

# DISARM is a widespread bacterial defence system with broad anti-phage activities

Gal Ofir<sup>1</sup>, Sarah Melamed<sup>1</sup>, Hila Sberro<sup>1,4</sup>, Zohar Mukamel<sup>2,3</sup>, Shahar Silverman<sup>1</sup>, Gilad Yaakov<sup>1</sup>, Shany Doron<sup>1</sup> and Rotem Sorek<sup>1\*</sup>

**The evolutionary pressure imposed by phage predation on bacteria and archaea has resulted in the development of effective anti-phage defence mechanisms, including restriction-modification and CRISPR-Cas systems. Here, we report on a new defence system, DISARM (defence island system associated with restriction-modification), which is widespread in bacteria and archaea. DISARM is composed of five genes, including a DNA methylase and four other genes annotated as a helicase domain, a phospholipase D (PLD) domain, a DUF1998 domain and a gene of unknown function. Engineering the *Bacillus paralicheniformis* 9945a DISARM system into *Bacillus subtilis* has rendered the engineered bacteria protected against phages from all three major families of tailed double-stranded DNA phages. Using a series of gene deletions, we show that four of the five genes are essential for DISARM-mediated defence, with the fifth (PLD) being redundant for defence against some of the phages. We further show that DISARM restricts incoming phage DNA and that the *B. paralicheniformis* DISARM methylase modifies host CCWGG motifs as a marker of self DNA akin to restriction-modification systems. Our results suggest that DISARM is a new type of multi-gene restriction-modification module, expanding the arsenal of defence systems known to be at the disposal of prokaryotes against their viruses.**

The arms race between prokaryotes and the viruses that infect them—bacteriophages (phages)—is driving a continuous and intensive evolution of attack and defence mechanisms<sup>1–3</sup>. Bacterial defence systems target various stages of viral infection in order to thwart the attack, and are rapidly evolving and diverging to answer the fast evolutionary response of phages to these defence strategies<sup>4</sup>. The multiple defence strategies of bacteria include surface modifications to prevent adsorption of phages<sup>3</sup>, restriction-modification (R/M) systems that modify the bacterial genome and degrade unmodified foreign DNA<sup>5</sup>, abortive infection (Abi) systems that trigger cell death or metabolic arrest upon infection<sup>6</sup>, CRISPR-Cas systems that memorize viral genetic material as probes to target future infection attempts<sup>7,8</sup> and more newly discovered defence systems such as the prokaryotic argonaute<sup>9</sup> and BREX<sup>10</sup>. Multiple lines of evidence suggest that many new defence mechanisms are yet to be discovered<sup>2,11</sup>.

R/M systems are the most common form of active defence against phages used by bacteria and archaea<sup>1,5</sup>. Such systems, generally classified into four types, modify the self-genome on specific sequence motifs and degrade the DNA in which such motifs are non-modified<sup>12</sup>. R/M systems contain a restriction endonuclease activity, a DNA modification activity (most commonly methylation, not present in type IV systems) and target recognition capabilities. These three activities could either be encoded on three different protein subunits (as in type I R/M systems), two different subunits of restriction and modification, each carrying the target recognition domain (type III and most of the type II R/M systems), or on a single polypeptide (type IIG)<sup>5,12</sup>.

Many genes involved in bacterial and archaeal defence have been shown to be physically clustered in genomic loci termed 'defence islands'<sup>11,13</sup>. Genes of unknown function that are enriched in defence

islands have been hypothesized to be involved in defence as well, a concept that led to the recent discovery of the BREX defence system<sup>10</sup>. In this study we describe a new cassette of genes that is associated with defence islands in general and R/M systems in particular. This five-gene cassette, which we found in >350 sequenced genomes of bacteria and archaea, generally appears in two classes in nature and encodes genes with helicase-related domains, a DNA methyltransferase and a phospholipase D domain-containing gene. A representative system from *Bacillus paralicheniformis* 9945a was experimentally demonstrated to confer defence against viruses of all three families of tailed phages, in a yet unknown mechanism of action that involves all five genes, methylates the self-genome on specific sequence motifs, and halts phage propagation in the early stages of infection.

## Results

**Identification of DUF1998-containing gene cassettes in defence islands.** We have focused our attention on the protein domain DUF1998 (pfam09369), which has no known function, as this domain was previously found to be enriched in prokaryotic defence islands and hence suggested to participate in anti-phage defence<sup>13</sup>. As defence systems are frequently encoded by multiple genes working in concert that are co-localized in the genome, we attempted to characterize the genetic context in which DUF1998 genes are found.

To this end we searched the Integrated Microbial Genomes (IMG) database<sup>14</sup> for genes whose only annotated domain is DUF1998. Within the 35,893 genomes we scanned, we found 1,369 such genes, encoded in 1,273 different genomes (3.5% of the scanned genomes; Supplementary Table 1). We then analysed the immediate genomic vicinity of these genes and found that, in the vast majority of cases (1,095/1,369, 80%) the DUF1998-containing gene was

<sup>1</sup>Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. <sup>2</sup>Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. <sup>3</sup>Department of Biological Regulation, Weizmann Institute of Science, Rehovot, Israel. <sup>4</sup>Departments of Medicine and Genetics, Stanford University, Stanford, CA, USA. Gal Ofir and Sarah Melamed contributed equally to this work.

\*e-mail: [rotem.sorek@weizmann.ac.il](mailto:rotem.sorek@weizmann.ac.il)

preceded by a large (~1,100–1,300 amino acids (aa)) gene with a pfam00271 domain, a domain that is a part of the catalytic core of DExx-box helicases<sup>15</sup>. Further structural modelling of this protein using Phyre2 (ref. <sup>16</sup>) confirmed structural homology to DExx-box helicases including the two essential catalytic domains<sup>17</sup>. The only abundant annotated gene downstream to the DUF1998 gene was a gene with a phospholipase D (PLD) domain (pfam13091), appearing in 370 of the 1,369 cases (27%). The PLD domain is associated with enzymes that manipulate phosphoester bonds, such as kinases, phospholipases and endonucleases<sup>18</sup>, and has been shown to be the catalytic domain of some restriction endonucleases<sup>19,20</sup>. As the combination of these three genes is the most abundant conserved genetic context of DUF1998, we defined these three consecutive genes as the core of our hypothesized system. We identified 351 such triplets in the analysed genomes (1% of the scanned genomes).

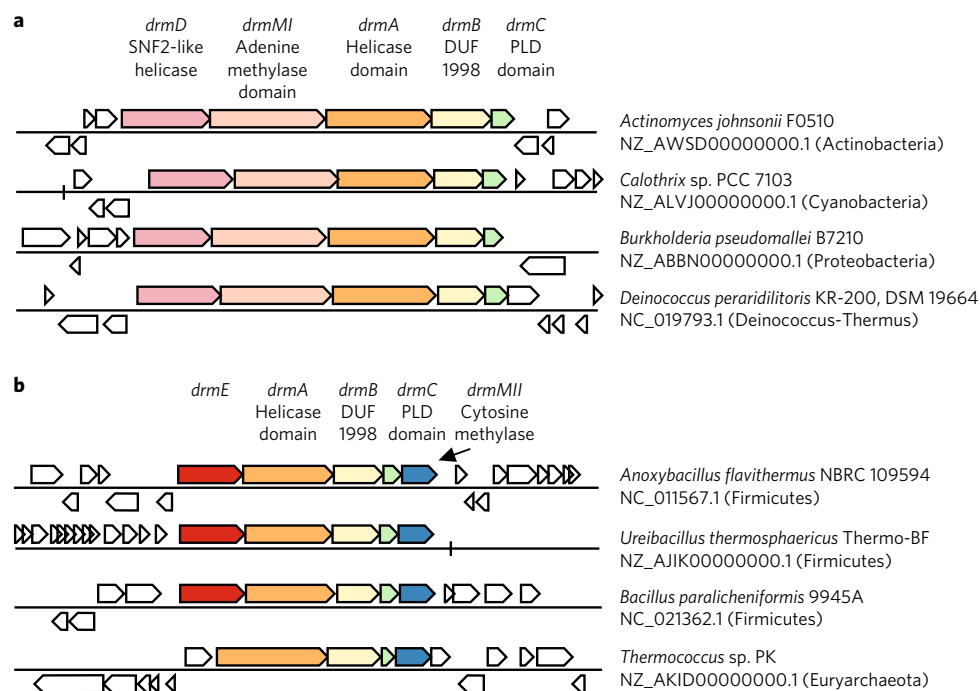
**An abundant 5-gene cassette represents a putative new defence system.** We then characterized the genomic neighbourhood of this three-gene core and found that it was almost always (324/351 of cases, 92%) associated with a gene containing a DNA methyltransferase domain, marking it as a possible new type of R/M system. We therefore named this system DISARM ('defence island system associated with restriction modification'). In most cases, the core gene triplet was adjacent to a DNA adenine N6 methyltransferase gene (pfam13659) that is usually annotated in the restriction enzymes database REBASE<sup>21</sup> as a putative type IIG R/M gene. In these cases, which we define as Class 1 DISARM, the system is comprised of the core triplet, the methyltransferase, and a fifth gene annotated as COG0553 SNF2 family helicase, containing SNF2-like ATPase (pfam00176) and helicase C-terminal (pfam00271) domains<sup>22</sup>. Gene cassettes containing the Class 1 DISARM system were found in 11 bacterial and 3 archaeal phyla (Fig. 1a, Table 1 and Supplementary Table 2).

**Table 1 | DISARM genes and their domain annotations**

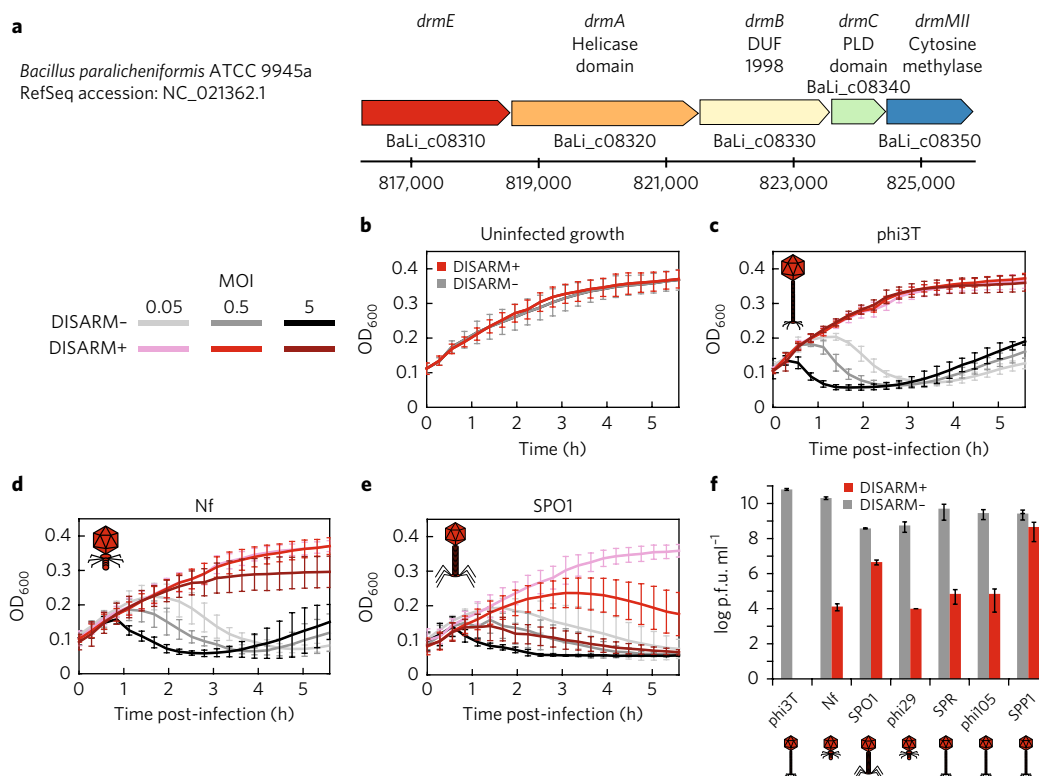
Gene	Signature domain	Putative function
<i>drmA</i>	pfam00271	Putative helicase domain
<i>drmB</i>	pfam09369	DUF1998, helicase-associated domain
<i>drmC</i>	pfam13091	Phospholipase D/nuclease domain
<i>drmD</i>	pfam00176	SNF2 family helicase
<i>drmE</i>	None	Unknown function
<i>drmMI</i>	pfam13659	N6-adenine DNA methyltransferase
<i>drmMII</i>	pfam00145	5-cytosine DNA methyltransferase

A second subset of systems did not contain the SNF2 helicase and the type IIG R/M enzyme pair mentioned above, but instead contained a DNA 5-cytosine methyltransferase (pfam00145). We refer to these systems as Class 2 DISARM. These systems are mostly found in extremophilic bacteria and archaea and in Firmicutes, especially Bacilli. Unlike the larger Class 1 systems, these Class 2 systems are more compact. In most Bacilli and halophilic archaea, the systems contain another protein of unknown function sized ~800 aa (Fig. 1b).

**DISARM confers protection against multiple phage types.** We selected a Class 2 DISARM system for experimental validation, as this class is more compact (spanning an average DNA size of 9.6 kb versus 16.5 kb spanned by Class 1 systems; Supplementary Table 2), facilitating easier engineering into a heterologous genome. To test whether the hypothesized DISARM system provides protection against phage infection, we cloned the DISARM locus of *B. paralicheniformis* 9945a, including the upstream and downstream intergenic regions, into the *Bacillus subtilis* BEST7003 genome (Fig. 2a).



**Fig. 1 | Two common classes of DISARM systems occur in bacteria and archaea. a**, Class 1 DISARM systems are composed of the core gene triplet of *drmA*, a gene with a helicase domain (orange); *drmB*, DUF1998 domain-containing gene (yellow); and *drmC*, containing a PLD domain (green). This core gene triplet is preceded by *drmD*, an SNF2-like helicase (pink), and *drmMI*, an adenine methylase (peach). RefSeq genome accessions are indicated. System positions are provided in Supplementary Table 2. **b**, Class 2 DISARM systems contain, in addition to the core triplet of *drmABC*, also *drmMII*, a cytosine methylase (blue). In Bacilli and some haloarchaea the systems also include an ~800 aa gene of unknown function, here named *drmE* (red).



**Fig. 2 | DISARM provides protection against phages.** **a**, The DISARM locus of *B. paralicheniformis* ATCC 9945a. Numbers below the axis represent the position on the *B. paralicheniformis* genome. Locus tags are provided for each gene. **b**, Insertion of the DISARM locus into the *B. subtilis* BEST7003 genome does not impair growth. Curves show the mean of two independent experiments with three technical repeats each. Error bars are 95% confidence interval of the mean. **c–e**, DISARM provides protection against phi3T (**c**), Nf (**d**) and SPO1 (**e**) phages. Bacteria were infected at  $t = 0$  at MOI of 0.05, 0.5 and 5. Curves show the mean of two independent experiments with three technical repeats each. Error bars are 95% confidence interval of the mean. **f**, Plaque formation of seven phages on DISARM-containing strains. The y axis represents concentration of p.f.u. Shown are the means of three replicates, and error bars are s.d. of the mean. Grey bars represent efficiency of plating (EOP) on DISARM– cells and red bars are EOP in DISARM+ cells.

We verified in advance that *B. subtilis* BEST7003 does not contain a DISARM system of its own by searching for homologues of each of the DISARM genes in its genome—none was found. The correct insertion of the system into the *B. subtilis* genome was verified by whole-genome sequencing (Supplementary Fig. 1). No change in growth dynamics was observed in the DISARM-containing bacteria compared to the control strain transformed with an empty plasmid (Fig. 2b).

We then challenged the DISARM-containing *B. subtilis* with phages from all three morphological families of the Caudovirales: the Siphophages phi3T, SPR, SPP1 and phi105, the Myophage SPO1 and the Podophages Nf and phi29. SPO1, SPP1, phi29 and Nf are obligatory lytic, while phi3T, SPR and phi105 are temperate phages. Infections were performed at three orders of magnitude of multiplicity of infection (MOI): 0.05, 0.5 and 5 phages per bacterium. Our results show that DISARM provided anti-phage protection against all phages, manifested by the delay or absence of culture collapse upon infection with phage (Fig. 2c–e and Supplementary Fig. 2). To quantify the protection, we measured the phage efficiency of plating (EOP) on DISARM-containing bacteria versus control cells. DISARM provided strong protection against most of the phages, with up to seven orders of magnitude of protection observed. For two of the phages we observed intermediate levels of DISARM defence, with two orders of magnitude against SPO1 and one order of magnitude against SPP1 (Fig. 2f).

To test whether the partial protection observed for some of the phages is due to a heritable trait such as resistance mutations or epigenetic modification in a subpopulation of the infecting phages, we isolated Nf and SPO1 phages from single plaques that appeared

on DISARM+ cells. These isolated phages did not show increased resistance (measured via plaque assays) against DISARM+ cells as compared to their ancestor phages (Supplementary Fig. 3), suggesting that for these phages escape from DISARM is not due to genetic or epigenetic traits. Instead, a small proportion of these phages seem naturally to be able to propagate inside at least a fraction of the DISARM+ cell population.

One of the known phage defence paradigms is abortive infection, in which bacteria commit suicide upon infection, thus preventing completion of the phage replication cycle<sup>6</sup>. This prevents the release of phage progeny and the spread of infection to neighbouring cells. In this scenario, infection with MOI > 1 is expected to cause death or bacteriostasis of a large fraction of cells in the culture and an immediate stasis or reduction of optical density (OD) upon infection. As DISARM-containing cultures infected with MOI = 5 did not stop growing or collapse upon infection (Fig. 2c), we infer that DISARM does not provide protection through an abortive infection mechanism.

We also performed transformation efficiency experiments using an episomal *Bacillus* plasmid. We could not detect a reduced transformation efficiency in DISARM+ cells, suggesting that DISARM does not interfere with DNA import of this plasmid through the natural competence system of *B. subtilis* (Supplementary Fig. 4).

**DISARM allows phage adsorption but prevents phage replication.** To test if DISARM protects the bacteria by preventing phage attachment, we compared the rate of phage adsorption to DISARM-containing versus DISARM-lacking bacteria. No significant difference was observed in the adsorption rates, indicating that

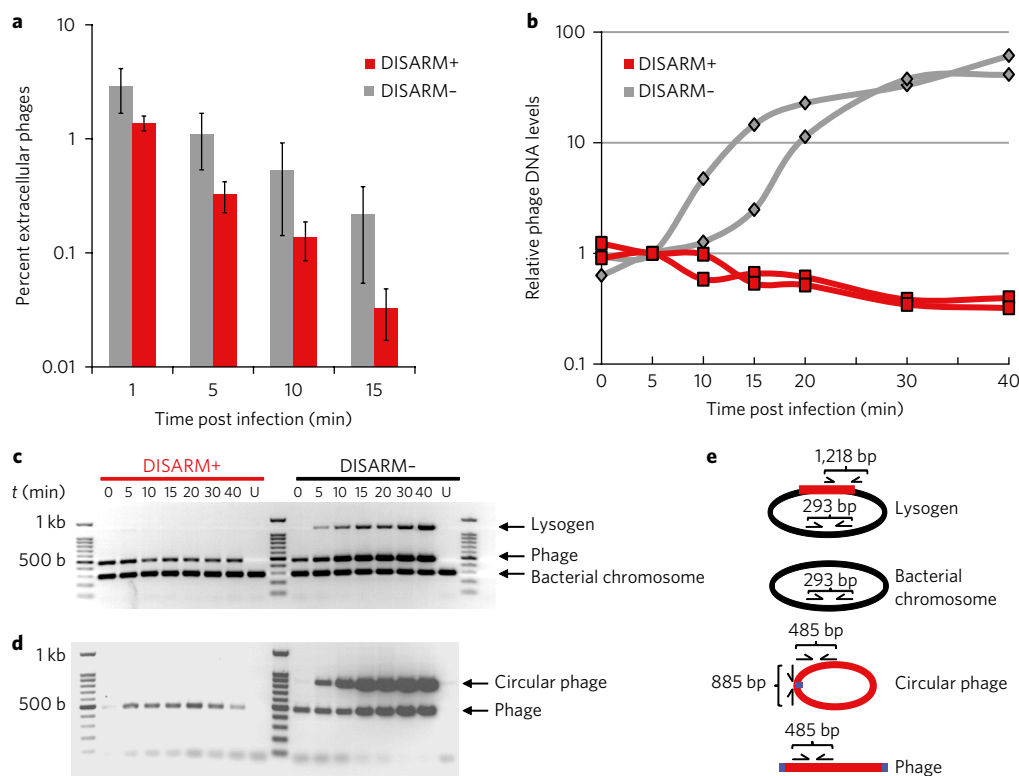
DISARM provides protection without hampering phage adsorption (Fig. 3a). We also tested if the phage genome is replicated in DISARM-containing cells. We used Illumina sequencing to quantify the amount of phage DNA in comparison to bacterial DNA during infection with phi3T at MOI = 1. As the bacterial genome is not degraded during phi3T infection<sup>10</sup>, the ratio between phage reads and bacterial reads can be used to quantify the number of phage genome equivalents per infected cell. The results show that although in control cells the phage DNA replicates during the infection process, in DISARM-containing cells it is not replicated but is rather depleted over time in comparison to the bacterial genome (Fig. 3b). In addition, phi3T was not able to circularize its genome or form detectable lysogens in DISARM-containing cells (Fig. 3c,d). This indicates that DISARM prevents phage DNA replication and lysogeny, and probably also causes phage DNA degradation. Moreover, as DNA circularization occurs soon after injection and is essential for both lytic and lysogenic cycles<sup>23</sup>, our results indicate that DISARM stops the infection at a very early stage.

To further examine the dynamics of phage DNA decay in DISARM-containing cells, we used a previously established system that allows imaging of phage SPP1 infection<sup>24,25</sup>. In this system, the SPP1 genome is modified to include a *lacO* array and the infected cells express a LacI-CFP fusion protein. Upon phage DNA injection to infected cells, LacI-CFP proteins bind to the phage *lacO* array, resulting in a clear focus. Foci were clearly observed on wild-type (WT), DISARM-lacking cells and, once established, the foci grew in

size during phage DNA replication (Fig. 4). Foci were also observed in infected DISARM-containing cells, but these foci did not expand and rapidly disappeared during the time course of infection (Fig. 4). These results further substantiate that DISARM does not block phage DNA injection into the infected cell, but causes intracellular phage DNA decay.

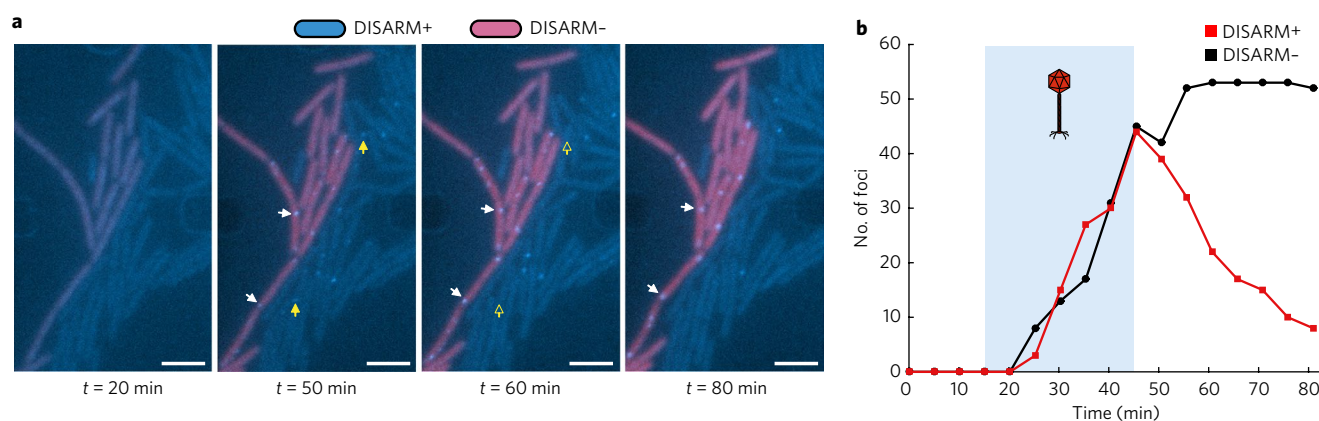
**Essential components for DISARM anti-phage activity.** To map the essential components of the DISARM system, we engineered a series of *B. subtilis* BEST7003 strains, each containing a DISARM system with a scarless deletion of each of the DISARM genes (Fig. 5). These deletions were verified by whole-genome sequencing (Supplementary Fig. 1). No growth impairment was observed in the deletion strains of *drmE*, *drmA*, *drmb* and *drmc* as compared to control cells or cells containing the full DISARM system (Supplementary Fig. 5). We were not able to obtain a single-gene deletion for *drmmII* (see below).

We then tested the deletion strains by infecting each of them with phi3T, SPO1 and Nf phages. Deletions of *drmA* (helicase domain), *drmb* (DUF1998 domain) and *drmE* (unknown function) abolished DISARM protection against all phages tested, indicating that each of these three genes is essential for DISARM activity (Fig. 5a–c and Supplementary Fig. 6). In contrast, deletion of the PLD domain gene, *drmc*, had no effect on DISARM protection against phi3T, and the  $\Delta drmc$  cells remained fully protected against it (Fig. 5d and Supplementary Fig. 7a–c). However, a reduction in protection



**Fig. 3 | Phage phi3T adsorption and DNA replication in DISARM-containing cells.** **a**, Adsorption of phages to DISARM-containing cells (red) is not impaired compared to control cells (grey). After infection of logarithmic-stage cultures ( $OD_{600}=0.3$ ) with phi3T at MOI = 1, samples were taken at 5 min intervals and the extracellular (unadsorbed) phage concentration was measured and compared to the initial phage concentrations (see Methods). Bars represent means of three experiments, and error bars are s.e.m. **b**, Ratio of phage DNA to bacterial DNA during infection. Total DNA was extracted from infected bacteria (MOI = 1) at the indicated time points and sequenced using an Illumina sequencer. The y axis represents relative phage DNA concentrations compared to bacterial genome equivalents, normalized to the value at  $t=5$  min post-infection. Each curve represents an independent repeat of the experiment. **c**, DISARM prevents lysogeny of phi3T. Agarose gel of multiplex PCR with three primer sets, aiming to detect bacterial DNA, phage DNA and lysogen. Lanes are marked with minutes post-infection. U: uninfected control. **d**, DISARM prevents phage circularization. Outward-facing primers at the edges of the phi3T genome were used to detect phage genome circularization. **e**, Schematic representation of fragments amplified in **d**.





**Fig. 4 | Fluorescence microscopy of phage DNA in DISARM<sup>−</sup> and DISARM<sup>+</sup> cells.** **a**, DISARM-lacking, RFP-expressing cells (red cells) were co-incubated with DISARM-containing (light blue) cells in a microfluidic device that allows visualization of a single bacterial layer (see Methods). Both strains express LacI-CFP constitutively. SPP1 phages containing a LacO array ( $10^5$  p.f.u.  $\mu\text{l}^{-1}$ ) were flowed into the device from  $t = 15$  min to  $t = 45$  min and an image was taken every 5 min. Upon injection of phage DNA, a fluorescent focus of LacI-CFP is formed on the LacO array in the phage DNA. White arrows show foci in DISARM-lacking cells, which do not disappear, and grow in size through the time course. Foci on DISARM-containing cells appear (filled yellow arrows) but later disappear (open yellow arrows). Scale bars, 5  $\mu\text{m}$ . Results are representative of two independent experiments. **b**, Quantification of phage foci over time in the microscopy field, of which a subsection is shown in **a**. Phage foci appear, starting at  $t = 25$  min, and become established in DISARM-lacking cells. Similar numbers of foci appear in DISARM-containing cells, but these foci disappear over time. The shaded area represents the time where phages were continuously flowed in.

efficiency against SPO1 and Nf phages was observed, such that  $\Delta\text{drmC}$  cells were less resistant than cells containing the full DISARM system, but still more resistant than DISARM-lacking cells (Fig. 5e,f). Quantification via plaque assays further showed that the  $\Delta\text{drmC}$  reduced DISARM defence against Nf by two to three orders of magnitude (Supplementary Fig. 7). Variable protection in  $\Delta\text{drmC}$  cells was also observed in other phages (Supplementary Fig. 7). Therefore, *drmC* appears to be redundant for defence against some phages, but is required for defence against others (see Discussion).

As DISARM contains a predicted C5 cytosine-specific DNA methyltransferase (*drmMII*) we performed whole-genome bisulfite sequencing to look for its cognate methylation sequence motif. In control *B. subtilis* no significant motif for 5-methylcytosine (5mC) was identified, but in DISARM-containing cells the motif CCWGG (W = A or T) was methylated in the underlined cytosine (average methylation rate of 82% for sites covered by >5 reads). Bisulfite sequencing further validated that the same motif is methylated in *B. paralicheniformis* ATCC 9445a, in which the DISARM locus resides naturally. The same methylation motif was found in the genome of *B. subtilis* subsp. *spizizenii* str. W23 (ref. <sup>26</sup>), which also contains a similar Class 2 DISARM system. In that strain, this motif was attributed to the homologue of the DISARM methyltransferase (named M.BsuW23II), to which no cognate restriction enzyme was suggested<sup>26</sup>.

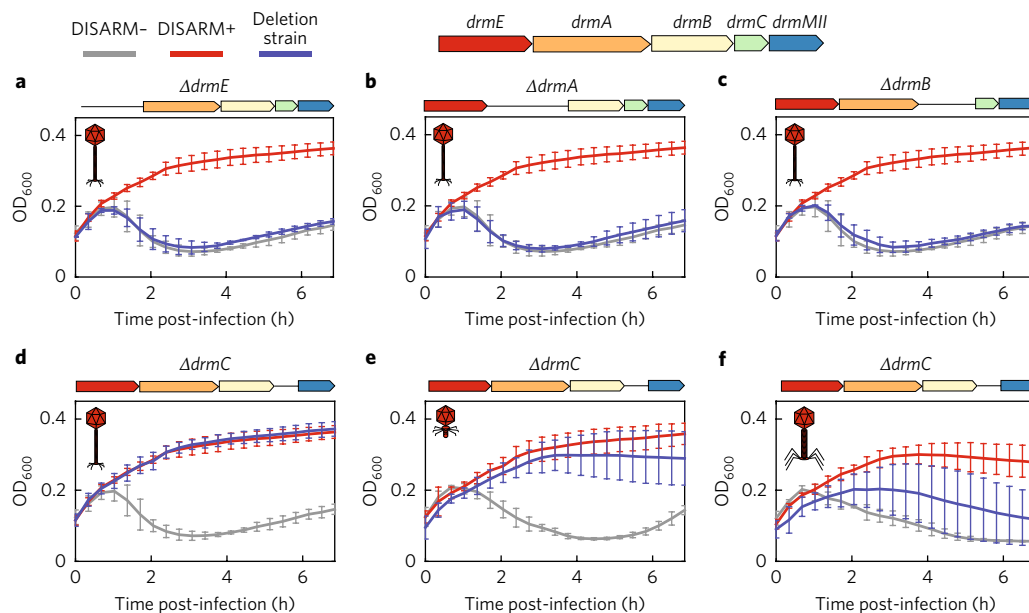
The above results suggest that *drmMII* methylates the DNA at CCWGG motifs. Presumably, other components of the DISARM system use non-methylated CCWGG motifs as a marker of foreign DNA akin to other known R/M systems. Consistent with this hypothesis, attempts to clone a *drmMII*-deleted DISARM system into *B. subtilis* yielded very low transformation efficiency (Supplementary Fig. 8). Whole-genome sequencing of three of the resulting transformed colonies showed massive deletions or frame-shift mutations in the DISARM locus in addition to the intended deletion of *drmMII*. This suggests that in the absence of *drmMII* the DISARM system is toxic to the cells and only cells with a defective DISARM locus can survive. It is likely that in the absence of CCWGG methylation in the bacterial chromosome, the restriction components in the DISARM system attack the chromosome, leading to the observed toxicity.

To further examine whether *drmMII* alone is sufficient for DNA methylation, we cloned this gene under a Pveg constitutive promoter in a WT *B. subtilis*. Bisulfite sequencing validated that the *drmMII*-expressing strain is methylated on 97.7% of the reads mapping to CCWGG motifs, validating *drmMII* as the system's methylase.

To test whether CCWGG methylation is sufficient to protect phages against DISARM interference, we propagated phi3T on the *drmMII*-expressing strain, yielding methylated phages. Bisulfite sequencing verified that 75 of the 78 CCWGG sites in phi3T became methylated after propagation in the methylase-expressing cells (Supplementary Fig. 9a). The three sites that were not modified overlapped with GGCC sites that are known to be methylated by the native methylase M.Phi3TI encoded by phi3T (the same methylase also methylates GCNGC sites<sup>27</sup>, while an additional methylase, M.Phi3TII, modifies TCGA motifs<sup>28</sup>). The DISARM<sup>+</sup> strain still protected against the modified phages despite their high level of methylation (Supplementary Fig. 9b). Moreover, DISARM also provided a high level of protection against phage Nf (Fig. 2d), although the genome of this phage (GenBank accession no. EU622808) is devoid of any CCWGG site, which we also verified by whole-genome sequencing of the Nf phage we used. These results suggest that DISARM probably uses an additional, yet unknown mechanism to identify invading phage DNA in addition to the methylation signature.

## Discussion

Our results establish DISARM as a new defence system, providing protection against diverse phages. The DISARM system is widespread in defence islands across the microbial world and contains three core genes (*drmA*, *drmB* and *drmC*) accompanied by a methyltransferase (*drmMI* or *drmMII*) and an additional gene (*drmD* or *drmE*). The existence of an R/M-related active methyltransferase, the toxicity caused by its deletion, and the depletion of phage DNA during infection (Figs. 3b and 4) suggest that DISARM represents a new composition of R/M system that differs from other known such systems. The modification module of DISARM is composed of a methyltransferase as in classic R/M systems, but the restriction module seems to be unique and requires multiple components. The interaction between these modules and the exact mechanism



**Fig. 5 | Deletion of DISARM components. a–c,** Deletion of *drmE*, *drmA* or *drmB* abolished DISARM defence against phi3T. **d,** Deletion of *drmC* has no effect on the defence against phi3T. **e,f,** Deletion of *drmC* reduces DISARM protection against Nf and SPO1, but the deletion strains are still protected compared to the control bacteria. Curves in **a–f** show means of two independent experiments with two technical repeats each, and error bars represent 95% confidence intervals of the mean. Infections were performed at MOI = 0.5.

through which DISARM restricts phage replication remain to be characterized. Moreover, because phage Nf does not contain a single CCWGG motif and is still restricted by DISARM, our results imply that the restriction module of DISARM may not rely solely on the presence of these motifs.

Although DISARM was able to strongly protect against most of the phages tested, for two of the phages (SPO1 and SPP1) protection only amounted to one to two orders of magnitude. Phage SPO1 is known to have a heavily modified genome, containing hydroxymethyluracil (hmU) in place of thymine in its genome<sup>29</sup>. As CCWGG sites must contain the modified base on one of the strands (W = A or T), it is possible that the modified base somehow partially interferes with DISARM restriction. SPP1, which also partially escapes DISARM protection, has no reported DNA modifications, but only contains one CCWGG motif. It is possible that the lower level of DISARM protection against this phage can be attributed to the single motif, although phi3T modified in all but three CCWGG motifs was unable to overcome DISARM (Supplementary Fig. 9), and so did phage Nf, which completely lacks this motif in its genome.

We have shown that four of the five genes comprising DISARM are absolutely essential for its activity in phage resistance. The fifth gene, *drmC*, is partially required for defence against SPO1 and Nf, but is redundant against phi3T. The *drmC* gene has a PLD domain, a domain that was previously shown to be the catalytic nuclease domain in R/M systems such as BfiI and NgoAVII<sup>19,20</sup>. However, due to its redundancy in defence against most of the tested phages, *drmC* is unlikely to function as the core restriction endonuclease of the system and is more likely to take an auxiliary role. Indeed, while *drmA* is almost always associated with the DUF1998-containing *drmB* gene, *drmC* appears associated with this pair in only one-third of cases.

In this study we have focused on the most prevalent genomic context of standalone DUF1998-containing genes (*drmB*) and showed that such domains preferentially occur in the context of DISARM systems. In DISARM, the DUF1998-containing *drmB* is found immediately downstream to *drmA*, a putative DEXX-helicase (pfam00271). Interestingly, a fusion protein between DEXX-helicase

and DUF1998, named *dpdJ*, was recently demonstrated to be involved in the Dpd system, a new R/M system comprising at least 12 genes that modifies the bacterial DNA with 7-deazaguanine derivatives and possibly restricts unmodified DNA<sup>30</sup>. In the Dpd system, the abovementioned fusion gene *dpdJ* is followed by a PLD-containing gene named *dpdK*. The roles of these genes in the Dpd systems are unknown, but in light of our results it is possible that they take part in the Dpd restriction module, similarly to the putative roles of *drmABC* in DISARM.

In line with the hypothesis that the DISARM components can interact with different modules of other R/M systems, it is worth noting that DISARM systems are often genomically associated with other methyltransferases and putative restriction endonucleases (Supplementary Fig. 10). For example, in many cases Class 1 systems are accompanied by an additional predicted cytosine methylase (pfam00145) (Supplementary Fig. 10a,b). DISARM systems are also frequently associated with Res and Mod genes of a type III R/M system (Supplementary Fig. 10a). In other cases, both a cytosine methylase and an HNH endonuclease (pfam13391) accompany Class 1 DISARM systems (Supplementary Fig. 10). The association with multiple other restriction systems may reflect the general tendency of defence systems to cluster in defence islands<sup>13</sup> or, alternatively, might suggest that the function of the core DISARM genes can be combined with additional R/M modules and provide a synergistic defensive advantage.

Bacterial and archaeal systems that provide defence against foreign genetic elements have demonstrated time and again their profound diversity and rapid evolution. Despite over four decades of extensive research, it seems that the full spectrum of R/M systems in nature is yet to be completely documented. The discovery of DISARM, most probably representing a new kind of R/M system, exemplifies this concept and suggests that additional new systems will be revealed in the future. A better understanding of the arsenal of defensive measures at the disposal of bacteria and archaea will bring us closer to understanding their arms race against their parasites, a major evolutionary driving force shaping prokaryotic genomes.

## Methods

**Genomic identification and analysis of DISARM systems.** The IMG database<sup>14</sup> was searched on June 2016 for genes containing the pfam09369 (DUF1998) domain with no additional domains. The pfam annotations of the neighbouring genes of these DUF1998 standalone genes were retrieved from IMG and examined to identify the most common neighbourhood. DUF1998 genes with a pfam00271 gene upstream and a pfam13091 gene downstream were collected as core DISARM genes and further analysed. DUF1998 genes with only a pfam00271 gene upstream were manually screened to identify possible misannotation of a downstream pfam13091 gene. The pfam00271-, DUF1998- and pfam13091-containing genes were termed *drmA*, *drmB* and *drmC*, respectively.

To study the genomic neighbourhood of the identified core DISARM systems, the pfam and COG annotations of 30 genes upstream and downstream of the DUF1998 gene were retrieved from IMG and manually inspected. This led to the discovery of the additional related DISARM genes. Class 1 systems were defined according to the presence of pfam00176 and pfam13659 up to 30 genes upstream or downstream of the DUF1998 gene. Systems that contained a pfam00176 gene and type III R/M genes (methylase with pfam01555 and restriction nuclease with pfam04851) were defined as Class 1, based on the assumption that the R/M system methylase replaces *drmMI*. Class 2 systems were defined according to the lack of pfam13659 and pfam00176 and the presence of the pfam00145 gene. In Class 1 systems, the closest gene with pfam13659 relative to the DUF1998 gene was defined as *drmMI* and the closest gene with pfam00176 relative to the DUF1998 gene was defined as *drmD*. In Class 2 systems, the closest gene with pfam00145 was defined as *drmMII*, and *drmE* genes were manually curated due to their lack of any annotated domain. As pfam13659 was recently cancelled and deleted from the Pfam database, some of the *drmMI* genes remained with no pfam annotation, but could be identified by their Superfamily SSF5335 (S-adenosyl-L-methionine-dependent methyltransferases) annotation and their size and so were manually curated as *drmMI*.

**Cloning of DISARM into *B. subtilis* BEST7003.** A cloning vector for large fragments was constructed by assembling the p15a origin of replication (*ori*) from pACYCDuet-1 and the *amyE* integration cassette from plasmid pDR110 (provided by I. Kolodkin-Gal). The p15a *ori*<sup>31</sup> was amplified using primers OGO174 + OGO175 (a list of all primers is provided in Supplementary Table 3). The *amyE* integration cassette was amplified using primers OGO176 + OGO185. The two fragments were assembled and transformed into *Escherichia coli* cells using a Gibson assembly cloning kit (NEB E5510S), and assembled plasmids were verified by restriction pattern and full sequencing.

*Bacillus licheniformis* (Weigmann) Chester ATCC 9945a was obtained from ATCC. The species designation for this strain was recently changed to *B. paralicheniformis*<sup>32</sup> with NCBI taxonomy ID 766760. The DISARM locus of *B. paralicheniformis* 9945a in coordinates 815,730–826,377 (RefSeq NC\_021362.1) was amplified using primers Hezi\_1\_F and Hezi\_2\_R. The vector backbone was amplified using primers OGO207 + OGO208 and the two fragments were assembled using Gibson assembly.

*B. subtilis* BEST7003 cells were previously provided by M. Itaya. Assembled plasmids were transformed into *B. subtilis* BEST7003 cells as described by Itaya<sup>33</sup>.

Scarless deletion strains were constructed by amplification of the DISARM system in two fragments, omitting the desired deletion region, and Gibson assembly with the vector backbone. The vector backbone was generated by primers OGO207 + OGO208. The following PCR fragments were used to generate deletion systems:  $\Delta\text{drmE}$  – OSM13 + SM3, SM4 + OGO175;  $\Delta\text{drmA}$  – OSM13 + SM5, SM6 + OGO175;  $\Delta\text{drmB}$  – OSM13 + SM2, SM7 + OGO175;  $\Delta\text{drmC}$  – OSM13 + SM9, SM10 + OGO175;  $\Delta\text{drmMII}$  – OSM13 + SM11, SM1 + OGO175. The constructed plasmids were then used for integration of the deletion-containing system into *B. subtilis* BEST7003. Deletion of each gene included the open reading frame (ORF) only, without damaging intergenic regions. Deleted regions were as follows (coordinates on RefSeq NC\_021362.1):  $\Delta\text{drmE}$  – 816,274–818,674;  $\Delta\text{drmA}$  – 818,671–821,971;  $\Delta\text{drmB}$  – 822,039–823,752;  $\Delta\text{drmC}$  – 823,776–824,487;  $\Delta\text{drmMII}$  – 824,499–825,847. The entire genome of each constructed strain was verified by Illumina whole-genome sequencing (Supplementary Fig. 1). Sequence analysis for strain verification was performed using *breseq*<sup>34</sup>. A control strain containing an empty integration cassette was constructed in parallel, sequenced, and used as a control in the following experiments.

For construction of the strain expressing *drmMII*, the ORF of *drmMII* was amplified from the genomic DNA of DISARM-containing *B. subtilis* using primers OGO425 + OGO426. The backbone of pJMP4 (provided by J. M. Peters) was amplified using primers OGO423 + OGO424 and the two fragments were assembled using Gibson assembly so that the *drmMII* gene is under the control of the plasmid's Pveg constitutive promoter.

**Phage cultivation.** Phages were previously received from the Bacillus Genetic Stock Center (BGSC). BGSC IDs for the phages used are 1L1 for phi3T, 1P4 for SPO1, 1L56 for SPR, 1L11 for phi105, 1P7 for SPP1 and 1P19 for Nf (1P19 is listed in the BGSC catalogue as phage phi29, but upon sequencing and assembly this phage was found to be 100% identical to the reference sequence of phage Nf, GenBank accession no. EU622808). Phage phi29 was received from DSMZ (DSM 5545).

Phages were propagated on *B. subtilis* BEST7003 (provided by M. Itaya) using the plate lysate method as described by Fortier and Moineau<sup>35</sup>. Lysate titre was determined using the small drop plaque assay method, as described by Mazzocco and co-authors<sup>36</sup>.

**Phage infection growth curves.** Overnight cultures of bacteria were diluted 1:100 into Magnesium-Manganese broth (MMB) (Lysogeny Broth (LB) + 0.1 mM MnCl<sub>2</sub> + 5 mM MgCl<sub>2</sub>). Then, 200 µl volumes of the diluted culture were dispensed into the wells of 96-well plates and grown at 37 °C with shaking for 1 h until early log phase. The number of bacterial cells in the culture was calculated according to an OD<sub>600</sub> to colony-forming unit (c.f.u.) calibration curve. Then, 20 µl phage lysate was added to the desired MOI and the growth was followed in a TECAN Infinite 200 plate reader with OD<sub>600</sub> measurement every 15 min at 37 °C with shaking.

**Adsorption assay.** Volumes (15 ml) of mid-log bacterial cultures in MMB medium at an OD<sub>600</sub> of 0.3 were infected with phage at an MOI of 1. During infection, the culture was incubated with shaking at 37 °C. At time points 1, 5, 10, 15, 20, 30 and 40 min post-infection, 0.5 ml samples were taken and mixed with 100 µl ice-cold chloroform. Samples were vortexed, incubated at 37 °C for 5 min, vortexed, incubated on ice for 5 min, vortexed again, and incubated at room temperature for 40 min. Samples were then briefly centrifuged and the phage concentration in the upper aqueous phase was determined by double-layer plaque assay using *B. subtilis* BEST7003 as an indicator strain. As a control, the same amount of phage lysate was mixed with 15 ml MMB without bacteria and a sample was processed through the same stages and measured by double-layer plaque assay to determine the reference phage concentration.

**Escape phages isolation and testing.** Overnight cultures of DISARM-containing and DISARM-lacking cells were diluted 1:100 and grown to an OD of 0.3. A 100 µl volume of the culture was mixed with 100 µl of phage lysate and incubated at room temperature for 5 min. Molten top agar (MMB + 0.5% agar, 4 ml) was added, vortexed, and poured over an MMB Petri dish. Plates were incubated at room temperature overnight. Isolated plaques from the DISARM-containing plates were picked into 100 µl phage buffer (50 mM Tris pH 7.4, 100 mM MgCl<sub>2</sub>, 10 mM NaCl). Serial dilutions in MMB were performed and the phages were plated using the small drop plaque assay on DISARM-containing and DISARM-lacking cells.

**DNA extraction during infection (used for sequencing, PCR for lysogeny and phage circularization detection).** Mid-log bacterial culture in MMB (50 ml) was infected with phi3T at an MOI of 1, and 5 ml samples were taken immediately after infection ( $t = 0$ ) and at 5, 10, 15, 20, 30 and 40 min after infection. During infection, the culture was incubated with shaking at 37 °C. An uninfected control sample was taken before addition of phage. Samples were immediately transferred to ice. Samples were then centrifuged and the pellet was washed three times in ice-cold Tris pH 7.4 buffer to remove unabsorbed phages. The washed pellets were frozen in liquid nitrogen. Total DNA was extracted using a Qiagen DNeasy Blood and Tissue kit (Qiagen 69504). Detection of phage lysogeny was performed using multiplex PCR as previously described by Goldfarb and co-authors<sup>10</sup>. Phage genome was detected using primers PTG83 + PTG84, bacterial genome was detected using primers PTG18 + PTG29, and the lysogeny junction was detected using primers PTG125 + PTG126. Detection of phage circularization was performed using primers PTG115 + PTG116. To determine the relative abundance of bacterial and phage DNA, Illumina libraries were prepared and sequenced using a modified Nextera protocol as described by Baym and co-authors<sup>37</sup>. Reads were aligned to the bacterial reference genome and the phi3T genome (GenBank accession no. KY030782) as previously described in ref. <sup>10</sup>. The numbers of reads aligned to the phage and host genomes at each time point were normalized to the genome sizes to calculate the number of phage genome equivalents per bacterial genome.  $T = 5$  min was used as a reference point for comparison of phage DNA levels, representing a time point until which phage adsorption continued but no phage replication initiated.

**Bisulfite sequencing.** Genomic DNA of DISARM-containing *B. subtilis* BEST7003, control *B. subtilis* BEST7003, the constitutive *drmMII* strain and *B. paralicheniformis* 9945a, as well as genomic DNA of phage phi3T, was used to construct PBAT libraries, using a modified version of the published protocol<sup>38</sup>. Briefly, 50 ng of genomic DNA was converted and purified according to the manufacturer's instructions (EZ DNA methylation lightning MagPrep, Zymo Research), using half of the recommended amount of each reagent. Bisulfite-converted products were subjected to second strand synthesis by Klenow fragment 3' to 5' exo- (10 units, M0212L, NEB) and the indexed random nonamer primer (0.8 µM)

5'ACACTCTTCCCTACACGACGCTCTCCGATCT-INDEX-GGNNNNNNNNN3'.

This primer includes the truncated Illumina P5 adaptor followed by 8 bp internal index. The excess of primer was removed at the end of the reaction by exonuclease I (M0293L, NEB) and the products were purified with 0.8x beads (Agencount Ampure XP beads, Beckman Coulter). DNA was denatured for 6 min



at 95°C and the second strand was synthesized by Klenow polymerase using the indexed random nonamer primer (0.8 µM) containing the P7 Illumina adaptor 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-INDEX-CCNNNNNNNN3'.

The products were purified with 0.8× beads and the library was generated by 12 cycles of PCR amplification using 2.5 units of GoTaq Hot Start polymerase (M5005, Promega) together with 0.4 µM Illumina Forward PE1.0 primer (5'-AATGATACGCGACACCAGGATCTCACTCTTCCCTACACGACGCTCTTCCGATCT-3') and 0.4 µM pre indexed Illumina Reverse primer (5'-CAAGCAGAAGACGGCATACGAGATXXXXXXGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3', where XXXXXX represents the barcode for multiplexing). Amplified libraries were purified with 0.7× Agencourt Ampure XP beads and were assessed by a Qubit dsDNA HS assay kit (Thermo Fisher Scientific) and bioanalyser (Agilent). The final quality-ensured libraries were pooled and sequenced on the NextSeq500 Illumina and generated 4–6 million reads per each library.

Before analysis, adaptor trimming and quality trimming were performed using Cutadapt<sup>40</sup>. Analysis of bisulfite-modified sequence reads was done using Bismark<sup>40</sup>. A whole-genome cytosine methylation report was generated and methylated positions were defined as positions with total coverage greater than 5 and methylation ratio greater than 2. The neighbourhood of the methylated positions was extracted from the reference genome and analysed for a recurrent motif. The positions of all CCWGG motifs in the genomes were extracted from the reference genomes and the methylation ratio of these positions was extracted from the cytosine methylation report.

**Fluorescence microscopy visualization of injected phage DNA.** *Strains.* The thrC::LacI-CFP cassette of strain ET3 (ref. <sup>41</sup>), provided by S. Ben-Yehuda, was amplified from the genomic DNA using primers OGO380 + OGO381 and transformed into the BEST7003 and DISARM-containing strains. The control cells were then transformed with a constitutive RFP construct (amyE::Pveg-RFP) using plasmid pJMP4, provided by J. M. Peters. Phage SPP1 containing an array of 64 repeats of LacO (refs. <sup>24,25</sup>) was provided by P. Tavares<sup>42</sup>.

**Fluorescence microscopy in a microfluidic device.** Overnight cultures of LacI-CFP DISARM-containing and RFP expressing control cells were diluted 1:100 and grown until an OD of 0.3. The culture was then diluted again 1:10, and equal amounts of both strains were mixed together. The mixed culture was loaded into the chamber of a CellASIC ONIX plate for bacterial cells (Mercury, B04A-03-5PK) according to the manufacturer instructions, then mounted on a Zeiss Axio Observer Z1 inverted microscope. The cells were grown under a constant flow of MMB medium and monitored periodically in bright-field and RFP channels to monitor cell division for ~1 h. Imaging then started and was performed in three channels: bright-field, CFP (filter set 47 HE) and RFP (filter set 64 HE). Images were captured every 5 min. After 15 min, SPP1-LacO phages in MMB (10<sup>5</sup> p.f.u. µl<sup>-1</sup>; p.f.u. are plaque-forming units) were flowed into the chamber for a period of 30 min. The infection was followed until the beginning of cell lysis of DISARM-lacking cells was observed.

**Image analysis.** Analysis was done using Imaris software (Bitplane). Background was subtracted from CFP and RFP channels. Fluorescent foci were segmented and tracked from the CFP channel using the Imaris spots object. Foci were allocated to DISARM-lacking cells according to RFP level at the same location. Segmentation and tracking were manually corrected. The number of foci within each strain was counted for each frame separately.

**Life Sciences Reporting Summary.** Further information on experimental design and reagents is available in the [Life Sciences Reporting Summary](#).

**Data availability.** The data that support the findings of this study are available in the Supplementary Tables and under accession [PRJEB22683](#) deposited to the European Nucleotide Archive (ENA).

Received: 22 November 2016; Accepted: 29 September 2017;  
Published online: 30 October 2017

## References

- Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).
- Stern, A. & Sorek, R. The phage–host arms race: shaping the evolution of microbes. *Bioessays* **33**, 43–51 (2011).
- Dy, R. L., Richter, C., Salmond, G. P. C. & Fineran, P. C. Remarkable mechanisms in microbes to resist phage infections. *Annu. Rev. Virol.* **1**, 307–331 (2014).
- Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
- Tock, M. R. & Dryden, D. T. F. The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.* **8**, 466–472 (2005).
- Dy, R. L., Przybiski, R., Semeijn, K., Salmond, G. P. C. & Fineran, P. C. A widespread bacteriophage abortive infection system functions through a type IV toxin–antitoxin mechanism. *Nucleic Acids Res.* **42**, 4590–4605 (2014).
- Sorek, R., Lawrence, C. M. & Wiedenheft, B. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* **82**, 237–266 (2013).
- Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
- Swarts, D. C. et al. DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261 (2014).
- Goldfarb, T. et al. BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* **34**, 169–183 (2015).
- Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* **41**, 4360–4377 (2013).
- Roberts, R. J. et al. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.* **31**, 1805–1812 (2003).
- 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).
- Markowitz, V. M. et al. IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res.* **40**, D115–D122 (2012).
- Caruthers, J. M. et al. Structure of the second domain of the *Bacillus subtilis* DEAD-box RNA helicase YxiN. *Acta Crystallogr. F* **62**, 1191–1195 (2006).
- 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).
- Caruthers, J. M. & McKay, D. B. Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* **12**, 123–133 (2002).
- Selvy, P. E., Lavieri, R. R., Lindsley, C. W. & Brown, H. A. Phospholipase D: enzymology, functionality, and chemical modulation. *Chem. Rev.* **111**, 6064–6119 (2011).
- Grazulis, S. et al. Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl Acad. Sci. USA* **102**, 15797–15802 (2005).
- Zaremba, M. et al. DNA cleavage by CgII and NgoAVII requires interaction between N- and R-proteins and extensive nucleotide hydrolysis. *Nucleic Acids Res.* **42**, 13887–13896 (2014).
- Roberts, R. J., Vincze, T., Posfai, J. & Macelis, D. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* **43**, D298–D299 (2015).
- Dürr, H., Flaus, A., Owen-Hughes, T. & Hopfner, K.-P. Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.* **34**, 4160–4167 (2006).
- Weigel, C. & Seitz, H. Bacteriophage replication modules. *FEMS Microbiol. Rev.* **30**, 321–381 (2006).
- Jakutyte, L. et al. Bacteriophage infection in rod-shaped Gram-positive bacteria: evidence for a preferential polar route for phage SPP1 entry in *Bacillus subtilis*. *J. Bacteriol.* **193**, 4893–4903 (2011).
- Jakutyte, L. et al. First steps of bacteriophage SPP1 entry into *Bacillus subtilis*. *Virology* **422**, 425–434 (2012).
- Blow, M. J. et al. The epigenomic landscape of prokaryotes. *PLoS Genet.* **12**, e1005854 (2016).
- Noyer-Weidner, M., Jentsch, S., Kupsch, J., Bergbauer, M. & Trautner, T. A. DNA methyltransferase genes of *Bacillus subtilis* phages: structural relatedness and gene expression. *Gene* **35**, 143–150 (1985).
- Noyer-Weidner, M., Walter, J., Terschüren, P. A., Chai, S. & Trautner, T. A. Mphi 3TII: a new monospecific DNA (cytosine-C5) methyltransferase with pronounced amino acid sequence similarity to a family of adenine-N6-DNA-methyltransferases. *Nucleic Acids Res.* **22**, 5517–5523 (1994).
- Stewart, C. R. et al. The genome of *Bacillus subtilis* bacteriophage SPO1. *J. Mol. Biol.* **388**, 48–70 (2009).
- Thiaville, J. J. et al. Novel genomic island modifies DNA with 7-deazaguanine derivatives. *Proc. Natl Acad. Sci. USA* **113**, E1452–E1459 (2016).
- Sathiamoorthy, S. & Shin, J. A. Boundaries of the origin of replication: creation of a pET-28a-derived vector with p15A copy control allowing compatible coexistence with pET vectors. *PLoS ONE* **7**, e47259 (2012).
- Dunlap, C. A., Kwon, S.-W., Rooney, A. P. & Kim, S.-J. *Bacillus paralicheniformis* sp. nov., isolated from fermented soybean paste. *Int. J. Syst. Evol. Microbiol.* **65**, 3487–3492 (2015).
- Itaya, M. Stable positional cloning of long continuous DNA in the *Bacillus subtilis* genome vector. *J. Biochem.* **134**, 513–519 (2003).
- 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).
- Fortier, L.-C. & Moineau, S. Phage production and maintenance of stocks, including expected stock lifetimes. *Methods Mol. Biol.* **501**, 203–219 (2009).



36. 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).
37. Baym, M. et al. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS ONE* **10**, e0128036 (2015).
38. Miura, F., Enomoto, Y., Dairiki, R. & Ito, T. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* **40**, e136 (2012).
39. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
40. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
41. Tzipilevich, E., Habusha, M. & Ben-Yehuda, S. Acquisition of phage sensitivity by bacteria through exchange of phage receptors. *Cell* **168**, 186–199 (2017).
42. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

## Acknowledgements

The authors thank M. Shamir, G. Amitai, S. Edelheit, A. Lopatina, T. Wein, Z. Erez and Z. Meir for discussions during the course of this study. The authors also thank A. Leavitt for discussions and assistance in adsorption assays and M. Keren for assistance with plaque assays. The authors thank I. Kolodkin-Gal for the pDR110 plasmid, J.M. Peters for the pJMP4 plasmid, S. Ben-Yehuda for the ET3 strain, P. Tavares for the *lacO*-containing SPP1 phage strain, T. Bucher, I. Pereman and E. Tzipilevich for their assistance with

microscopy and O. Golani from the Weizmann Life Sciences Core Facilities for assistance in image analyses. This study was supported, in part, by the Israel Science Foundation (personal grants 1303/12 and 1360/16 and I-CORE grant 1796/12), the European Research Council (grant ERC-CoG 681203), the Minerva Foundation and by a research grant from the David and Fela Shapell Family Foundation.

## Author contributions

G.O. and S.M. designed the experiments, performed the experiments and analysed the results. H.S. identified, bioinformatically, the DISARM system. S.D. and G.O. performed bioinformatics analysis of DISARM systems. Z.M. performed bisulfite sequencing. S.S. assisted with experiments. G.Y. and G.O. performed microscopy experiments. G.O. and R.S. wrote the manuscript. R.S. supervised the study.

## Competing interests

R.S. is a scientific founder of BiomX and a member of its scientific advisory board.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41564-017-0051-0>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to R.S.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

### ► Experimental design

#### 1. Sample size

Describe how sample size was determined.

No sample size calculations were performed.

#### 2. Data exclusions

Describe any data exclusions.

No data was excluded.

#### 3. Replication

Describe whether the experimental findings were reliably reproduced.

The experimental findings were successfully reproduced.

#### 4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No preallocation to experimental groups was performed.

#### 5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

No blinding was necessary.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

#### 6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

- | n/a                                 | Confirmed  |
|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The <u>exact sample size</u> ( <i>n</i> ) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement indicating how many times each experiment was replicated   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as an adjustment for multiple comparisons   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The test results (e.g. <i>P</i> values) given as exact values whenever possible and with confidence intervals noted  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A clear description of statistics including <u>central tendency</u> (e.g. median, mean) and <u>variation</u> (e.g. standard deviation, interquartile range)  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Clearly defined error bars   |

See the web collection on statistics for biologists for further resources and guidance.

## ► Software

Policy information about availability of computer code

### 7. Software

Describe the software used to analyze the data in this study.

Data was analyzed using Excel and Matlab

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

## ► Materials and reagents

Policy information about availability of materials

### 8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

All materials used are available from commercial vendors. Bacterial and phage strains are available from repositories as described in the Methods section.

### 9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used

### 10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used have been authenticated OR state that no eukaryotic cell lines were used.*

c. Report whether the cell lines were tested for mycoplasma contamination.

*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination OR state that no eukaryotic cell lines were used.*

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

*Provide a rationale for the use of commonly misidentified cell lines OR state that no commonly misidentified cell lines were used.*

## ► Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

### 11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

No animals were used.

Policy information about studies involving human research participants

### 12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

No human participants were examined.