Cyclic GMP-AMP signalling protects bacteria against viral infection

Cyclic GMP-AMP signaling protects bacteria against viral infection

Daniel Cohen [1,*], Sarah Melamed [1,*], Adi Millman [1], Gabriela Shulman [1], Yaara Oppenheimer-Shaanan [1], Assaf Kacen [2], Shany Doron [1], Gil Amitai [1,#], Rotem Sorek [1,#]

[1] Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel

[2] Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel

* These authors contributed equally

Corresponding authors: gil.amitai@weizmann.ac.il , rotem.sorek@weizmann.ac.il

Summary

The cGAS-STING pathway is a central component of the cell-autonomous innate immune system in animals[1,2]. The cGAS protein (cyclic GMP-AMP synthase) is a sensor of cytosolic viral DNA and, upon DNA sensing, produces a cyclic GMP-AMP (cGAMP) signaling molecule that binds the STING protein and activates the immune response[3–5]. cGAMP production was also detected in bacteria[6], and in *Vibrio cholerae* it was shown to activate a phospholipase that degrades the inner bacterial membrane[7], but its biological role was unknown. Here we show that cGAMP signaling is part of an anti-phage defense system common in bacteria. This system is composed of a four-gene operon that codes for the bacterial cGAS, the associated phospholipase, and two enzymes with the eukaryotic-like domains E1/E2 and JAB. We show that this operon confers resistance against a wide array of phages. Phage infection triggers cGAMP production, which, in turn, activates the phospholipase leading to loss of membrane integrity and cell death prior to completion of phage reproduction. Diverged versions of this system appear in over 10% of prokaryotic genomes and we show that variants with effectors other than phospholipase also protect against phage infection. Our results suggest that the eukaryotic cGAS-STING antiviral pathway has ancient evolutionary roots in microbial defense against phages.

Bacterial anti-phage immune systems, such as CRISPR-Cas and restriction modification systems, tend to concentrate in so-called "defense islands" in bacterial genomes[8], a property that facilitated the discovery of new defense systems based on their co-localization with known ones[9–11]. We noticed that homologs of the *V. cholerae* cGAS gene (*dncV*) tend to frequently appear near defense genes. Out of 637 homologs of this protein identified through a homology search in 38,167 microbial genomes, we found that 417 (65.5%) are located in the vicinity of known defense systems (Fig. 1a). We have previously shown that such a high propensity of co-localization with defense genes is a strong predictor that the gene under inspection has a role in phage resistance[11]. These results therefore suggest that the bacterial cGAS-encoding gene participates in defense against phages.

The bacterial cGAS (denoted DncV) and the effector phospholipase (CapV, acronym for cGAMP-activated phospholipase in *Vibrio*[7]) are encoded by adjacent genes in the *V. cholerae* genome and are likely expressed in a single operon[7]. The same operon also contains two additional genes whose presence next to the *capV-dncV* gene pair is conserved in the majority of cases (613/637, 96%), suggesting that the putative functional defense system involving the bacterial cGAS comprises these four genes (Fig. 1a). As previously noted[7,12], the two additional genes encode proteins with domains known to be associated with the eukaryotic ubiquitin system: E1 and E2 domains typical of ubiquitin transfer and ligation enzymes, and a JAB domain similar to de-ubiquitinase enzymes that removes ubiquitin from target proteins (Fig. 1b).

To test whether the cGAS-containing four-gene operon is an anti-phage defense system, we cloned the *V. cholerae* El Tor operon, as well as the homologous operon from *E. coli* TW11681, into the lab strain *E. coli* MG1655 that naturally lacks this system. In both cases, the putative four-gene system was cloned together with its upstream and downstream intergenic regions in order to preserve promoters, terminators, and other regulatory sequences. We then challenged the system-containing bacteria with an array of phages that span the three major families of tailed double-stranded DNA phages (*Myoviridae*: T2, T4, T6, P1; *Siphoviridae*: Lambda-vir, T5, SECphi18, SECphi27; and *Podoviridae*: T7), as well as a single-stranded DNA phage of the *Microviridae* family, SECphi17.

Both the *V. cholerae* and the *E. coli* derived four-gene operons conferred defense against multiple phages. The system from *E. coli* provided 10-1,800 fold protection against 6 of the 10 phages tested, while the *V. cholerae* system protected against two of the phages (Fig. 1c; Extended Data Fig. 1). None of the systems protected against transformation of a multi-copy plasmid (Extended Data Fig. 2). The stronger protective effect observed for the *E. coli*-derived system as compared to that from *V. cholerae* possibly stems from its higher compatibility with the *E. coli* host and the coliphages we used. We therefore proceeded with the *E. coli*-derived system for further experiments.

To determine whether the ability to produce and sense cGAMP affects defense, we experimented with mutated forms of the system. Deletion of either the cGAS or the phospholipase genes resulted in complete loss of protection against phage infection (Fig. 1d-e; Extended Data Fig. 3a-d). Moreover, point mutations in two essential aspartate residues in the cGAMP-producing active site of the cGAS gene (D129A/D131A)[7,13] abolished defense, suggesting that cGAMP production is absolutely necessary for the phage-defensive properties of the system (Fig. 1d-e;

Extended Data Fig. 3a-d). A S60A point mutation inactivating the catalytic site of the CapV phospholipase also rendered the system completely inactive against all phages tested, demonstrating that the cGAMP-controlled phospholipase activity of CapV is essential for defense (Fig. 1d-e; Extended Data Fig. 3a-d). These results indicate that as in eukaryotes, the cGAS-cGAMP signaling pathway in bacteria participates in anti-viral defense.

We next examined mutations in the two additional genes in the four-gene operon. Defense against the myophage P1 was not affected when either the E1/E2 domain or the JAB domain genes were deleted, suggesting that these two genes are not essential for protection against P1 (Fig. 1d). Accordingly, a construct containing only the *capV-dncV* cGAS-phospholipase gene pair showed complete defense against P1 (Extended Data Fig. 4), demonstrating that the cGAS-phospholipase pair can work as a standalone defense system against some phages. However, deletion of the E1/E2 domain gene abolished defense against all of the other phages that were tested (T2, T4, T5, T6 and Lambda-vir), while the presence of an intact JAB-domain gene was necessary for protection against some phages (T4, T5, T6) but not against others (P1, T2 and Lambda-vir). Point mutations predicted to inactivate the E1 active site (C493A/C496A) or the peptidase active site in the JAB domain (D38A) recapitulated the results of the full gene deletions, suggesting that the predicted enzymatic activities of these two proteins are necessary for their roles in phage defense (Fig. 1d-e; Extended Data Fig. 3a-d). Altogether, these mutational analyses suggest that the four-gene operon forms a bacterial system that relies on cGAMP signaling and phospholipase activity to defend against a broad range of phages, with the activities of the two additional genes necessary for defense against some, but not all, phages. We denote this system CBASS (Cyclic-oligonucleotide-Based Anti-phage Signaling System).

In animals, sensing of viral infection is the trigger that activates cGAMP production[3]. We therefore sought to examine whether phage infection triggers cGAMP production in the bacterial system as well. To this end, we took advantage of the fact that the CapV phospholipase is activated by cGAMP, and used this protein as a reporter for the presence of cGAMP in infected cells. We first expressed and purified the CapV phospholipase from the *E. coli* CBASS system and measured its activity using an *in vitro* phospholipase activity assay. As expected, chemically synthesized 3'3'-cGAMP induced CapV phospholipase activity in a concentration-dependent manner (Fig. 2a). Next, we exposed the purified CapV to cell lysates derived from Δ*capV* CBASS-containing cells that were filtered to include only small molecules (Methods). The CapV phospholipase became markedly activated when exposed to cell lysates collected 40 minutes after infection with phage P1, suggesting that phage infection triggers cGAMP accumulation in CBASS-containing cells (Fig. 2b). These results were corroborated by targeted mass spectrometry analysis, which detected cGAMP concentrations of 1.4-1.9 μM in lysates collected 40 minutes after infection (Fig. 2c). CapV was not activated when exposed to lysates collected from uninfected cells, and only slightly activated by lysates collected from cells 30 minutes after infection, implying that the phage component or cell state sensed by the CBASS system appears in the cell relatively late in the phage infection cycle (Fig. 2b).

When stimulated by cGAMP, the CapV phospholipase was shown to degrade the membrane of *V. cholerae*, leading to loss of membrane integrity and cell growth arrest or death[7]. This, combined with our findings that cGAMP production is triggered by

phage infection, led us to hypothesize that this defense system executes its defense via abortive infection. Abortive infection defense systems exert their activity by causing the infected bacterial cell to "commit suicide" before the phage replication cycle can be completed. This strategy eliminates infected cells from the bacterial population and protects the culture from a viral epidemic[14]. Such a phenotype predicts that with a high multiplicity of infection (MOI), where nearly all bacteria are infected in the first cycle, massive cell death will be observed in the culture even for cells that harbor the defense system.

To test this hypothesis, we infected bacteria growing in liquid cultures with phage P1 at varying MOIs and examined the culture dynamics. At an MOI of 0.2, where only ~20% of bacteria are initially infected by phage, the culture of the wild type cells collapsed due to phage propagation, while the culture of the CBASS-containing cells was viable (Fig. 2d). Conversely, at an MOI of 2, where almost all bacteria are infected by phage, the culture of the CBASS-containing cells collapsed as well, indicating cell lysis of the infected, system-containing cells. Moreover, while at an MOI of 2 the wild type culture started collapsing 70 minutes after initial infection (consistent with time to lysis of 60-70 minutes post infection reported for phage P1[15]), the CBASS-containing culture started collapsing as early as 45 minutes after infection, a time insufficient for completion of phage P1 replication cycle (Fig. 2d). These results were reproduced also for other phages where, when phages were applied at high MOI, the CBASS-containing cells showed lysis earlier than the time necessary for the completion of the phage replication cycle (Extended Data Fig. 5a). The same phenotype was maintained, in the case of phage P1, for constructs that included only the cGAS and phospholipase gene-pair, without the E1/E2 and JAB genes, indicating that the latter two genes are not necessary for defense by abortive infection against this phage (Extended Data Fig. 5b).

These results were further corroborated by staining of infected cells with propidium iodide, a fluorescent DNA-binding agent that is not membrane-permeable and cannot normally enter cells. Fluorescence activated cell sorting (FACS) analysis showed that at 40 minutes after infection by phage P1, a substantial population of CBASS-containing cells became stained by propidium iodide, indicating loss of membrane integrity (Extended Data Fig. 6). This was associated with loss of normal cell shape, as observed via microscopy (Extended Data Fig. 7).

Combined, the above results suggest a model in which the cGAS-phospholipase defense system somehow senses a phage infection, and this triggers production of cGAMP by the cGAS protein. The cGAMP molecule, in turn, activates the phospholipase which degrades the bacterial membrane and causes cell death (Fig. 2e). Such cell death prior to completion of the phage replication cycle aborts the infection and prevents further phage propagation.

It was recently shown that the bacterial cGAS (*dncV*) gene is a member of a large family of oligonucleotide cyclases that are found in ~10% of all sequenced bacterial genomes[16]. Other members of this family can synthesize a diverse set of cyclic oligonucleotide molecules, including cyclic UMP-AMP, cyclic di-UMP, and even the cyclic trinucleotide AMP-AMP-GMP[16]. This family of genes was divided into several clades based on sequence similarity between its members[16]. We found that all major clades have a high propensity to be genomically associated with other known defense genes in defense islands (Fig. 3a). Between 44% and 74% of the genes in each clade

are located near known defense systems, suggesting that this entire family of oligonucleotide cyclases participates in phage defense (Fig. 3a).

Examining the genomic environment of the 6,232 genes we identified as belonging to this family (Supplementary Table 1; Methods), we found that in 26% of the cases (1,612 instances) they appeared as part of a four-gene operon that resembled the *Vibrio* CBASS system, including the E1/E2 and the JAB domain genes. In 683 (42%) of these operons the predicted effector gene had a phospholipase domain, but, interestingly, in the remaining cases the phospholipase domain in the effector gene was replaced by another domain (Fig. 3b). Common alternative effector domains included endonucleases, usually of the HNH type; a domain comprising predicted transmembrane helices; a domain of unknown function DUF4297 that was shown to participate in the Lamassu phage resistance system[11]; and a TIR (Toll Interleukin Receptor) domain that was also previously shown to participate in anti-phage defense (Fig. 3b)[11]. These observations imply that in these variants of the CBASS defense system, the alternative domains replace the phospholipase in exerting the cell-suicide effector activity. For example, the endonuclease may degrade cellular DNA, while the transmembrane-domain effector may oligomerize to form membrane pores as in the case of the RexAB abortive infection system[17].

In addition, and in agreement with previous reports[18], we found 2,745 cases in which the genes belonging to the family of oligonucleotide cyclases appeared in the context of a two-gene operon that lacked the E1/E2 and JAB domain genes. The second gene in the operon also frequently included phospholipase, endonuclease, or transmembrane helices domains, suggesting that these operons represent a minimal CBASS defense system that comprises only two genes (Fig. 3b). The most common predicted effector domain in these operons was a domain that included either two or four transmembrane helices, overall accounting for over 2,000 instances of such two-gene operons.

To test whether predicted two-gene CBASS systems have a role in phage resistance, we examined a two-gene system from *Bacillus cereus* VD146 that contains a predicted effector gene with four transmembrane helices (Fig. 4a). We engineered this system into the lab strain *Bacillus subtilis* BEST7003 and challenged the engineered strain with an array of 11 *Bacillus* phages as previously described[11] (Methods). The system conferred strong defense against one of these phages, the myophage SBSphiC, verifying its ability to defend against phages (Fig. 4b). Deletion of either the cGAS-like gene or the effector gene rendered the system inactive (Fig. 4b). Finally, infection assays of bacteria that contained the two-gene system in liquid culture showed culture collapse at high MOI, consistent with an abortive infection system (Fig. 4c).

Interestingly, we observed 52 instances of a two-gene CBASS system where in addition to an N-terminal TIR domain, the effector gene also contained a C-terminal STING domain similar to the eukaryotic STING (Fig. 3b; Extended Data Fig. 8). Such a domain arrangement of STING fused to a TIR domain is also found in primitive eukaryotes including the oyster *Crassostrea gigas* and the annelid worm *Capitella teleta*[1]. It is therefore possible that this prokaryotic cGAS-STING represents the ancient evolutionary origin of the eukaryotic cGAS-STING system.

This study reveals the biological role of a large family of defense systems widespread in microbial genomes, but much is still unknown. The exact phage component sensed

by the system is yet to be identified; it is unlikely that this component is cytoplasmic double-stranded DNA (dsDNA) as in the animal cGAS-STING system, because bacteria do not have a nucleus and hence their cytoplasm always contains dsDNA. The role of the E1/E2 and JAB genes also remains a riddle. These genes could be involved in the sensing of some phages, or in mitigation of anti-cGAS activities that phages likely encode.

CBASS systems are widespread in microbial genomes and can be found in representatives of nearly every sequenced bacterial phylum[16] as well as in archaea. It is therefore likely that phages have developed means to mitigate CBASS systems, akin to anti-CRISPR and anti-restriction genes encoded by some phages[19]. These anti-defense genes could be, for example, enzymes that specifically cleave the cGAMP dinucleotide similar to those encoded by some eukaryotic viruses[20]. Such phage-encoded anti-cGAMP enzymes, if they exist, would explain the necessity of CBASS systems to diversify their signaling molecules in order to escape phage counter-defense.

Accumulating evidence suggests that important components of the eukaryotic innate immune system have counterparts in bacterial immune systems. Argonaute, a central protein in the antiviral RNAi machinery of plants, insects and animals[21], was also reported to have immunity roles in bacteria and archaea[22,23]. Toll Interleukin Receptor (TIR) domains, which are essential components of the pathogen-recognizing Toll-like receptors[24], are abundant in bacteria and were recently shown to play a primary role in anti-phage defense[11]. Foreign RNA is sensed in eukaryotic cells by the oligoadenylate synthase (OAS) protein leading to the activation of a non-specific RNAse[25], a process that was recently shown to have parallels in type III CRISPR-Cas immunity upon sensing of phage RNA[26]. All these processes, in addition to our current report that cGAS signaling plays a similar antiviral role in both eukaryotes and prokaryotes, are unlikely to have been the result of parallel evolution. Rather, these observations cumulatively point to a scenario in which these defense systems first evolved in prokaryotes as means of defense against phages, and that the ancient eukaryote that was likely formed by fusion of a bacterium and an archaeon[27] inherited a primordial version of these systems from the prokaryotes that formed it. Under this hypothesis, these systems became the basis for the primitive immune system of the ancient eukaryote, and have evolved into the cell-autonomous immune system that we know today. If this hypothesis is correct, it is not unlikely that future studies would find homologs of additional components of the human immune system functioning as phage resistance systems in bacteria.

References

1. Kranzusch, P. J. *et al.* Ancient Origin of cGAS-STING Reveals Mechanism of Universal 2',3' cGAMP Signaling. *Mol. Cell* **59**, 891–903 (2015).

2. Margolis, S. R., Wilson, S. C. & Vance, R. E. Evolutionary Origins of cGAS-STING Signaling. *Trends in Immunology* **38**, 733–743 (2017).

3. Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786–791 (2013).

4. Ablasser, A. *et al.* cGAS produces a 2′-5′-linked cyclic dinucleotide second messenger that activates STING. *Nature* **498**, 380–384 (2013).

5. Ishikawa, H., Ma, Z. & Barber, G. N. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461**, 788–792 (2009).

6. Davies, B. W., Bogard, R. W., Young, T. S. & Mekalanos, J. J. Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for V. cholerae virulence. *Cell* **149**, 358–370 (2012).

7. Severin, G. B. *et al.* Direct activation of a phospholipase by cyclic GMP-AMP in El Tor Vibrio cholerae . *Proc. Natl. Acad. Sci.* **115**, E6048–E6055 (2018).

8. Makarova, K. S., Wolf, Y. I., Snir, S. & Koonin, E. V. Defense Islands in Bacterial and Archaeal Genomes and Prediction of Novel Defense Systems. *J. Bacteriol.* **193**, 6039–6056 (2011).

9. Goldfarb, T. *et al.* BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* **34**, 169–183 (2015).

10. Ofir, G. *et al.* DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nature Microbiology* **3**, 90–98 (2018).

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

12. Iyer, L. M., Burroughs, A. M. & Aravind, L. The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like β-grasp domains. *Genome Biol.* **7**, R60 (2006).

13. Kato, K., Ishii, R., Hirano, S., Ishitani, R. & Nureki, O. Structural basis for the catalytic mechanism of DncV, bacterial homolog of cyclic GMP-AMP synthase. *Structure* **23**, 843–850 (2015).

14. Molineux, I. J. Host-parasite interactions: recent developments in the genetics of abortive phage infections. *New Biol.* **3**, 230–236 (1991).

15. Walker, J. T. & Walker, D. H. Mutations in coliphage P1 affecting host cell lysis. *J. Virol.* **35**, 519–530 (1980).

16. Whiteley, A. T. *et al.* Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* **567**, 194–199 (2019).

17. Snyder, L. Phage-exclusion enzymes: a bonanza of biochemical and cell biology reagents? *Molecular Microbiology* **15**, 415–420 (1995).

18. Burroughs, A. M., Zhang, D., Schäffer, D. E., Iyer, L. M. & Aravind, L. Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res.* **43**, 10633–10654 (2015).

19. Pawluk, A., Davidson, A. R. & Maxwell, K. L. Anti-CRISPR: Discovery, mechanism and function. *Nat. Rev. Microbiol.* (2018).

20. Tock, M. R. & Dryden, D. T. F. The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.* **8**, 466–472 (2005).

21. Joshua-Tor, L. & Hannon, G. J. Ancestral roles of small RNAs: An ago-centric perspective. *Cold Spring Harb. Perspect. Biol.* **3**, 1–11 (2011).

22. Swarts, D. C. *et al.* DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261 (2014).

23. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. Bacterial Argonaute Samples the Transcriptome to Identify Foreign DNA. *Mol. Cell* **51**, 594–605 (2013).

24. Akira, S. & Takeda, K. Toll-like receptor signalling. *Nature Reviews Immunology* **4**, 499–511 (2004).

25. Zhou, A. *et al.* Interferon action and apoptosis are defective in mice devoid of 2',5'-oligoadenylate-dependent RNase L. *EMBO J.* **16**, 6355–6363 (1997).

26. Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G. & Siksnys, V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605–609 (2017).

27. Margulis, L. Archaeal-eubacterial mergers in the origin of Eukarya: phylogenetic classification of life. *Proc. Natl. Acad. Sci.* **93**, 1071–1076 (1996).

Figure legends

Figure 1. Bacterial cGAS-containing systems protect against phage infection. (a) Bacterial cGAS genes are part of a conserved four-gene operon genomically associated with anti-phage defense systems. Representative instances of the cGAS-containing four-gene putative defense system and their genomic environment are presented. Genes known to be involved in defense are in yellow. Genes of mobile genetic elements are in dark grey. RM, restriction-modification; TA, toxin-antitoxin. BREX, Hachiman and Wadjet are recently described defense systems[9,11]. **(b)** Domain organization of the genes in the cGAS-containing operon. **(c)** The four-gene operon from either *V. cholerae* sv. O1 bv. El Tor or *E. coli* TW11681 was cloned into *E. coli* MG1655 (Methods). Shown is efficiency of plating (EOP) of ten phages infecting the control *E. coli* MG1655 strain ("no system"), the strain with the four-gene operon cloned from *E. coli* TW11681 ("system from *E. coli*") and the strain with the four-gene operon cloned from *V. cholerae* El Tor ("system from *V. cholera*"). Data represent plaque forming units (PFU) per ml; bar graph represents average of three independent replicates, with individual data points overlaid. **(d-e)** EOP of phages infecting the wild-type (WT) *E. coli*-derived four-gene system, deletion strains, and strains with point mutations. Data represent plaque-forming units (PFU) per ml; bar graph represents average of three independent replicates, with individual data points overlaid. Empty vector represents a control *E. coli* MG1655 strain that lacks the system and contains an empty vector instead. **(d)** Infection with phage P1. **(e)** Infection with phage T2.

Figure 2. Phage infection triggers cGAMP accumulation and cell death. (a) Purified CapV protein was incubated *in vitro* with synthetically produced 3'3'-cGAMP in the presence of resorufin butyrate, a phospholipase substrate that emits fluorescence when hydrolyzed. X-axis, concentration of added cGAMP in μM; y-axis, enzyme activity rate measured by the accumulation rate of fluorescence units (FU) per second (Methods). Bar graph represents average of 3 technical replicates, with individual data points overlaid (except for the 2.5μM concentration, where data represent average of 2 replicates). **(b)** Purified CapV protein was incubated *in vitro* with filtered cell lysates derived from bacterial cultures infected by phage P1 at multiplicity of infection (MOI) of 2. Lysates were extracted from cells containing the CBASS system with the *capV* gene deleted ("with system") or from control cells lacking the CBASS system and containing an empty vector instead ("no system"). Lysates were collected either 30 minutes or 40 minutes after initial infection. Bar graph represents average of 3 technical replicates, with individual data points overlaid. **(c)** Targeted mass spectrometry (MS) analysis was performed on the filtered cell lysates collected 40 minutes after initial infection by phage P1 at MOI of 2. Lysates from uninfected samples were taken as control. Y axis represents the concentration of 3'3'-cGAMP in the cell lysate measured by MS, as calculated based on a calibration curve with synthetically produced 3'3'-cGAMP (Methods). Shown are three independent replicates for each condition, with each bar representing an individual replicate. **(d)** Growth curves in liquid culture for CBASS-containing and CBASS-lacking (empty vector) bacteria infected by phage P1 at 37°C. Bacteria were infected at t=0 at an MOI of 0.2 or 2. Shown are three independent replicates for each MOI, with each curve showing an individual replicate. **(e)** Model for the CBASS-based anti-phage activity in bacteria.

Figure 3. Widespread occurrence of CBASS systems in prokaryotic genomes. (a) Phylogenetic tree of proteins with predicted oligonucleotide cyclase domains that were identified by searching a set of 38,167 microbial genomes (Methods). Clades are colored following the color coding in ref. 16. Percentage of genes that are located near known defense systems is indicated for each clade. The outermost ring indicates the type of effector gene associated with the oligonucleotide cyclase gene, with the color code for effector types appearing on the lower left. The inner ring indicates the type of system – either two gene system (comprised of an oligonucleotide cyclase and an effector) or four-gene system (that also includes E1/E2 and JAB domain genes). Color code for system type appears on the upper left. Selected oligonucleotide cyclase genes that were biochemically characterized [16] are indicated, with the cyclic oligonucleotide they produce appearing in grey font. **(b)** Common configurations of four-gene and two-gene CBASS defense systems. The identified number of instances of each such system, and the percent occurrence next to known defense systems, are indicated for each configuration.

Figure 4. A two-gene CBASS system protects *Bacillus* against phage infection. (a) Domain organization of a 2-gene operon found in the *B. cereus* VD146 genome. Locus tags of the depicted genes are indicated below each gene. **(b)** The two-gene operon from *B. cereus* VD146 was cloned and genomically integrated in *Bacillus subtilis* BEST7003 that naturally lacks this system. Shown is EOP of phage SBSphiC infecting the system-lacking and system-containing strains, as well as strains in which one of the two genes was deleted. Bar graph represents average of 3 independent replicates, with individual data points overlaid. **(c)** Growth curves in liquid culture for *B. subtilis* containing the *B. cereus* two-gene CBASS-system or CBASS-lacking *B. subtilis* that contains an empty vector instead, infected by phage SBSphiC. Bacteria were infected at t= 0 at MOI of 0.2 or 2. Shown are three independent replicates for each MOI, with each curve showing an individual replicate.

Methods

Genomic identification and analysis of DncV homologs

The protein sequence of DncV (NCBI accession: NP_229836) was searched against the protein sequences of all genes in 38,167 bacterial and archaeal genomes downloaded from the IMG database[28] on October 2017, using the "search" option in the MMseqs2 package[29] (release 6-f5a1c) with default parameters. Hits with e-value less than 1e-20 were taken as homologs. The fraction of homologs found in the vicinity of known defense systems was calculated as described in Doron et. al.[11], using a positive set of defense gene families (pfams and COGs) that was updated to include the set of new defense genes discovered in ref. [11].

Genomic analysis of oligonucleotide cyclase genes

For the analysis in Fig. 3, the proteins from the database of 38,167 bacterial and archaeal genomes were first clustered using the "cluster" option of MMseqs2[29] (release 2-1c7a89), with default parameters. Clusters were further aggregated into larger clusters using four additional cycles of clustering, where, in each cycle, a representative sequence was taken from each cluster using the "createsubdb" option of MMseqs2 and representative sequences were clustered using the "cluster" option with the "--add-self-matches" parameter. For the first additional clustering cycle the "cluster" option was run with default parameters; for additional cycles 2-4 clustering was run with sensitivity parameter "-s 7.5", and for additional cycle 4, the "--cluster-mode 1" parameter was added.

The sequences of each cluster were aligned using Clustal Omega[30]. Each multiple sequence alignment was scanned with HHpred[31] using 60% gap rule (-M 60) against the PDB_mmCIF70[32] and pfam31[33] databases. Clusters with HHpred hits to one of the cGAS entries (PDBs: 4LEV, 4MKP, 4O67, 5VDR, 5V8H, 4LEW, 5VDP, 4KM5, 4O68, 4O69, 5V8J, 5V8N, 5V8O, 5VDO, 5VDQ, 5VDS, 5VDT, 5VDU, 5VDV, 5VDW, 4XJ5, 4XJ1, 4XJ6, 4XJ3, 4XJ4, and pfam PF03281) with >90% probability in the top 30 hits were taken for manual analysis. Clusters containing genes suspected as belonging to toxin-antitoxin gene pairs were discarded. Overall, this procedure has identified 30 clusters harboring 6,232 predicted oligonucleotide cyclase genes from 5,150 genomes.

The genomic environments spanning 10 genes upstream and downstream of each of the 6,232 predicted oligonucleotide cyclase genes were searched to identify conserved gene cassettes and known defense genes around the oligonucleotide cyclase genes, as described in ref. [11]. Predicted systems were manually reviewed and unrelated genes (e.g. mobilome genes and genes of other defense systems) were omitted.

To generate the phylogenetic tree in Fig. 3a, the "clusthash" option of MMseqs2 (release 6-f5a1c) was first used to remove protein redundancies (using the "--min-seq-id 0.9" parameter). Sequences shorter than 200aa were also removed. To further remove outlier sequences, an all-versus-all search was conducted using the "search" option in MMseqs2, and proteins with less than 20 hits were manually examined. Overall 17 outlier proteins were removed this way. The human cGAS protein (UniProt Q8N884) was added, as well as the human OAS genes (UniProt P00973, P29728, Q9Y6K5) that

were added as outgroup. Sequences were aligned using MAFFT[30] with gap open penalty of " --op 2". The FastTree software[34] was used to generate a tree from the multiple sequence alignment using default parameters. The iTOL[35] software was used for tree visualization.

Cloning of CBASS systems into *E. coli* MG1655

The following CBASS-containing full-length constructs were designed: The four-gene CBASS system from *V. cholerae* sv. O1 bv. El Tor N16961 (GenBank accession NC_002505) was identified as locus tags VC0178-VC0181, and the operon was taken together with its upstream and downstream intergenic regions, altogether spanning the nucleotide range 178424-183957 in GenBank accession NC_002505. Similarly, the four-gene system from *E. coli* TW11681 (GenBank accession AELD00000000) was identified as locus tags ESGDRAFT_00026-ESGDRAFT_00029 and the operon, together with its upstream and downstream intergenic regions, spanned the nucleotide range 21738-27072 in GenBank accession AELD00000000. These two constructs were commercially synthesized and cloned by Genscript Corp., directly into the plasmid pSG1-rfp[10,11] between the AscI and NotI sites of the multiple cloning site.

For strains with gene deletions and point mutations, plasmids containing systems with these deletions/mutations were also commercially synthesized and cloned into the plasmid pSG1-rfp by Genscript Corp., except for one construct where both the E1/E2 and JAB genes were deleted (used for Extended Data Fig. 4 and 5b). To build this construct, the pSG1-rfp backbone was amplified with primers P1-Fw + P1-Rv using KAPA HiFi HotStart ReadyMix (Kapa Biosystems KK2601) (Supplementary Table 2). Primers P2-Fw + P2-Rv were used to lift the gene pair *capV* and *dncV* with their native promoter from the plasmid that contained the full *E. coli* TW11681-derived system (Supplementary Table 2). After gel purification of the PCR products with Zymoclean Gel DNA Recovery Kit (cat # D4001), the two fragments were assembled using NEBuilder HiFI DNA Assembly cloning kit (NEB E5520S) to produce the construct that lacks both gene C (E1/E2) and gene D (JAB domain).

The plasmids were transformed into *E. coli* MG1655 cells by electroporation, and the resulting transformants of the wild type and mutated systems were verified by whole genome sequencing as described in ref. [10] to verify system integrity and lack of mutations. A negative control was constructed as a transformant containing an empty pSG1 plasmid.

Cloning of the CBASS system from *B. cereus* into *B. subtilis* BEST7003

The two-gene CBASS system from *B. cereus* VD146 (GenBank accession KB976672) was identified as locus tags IK1_05630- IK1_05631, and the operon was taken together with its upstream and downstream intergenic regions, altogether spanning the nucleotide range 60974-63493 in GenBank accession KB976672. The construct, as well as two additional constructs where one of the two genes was deleted, were commercially synthesized and cloned by Genscript Corp., directly into the plasmid pSG1-rfp[10,11] between the AscI and NotI sites of the multiple cloning site. The plasmid was transformed into *B. subtilis* and integrated into the amyE locus as previously described[11]. A negative control was constructed as a transformant containing an empty

pSG1 plasmid integrated in the *amyE* locus. The resulting transformants of the wild type and mutated systems were verified by whole genome sequencing as described in ref. [10] .

Phage cultivation

E. coli phages (P1, T4, T5, T7, and lambda-vir) were kindly provided by U. Qimron. Phages SECphi17, SECphi18 and SECphi27 were isolated in our lab[11]. T2 and T6 were ordered from the DSMZ (DSM 16352 and DSM 4622, respectively). The following *B. subtilis* phages were obtained from the BGSC: SPO1 (BGSCID 1P4), phi3T (BGSCID 1L1), SPβ (BGSCID 1L5), SPR (BGSCID 1L56), phi105 (BGSCID 1L11), rho14 (BGSCID 1L15), SPP1 (BGSCID 1P7), SP82G (BGSCID 1P5). Phage phi29 was obtained from the DSMZ (DSM 5546). Phages SBSphiJ and SBSphiC were isolated in our lab[11]. Phages were propagated on either *E. coli* MG1655 or *B. subtilis* BEST7003 in liquid culture, and their titer was determined using the small drop plaque assay method, as previously described[11].

Plaque assays

Bacteria were mixed with MMB agar (LB + 0.1 mM MnCl2 + 5 mM MgCl2 + 0.5% agar), and 10-fold serial dilutions of the phage lysate in MMB were dropped on top of them. After the drops were dry, plates were incubated overnight at 37°C for *E. coli* phage and at room temperature for *B. subtilis* phages. Plaques were counted to calculate Efficiency of plating (EOP) in plaque forming units (PFU) per ml. For phages showing fuzzy killing zone where single plaques could not be counted, the lowest phage concentration where a killing zone was observed was counted as 10 plaques. Fold defense was calculated as the EOP on control bacteria divided by the EOP value obtained on system-containing bacteria.

Phage infection dynamics in liquid media

Overnight cultures were diluted 1:100 in MMB medium and incubated at 37°C while shaking at 250 rpm until early log phase (OD600 = 0.3). 180 µl of the diluted culture were transferred into wells in a 96-well plate containing 20 µl of phage lysate for a final MOI of 2, 0.2, or 0.02 as applicable. Infections were performed in triplicate and optical density at a wavelength of 600 nm was followed using a TECAN Infinite 200 plate reader with measurement every 5 minutes.

Transformation efficiency assay

To prepare electro-competent cells, *E. coli* MG1655 cells (with or without the CBASS system) were diluted 1:100 in 100ml LB media supplemented with ampicillin, 100µg/ml. At OD_{600}=0.6, the cells were transferred to ice for 15 minutes and then centrifuged for 15 minutes at 4000 RPM. The supernatant was discarded and the pellet was resuspended in ultrapure ice-cold water. This step was repeated twice and the supernatant was removed. The resultant pellet was resuspended in 10% glycerol, centrifuged for another 10 minutes and the supernatant was discarded. The final pellet was mixed in 500 µl of 10% glycerol and aliquots of 50 µl were flash frozen in liquid nitrogen and transferred immediately to -80°C.

100ng of pRSFDuet-1 was added to 50µl electro-competent cells (prepared as described above) and the mixture was transferred to a Bio-Rad Gene Pulser Curvette (0.2 cm, cat # 165-2086). The cells were electroporated with a Bio-Rad Micropulser using the "Ec1" setting and then immediately transferred to 1ml LB media to recover at 37°C for 1 hour. After incubation cells were diluted 1:100 and 100µl were plated on LB plates containing ampicillin 100µg/ml and Kanamycin 50µg/ml, and incubated 37°C overnight. Transformation efficiency was calculated by dividing the number of transformants that grew on LB plates containing ampicillin and kanamycin by the live count grown on LB ampicillin 100µg/ml only.

Cloning, expression and purification of CapV from *E. coli* TW11681

The CapV protein from *E. coli* TW11681 was PCR amplified (primers P3-Fw + P3-Rv, and P4-Fw + P4-Rv), fused with a C-terminal His Tag (linker plus His Tag sequence: $SerGly_4His_6$) and cloned into pET28b. The plasmid was transformed into BL21(DE3)pLys cells and grown at 37°C, 250 RPM, until induction (1mM IPTG) at an OD of 0.6. After induction, bacterial cell culture growth was continued for an additional 14 hours at a temperature of 18°C. Cells were then centrifuged for 10 minutes (3900 g, 4°C) and pellets kept at -20°C. Pellets were thawed on ice and resuspended with an ice-cold buffer containing 50mM phosphate buffer (pH 7.4), 300mM NaCl, 10% glycerol (v/v). The buffer was supplemented with 10ug/ml lysozyme, and 1µl Benzonase®-nuclease (5KU Merck). The suspension was than mixed with Lysing matrix B (MP) beads and cells were disrupted mechanically using a FastPrep-24™ (MP) bead-beater device (2 cycles of [40 sec 6m/sec]). Cell lysate was centrifuged at 12,000 g for 10 minutes at 4°C and the supernatant was then mixed with Ni-NTA Magnetic Agarose Beads (Qiagen) for 2hr (4°C). CapV-6His proteins bound to Ni-NTA beads were washed 3 times with a 50mM phosphate buffer (pH 7.4), 300mM NaCl, 10% glycerol (v/v), 20mM imidazole and then eluted with the 50mM phosphate buffer (pH 7.4), 300mM NaCl, 10% glycerol (v/v), 250mM imidazole. The eluent was loaded onto an Amicon Ultra-0.5 Centrifugal Filter Unit 10KDa filter (Merck) in order to exchange the elution buffer with the reaction buffer (50mM phosphate buffer (pH 7.4), 300mM NaCl, 10% glycerol (v/v)). Buffer exchange was done by centrifuging the Centrifugal Filter Unit at 14,000 g for 10 minutes at 4°C 4 times with the reaction buffer. Eluted protein purification was assessed by running the sample on an SDS-PAGE gel and quantity was measured by using a Qubit Protein Assay Kit (Thermo Fisher Scientific).

Fluorogenic biochemical assay for CapV activity

The esterase activity of the His6-tagged CapV was probed with the fluorogenic substrate resorufin butyrate. His6-tagged CapV was diluted in 50 mM sodium phosphate (pH 7.4), 300 mM NaCl, 10% (v/v) glycerol to a final concentration of 1.77µM. To measure the linear range of CapV activation, purified His6-tagged CapV was incubated for 5 minutes with increasing concentrations of 3'3'-cGAMP (Merck), ranging from 0.04-2.5µM of 3'3'-cGAMP. In parallel, His6-tagged CapV was incubated for 5 minutes with cell lysates derived from *E.coli* cells infected by phage P1 or uninfected. Subsequently, the enzyme-cGAMP or enzyme-lysate solution was added to DMSO-solubilized resorufin butyrate [stock of 20mM mixed with 50 mM sodium phosphate (pH 7.4), 300 mM NaCl, 10% (v/v) glycerol reaching a final concentration of 64µM (final assay DMSO concentration was 0.32%)] to a final assay volume of 50µl,

and fluorescence was measured in a 96-well plate (Corning 96 well half area black non-treated with a flat bottom). Plates were read once every 30 seconds for 10 minutes at 37°C using an Infinite-200 (Tecan) with excitation/emission wavelengths of 550/591 nm. Enzymatic reaction velocity was measured as described in ref. [36]. For this, a regression fit was calculated for the output values of each reaction over time, and the slope of this linear regression fit was used for determining the initial reaction velocity (FU/sec).

Cell lysate preparation

E.coli MG1655 cells containing the *E. coli* TW11681-derived CBASS system where the *capV* gene was deleted, were used for preparation of cell lysates. Cells containing an empty vector (pSG1) were used as control. Cells were grown in 100ml MMB medium (flask size 250ml) at 37°C (250RPM) until reaching an OD of 0.3. Cells were then infected with 5ml of phage P1 (titer of 10^{10} infective particles per ml, estimated MOI of 2). After 30 or 40 minutes from initial infection, samples were collected and centrifuged at 3900 g for 5 minutes at 4°C. Following centrifugation, pellets were kept on ice until resuspended in 600µl buffer containing 50 mM sodium phosphate (pH 7.4), 300 mM NaCl and 10% (v/v) glycerol. The resuspended pellet was supplemented with 1µl hen-lysozyme (Merck) (final hen-lysozyme concentration of 16µg/ml). The resuspended cells were then mixed with Lysing matrix B (MP) beads and cells were disrupted mechanically using a FastPrep-24™ (MP) bead-beater device (2 cycles of [40 sec 6m/sec] at 4°C). Cell lysate was then centrifuged at 12,000 g for 10 min at 4°C and the supernatant was loaded onto a 3KDa filter Amicon Ultra-0.5 Centrifugal Filter Unit (Merck) and centrifuged at 14,000 g for 30 min at 4°C. The flowthrough, containing substance smaller than 3KDa, was used as the lysate sample for evaluating cGAMP production within the cells (Fig. 2).

Quantification of 3'3'-cGAMP by HPLC and MS

Cell lysates were prepared as described above. Lysates collected at 40 minutes post-infection were analyzed by MS-Omics (Copenhagen, Denmark) using UPLC (Vanquish, Thermo Fisher Scientific) coupled with a high-resolution quadrupole-orbitrap mass spectrometer (Q Exactive™ HF Hybrid Quadrupole-Orbitrap, Thermo Fisher Scientific). An electrospray ionization interface was used as an ionization source. Analysis was performed in positive ionization mode. A calibration series of 3'3'-cGAMP (SML1232, Sigma-Aldrich) ranging 0.001 to 50 µM was prepared and a linear regression from 0.001 to 5 µM was used for cGAMP quantification.

Fluorescent activated cell sorting (FACS) analysis of infected cells

Overnight cultures of *E.coli* MG1655 cells containing the *E. coli* TW11681-derived CBASS-system and *E.coli* MG1655 cells containing an empty vector (pSG1) were diluted 1:100 in 0.5ml MMB and grown at 37°C and 500 rpm to an OD of 0.3. Cells were then infected with phage P1 (titer of 10^{10} infective particles per ml, estimated MOI of 2). At 40 minutes post-infection, 40 µl of each culture were diluted into 2 ml of filtered PBS containing 1 µl of propidium iodide (Invitrogen LIVE/DEAD *Bac*Light Bacterial Viability Kit (L7007)). The diluted bacteria were incubated in the dark in room temperature for five minutes and were then analyzed by the BIORAD ZE5 Cell Analyzer. A 5 sec agitation was performed and then 25,000 ungated events were

recorded for both the CBASS-containing and CBASS-lacking cultures at 0.2 µl/sec. Forward scatter and propidium iodide fluorescence were measured using the height (H) parameter. FlowJo v10 was used to analyze and visualize the data.

Microscopy of infected cells

E.coli MG1655 cells containing the *E. coli* TW11681-derived CBASS-system or the same CBASS system with a phospholipase catalytic site point mutation (S60A) were grown in MMB medium 37ºC. When growth reached an OD of 0.3, bacteria were infected with P1 phage (MOI ~2). 500µl of the infected samples were centrifuged at 10,000g for 2 min at 25ºC and resuspended in 5 µl of PBS x 1 (Phosphate-Buffered Saline) supplemented with 1 µg/ml membrane stain FM4-64 (Thermo Fisher Scientific T-13320) and 2 µg/ml DNA stain 4,6-Diamidino-2-phenylindole (DAPI) (Sigma-Aldrich Cat#: D9542-5MG). Cells were visualized and photographed using an Axioplan2 microscope (ZEISS) equipped with ORCA Flash4.0 camera (HAMAMATSU). System control and image processing were carried out using Zen software version 2.0 (Zeiss).

References

28. 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).
29. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
30. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
31. 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).
32. Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nature Structural Biology* (2003).
33. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).
34. 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).
35. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
36. Lam, V. *et al.* Resorufin butyrate as a soluble and monomeric high-throughput substrate for a triglyceride lipase. *J. Biomol. Screen.* **17**, 245–251 (2012).
37. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).

Acknowledgements

We thank Azita Leavitt and Saar Sharir for assistance in DNA extraction, library preparation and sequencing, Aude Bernheim for assistance in data visualization, as well as the Sorek lab members for fruitful discussions. This study was supported, in part, by the Israel Science Foundation (personal grant 1360/16), the European Research Council (grant ERC-CoG 681203), the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine, and the Knell Family Center for Microbiology. Adi Millman was supported by a fellowship from the Ariane de Rothschild Women Doctoral Program.

Author contributions

DC, SM and GA led the study and performed all experiments unless otherwise indicated. AM performed the computational analyses appearing in Fig. 1 and 3. YOS performed the microscopy analysis appearing in Extended Data Fig. 6. GS assisted in the plaque assays appearing in Fig. 2 and 4. AK and SD performed the computational analyses leading to Fig. 4. RS supervised the study and wrote the paper together with the team.

Author information

Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel

Daniel Cohen, Sarah Melamed, Adi Millman, Gabriela Shulman, Yaara Oppenheimer-Shaanan, Shany Doron, Gil Amitai, Rotem Sorek

Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel
Assaf Kacen

Corresponding authors: gil.amitai@weizmann.ac.il , rotem.sorek@weizmann.ac.il

Competing interests
R.S. is a scientific cofounder and consultant of BiomX Ltd, Pantheon Ltd and Ecophage Ltd

Data availability statement
Data supporting the findings of this study are available within the manuscript, extended data and supplementary tables. GenBank accessions, locus tags and nucleotide ranges of the CBASS systems appear in the Methods section. Integrated Microbial Genomes (IMG) gene and genome ID number, contig ID, and system and effector classification appear in Supplementary Table 1. Primer sequences of the CBASS systems herein described are available in Supplementary Table 2.

Extended data legends

Extended Data Figure 1. Fold anti-phage defense conferred by 4-gene defense systems against various phages. The four-gene operon from either *V. cholerae* sv. O1 bv. El Tor or *E. coli* TW11681 was cloned into *E. coli* MG1655 (Methods). Shown is fold anti-phage defense as measured by plaque assays. Fold defense was calculated as the ratio between the efficiency of plating (EOP) of the phage on the system-lacking control strain and the EOP on the system-containing strain (Methods; see also Figure 2B). Bar graph represents average of 3 independent replicates, with individual data points overlaid.

Extended Data Figure 2. Transformation efficiency assays. Transformation efficiency of plasmid pRSFDuet-1 into strains containing the 4-gene operon derived either from *E. coli* TW11681 or from *V. cholerae* sv. O1 bv. El Tor, presented as a percentage of the transformation efficiency to *E. coli* MG1655 carrying an empty vector instead of the 4-gene operon. Bar graph represents average of 3 independent replicates, with individual data points overlaid.

Extended Data Figure 3. Efficiency of plating of coliphages on defense systems with whole-gene deletions or point mutations. EOP of phages infecting the wild-type (WT) *E. coli*-derived four-gene system, deletion strains, and strains with point mutations. Data represent plaque forming units (PFU) per ml; bar graphs represent average of 3 independent replicates, with individual data points overlaid. Empty vector represents a control *E. coli* MG1655 strain that lacks the system and has an empty vector instead. (a) Infection with phage T4. (b) Infection with phage T5. (b) Infection with phage T6. (d) Infection with phage Lambda-vir.

Extended Data Figure 4. Efficiency of plating of phage P1 on a double deletion strain. Shown is EOP of phage P1 infecting the wild-type (WT) *E. coli*-derived four-gene system, strains with individual genes deleted, and a strain with two genes deleted. Data represent plaque forming units (PFU) per ml; bar graphs represent average of 3 independent replicates, with individual data points overlaid. Empty vector represents a control *E. coli* MG1655 strain that lacks the system and has an empty vector instead.

Extended Data Figure 5. **The bacterial CBASS system functions through abortive infection.** (a) Growth curves in liquid culture for CBASS-containing and CBASS-lacking (empty vector) bacteria infected by phage SECphi18 at 25°C. Bacteria were infected at t=0 at an MOI of 0.02 or 2. Shown are three independent replicates for each MOI, with each curve showing an individual replicate. (b) Growth curves in liquid culture for cells containing a minimal CBASS system comprised of cGAS-phospholipase (*capV-dncV*) only. Bacteria were infected at t= 0 at MOI of 2 by phage P1. Shown are three independent replicates for each MOI, with each curve showing an individual replicate. Data for empty vector is the same as in panel a.

Extended Data Figure 6. Cell sorting of infected cells stained with propidium iodide. Cells containing the *E. coli* TW11681-derived CBASS-system and control cells containing an empty vector were stained with propidium iodide, a fluorescent DNA-binding agent that penetrates cells which have impaired membrane integrity. Cells were infected by phage P1 (MOI of 2) and sorted based on propidium iodide fluorescence intensity (Y axis). X axis represents forward scatter. (**a**) CBASS-lacking cells, uninfected; (**b**) CBASS-containing cells, uninfected; (**c**) CBASS-lacking cells, 40 minutes post infection; (**d**) CBASS-containing cells, 40 minutes post infection. A large population of cells with high propidium iodide fluorescence intensity is observed. For panels a-d, data for representative replicate are shown out of two independent replicates.

Extended Data Figure 7. Microscopy of infected cells. Shown are phase contrast and overlay of membrane stain (red) and DAPI (blue) images captured at (a) 20 minutes; (b) 40 minutes; and (c) 60 minutes of after infection with phage P1 at MOI of 2. Left panels, *E.coli* MG1655 cells containing the *E. coli* TW11681-derived CBASS-system; right panels, *E.coli* MG1655 cells containing the *E. coli* TW11681-derived CBASS-system with a single point mutation inactivating the CapV phospholipase (CapV:S60A). Cell shape is deformed after 40 minutes in CBASS-containing cells but not in cells where the CBASS is mutated. After 60 minutes, phage-mediated cell lysis is observed in cells where the CBASS is mutated. Shown are representative images from a single replicate out of two independent replicates.

Extended Data Figure 8. Domain analysis and homology-based structure prediction of a bacterial TIR-STING protein. (**a**) Schematics of HHpred homology-based search results of the *Prevotella corporis* TIR-STING protein (Table S1). (**b**) Phyre2 [37] secondary structure prediction of the TIR domain in the *Prevotella corporis* TIR-STING protein as compared to the solved crystal structure of the human TIR domain protein MyD88 (pdb accession 2Z5V_A). (**c**) Phyre2 [37] secondary structure prediction of the STING domain in the *Prevotella corporis* TIR-STING protein as compared to the solved crystal structure of the human STING protein (pdb accession 5BQX_A). Black: identical residues; grey: similar residues. Secondary structure prediction for the bacterial protein appears above the alignment; secondary structure of solved human domain appears below. (**d**) Structural alignment of human TIR domain protein MyD88 and the modeled bacterial TIR domain. (**e**) Structural alignment of human STING domain and the modeled bacterial STING domain. For panels d and e, blue and red represent the structure of the human protein and the model of the bacterial domain structure, respectively.