Here, we report that bacteria use a similar strategy to defend against bacteriophage infection. We identify a family of defensive bacterial dCTP deaminase proteins that convert dCTP into deoxy-uracil nucleotides in response to phage infection. We also identify a family of phage resistance genes that encode dGTPase enzymes, which degrade dGTP into phosphate-free deoxy-guanosine (dG) and are distant homologs of human SAMHD1. Our results suggest that bacterial defensive proteins deplete specific deoxynucleotides (either dCTP or dGTP) from the nucleotide pool during phage infection, thus starving the phage of an essential DNA building block and halting its replication. Our study shows that manipulation of the deoxynucleotide pool is a potent antiviral strategy shared by both prokaryotes and eukaryotes.

Introduction

Bacteria use multiple immune mechanisms to defend against phage infection [1,2]. Restriction-modification (RM) and CRISPR-Cas systems have long been recognized as major lines of defence against phage [1]. Abortive infection systems, in which infected cells commit suicide and thus abort phage propagation, have also been reported in bacteria since the early days of phage research [3].

In recent years, it has become clear that bacteria encode a plethora of additional immune systems that escaped early detection [1,2]. These include defence systems that produce small molecules which block phage replication [4,5], systems that rely on secondary-messenger signalling molecules that activate immune effectors [6–9], and retron systems that employ reverse-transcribed non-coding RNAs as part of their anti-phage activity [10,11]. Many of these recently identified bacterial immune systems were shown to be functionally and structurally homologous to immune genes that protect eukaryotic cells from infection [5,7,9,12]. Despite this recent progress in identifying bacterial defence systems, it has been hypothesized that many such systems still await discovery [13].

Main text

Cytidine deaminases confer phage defence in *Escherichia coli*

We initiated this study by focusing on a family of homologous genes, found in a broad set of bacterial genomes, which contain a predicted cytidine deaminase domain (Fig. 1A, Extended Data Fig. 1A). This family of genes caught our attention due to their frequent localization next to known anti-phage defence systems in diverse bacterial genomes, a strong predictor for a role in anti-phage defence[14] (Extended Data Fig.1B; Supplementary Table S1). We cloned two genes from this family, one from *Escherichia coli* U09 and the other from *E. coli* AW1.7, into a lab strain of *E. coli* (MG1655) that does not naturally encode such genes. Infection assays with a panel of phages showed that both genes conferred substantial defense against a diverse set of phages (Fig. 1B; Extended Data Fig. 1C), and that the expression of dCTP deaminase from its native promoter did not impair bacterial growth (Extended Data Fig.1D). Since it conferred wider defence against phages, we further functionally characterized the gene from *E. coli* AW1.7 (Fig. 1B-C). Point mutations in the active site of the deaminase domain (C370A and C401A, predicted to disrupt zinc binding) abolished defense (Fig. 1C). In addition, point mutations in the nucleotide-binding motif of a predicted kinase domain found at the N-terminus of the protein also resulted in impaired defense (Fig. 1C), suggesting that both domains are essential for the defence against phage.

Cytosine deaminase enzymatic activities have been reported for immune proteins that protect human cells from viral infection by inducing deoxycytidine-to-deoxyuridine substitutions in the DNA of the viral genome, causing hyper-mutations that destroy the coding capacity of the virus[15]. To examine whether the bacterial cytosine deaminase has a similar function, we extracted and sequenced total DNA and RNA from deaminase-containing and control strains of *E. coli* MG1655 infected with phage T7. We did not

observe elevated rates of C-to-T mismatches (or the expected corresponding G-to-A in the complementary strand) as compared to other mismatches in phage DNA or RNA in cells encoding the cytidine deaminase, or as compared to control cells lacking the gene (Extended Data Fig.2). In addition, no hyper-editing of cytosines in the DNA or RNA molecules of the bacterial host was observed (Extended Data Fig.2). We concluded that the bacterial cytosine deaminase genes that protect against phages do so via a mechanism that does not involve hyper-editing of the nucleic acid polymers.

Deamination of dCTP depletes it from the deoxynucleotide pool

We next examined whether the bacterial defensive protein modifies single nucleotides rather than DNA or RNA polymers. To this end, we filtered extracts from cells infected by phage T7 and used liquid chromatography followed by mass spectrometry (LC-MS) to monitor the nucleotide content in these extracts (Fig. 1D). Remarkably, deoxycytidine triphosphate (dCTP), which naturally accumulates in control cells infected by phage T7, was completely absent in infected cells expressing the deaminase gene from *E. coli* AW1.7 (Fig. 1E). The depletion of dCTP was associated with substantial elevation of deoxyuridine monophosphate (dUMP), recorded as early as five minutes after initial infection, suggesting that the observed loss of dCTP was caused by its deamination into deoxyuridine compounds (Fig. 1F). Similar depletion was observed in deoxycytidine monophosphate (dCMP) and deoxycytidine diphosphate (dCDP), but not in the ribonucleotides CTP, CDP, and CMP, suggesting that deamination is limited to deoxycytidines (Extended Data Fig.3A-E). Although the deamination of dCTP is expected to generate dUTP molecules, it is known that *E. coli* expresses housekeeping dUTPase enzymes that rapidly convert dUTP to dUMP molecules[16], likely explaining the observed accumulation of dUMP and not of dUTP (Fig. 1F). Indeed, an accompanying paper by Waters and colleagues demonstrated the *in vitro* conversion of dCTP to dUTP by a dCTP deaminase enzyme of the same family[17]. Depletion of dCTP was not observed in cells expressing dCTP deaminase mutants (Extended Data Fig.4).

These results suggest that during infection by phage T7, the defensive deaminase protein converts deoxycytidine nucleotides into deoxyuracils, depleting the cell of the dCTP building blocks that are essential for phage DNA replication. Indeed, monitoring phage DNA during infection showed impaired phage DNA replication in deaminase-containing cells (Fig. 1G). Moreover, we observed that the other three deoxynucleotides, dATP, dGTP and dTTP, all accumulated in the cell at 10 and 15 minutes post infection in cells expressing the dCTP deaminase (Extended Data Fig. 3F-H). These results suggest that the depletion of dCTP limits phage DNA replication, resulting in a buildup of the other deoxynucleotide building blocks that would have otherwise been incorporated into the replicating polynucleotide chain of the phage.

dGTPases deplete dGTP during phage infection

Our discovery of a defence mechanism that uses nucleotide depletion via dCTP deamination prompted us to investigate whether there are additional phage resistance mechanisms that defend the cell by depleting deoxynucleotides. Depletion of the nucleotide pool was reported as an antiviral strategy in the human innate immune system,

where it is manifested by the SAMHD1 antiviral protein that restricts HIV infection in non-replicating cells[18]. SAMHD1 was found to remove the triphosphate from dNTPs, breaking the nucleotides into phosphate-free deoxynucleosides and an inorganic triphosphate (PPPi)[18]. The massive degradation of dNTPs by SAMHD1 depletes the nucleotide pool of the cells and inhibits replication of the viral genome[19].

Several dNTP triphosphohydrolases (dNTPases), and specifically dGTPases, have previously been described in bacteria[20–22]. Bacterial dGTPases cleave dGTP *in vitro*, and their structures show substantial homology to the active site architecture of SAMHD1[23,24]. Although dGTPase homologs are abundant in bacteria, their physiological roles mainly remain unknown[23]. We hypothesized that bacterial dGTPases might have a role in defence against phage through dGTP depletion.

We identified multiple genes with predicted dGTPase domains that were localized within operons of type I restriction-modification systems, or near other known defence systems, suggesting a possible role in anti-phage defence[14] (Fig. 2A, Extended Data Fig. 5A). These genes could not be aligned to human SAMHD1, but structural homology modeling using AlphaFold2[25] revealed significant similarity to the structure of SAMHD1 (Extended Data Fig.5B). To evaluate whether these genes confer phage resistance, we cloned eight of these dGTPases into *E. coli* MG1655. Following infection with a diverse set of phages, we found that six of the cloned genes substantially protected the *E. coli* host from phage infection, with the most prominent defence observed for the dGTPase from *Shewanella putrefaciens* CN-32 (*Sp*-dGTPase) (Fig. 2A-B; Extended Data Fig.6,7; Table S5). A D101A point mutation in the histidine/aspartate (HD) motif in the predicted active site of *Sp*-dGTPase abolished defense, suggesting that the dGTPase functionality is essential for anti-phage activity (Fig. 2C). Overexpression of dGTPases did not impair bacterial growth (Extended Data Fig.6B), and two dGTPases protected against phages when cloned under the control of their native promoters, suggesting that defence is not an artifact of overexpression (Fig. 2B, Extended Data Fig. 7A-C).

We next used LC-MS to analyze the nucleotide content in cells expressing *Sp*-dGTPase. In control cells that do not express the defensive gene, the concentration of dGTP was substantially elevated after 10 minutes from the onset of infection by phage T7, as expected for this phage[26] (Fig. 2D). However, in cells expressing the defensive gene, dGTP was present in lower amounts throughout the infection time course (Fig. 2D). In parallel to the depletion of dGTP, we observed substantial elevation in the concentrations of deoxyguanosine (dG), consistent with the hypothesis that the defensive protein removes the triphosphate from dGTP (Fig. 2E). Similar reduction in dGTP levels was observed for a dGTPase from *E. coli* STEC 2595, which was cloned under the control of its native promoter (Extended Data Fig.7). In parallel to the reduction of dGTP, the other three deoxynucleotides accumulated in dGTP-expressing cells, and phage DNA replication was impaired (Fig. 2F, Extended Data Fig. 6E-G, 7H-K). These results reveal a family of bacterial antiviral enzymes that deplete dGTP in phage-infected cells and limit phage replication.

Numerous species encode antiviral nucleotide-depletion genes

We found 977 homologs of the defensive dCTP deaminase in 952 genomes, representing 2.5% of the set of 38,167 genomes that we analyzed (Supplementary Table S1). The dCTP deaminases were found in genomes of a diverse set of species from the Proteobacteria phylum, as well as other phyla including Firmicutes, Actinobacteria and Bacteroidetes (Fig. 3A; Supplementary Table S1). In most cases, only a small fraction of the sequenced genomes of a given species harbored the dCTP deaminase. For example, the gene was found in 5%, 10%, and 12% of analyzed *E. coli*, *Legionella pneumophila*, and *Burkholderia pseudomallei* genomes, respectively (Fig. 3B). Such a sparse pattern of gene presence/absence is typical of bacterial defense systems and is indicative of extensive horizontal gene transfer between genomes[2]. An exception for this pattern was observed in *Vibrio cholerae*, where we found the defensive dCTP deaminase in most of the analyzed genomes (200 of 291, 69%) (Fig. 3B).

Homologs of the defensive *Sp*-dGTPase were abundant in the set of bacterial genomes that we analyzed, appearing in >2,300 of the genomes (at least 25% sequence identity to *Sp*-dGTPase over an alignment overlap of >= 90%) (Supplementary Table S2). Among these, proteins that we originally tested, which showed high similarity to *Sp*-dGTPase, were preferentially localized next to known defense systems, implying that the main role of these direct homologs is to defend against phages (Fig. 4, Supplementary Table S2). However, more distant homologs did not show high propensity to co-localize with defense systems, with only 10% of these homologs found next to known defensive genes (Fig. 4B, Supplementary Table S2). Experimentally examining 20 such distant homologs from across the phylogenetic tree showed that many of them (9 of 20) provided defense against phage when cloned into *E. coli* MG1655 (Fig. 4B, Extended Data Fig.8). These findings suggest that these more distant homologs function in phage defense, but may also have housekeeping roles in bacterial physiology and are thus not encoded in defense islands. In contrast to the sparse distribution observed for dCTP deaminases, dGTPases were present in most strains of encoding species, further supporting a housekeeping role for these genes (Fig. 3C-D). In this context it is noteworthy that SAMHD1 was found to be involved in housekeeping DNA repair, a role that has also been suggested for bacterial dGTPases[23,27].

Phages can evolve resistance to nucleotide depletion

To gain more insight into how phage infection triggers nucleotide-depletion, we attempted to isolate phage mutants that escape this mode of defence. We were able to obtain four mutants of phage T7 that could partially overcome the defence conferred by the dCTP deaminase gene, as well as seven T7 mutants that overcame dGTPase defense (Supplementary Table S4). In all four mutants that escaped the dCTP deaminase, we found point mutations in gene 5.7 of phage T7, leading to amino-acid substitutions (Fig. 5A). Surprisingly, three of the T7 mutants that escaped *Sp*-dGTPase defence were also mutated in this gene, with mutations either causing frameshift in Gp5.7 or amino-acid substitutions (Fig. 5A-B). We verified that Gp5.7 is not produced during infection by a mutant phage with a frameshift in gene 5.7 using protein mass spectrometry (Extended Data Fig.9A-B). In an additional three mutants that overcame *Sp*-dGTPase, a frameshift-causing mutation was identified in gene 5.5, which is directly upstream to gene 5.7 and is thought to be co-translated with gene 5.7, forming a 5.5-5.7 fusion protein[28]. A single T7 escapee was not

mutated in either of these genes, but had a point mutation in the endonuclease I gene, encoding a protein responsible for Holliday junction resolution[29] (Supplementary Table S4). These results suggest that mutations in Gp5.7 allow phage T7 to escape both the dCTP deaminase and the dGTPase defense. Indeed, escape mutants isolated on the dCTP deaminase were able to overcome *Sp*-dGTPase defense, and vice versa (Fig. 5C-D).

We infected dCTP deaminase-expressing cells with the Gp5.7-mutated T7 phages, and measured dCTP concentrations. A reduction in dCTP was observed post infection, but this reduction caused only partial, yet not full, depletion of dCTP in the infected cells (Fig. 5E). These results suggest that the nucleotide-depleting enzyme is only partially active against the mutated phages and that the remaining dCTP in the cell still allows phage DNA replication, enabling their escape from the nucleotide-depletion defense.

Gp5.7 of phage T7 is responsible for shutting down σ^S-dependent host RNA polymerase (RNAP) transcription, which would have otherwise interfered with phage propagation[30]. A second T7 protein, Gp0.7, is also known to modify the host RNAP at the early stages of infection[31]. Phages deleted in genes 0.7 or 5.7 were previously reported to be viable but propagate sub-optimally on *E. coli* cells[32,33]. Intriguingly, two of the T7 mutants that escaped the *Sp*-dGTPase defense contained, in addition to the mutation in gene 5.7, also mutations in gene 0.7 (Supplementary Table S4).

Our observation that mutations in phage RNAP-modifying proteins allow T7 to escape nucleotide depletion supports a hypothesis that both the dCTP deaminase and dGTPase may be sensitive to transcription inhibition imposed during T7 phage infection. To test this hypothesis, we applied rifampicin, an antibiotic that inhibits bacterial DNA-dependent RNA polymerase, on cells expressing the dCTP deaminase or *Sp*-dGTPase. In both cases, the respective nucleotide became depleted while the other DNA nucleotides did not (Fig. 5F-G, Extended Data Fig. 9C-H). Together, these results suggest that phage-mediated inhibition of host transcription may be involved in triggering the activation of bacterial dNTP-depletion defense (Fig. 5H).

Discussion

In this study we describe a previously unknown mechanism of phage resistance in which phage infection triggers the activity of defensive nucleotide-manipulating enzymes that deplete one of the deoxynucleotides from the infected cell. As phage DNA replication requires a huge supply of all of the dNTPs, reduction in or elimination of one of the deoxynucleotide pool members impairs the ability of the phage to replicate its genome. We hypothesize that nucleotide depletion might be activated when host bacteria transcription is inhibited. Many phage manipulate host transcription[34], so a mechanism of defence that relies on recognition of infection by modulation of transcription, could explain the breadth of defence against multiple phage families that is conferred by the two defence genes we studied.

It is intriguing that neither of the nucleotide depletion proteins that we studied were toxic when expressed in the host in the absence of phage infection. This suggests that while the proteins are present in the cell prior to phage infection, they may only be activated when a

phage is recognized. Indeed this is the case for many anti-phage defence systems, such as CBASS[7], Pycsar[35], toxin-antitoxin systems and other abortive infection systems[3], which respond to phage infection by applying lethal measures.

The dCTP deaminase analysed in our work includes a predicted N-terminal kinase domain. Point mutation in this domain abolished defence, suggesting that it is important for the defensive activity, but its specific role remains unclear. We hypothesize that this domain may participate in sensing phage infection, but its exact function remains to be defined. Housekeeping dCTP deaminase (dcd) proteins of bacteria play a role in dTTP synthesis needed for bacterial replication. The dCTP deaminase family studied here shares very little protein sequence similarity with dcd, but the two proteins share structural homology in the dCTP deaminase domain. The N-terminal kinase domain is not present in dcd proteins.

Enzymes with cytidine deaminase activity are known to be involved in multiple aspects of immunity[15]. The human APOBEC3G protects against viruses by deaminating cytosines in the viral genome, thus destroying its coding capacity[36,37]. In addition, cytosine deamination performed by activation-induced cytidine deaminase (AID) enzymes introduces somatic hypermutations which are essential for antibody diversification and maturation[38]. We now report a third role for cytidine deaminases in immunity, where the deamination is performed on the mononucleotide building blocks in order to eliminate them during viral infection. Antiviral defence by deamination of adenines was also recently demonstrated to function in defence systems that protect bacteria from phage[13].

The model *E. coli* strain MG1655 encodes a dGTPase gene called *dgt*, which shows distant homology (24% identity) to the defensive *Sp*-dGTPase studied here[39]. The function of *dgt* was linked to maintenance of genome integrity[40], although some evidence hints at a role in phage defence, as T7 was shown to encode a specific inhibitor of *dgt*[41]. It is therefore possible that families of dGTPase enzymes other than the one that we studied here are also involved in phage defense.

Multiple proteins in T7 modify the host RNA polymerase (RNAP) to allow optimal infection. These include Gp0.7, Gp2, and Gp5.7. Gp0.7 is a kinase that phosphorylates σ^{70}-bound host RNAP and alters its functionality to suit phage needs[30]. At later stages of infection, Gp2 shuts off the ability of σ^{70}-RNAP to transcribe the phage genome[30]. Finally, Gp5.7 binds and inhibits transcription by σ^{S}-RNAP, which appears in the host when the cell activates the stringent response during phage infection[30]. It is intriguing that mutations in two of these RNAP-modifying proteins allow the phage to escape both mechanisms of nucleotide depletion that are characterized here (notably, Gp2 is indispensable for T7 growth and hence cannot be mutated). This implies that both the dCTP deaminase and dGTPase might somehow monitor the integrity of the host RNAP or its activity and become activated when it is tampered with. Previously identified anti-phage toxin-antitoxin systems like ToxIN[42] and the hok/sok[43] killer locus were suggested to be activated by transcriptional inhibition as well. In both cases the regulation over the toxin was achieved through an RNA molecule that, upon inhibition of transcription, declined in concentration, allowing the activity of the toxin. Indeed, an accompanying paper by Waters and colleagues

shows that a *Vibrio*-encoded dCTP deaminase is regulated by a cis-acting non-coding RNA[17].

Our data show that nucleotide-depleting enzymes protect against various types of phage. The well-characterized phage examined in this study, including T4, T5 and T7, are known to degrade host DNA early during infection. Such DNA degradation would ultimately result in cell death, even if phage replication was prevented by the defensive enzymes. It remains to be seen whether a nucleotide-depletion mechanism could be reversible in bacteria, allowing them to survive post phage infection, in cases the infecting phage does not degrade host DNA.

It is plausible that additional nucleotide-depletion mechanisms exist as part of the microbial antiviral arsenal. Enzymes that degrade all four dNTPs are known to exist in multiple bacteria including *Thermus thermophilus*, *Enterococcus faecalis*, and *Pseudomonas aeruginosa*[20,22,44]. These enzymes, which have *in vitro* activities similar to SAMHD1, may also play a role in anti-phage defense. Moreover, nucleotide-depleting antiviral proteins may have enzymatic reactions other than nucleotide deamination or triphosphohydrolysis. We envision that nucleotide methylases, ribosyltransferases, or other nucleotide modifying enzymes will be discovered in the future as having antiviral activities via manipulation of the nucleotide pool.

In the past few years, multiple components of the human innate immune system have been shown to have evolutionary roots in bacterial defence against phage. These include the cGAS-STING pathway[45], the viperin antiviral protein[5], the argonaute protein of the RNAi machinery[46], TIR domains[9], and gasdermin proteins[12]. The defensive dGTPases characterized in our study show distant structural homology to the SAMHD1 active site, but this homology is too limited to explain whether SAMHD1 is evolutionarily derived from a bacterial dNTPase. Regardless of the evolutionary trajectory, we find it remarkable that depletion of the nucleotide pool is a successful antiviral defense strategy shared by eukaryotes and prokaryotes alike.

Online Methods

Detection of nucleotide-depleting systems in defence islands

Protein sequences of all genes in 38,167 bacterial and archaeal genomes were downloaded from the Integrated Microbial Genomes (IMG) database[47] in October 2017. These proteins were filtered for redundancy using the 'clusthash' option of MMseqs2 (release 2-1c7a89)[48] using the '-min-seq-id 0.9' parameter and then clustered using the 'cluster' option, with default parameters. Each cluster with >10 genes was annotated with the most common pfam, COG and product annotations in the cluster. Defense scores were calculated as previously described[14], recording the fraction of genes in each cluster that have known defense genes in their genomic environment spanning 10 genes upstream and downstream the inspected gene. In addition, each cluster was processed using Clustal-Omega (version 1.2.4)[49] to produce a multiple sequence alignment. The alignment of each cluster was searched using the 'hhsearch' option of hhsuite (version 3.0.3)[50] against the PDB70 and

pfamA_v32 databases, using the '-p 10 -loc -z 1 -b 1 -ssm 2 -sc 1 -seq 1 -dbstrlen 10000 -maxres 32000 -M 60 -cpu 1' parameters. Clusters that had HHsearch hits[51] both to an N-terminal kinase and a C-terminal dCTP deaminase, sized >350aa, and which had a defense score >0.33 were included in the family of the dCTP deaminase genes studied here. Clusters sized <=10 genes, whose representative sequence[48] aligned to the representative sequences of selected dCTP deaminase clusters with an e-value <0.01, and whose size was >350aa, were also included in the family, adding 30 genes to the list in Supplementary Table S1. To generate the list of dGTPase homologs in Supplementary Table S2 we aligned the studied *Sp*-dGTPase protein to all proteins in the database using the 'search' option of MMseqs2 (release 12-113e3) using the '-s 7 -threads 20 -max-seqs 100000' parameters, and retained all hits sized >400aa that had >=25% sequence identity to *Sp*-dGTPase, with alignment overlap of at least 90% on both query and subject sequences.

Bacterial strains and phages

E. coli strain MG1655 (ATCC 47076) was grown in MMB (LB + 0.1 mM MnCl2 + 5 mM MgCl2, with or without 0.5% agar) at 37 °C or room temperature (RT). Whenever applicable, media were supplemented with ampicillin (100 μg/ml), to ensure the maintenance of plasmids. Infection was performed in MMB media at 37°C or RT as detailed in each section. Phages used in this study are listed in Supplementary Table S3.

Plasmid and strain construction

dCTP deaminase and dGTPase genes used in this study were synthesized by Genscript Corp. and cloned into the p15a-origin-containing pSG1 plasmid[14] with their native promoters, or into the pBad plasmid (Thermofisher, cat. #43001), respectively, as previously described[5,14]. Mutants of the dCTP deaminase gene of *E. coli* AW1.7 were also synthesized and cloned by Genscript. All synthesized sequences are presented in Table S5. Mutant of the dGTPase gene of *Shewanella putrefaciens* CN-32 was constructed using Q5 Site directed Mutagenesis kit (NEB, cat. #E0554S) using primers presented in Supplementary Table S6.

Plaque assays

Phages were propagated by picking a single phage plaque into a liquid culture of *E. coli* MG1655 grown at 37°C to OD_{600} of 0.3 in MMB medium until culture collapse. The culture was then centrifuged for 10 minutes at 15,000 x g and the supernatant was filtered through a 0.2 μm filter to get rid of remaining bacteria and bacterial debris. Lysate titer was determined using the small drop plaque assay method as described before[52].

Plaque assays were performed as previously descried[52]. Bacteria (*E. coli* MG1655 with pSG1-dCTP deaminase or pBad-dGTPase) and negative control (*E. coli* MG1655 with empty pSG1 or pBad-GFP) were grown overnight at 37°C. Then 300 μl of the bacterial culture was mixed with 30 ml melted MMB agar (LB + 0.1 mM MnCl2 + 5 mM MgCl2 + 0.5% agar, with or without 0.02% arabinose) and left to dry for 1 hour at room temperature. 10-fold serial dilutions in MMB were performed for each of the tested phages and 10 μl

drops were put on the bacterial layer. Plates were incubated overnight at RT. Plaque forming units (PFUs) were determined by counting the derived plaques after overnight incubation.

Liquid culture growth

Overnight cultures of bacteria harboring the *Sp*-dGTPase gene or a control strain harboring a pBAD-GFP plasmid were diluted 1:100 in MMB medium supplemented with 0.2% arabinose. Cells were incubated at 37°C while shaking at 200 rpm for 1 hour. 180 μL of the bacterial culture were transferred into wells in a 96-well plate and incubated at 37°C with shaking in a TECAN Infinite200 plate reader. OD_{600} was followed with measurement every 10 min using the TECAN iControl v3.8.2.0 software.

DNA and RNA editing assays

Overnight cultures of bacteria (*E. coli* MG1655 harboring pSG1-dCTP deaminase plasmid) or negative control (*E. coli* MG1655 with the pSG1 plasmid) were diluted 1:100 in 60 ml of MMB medium and incubated at 37 °C while shaking at 200 rpm until early log phase (OD_{600} of 0.3). 10 ml samples of each bacterial culture were taken and centrifuged at 15,000 x g for 5 min at 4 °C. The pellets were flash frozen using dry ice and ethanol. The remaining cultures were infected by phage T7 at a final MOI of 2. 10 ml samples were taken throughout infection at 5, 10 and 15 minutes post infection, and centrifuged and flash frozen as described above.

DNA of all samples was extracted using the QIAGEN DNeasy blood and tissue kit (cat #69504), using the gram negative bacteria protocol. Libraries were prepared for Illumina sequencing using a modified Nextera protocol as previously described[53]. Following sequencing on an Illumina NextSeq500 the sequence reads were aligned to bacterial and phage reference genomes (GenBank accession numbers: NC_000913.3, NC_001604.1, respectively) by using NovoAlign (Novocraft) v3.02.02 with default parameters, and mutations were identified and quantified by counting each mismatch across the genome. Frequency of mismatches was compared between control and dCTP deaminase samples throughout the infection time course.

RNA extraction was performed as described previously[54]. Briefly, frozen pellets were re-suspended in 1 ml of RNA protect solution (FastPrep) and lysed by Fastprep homogenizer (MP Biomedicals). RNA was extracted using the FastRNA PRO blue kit (MP Biomedicals, 116025050) according to the manufacturer's instructions. DNase treatment was performed using the Turbo DNA free kit (Life Technologies, AM2238). RNA was subsequently fragmented using fragmentation buffer (Ambion-Invitrogen, cat. #10136824) at 72°C for 1 min and 45 sec. The reactions were cleaned by adding ×2.5 SPRI beads (Agencourt AMPure XP, Beckman-Coulter, A63881). The beads were washed twice with 80% ethanol and air dried for 5 minutes. The RNA was eluted using H_2O. Ribosomal RNA was depleted by using the Ribo-Zero rRNA Removal Kit (epicentre, MRZB12424). Strand-specific RNA-seq was performed using the NEBNext Ultra Directional RNA Library Prep Kit (NEB, E7420) with the following adjustments: all cleanup stages were performed using ×1.8 SPRI beads, and only one cleanup step was performed after the end repair step. Following sequencing on an Illumina NextSeq500, sequenced reads were demultiplexed

and adapters were trimmed using 'fastx clipper' software with default parameters. Reads were mapped to the bacterial and phage genomes by using NovoAlign (Novocraft) v3.02.02 with default parameters, discarding reads that were non-uniquely mapped as previously described[54]. Mutations from reference genomes were identified and quantified by counting each mismatch across the transcriptome. Frequency of mismatches was compared between control and dCTP deaminase samples throughout the infection time course.

Cell lysate preparation

Overnight cultures of *E. coli* harboring the defensive gene and negative controls were diluted 1:100 in 250 ml MMB medium (with or without 0.02% arabinose, as described in Table S5) and grown at 37 °C (250 rpm) until reaching OD_{600} of 0.3. The cultures were infected by T7 (WT phage or mutants where indicated) at a final MOI of 2. Following the addition of phage, at 5, 10 and 15 minutes post infection (plus an uninfected control sample), 50 ml samples were taken and centrifuged for 5 minutes at 15,000 x g. Pellets were flash frozen using dry ice and ethanol. The pellets were re-suspended in 600 µl of 100 mM phosphate buffer at pH=8 and supplemented with 4 mg/ml lysozyme. The samples were then transferred to a FastPrep Lysing Matrix B 2 ml tube (MP Biomedicals cat. #116911100) and lysed using a FastPrep bead beater for 40 seconds at 6 m/s (two cycles). Tubes were then centrifuged at 4°C for 15 minutes at 15,000g. Supernatant was transferred to Amicon Ultra-0.5 Centrifugal Filter Unit 3 kDa (Merck Millipore cat. #UFC500396) and centrifuged for 45 minutes at 4°C at 12,000g. Filtrate was taken and used for LC-MS analysis.

Quantification of nucleotides by HPLC-MS

Cell lysates were prepared as described above and sent for analysis at the Targeted Metabolomics unit of the Weizmann institute. Quantification of nucleotides was carried out using an Acquity I-class UPLC system coupled to Xevo TQ-S triple quadrupole mass spectrometer (both Waters, US). The UPLC was performed using a SeQuant ZIC-pHILIC column as described previously[55], with linear gradient decrease of acetonitrile in 20mM ammonium carbonate during 10 minutes. Mass spectrometry analysis was performed using electrospray interface in positive ionization mode for all metabolites, except for uracil-containing metabolites. Metabolites were detected using multiple-reaction monitoring (MRM), using argon as the collision gas. Quantification was made using standard curve in 0-10µg/mL concentration range. $^{15}N_5$-AMP and $^{13}C_{10}$-ATP (both Sigma) were added to standards and samples as internal standards (0.5uM and 100 uM, respectively). TargetLynx (Waters) was used for data analysis.

Isolation of mutant phages

To isolate mutant phages that escape dCTP deaminase and dGTPase defense, phages were plated on bacteria expressing each defense system using the double-layer plaque assay[56]. Bacterial cells expressing each defense system were grown in MMB (supplemented with

0.2% arabinose for the dGTPase) to an OD_{600} of 0.3. Then, 100 µl of phage lysate was mixed with 100 µl bacterial cells expressing each defense gene, and left at room temperature for 10 minutes. 5 ml of pre-melted 0.5% MMB (supplemented with 0.2% arabinose for the dGTPase) was added and the mixture was poured onto a bottom layer of 1.1% MMB. For the dGTPase, the double layer plates were incubated overnight at 37ºC, and single plaques were picked into 90 ml phage buffer. For the dCTP deaminase, double layer plates were incubated overnight at room temperature and the entire top layer was scraped into 2 ml of phage buffer to enrich for phages that escape the defense genes. Phages were left for 1 h at room temperature during which the phages were mixed several times by vortex to release them from the agar into the phage buffer. The phages were centrifuged at 3200 g for 10 min to get rid of agar and bacterial cells, and the supernatant was transferred to a new tube. For both defense genes, in order to test the phages for the ability to escape from the defense system, the small drop plaque assay was used[52]. 300 ml bacteria harboring the defensive gene or negative control (*E. coli* MG1655 with plasmid pSG1 or pBad-GFP lacking the system) were mixed with 30 ml melted MMB 0.5% agar and left to dry for 1 hour at room temperature. 10-fold serial dilutions in phage buffer were performed for the ancestor phages (WT phage used for the original double layer plaque assay) and the phages formed on the strain expressing the defense genes. 10 µl drops were put on the bacterial layer. The plates were incubated overnight at room temperature or 37ºC for dCTP deaminase or dGTPase, respectively.

Amplification of mutant phages

Isolated phages for which there was decreased defense compared to the ancestor phage were further propagated by picking a single plaque formed on the defense gene in the small drop plaque assay into a liquid culture of *E. coli* harboring the defensive gene, which was grown at 37°C in MMB with shaking at 200 rpm to an OD_{600} of 0.3. The phages were incubated with the bacteria at 37°C with shaking at 200 rpm for 3 hr, and then an additional 9 ml of bacterial culture grown to OD_{600} of 0.3 in MMB was added, and incubated for additional 3 h in the same conditions. For the dGTPase system the growth media was supplemented with 0.2% arabinose throughout the entire amplification process. The lysates were then centrifuged at 15,000 x g for 10 min and the supernatant was filtered through a 0.2 µm filter to get rid of remaining bacteria. Phage titer was then checked using the small drop plaque assay on the negative control strain as described above.

Sequencing and genome analysis of phage mutants

High titer phage lysates ($>10^7$ PFU/ml) of the ancestor phage and isolated phage mutants were used for DNA extraction. 0.5 ml of phage lysate was treated with DNase-I (Merck cat #11284932001) added to a final concentration of 20 ug/ml and incubated at 37°C for 1 hour to remove bacterial DNA. DNA was extracted using the QIAGEN DNeasy blood and tissue kit (cat. #69504) starting from a proteinase-K treatment to degrade the phage capsids. Libraries were prepared for Illumina sequencing using a modified Nextera protocol as described above. Following sequencing on Illumina NextSeq500, reads were aligned to the phage reference genome (GenBank accession number: NC_001604.1) and mutations compared to the reference genome were identified using Breseq (versions 0.29.0 or 0.34.1

for mutant phages that escape dCTP deaminase and dGTPase, respectively) with default parameters[57]. Only mutations that occurred in the isolated mutants, but not in the ancestor phage, were considered. Silent mutations within protein coding regions were disregarded as well.

Protein mass spectrometry

For mass spectrometry of mutant gp5.7 phage, overnight cultures of *E. coli* MG1655 were diluted 1:100 in 250 ml MMB medium and grown at 37 °C (250 rpm) until reaching OD_{600} of 0.3. The cultures were infected by T7 (WT phage or T7 mutant n.5) at a final MOI of 2. Following the addition of phage, at 15 minutes post infection, 50 ml samples were taken and centrifuged for 5 minutes at 15,000 x g. For protein mass spectrometry of cells expressing the dGTPase proteins, overnight cultures *E. coli* MG1655 expressing the *Sp*-dGTPase (as well as control cells expressing GFP) were diluted 1:100 in 100 ml MMB medium supplemented with 0.02% arabinose, and grown at 37 °C (250 rpm) until reaching OD_{600} of 0.3. 50 ml samples were taken and centrifuged for 5 minutes at 15,000 x g. Pellets were flash frozen using dry ice and ethanol. The cell pellets were subjected to lysis and in solution tryptic digestion using the S-Trap method (Protifi). The resulting peptides were analyzed using nanoflow liquid chromatography (nanoAcquity) coupled to high resolution, high mass accuracy mass spectrometry (Q-Exactive HF). Each sample was analyzed on the instrument separately in a random order in discovery mode. Raw data were processed with MaxQuant v1.6.6.0. The data were searched with the Andromeda search engine against a database containing protein sequences of the WT and mutant protein sequences provided, the *E. coli* K12 and T7 protein databases as downloaded from Uniprot, and a list of common lab contaminants. The following modifications were defined for the search: Fixed modification- cysteine carbamidomethylation. Variable modifications- methionine oxidation, asparagine and glutamine deamidation. The quantitative comparisons were calculated using Perseus v1.6.0.7. Decoy hits were filtered out.

Rifampicin assay

Overnight cultures of *E. coli* harboring the defensive gene and negative controls were diluted 1:100 in MMB medium (with or without 0.02% arabinose, as described in Table S5) and grown at 37 °C (250 rpm) until reaching OD_{600} of 0.3 (for dCTP deaminase expressing cells) or OD_{600} of 0.6 (for dGTPase expressing cells). The cultures were supplemented with rifampicin to a final concentration of 100 ug/ml. 60 minutes post the addition of rifampicin (as well as an untreated control), 50 ml samples were taken and centrifuged for 5 minutes at 15,000 x g. Pellets were flash frozen using dry ice and ethanol. The pellets were re-suspended in 600 µl of 100 mM phosphate buffer at pH=8 and supplemented with 4 mg/ml lysozyme. The samples were then transferred to a FastPrep Lysing Matrix B 2 ml tube (MP Biomedicals cat. #116911100) and lysed using a FastPrep bead beater for 40 seconds at 6 m/s (two cycles). Tubes were then centrifuged at 4°C for 15 minutes at 15,000g. Supernatant was transferred to Amicon Ultra-0.5 Centrifugal Filter Unit 3 kDa (Merck Millipore cat. #UFC500396) and centrifuged for 45 minutes at 4°C at 12,000g. Filtrate was taken and used for LC-MS analysis as described above.

DNA replication assay

Overnight cultures of *E. coli* harboring the defensive gene and negative controls were diluted 1:100 in MMB medium (with or without 0.02% arabinose, as described in Table S5) and incubated at 37 °C while shaking at 200 rpm until early log phase (OD_{600} of 0.3). For dCTP deaminase, *Sp*-dGTPase expressing cells and their controls 10 ml samples of each bacterial culture were taken and spiked with 0.5 ml of *Bacillus subtilis* BEST7003 (stationary phase, diluted 1:10), and centrifuged at 15,000 x g for 5 min at 4 °C. The pellets were flash frozen using dry ice and ethanol. The remaining cultures were infected by phage T7 at a final MOI of 2. 10 ml samples were taken at 10 and 15 minutes post infection, and spiked with 0.5 ml *Bacillus subtilis* BEST7003 prior to centrifugation. Samples were flash frozen as described above. For *Ec S*-dGTPase expressing cells and their controls, 10 ml samples of each bacterial culture were taken and centrifuged at 15,000 x g for 5 min at 4 °C. The pellets were flash frozen using dry ice and ethanol. The remaining cultures were infected by phage T7 at a final MOI of 2. 10 ml samples were taken at 10 and 15 minutes post infection. Samples were flash frozen as described above.

DNA of all samples was extracted using the QIAGEN DNeasy blood and tissue kit (cat #69504), using the gram negative bacteria protocol. Libraries were prepared for Illumina sequencing using a modified Nextera protocol as previously described[53]. For *Ec S*-dGTPase expressing cells and their controls, 20ng of *Salinispora pacifica* 45547 purified DNA was spiked in following DNA extraction for normalization purposes. Following sequencing on an Illumina NextSeq500 the sequence reads were aligned to *E. coli* MG1655, *B. subtilis* BEST7003 and phage reference genomes (GenBank accession numbers: NC_000913.3, AP012496, NC_001604.1, respectively) by using NovoAlign (Novocraft) v3.02.02 with default parameters. The amount of reads mapped to *E. coli* and T7 genomes were normalized to the amount of spike reads mapped to *B. subtilis* or *Salinispora pacifica* 45547 genome, as mentioned above.

Phylogenetic analysis

To generate the phylogenetic trees in Fig. 4 the 'clusthash' option of MMseqs2[48] (release 6-f5a1c) was used to remove protein redundancies (using the '--min-seq-id 0.9' parameter). Sequences of the defensive genes were aligned using clustal-omega v.1.2.4 with default parameters[58]. FastTree was used to generate a tree from the multiple-sequence alignment using default parameters[59]. iTOL was used for tree visualization[60]. For dCTP deaminase, 11 sequences of dcd housekeeping proteins were added (IMG ID: 2626483291, 2647195735, 2630603045, 2650338752, 2687266570, 2686871019, 2513211955, 2506703112, 2659712595, 2687072778, 2671786882) and used as an outgroup. For dGTPase, 6 sequences of dgt housekeeping proteins were added (IMG ID: 2875679221, 2908557952, 2908562675, 2909293708, 2913142694, 2913158190) and used as an outgroup.

Data availability

Data that support the findings of this study are available in the Article and its Extended Data. Gene accessions appear in the Methods section of the paper. DNA and RNA sequencing data used in Extended Data Fig. 2 can be found in the European Nucleotide Archive (ENA) ID: ERA11772567. Additional data are available from the corresponding authors upon request.

Statistics & Reproducibility

No statistical method was used to predetermine sample size. Experiments were performed in triplicates unless stated otherwise. Randomization was used for sample injection order in mass spectrometry measurements. No data were excluded from the analyses unless stated otherwise in figure legends.

Acknowledgements

We would like to thank Pascale Cossart for pointing us to the SAMHD1 antiviral mechanism, and Aude Bernheim and the Sorek laboratory members for comments on earlier versions of this manuscript and fruitful discussion. R.S. was supported, in part, by the European Research Council (grant ERC-CoG 681203), Israel Science Foundation (grant ISF 296/21), the Deutsche Forschungsgemeinschaft (SPP 2330, grant 464312965), the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine, the Minerva Foundation with funding from the Federal German Ministry for Education and Research, the Knell Family Center for Microbiology, the Yotam project and the Weizmann Institute Sustainability And Energy Research (SAERI) initiative, and the Dr. Barry Sherman Institute for Medicinal Chemistry. A.M. was supported by a fellowship from the Ariane de Rothschild Women Doctoral Program and, in part, by the Israeli Council for Higher Education via the Weizmann Data Science Research Center, and by a research grant from Madame Olga Klein-Astrachan. Protein mass spectrometry was performed at the Weizmann De Botton Protein Profiling Institute.

Author Contributions Statement

N.T and R.S led the study and performed all analyses and experiments unless otherwise indicated. N.T performed the genetic analyses and the in vivo experimental assays, and analyzed the data. A.M., E.Y and N.T performed the computational analyses of systems prediction. A.S-A. designed and executed the mutant phage isolation experiments and their analysis. C.A. assisted with plaque assays, and DNA isolation of mutant phages. T.F. assisted with isolation of mutant phages, and in conducting plaque assays. A.L and S.M assisted in conducting plaque assays and preparing DNA sequencing libraries. G.A assisted with sequence analysis and prediction of protein domain functions. A.B, and T.M performed mass spectrometry and data analysis. The manuscript was written by N.T. and R.S. All authors contributed to editing the manuscript, and support the conclusions.

Competing Interests Statement

R.S. is a scientific cofounder and advisor of BiomX and Ecophage.

Supplementary Tables

Table S1. Homologs of the dCTP deaminases in this.

Table S2. Homologs of the dGTPases in this study.

Table S3. Phages used in this study.

Table S4. Mutations observed in phages that escape nucleotide-depletion defense.

Table S5. Defense genes synthetized and cloned in this study.

Table S6. Primers used in this study.

Figures

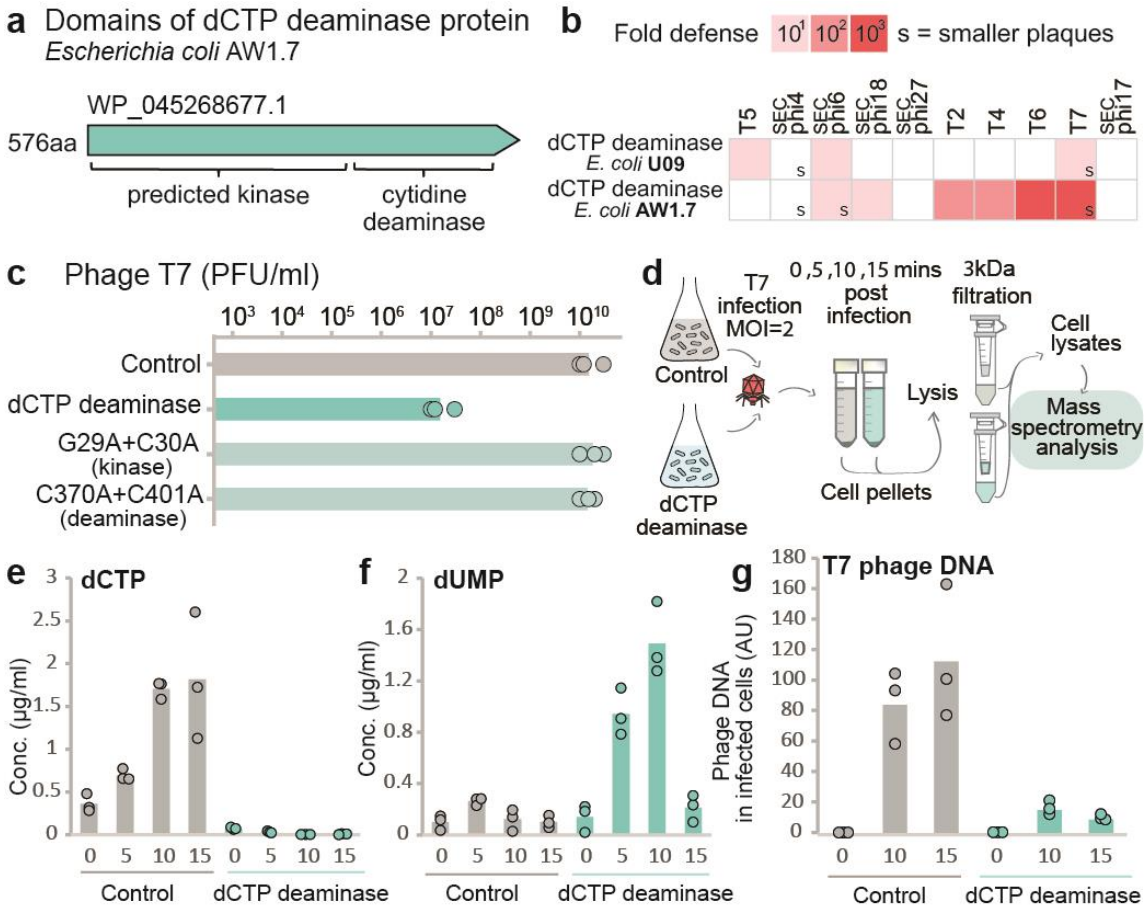


Fig. 1. A family of cytidine deaminases provide defense against phages. (A) Domain organization of the cytidine deaminase of *E. coli* AW1.7. Protein accession in NCBI is indicated above the gene. (B) Cytidine deaminases defend against phages. Cytidine deaminases from two *E. coli* strains were cloned, together with their native promoter regions, and transformed into *E. coli* MG1655. Fold defense was measured using serial dilution plaque assays. Data represent an average of three replicates (see detailed data in Extended Data Fig.1). The designation 's' stands for a marked reduction in plaque size. (C) Effect of point mutations on the defensive activity of dCTP deaminase from *E. coli* AW1.7. Data represent plaque-forming units per ml (PFU/ml) of T7 phages infecting control cells, dCTP deaminase-expressing cells, and two strains mutated in the predicted kinase or deaminase domains. Shown is the average of three replicates, with individual data points overlaid. (D) Schematic representation of the MS experiment. (E-F) Concentrations of deoxynucleotides in cell lysates extracted from T7-infected cells, as measured by LC-MS with synthesized standards. X-axis represents minutes post infection, with zero representing non-infected cells. Cells were infected by phage T7 at a multiplicity of infection (MOI) of 2 at 37°C. Each panel shows data acquired for cells expressing the dCTP deaminase from *E. coli* AW1.7 or for control cells that contain an empty vector. Bar graphs represent the average of three biological replicates, with individual data points overlaid. (G) Effect of dCTP-deaminase on T7 DNA replication throughout infection. Cells were infected by phage T7 at an MOI of 2 at 37°C. A fixed amount of *Bacillus subtilis* BEST7003 cells was spiked into each sample prior to centrifugation for normalization purposes. Total DNA was extracted from each sample and DNA was Illumina-sequenced. Each panel shows data acquired for dCTP deaminase-expressing cells or for control cells that contain an empty vector. Y axis represents phage DNA sequence reads normalized to spike reads. Bar graphs represent the average of three biological replicates, with individual data points overlaid.

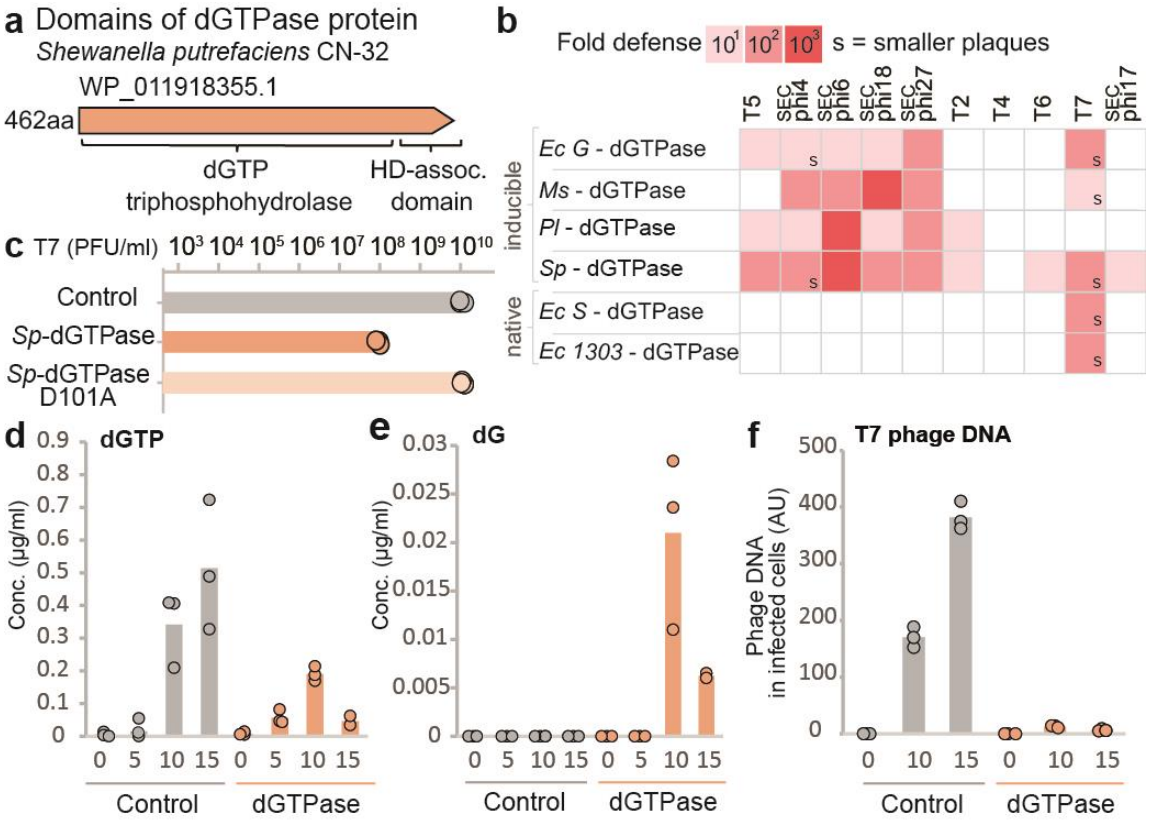


Fig. 2. dGTPases defend against phages by depleting dGTP during phage infection. (A) Domain organization of *Sp*-dGTPase. Protein accession in NCBI is indicated above the gene. (B). Multiple dGTPases defend against phages. dGTPases were either cloned under the control of an arabinose-inducible promoter, or cloned together with their native promoter regions, and transformed into *E. coli* MG1655. Fold protection was measured using serial dilution plaque assays. Data represent the average of three replicates (see data in Extended Data Fig. 6, 7). *Ec G*, *E. coli* G177; *Ms*, *Mesorhizobium sp*. URHA0056; *Pl*, *Pseudoalteromonas luteoviolacea* DSM6061; *Sp*, *Shewanella putrefaciens* CN-32; *Ec S*, *E. coli* STEC 2595; *Ec 1303*, *E. coli* 1303. (C) Plating efficiency of phage T7 on control cells, *Sp*-dGTPase-expressing cells, and a strain mutated in the predicted HD motif. Data represent plaque-forming units per ml (PFU/ml) in three replicates. (D-E) Concentrations of deoxynucleotides in cell lysates extracted from T7-infected cells, as measured by LC-MS with synthesized standards. X-axis represents minutes post infection, with zero representing non-infected cells. Cells were infected by phage T7 at an MOI of 2. Each panel shows data acquired for *Sp*-dGTPase-expressing cells or for control cells that contain a GFP-expressing vector. Bar graphs represent the average of three biological replicates (or two replicates for dGTPase-expressing cells at 15 minutes), with individual data points overlaid. (F) Effect of dGTPase expression on T7 DNA replication throughout infection. Cells were infected by phage T7 at an MOI of 2 at 37°C. A fixed amount of *Bacillus subtilis* BEST7003 cells was spiked into each sample prior to centrifugation for normalization purposes. Total DNA was extracted from each sample and DNA was Illumina-sequenced. Each panel shows data acquired for *Sp*-dGTPase-expressing cells or for control cells that contain a GFP-expressing vector. Y axis represents phage DNA sequence reads normalized to spike reads. Bar graphs represent the average of three biological replicates, with individual data points overlaid.

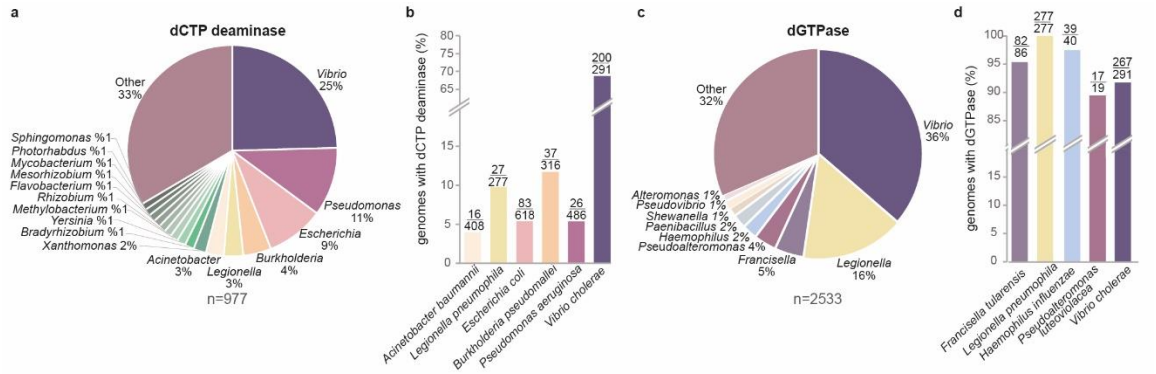


Fig. 3. Distribution of homologs of nucleotide-depleting defense genes in bacterial genomes. (A) Distribution of 977 homologs of dCTP deaminase among bacterial genomes in the analyzed dataset. (B) Presence of dCTP deaminase homologs in selected species. The values above each bar represent the number of sequenced genomes that contain the dCTP deaminase homolog out of the total number of genomes of the species present in the analyzed database. (C) Distribution of 2533 homologs of *Sp*-dGTPases among bacterial genomes in the analyzed dataset. (D) Presence of dGTPase homologs in selected species. The values above each bar are as in panel C.

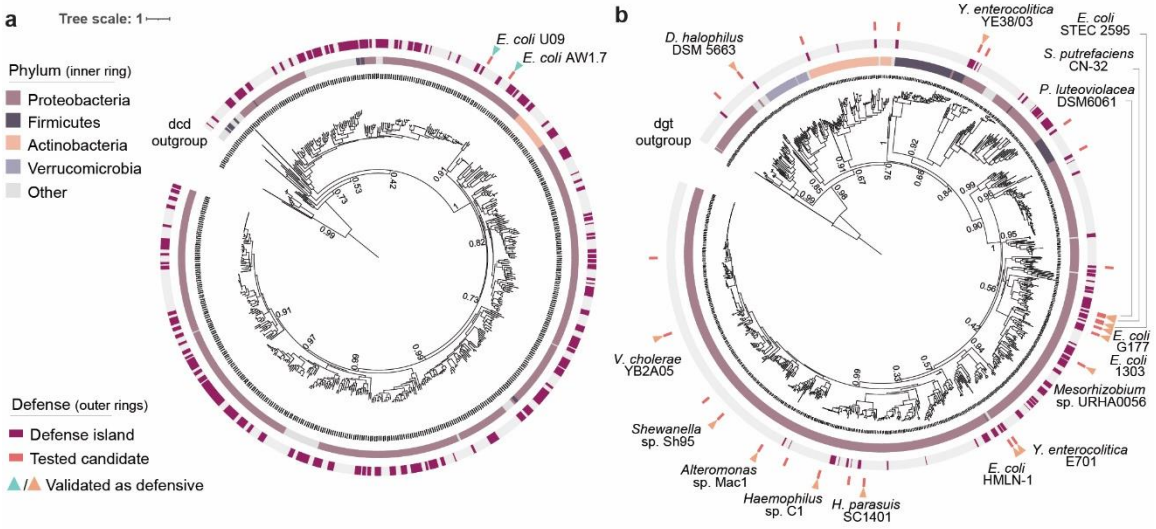


Fig. 4. Diversity of nucleotide-depleting defence genes in bacterial genomes. (A-B) Phylogenetic tree of dCTP deaminase (A) and dGTPase (B) homologs in prokaryotic genomes. Only non-redundant sequences were used to build the trees. The rings represent, from innermost to outermost: the phylogenetic distribution of the homologs, presence next to known defense genes, genes that were tested in this study, and genes that provided defense against phages.

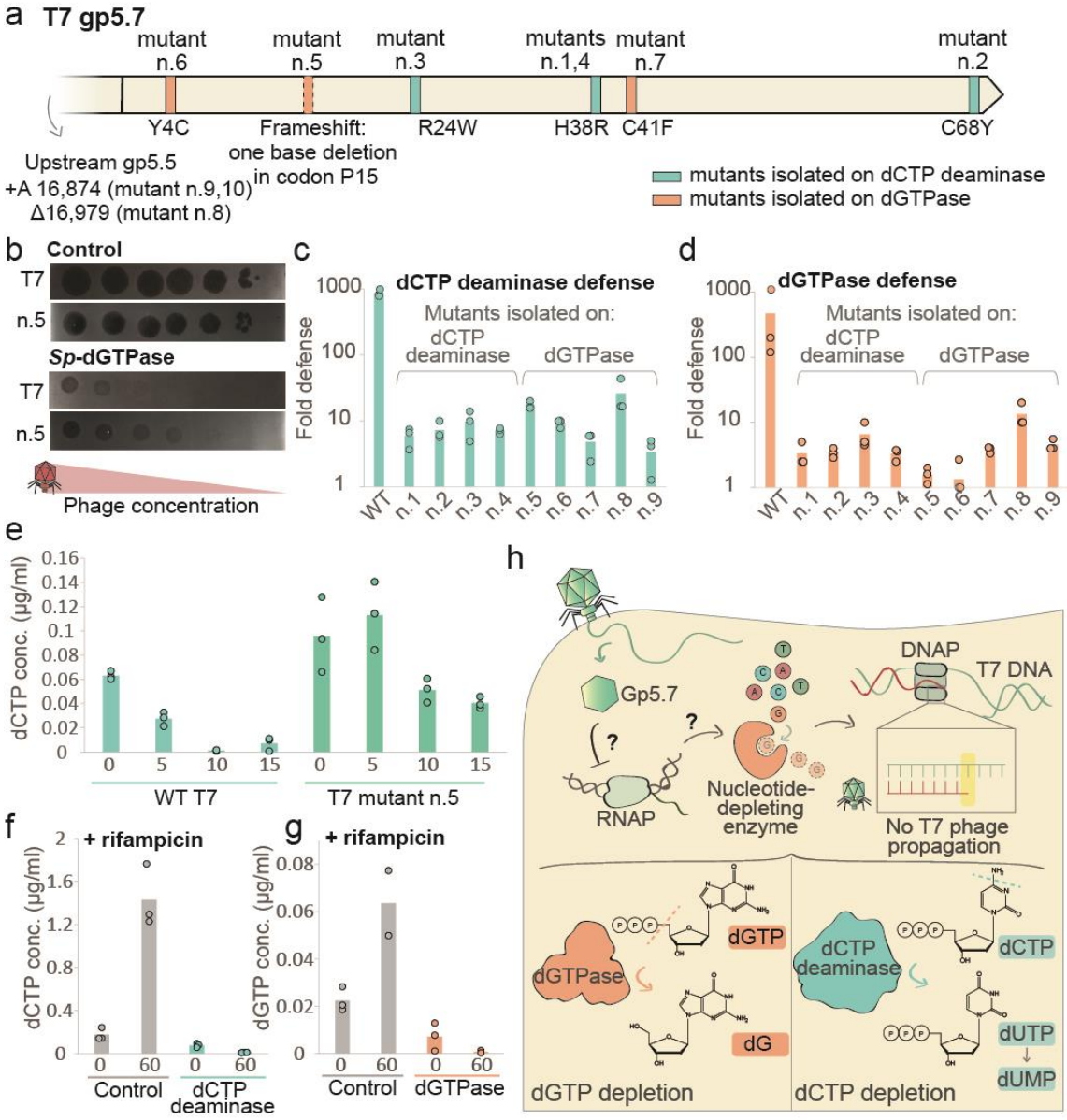


Fig. 5. Phage mutants can overcome nucleotide-depletion defense. (A) Positions in the T7 gene 5.7 that are mutated in phages that escape defense by dCTP deaminase or *Sp*-dGTPase. The full list of mutations for each phage in this Fig. is detailed in Table S4. (B) A representative phage mutant capable of escaping *Sp*-dGTPase defense. Shown are ten-fold serial dilution plaque assays, comparing the plating efficiency of WT and mutant phages on bacteria that express the *Sp*-dGTPase and a control strain that lacks the system and contains a GFP-expressing vector instead. Images are representative of three replicates. (C-D) Reduction of defense of dCTP deaminase (C) or dGTPase (D) when infected by the isolated phage mutants. Bacteria expressing dCTP deaminases and dGTPase, as well as a negative controls, were grown on agar plates in room temperature. Tenfold serial dilutions of the phage lysate were dropped on the plates. Y axis represents fold defense, calculated as the ratio between plaque-forming units per milliliter on control strains and on system-expressing strains. Each bar graph represents average of three replicates, with individual data points overlaid. (E) Concentrations of dCTP in cell lysates extracted from cells expressing dCTP deaminase infected by WT T7 or T7 mutant n.5, at an MOI of 2. X-axis represents minutes post infection, with zero representing non-infected cells. Bar graphs represent average of three biological replicates, with individual data points overlaid. Data for WT T7 is same as in Fig. 1E and presented here for

comparison to mutant. (F-G) dCTP or dGTP concentration during rifampicin treatment on dCTP deaminase-expressing or dGTPase-expressing cells, respectively, as measured by LC-MS with synthesized standards. Early-log cells were supplemented with rifampicin (conc. 100ug/ml) and grown for 60 minutes. Lysates were then extracted and dCTP or dGTP concentrations were measured. Bar graphs represent average of three biological replicates (or two replicates in panel G at 60 mins), with individual data points overlaid. (H) A model for the mechanism of action of nucleotide-depleting anti-phage systems. Phage infection may be sensed by transcription shut-off, which activates a defensive enzyme to deplete a deoxy-nucleotide, hence preventing further phage DNA replication and propagation.

References

1. Hampton, H. G., Watson, B. N. J. & Fineran, P. C. The arms race between bacteria and their phage foes. *Nature* vol. 577 327–336 (2020).

2. Bernheim, A. & Sorek, R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.* **18**, 113–119 (2020).

3. Lopatina, A., Tal, N. & Sorek, R. Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. *Annual Review of Virology* vol. 7 371–384 (2020).

4. Kronheim, S. *et al.* A chemical defence against phage infection. *Nature* **564**, 283–286 (2018).

5. Bernheim, A. *et al.* Prokaryotic viperins produce diverse antiviral molecules. *Nature* **589**, 120–124 (2021).

6. Lowey, B. *et al.* CBASS Immunity Uses CARF-Related Effectors to Sense 3′–5′- and 2′–5′-Linked Cyclic Oligonucleotide Signals and Protect Bacteria from Phage Infection. *Cell* **182**, 38-49.e17 (2020).

7. Cohen, D. *et al.* Cyclic GMP–AMP signalling protects bacteria against viral infection. *Nature* **574**, 691–695 (2019).

8. Lau, R. K. *et al.* Structure and Mechanism of a Cyclic Trinucleotide-Activated Bacterial Endonuclease Mediating Bacteriophage Immunity. *Mol. Cell* **77**, 723-733.e6 (2020).

9. Ofir, G. *et al.* Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nat. 2021 6007887* **600**, 116–120 (2021).

10. Millman, A. *et al.* Bacterial Retrons Function In Anti-Phage Defense. *Cell* **183**, 1551-1561.e12 (2020).

11. Bobonis, J. *et al.* Bacterial retrons encode tripartite toxin/antitoxin systems. *bioRxiv* 2020.06.22.160168 (2020) doi:10.1101/2020.06.22.160168.

12. Johnson, A. G. *et al.* Bacterial gasdermins reveal an ancient mechanism of cell death. *Science (80-.).* **375**, 221–225 (2022).

13. Gao, L. *et al.* Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science (80-.).* **369**, 1077–1084 (2020).

14. Doron, S. *et al.* Systematic discovery of antiphage defense systems in the microbial pangenome. *Science (80-.).* **359**, eaar4120 (2018).

15. Harris, R. S. & Dudley, J. P. APOBECs and virus restriction. *Virology* vols 479–480 131–145 (2015).

16. Vértessy, B. G. & Tóth, J. Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc. Chem. Res.* **42**, 97–106 (2009).

17. Severin, G. B. *et al.* A Broadly Conserved Deoxycytidine Deaminase Protects

Bacteria from Phage Infection. *bioRxiv* 2021.03.31.437871 (2021) doi:10.1101/2021.03.31.437871.

18. Goldstone, D. C. *et al.* HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* **480**, 379–382 (2011).

19. Ayinde, D., Casartelli, N. & Schwartz, O. Restricting HIV the SAMHD1 way: Through nucleotide starvation. *Nat. Rev. Microbiol.* **10**, 675–680 (2012).

20. Quirk, S. & Bessman, M. J. dGTP triphosphohydrolase, a unique enzyme confined to members of the family Enterobacteriaceae. *J. Bacteriol.* **173**, 6665–6669 (1991).

21. Kondo, N. *et al.* Structure of dNTP-inducible dNTP triphosphohydrolase: Insight into broad specificity for dNTPs and triphosphohydrolase-type hydrolysis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 230–239 (2007).

22. Mega, R., Kondo, N., Nakagawa, N., Kuramitsu, S. & Masui, R. Two dNTP triphosphohydrolases from Pseudomonas aeruginosa possess diverse substrate specificities. *FEBS J.* **276**, 3211–3221 (2009).

23. Singh, D. *et al.* Structure of Escherichia coli dGTP Triphosphohydrolase: A hexameric enzyme with DNA effector molecules. *J. Biol. Chem.* **290**, 10418–10429 (2015).

24. Barnes, C. O. *et al.* The crystal structure of dGTPase reveals the molecular basis of dGTP selectivity. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 9333–9339 (2019).

25. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 1–11 (2021) doi:10.1038/s41586-021-03819-2.

26. Lee, S. J. & Richardson, C. C. Choreography of bacteriophage T7 DNA replication. *Current Opinion in Chemical Biology* vol. 15 580–586 (2011).

27. Daddacha, W. *et al.* SAMHD1 Promotes DNA End Resection to Facilitate DNA Repair by Homologous Recombination. *Cell Rep.* **20**, 1921–1935 (2017).

28. Dunn, J. J., Studier, F. W. & Gottesman, M. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **166**, 477–535 (1983).

29. Hadden, J. M., Déclais, A. C., Carr, S. B., Lilley, D. M. J. & Phillips, S. E. V. The structural basis of Holliday junction resolution by T7 endonuclease I. *Nature* **449**, 621–624 (2007).

30. Tabib-Salazar, A. *et al.* T7 phage factor required for managing RpoS in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5353–E5362 (2018).

31. Severinova, E. & Severinov, K. Localization of the Escherichia coli RNA polymerase β′ subunit residue phosphorylated by bacteriophage T7 kinase Gp0.7. *J. Bacteriol.* **188**, 3470–3476 (2006).

32. Tabib-Salazar, A. *et al.* Full shut-off of Escherichia coli RNA-polymerase by T7 phage requires a small phage-encoded DNA-binding protein. *Nucleic Acids Res.*

45, 7697–7707 (2017).

33. Hirsch-Kauffmann, M., Hherrlich, P., Ponta, H. & Schweiger, M. Helper function of T7 protein kinase in virus propagation. *Nature* **255**, 508–510 (1975).

34. Nechaev, S. & Severinov, K. Bacteriophage-Induced Modifications of Host RNA Polymerase. *Annu. Rev. Microbiol.* **57**, 301–322 (2003).

35. Tal, N. *et al.* Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell* **184**, 1–12 (2021).

36. Okada, A. & Iwatani, Y. APOBEC3G-Mediated G-to-A Hypermutation of the HIV-1 Genome: The Missing Link in Antiviral Molecular Mechanisms. *Front. Microbiol.* **07**, 2027 (2016).

37. Stavrou, S. & Ross, S. R. APOBEC3 Proteins in Viral Immunity. *J. Immunol.* **195**, 4565–4570 (2015).

38. Kumar, R., DiMenna, L. J., Chaudhuri, J. & Evans, T. Biological function of activation-induced cytidine deaminase (AID). *Biomedical Journal* vol. 37 269–283 (2014).

39. Wurgler, S. M. & Richardson, C. C. Structure and regulation of the gene for dGTP triphosphohydrolase from Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2740–2744 (1990).

40. Gawel, D., Hamilton, M. D. & Schaaper, R. M. A novel mutator of Escherichia coli carrying a defect in the dgt gene, encoding a dGTP triphosphohydrolase. *J. Bacteriol.* **190**, 6931–6939 (2008).

41. Myers, J. A., Beauchamp, B. B. & Richardson, C. C. Gene 1.2 protein of bacteriophage T7. Effect on deoxyribonucleotide pools. *J. Biol. Chem.* **262**, 5288–5292 (1987).

42. Fineran, P. C. *et al.* The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 894–899 (2009).

43. Pecota, D. C. & Wood, T. K. Exclusion of T4 Phage by the hok/sok Killer Locus from Plasmid R1. *J. Bacteriol.* **178**, 2044–2050 (1996).

44. Kondo, N. *et al.* Insights into different dependence of dNTP triphosphohydrolase on metal ion species from intracellular ion concentrations in Thermus thermophilus. *Extremophiles* **12**, 217–223 (2008).

45. Morehouse, B. R. *et al.* STING cyclic dinucleotide sensing originated in bacteria. *Nature* **586**, 429–433 (2020).

46. Kuzmenko, A. *et al.* DNA targeting and interference by a bacterial Argonaute nuclease. *Nature* **587**, 632–637 (2020).

47. Chen, I. M. A. *et al.* IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic*

Acids Res. **47**, D666–D677 (2019).

48. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* vol. 35 1026–1028 (2017).

49. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

50. Soding, J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951–960 (2005).

51. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J. Mol. Biol.* **430**, 2237–2243 (2018).

52. Mazzocco, A., Waddell, T. E., Lingohr, E. & Johnson, R. P. Enumeration of bacteriophages using the small drop plaque assay system. *Methods Mol. Biol.* **501**, 81–85 (2009).

53. Baym, M. *et al.* Inexpensive Multiplexed Library Preparation for Megabase-Sized Genomes. *PLoS One* **10**, e0128036 (2015).

54. Dar, D. *et al.* Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science (80-.).* **352**, aad9822–aad9822 (2016).

55. JS, L. *et al.* Urea Cycle Dysregulation Generates Clinically Relevant Genomic and Biochemical Signatures. *Cell* **174**, 1559-1570.e22 (2018).

56. Kropinski, A. M., Mazzocco, A., Waddell, T. E., Lingohr, E. & Johnson, R. P. Enumeration of bacteriophages by double agar overlay plaque assay. *Methods Mol. Biol.* **501**, 69–76 (2009).

57. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).

58. Sievers, F. & Higgins, D. G. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* **27**, 135–145 (2018).

59. Price, M. N., Dehal, P. S. & Arkin, A. P. Fasttree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).

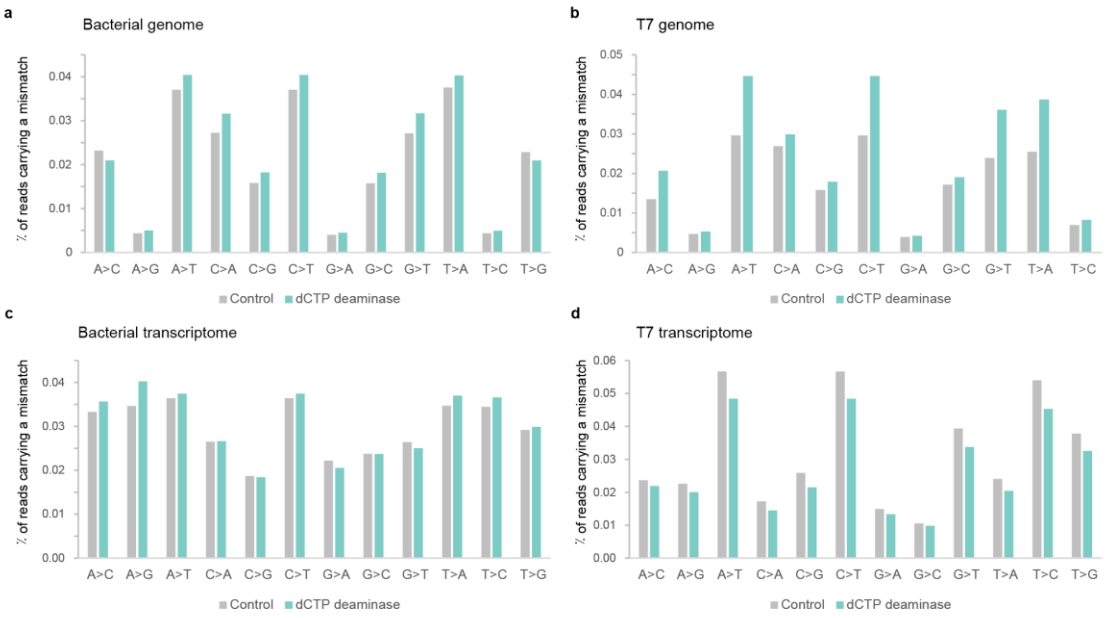
60. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).

61. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2256–2268 (2004).

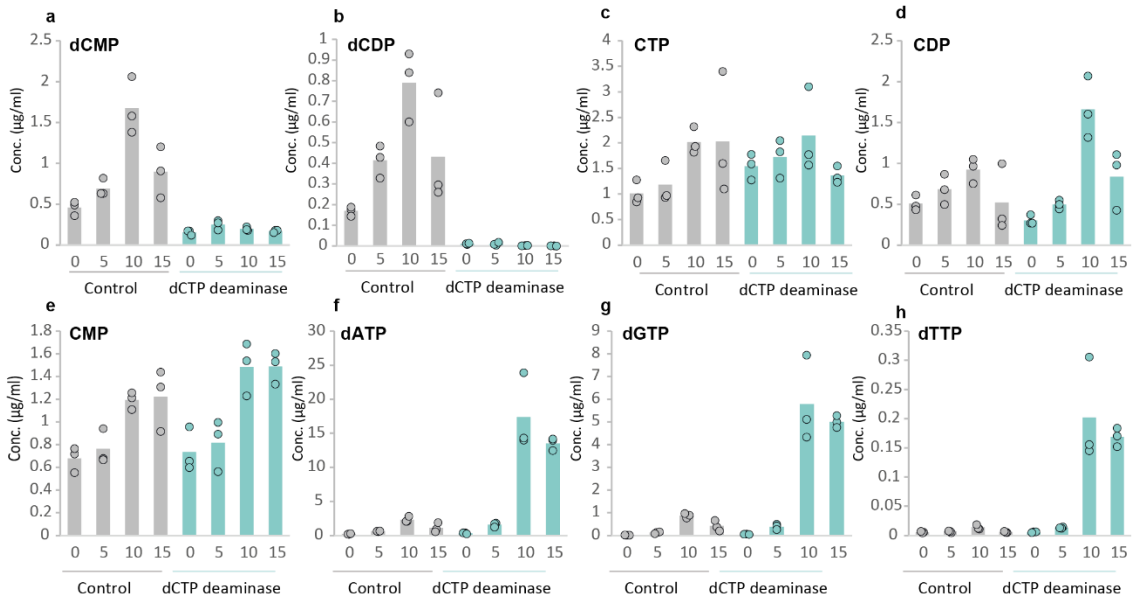
Supplementary Figs

Extended Data Fig.1. dCTP deaminases protect against phage infection. (A) Superposition of the C-terminal region (residues 351-550) of the AlphaFold-predicted structure of *E. coli* AW1.7 dCTP deaminase (turquoise), aligned with *Streptococcus mutans* cytosine deaminase (PDB: 2hvw) (grey). Zn^{2+} ions are depicted as blue spheres. Alignment was performed by PDBeFOLD[61], with a Q score = 0.33, Z score = 9.0 and RMSD = 2.72 Å between the two structures. (B) Representative instances of cytidine deaminase genes (in turquoise) and their genomic neighborhoods. Genes known to be involved in defense are shown in yellow. RM, restriction modification; TAs, toxin-antitoxin systems; Septu and Hachiman are recently described defense systems[14]. The bacterial species and the accession of the relevant genomic scaffold in the Integrated Microbial Genomes (IMG) database[47] are indicated on the left. (C) Bacteria expressing dCTP deaminases from *E. coli* U09 or *E. coli* AW1.7, as well as a negative control that contains an empty vector,

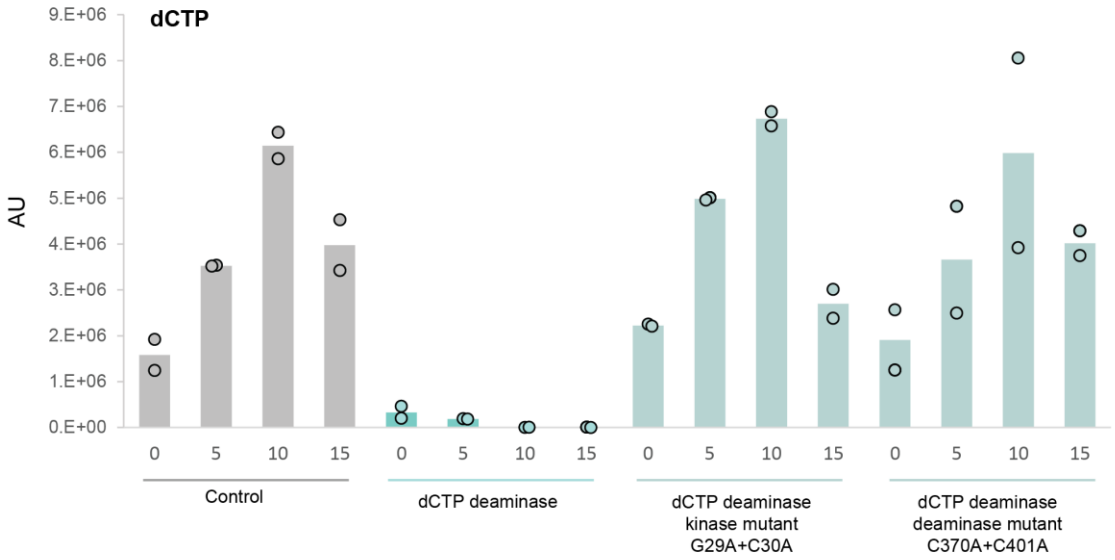
were grown on agar plates in room temperature. Tenfold serial dilutions of the phage lysate were dropped on the plates. Data represent plaque-forming units per milliliter for ten phages tested in this study. Each bar graph represents average of three replicates, with individual data points overlaid. (D) Growth curves of cells expressing *E. coli* AW1.7 dCTP deaminase (turquoise), and control cells (grey). Results of three replicates are presented as individual curves.



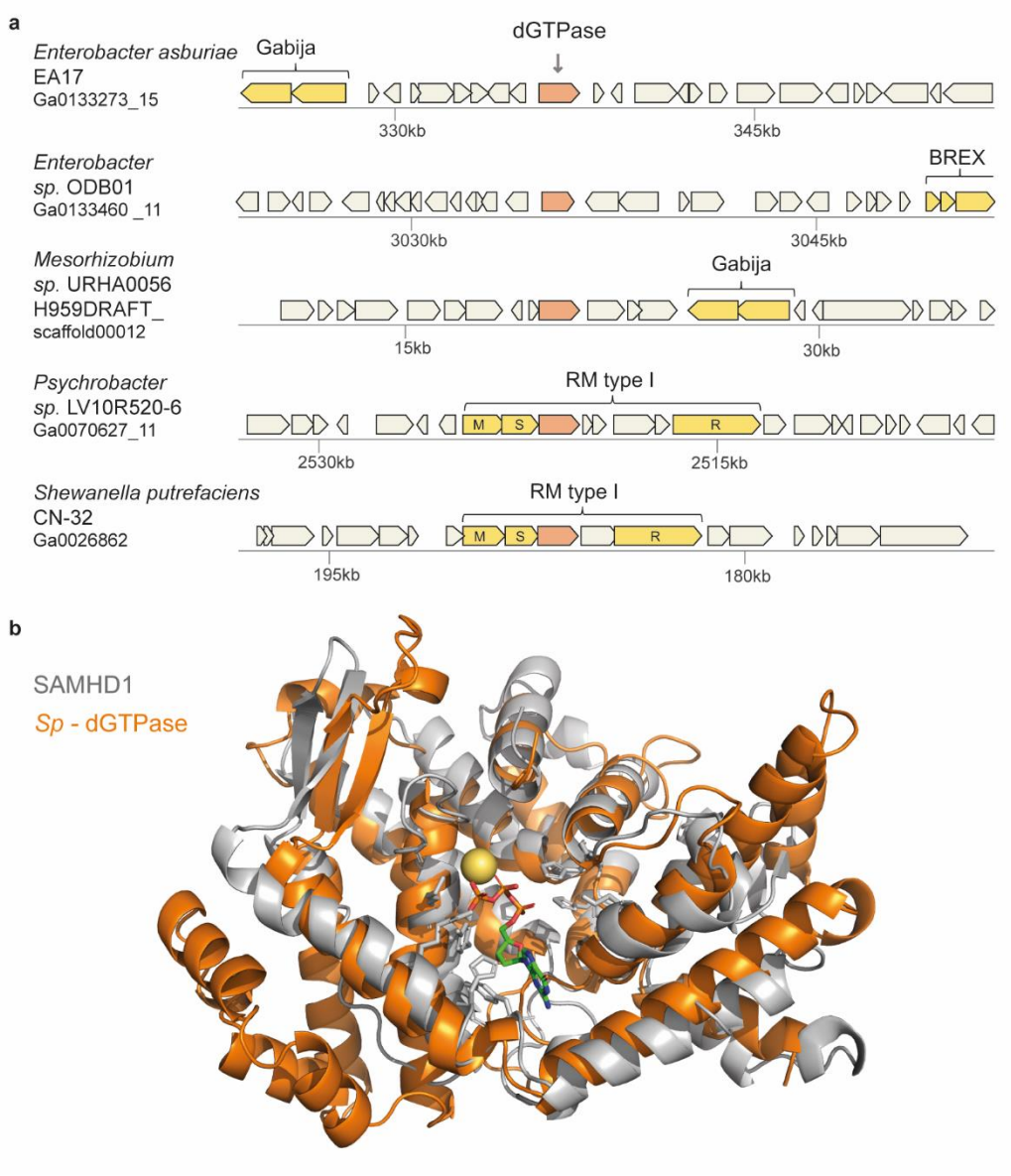
Extended Data Fig.2. No evidence for editing of genome and transcriptome by the dCTP deaminase. Cells expressing the deaminase from *E. coli* AW1.7 were infected by phage T7 at a multiplicity of infection (MOI) of 2 at 37°C. Total DNA and total RNA were extracted after 15 minutes from the onset of infection, and were subjected to DNA-seq and RNA-seq, respectively. Panels A and B show the abundance of DNA reads with specific mismatches for reads aligned to the bacterial genome (A) or the phage genome (B). Panels C and D show the abundance of RNA-seq reads with specific mismatches for reads aligned to the bacterial genome (C) or the phage genome (D).



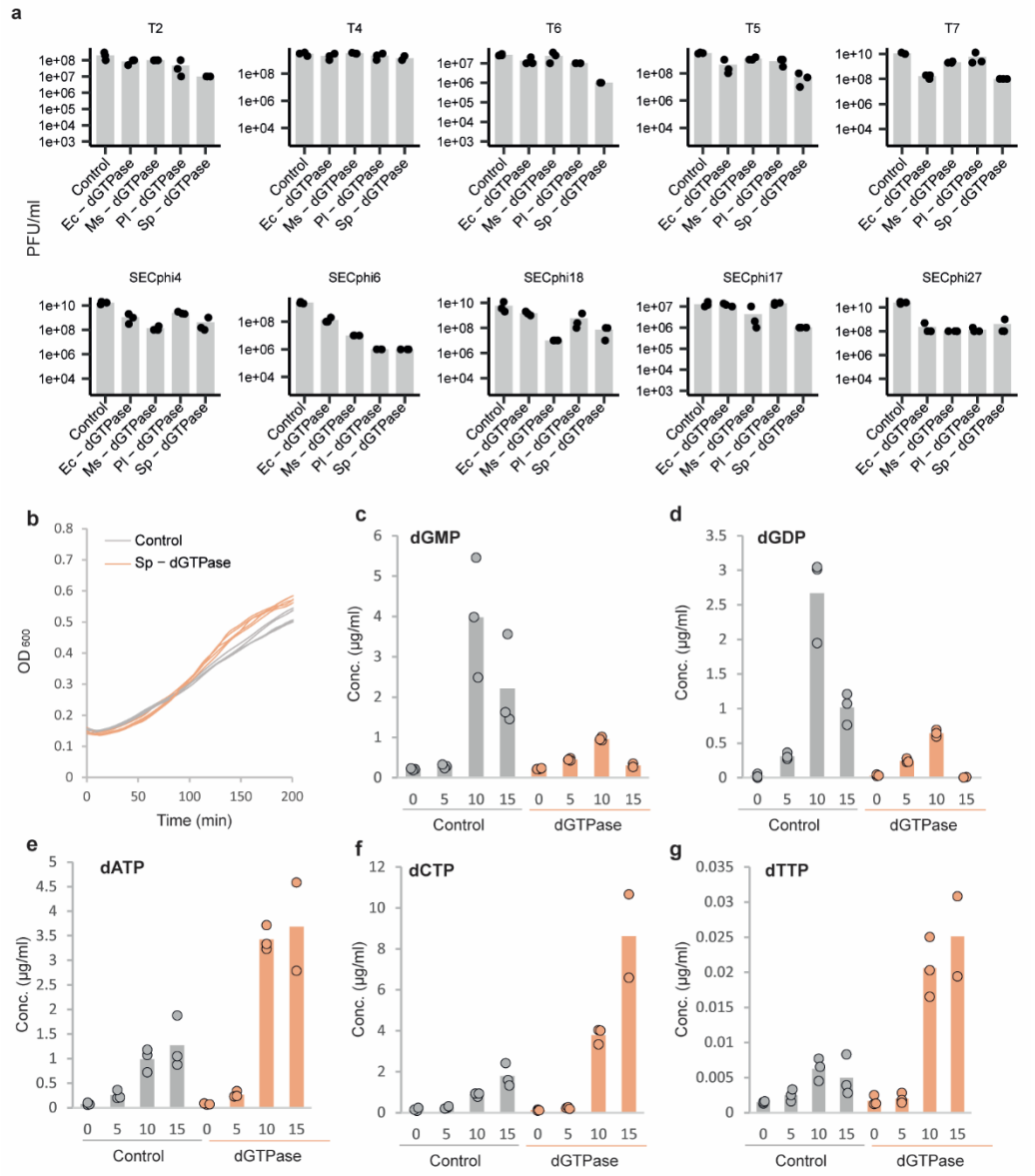
Extended Data Fig.3. Cellular nucleotides during phage infection. (A-H) Concentrations of various nucleotides in cell lysates extracted from T7-infected cells, as measured by LC-MS with synthesized standards. X axis represents minutes post infection, with zero representing non-infected cells. Cells were infected by phage T7 at an MOI of 2. Each panel shows data acquired for dCTP deaminase-expressing cells or for control cells that contain an empty vector. Bar graphs represent average of three biological replicates, with individual data points overlaid.



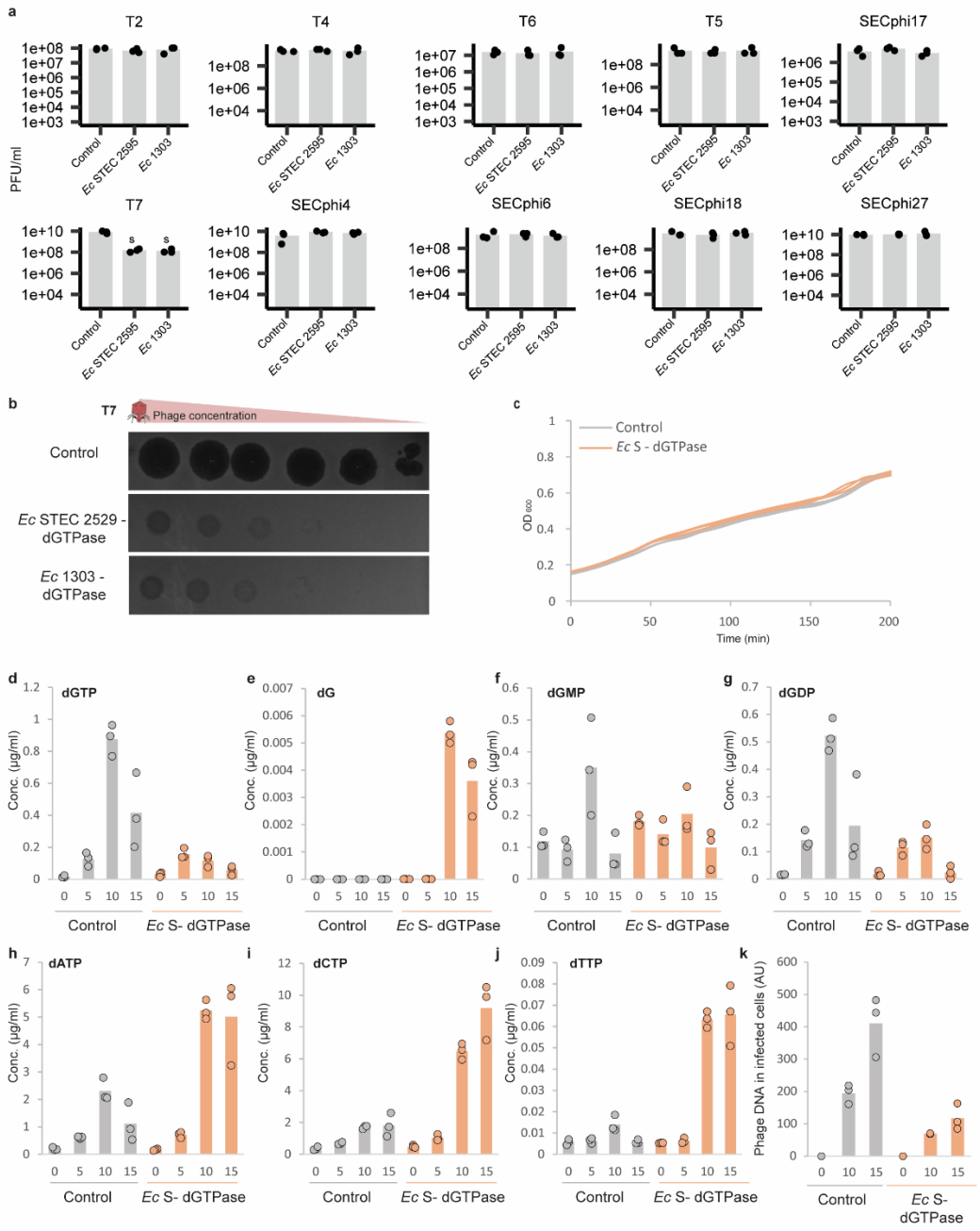
Extended Data Fig.4. Mutated dCTP deaminase does not elicit dCTP depletion. Relative abundance of dCTP in cell lysates extracted from T7-infected cells, as measured by LC-MS. X axis represents minutes post infection, with zero representing non-infected cells. Y axis represents the area under the peak for dCTP, in arbitrary units (AU). Cells were infected by T7 at an MOI of 2. Presented are data acquired for cells expressing the dCTP deaminase from *E. coli* AW1.7, control cells that contain an empty vector, or cell expressing mutated forms of the dCTP deaminase. Bar graphs represent average of two biological replicates, with individual data points overlaid.



Extended Data Fig.5. A family of dGTPases in defense islands. (A) Representative instances of dGTPase genes (in orange) and their genomic neighborhoods. Colors and annotations are as in Extended Data Fig.1A. M, S and R designations within type I RM operon genes represent the methylase, specificity, and restriction subunits, respectively. (B) Superposition of the AlphaFold predicted structural model of *Sp*-dGTPase (orange) aligned with the N-terminus of the human SAMHD1 (PDB: 4bzb, chain D) (grey). Mg^{2+} ion is depicted as a yellow sphere. SAMHD1 catalytic site is depicted with grey sticks, and is bound to the dGTP ligand. Alignment was performed by PDBeFOLD[61], with a Q score = 0.11, Z score = 5.5 and RMSD = 2.89 Å between the predicted dGTPase AlfaFold2 model and the human SAMHD1. The presented structural alignment includes residues 23-161 & 248-452 from the dGTPase AlphaFold model, aligned to residues 129-275 & 299-439 of the human SAMHD1 structure.

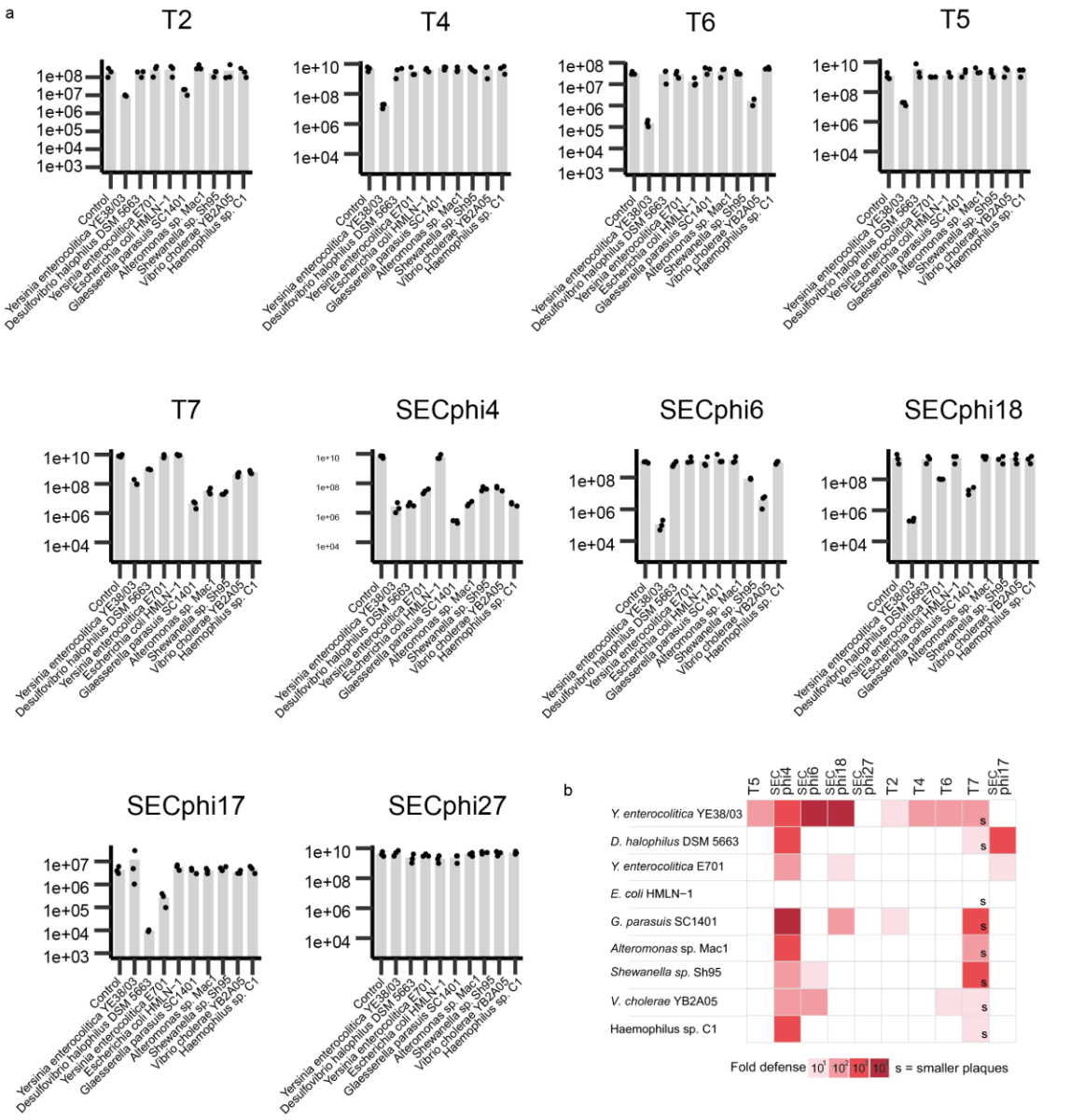


Extended Data Fig.6. dGTPases protect against phage infection. (A) *E. coli* MG1655 cells expressing dGTPases cloned under an arabinose-inducible promoter from several species (*Ec*, *E. coli* G177; *Ms*, *Mesorhizobium sp*. URHA0056; *Pl*, *Pseudoalteromonas luteoviolacea* DSM6061; *Sp*, *Shewanella putrefaciens* CN-32), as well as a negative control, were grown on agar plates in room temperature in the presence of 0.2% arabinose. Tenfold serial dilutions of the phage lysate were dropped on the plates. Data represent plaque-forming units per milliliter for phages tested in this study. Each bar graph represents average of three replicates, with individual data points overlaid. (B) Growth curves of cells over-expressing the *Sp*-dGTPase gene from *Shewanella putrefaciens* CN-32 (orange) and control cells (grey) following 0.2% arabinose induction. Results of four replicates are presented as individual curves. Expression of the *Sp*-dGTPase protein was verified via protein mass spectrometry. (C-G) Concentrations of deoxynucleotides in cell lysates extracted from T7-infected cells, as measured by LC-MS with synthesized standards. X axis represents minutes post infection, with zero representing non-infected cells. Cells were infected by phage T7 at an MOI of 2. Each panel shows data acquired for dGTPase-expressing cells or for control cells that express GFP. Bar graphs represent average of three biological replicates (or two replicates for the dGTPase samples at t=15 mins), with individual data points overlaid.

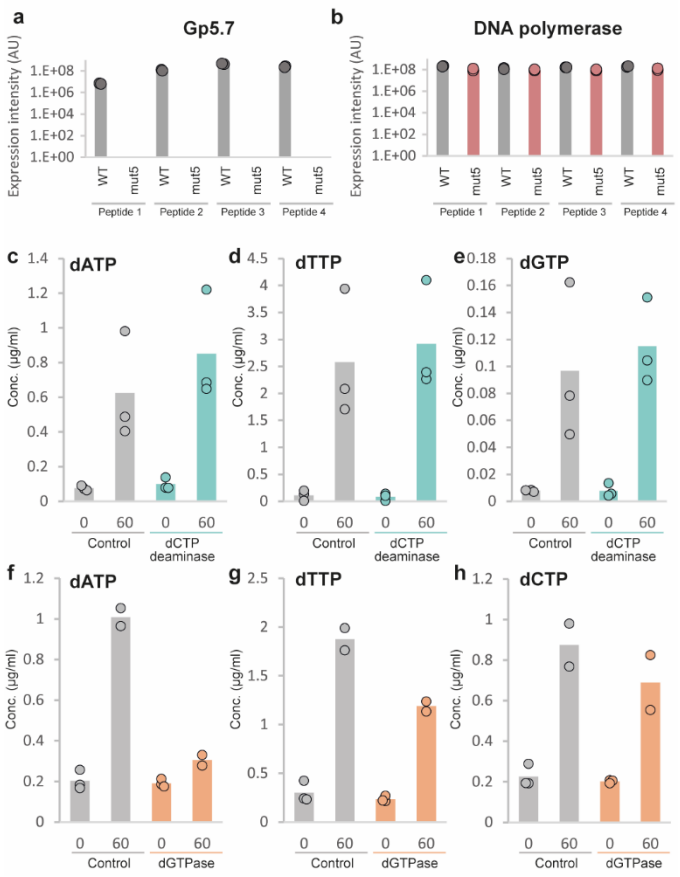


Extended Data Fig.7. dGTPases cloned under native promoters protect against phage infection. (A) *E. coli* MG1655 cells containing dGTPases cloned, together with their native promoters, from two different *E. coli* strains (*Ec S*, *E. coli* STEC 2595; *Ec 1303*, *E. coli* 1303), as well as a negative control, were grown on agar plates in room temperature. Tenfold serial dilutions of the phage lysate were dropped on the plates. Data represent plaque-forming units per milliliter for tested phages. Each bar graph represents average of three replicates, with individual data points overlaid. (B) T7 defense by dGTPases expressed from native promoters. Shown are ten-fold serial dilution plaque assays, comparing the plating efficiency of T7 phage on bacteria that express the *Ec* STEC 2529*, Ec* 1303 or a control strain that lacks the gene. Images are representative of three replicates. (C) Growth curves of cells harboring the *Ec* S - dGTPase gene from *E. coli* STEC 2595 (orange) and control cells with an empty plasmid (grey). Results of three replicates are presented

as individual curves. (D-J) Concentrations of deoxynucleotides in cell lysates extracted from T7-infected cells, as measured by LC-MS with synthesized standards. X axis represents minutes post infection, with zero representing non-infected cells. Cells were infected by phage T7 at an MOI of 2. Each panel shows data acquired for cells expressing the *Ec S* dGTPase or for control cells containing an empty pSG1 vector. Bar graphs represent average of three biological replicates, with individual data points overlaid. (K) Effect of native dGTPase expression on T7 DNA replication throughout infection. Cells were infected by phage T7 at an MOI of 2 at 37°C. Total DNA was extracted from each sample and DNA was Illumina-sequenced. Each panel shows data acquired for *Ec-S*-expressing cells or for control cells that contain an empty vector. Y axis represents phage DNA sequence reads normalized to reads from spiked-in DNA. Bar graphs represent the average of three biological replicates, with individual data points overlaid.



Extended Data Fig.8. Distant homologs of *Sp*-dGTPase protect against phage infection. (A) Bacteria expressing dGTPase cloned from multiple species (*Yersinia enterocolitica* YE38/03, *Desulfovibrio halophilus* DSM 5663, *Yersinia enterocolitica* E701, *Escherichia coli* HMLN-1, *Glaesserella parasuis* SC1401, *Alteromonas sp.* Mac1, *Shewanella sp.* Sh95, *Vibrio cholerae* YB2A05, *Haemophilus sp.* C1), as well as a negative control, were grown on agar plates in room temperature in the presence of 0.2% arabinose. Tenfold serial dilutions of the phage lysate were dropped on the plates. Data represent plaque-forming units per milliliter for tested phages. Each bar graph represents average of three replicates, with individual data points overlaid. (B) A summary of the defense results from the presented bar graphs.



Extended Data Fig.9. Mutation verification of Gp5.7 and rifampicin treatment. (A) Verification of the absence of Gp5.7 in T7 mutant n. 5 using mass spectrometry. Peptide fragments of Gp5.7 identified by protein mass spectrometry of cells 15 minutes post infection by WT T7 and T7 mutant n.5 (MOI=2). Multiple peptides of Gp5.7 are observed in the WT T7, but no peptides are detected in the mutant T7, supporting that the mutant does not express Gp5.7. Peptide fragments are as follows: peptide 1: GHISCLTTSGR, peptide 2: NGGAWEITASGTR, peptide 3: NNASLVAEAASR, peptide 4: TFQSNYVR. Bar graphs represent average of two biological replicates, with individual data points overlaid. (B) As control to the measurements in panel A, shown are peptide fragments of the T7 RNA polymerase identified by protein mass spectrometry in cells 15 minutes post infection. The T7 RNA polymerase is readily identified in both WT and mutant phages. Peptide 1: EQLALEHESYEMGEAR, peptide 2: MNTINIAK, peptide 3: SVMTLAYGSK, peptide 4: VLAVANVITK. Bar graphs represent average of two biological replicates, with individual data points overlaid. (C-H) Concentrations of dNTP nucleotides in cell lysates extracted from rifampicin-treated cells, as measured by LC-MS with synthesized standards. X axis represents minutes post treatment, with zero representing non-treated cells. (C-E) Each panel shows data acquired for dCTP deaminase-expressing cells or for control cells that contain an empty vector. Bar graphs represent average of three biological replicates, with individual data points overlaid. (F-H) Each panel shows data acquired for dGTPase-expressing cells or for control cells that express GFP. Bar graphs represent average of three biological replicates, with individual data points overlaid. In panels D-F, 60 minute data represent the average of two biological replicates.