Synergistic anti-phage activity of bacterial defence systems

Yi Wu[1,7], Sofya K. Garushyants[2,7], Anne van den Hurk[1], Cristian Aparicio-Maldonado[1], Simran Krishnakant Kushwaha[1,3], Claire M. King[1], Yaqing Ou[4], Thomas C. Todeschini[1], Martha R.J. Clokie[5], Andrew D. Millard[5], Yilmaz Emre Gençay[6], Eugene V. Koonin[2], Franklin L. Nobrega[1]*

[1] School of Biological Sciences, University of Southampton, SO17 1BJ Southampton, United Kingdom

[2] National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

[3] Department of Biological Sciences, Birla Institute of Technology & Science (BITS), Pilani, Rajasthan, India

[4] Wellcome Centre for Cell-Matrix Research, Faculty of Biology, Medicine and Health, University of Manchester, United Kingdom

[5] Department of Genetics and Genome Biology, University of Leicester, Leicester, UK

[6] SNIPR Biome, Copenhagen, Denmark

[7] These authors contributed equally

* Correspondence: F.Nobrega@soton.ac.uk

ABSTRACT

Bacterial defence against phage predation involves diverse defence systems acting both individually and concurrently, yet their interactions remain poorly understood. We investigated >100 defence systems in 42,925 bacterial genomes and identified numerous instances of their non-random co-occurrence and mutual exclusion. For several pairs of defence systems significantly co-occurring in *Escherichia coli* strains, we demonstrate synergistic anti-phage activity, including tmn synergising with defence systems containing sensory switch ATPase domains. Some of the defence system pairs that are mutually exclusive in *E. coli* significantly co-occur in other taxa, and one tested mutually exclusive pair showed synergy. These findings demonstrate compatibility and synergy between defence systems, allowing bacteria to adopt flexible strategies for phage defence, driven by specific environmental pressures. The evolution of bacterial immune repertoires seems to be shaped largely by selection for resistance to host-specific phages rather than by negative epistasis.

Keywords: Prokaryotic immunity, co-occurrence, phylogroups, Gabija, tmn, Druantia, Zorya, Kiwa, ietAS, Septu

INTRODUCTION

Bacteria evolved numerous, diverse lines of active immunity as well as abortive infection mechanisms to withstand phage predation[1]. Recent systematic screening uncovered numerous anti-phage defence systems that widely differ in protein composition and modes of action[2-7]. The mechanisms employed by bacterial defence systems include phage genome or protein sensing followed by degradation[8-10], introduction of modified nucleotides that abrogate phage replication[11,12], as well as multiple sensing mechanisms leading to abortive infection that results in the host cell dormancy or death[4,13-21]. However, for many, perhaps, the majority of the bacterial defence systems, the mechanism of action remains unknown.

A bacterial genome carries, on average, about five distinct (currently identifiable) defence systems[22]. The remarkable variability of immune repertoires was observed even within the same species[22-24]. Genes encoding components of these systems tend to cluster together in specific genomic regions known as defence islands, sometimes associated with mobile genetic elements (MGEs) integrated into distinct hotspots in the bacterial genome[24-26]. Defence systems are often horizontally transferred between bacteria, and close proximity of the respective genes could facilitate simultaneous transfer of multiple systems.

Despite the recent burst of bacterial defence system discovery, the causes of their clustering in defence islands remain poorly understood. It has been argued that co-localisation of defence systems in MGEs and the resulting joint horizontal gene transfer (HGT) could provide fitness advantages to recipient bacteria, especially in phage-rich environments[27]. Additionally, it has been suggested that synergistic interactions between defence systems could drive their co-localization and favour their joint transfer[28,29], as supported by the conservation of certain sets of defence systems[30]. For example, CRISPR-Cas systems of different subtypes often co-occur and the CRISPR arrays interact with Cas proteins across different systems[31]. Furthermore, toxin-antitoxin (TA) RNA pairs[32] and possibly other TA modules[33] safeguard CRISPR immunity by making cells dependent on CRISPR-Cas for survival. CRISPR-Cas and

restriction-modification (RM) systems[34], as well as BREX and the restriction enzyme BrxU[29], co-occur resulting in expanded phage protection. However, these examples of interaction between bacterial defence systems notwithstanding, their co-occurrence in bacteria and the connections between co-occurrence and co-localization in bacterial genomes have not been analysed on a large scale, and the underlying factors contributing to this phenomenon, such as synergistic interactions, remain largely unexplored. The possibility remains, notwithstanding all the adaptive explanations, that defence islands evolve neutrally through a preferential attachment process whereby multiple defence systems are incorporated into genomic regions devoid of essential genes where insertions are tolerated.

Here we report a comprehensive analysis of the co-occurrence of defence systems in 26,362 *Escherichia coli* genomes, as well as in complete genomes from four bacterial orders, Enterobacterales, Bacillales, Burkholderiales, and Pseudomonadales, to investigate the role of interactions between different defence systems in anti-phage immunity. Our findings show that defence system co-occurrence varies substantially across *E. coli* phylogroups and taxa and is not directly related to their co-localisation in the genome. For several pairs of non-randomly co-occurring and mutually exclusive defence systems in *E. coli*, we experimentally demonstrated synergistic interactions that provided an evolutionary advantage to the bacterial population. Moreover, some of the defence systems that are mutually exclusive in *E. coli* were found to co-occur in other bacterial taxa, and can also protect synergistically against specific phages. These findings imply that selection for robust immunity, rather than mechanistic incompatibility, is the primary driving force that shapes the defence system repertoire in bacteria.

RESULTS

Distinct defence system repertoires across *E. coli* phylogroups

To explore the variation in the immune repertoires among closely related bacteria, we analysed the defence system content in a comprehensive dataset of 26,362 *E. coli* genomes from the NCBI Reference Sequence (RefSeq) database[35,36]. *E. coli* is an ideal model organism for this research due to its wide distribution in diverse environments, high genetic diversity, the availability of numerous, well-characterised, complete genomes as well as a large panel of well-studied phages[37,38]. We found that, in agreement with previous observations, on average, *E. coli* genomes carry 5-7 defence systems, but some clades, such as those in phylogroup B2.1, harbour a greater diversity of such systems (**Figure S1A**, **Table S1**). The majority of the defence systems are encoded in the chromosomes, but there are some clades where additional systems, especially Gabija, tmn, PifA, ppl, AbiQ, and viperins, are carried on plasmids (**Figure S1B**).

We sought to investigate in greater detail the clade-specific patterns and the differences among the six major *E. coli* phylogroups that vary in their ecology. Our dataset prominently represented phylogroups A (25%), B1 (29%), and B2 (19%), which are highly prevalent in the human (A, B2) or domestic/wild animal microbiomes (B1)[39] (**Figure 1A**). Our analysis revealed substantial differences in the number (**Figure 1B**) and types (**Figure 1C**) of defence systems among the phylogroups. Phylogroup B2.1, which includes extra-intestinal pathogenic (ExPEC) strains[40], was particularly noteworthy, with the highest average number of defence systems (8) among the examined phylogroups (**Figure 1B**). The genomes in phylogroup B2 accumulate virulence factors[41,42] as well as antibiotic resistance genes[43], suggesting that this phylogroup is specifically prone to HGT mediated by MGEs such as pathogenicity islands and plasmids. Of particular interest in phylogroup B2 is the already reported absence of CRISPR-Cas type I-E that is common in other *E. coli* phylogroups[44] (**Figure 1C**). Conversely, significant enrichment (Chi Squared test for homogeneity, $p < 0.001$, **Table S2**) of Retron I-C (odds ratio (OR) = 9.04) and AbiE (OR = 4.77) was detected in the B2-1 subgroup, and high prevalence of CRISPR-Cas type I-F (OR = 5.12), Thoeris I (OR = 6.16), Septu I (OR = 3.17), and qatABCD (OR = 4.63) was observed in the B2-2 subgroup.

Phylogroup C showed enrichment of BREX I (OR = 3.54), and phylogroups E1 and E2 exhibited a much higher prevalence of Zorya II (OR = 4.62 and 10.25, respectively) and Druantia III (OR = 4.86 and 6.92, respectively) compared to the other phylogroups. Phylogroup E2 additionally showed a higher prevalence of PsyrTA (OR = 7.81), and a reduced prevalence of RM IV (OR = 0.02) that seems to be compensated by an increase in RM IIG (OR = 7.86).

For most phylogroups, we observed a relatively strong positive correlation ($r = 0.34\text{-}0.66$, $p < 2.2\times10^{-16}$) between the types of defence systems found in *E. coli* genomes within phylogroups and the genetic relatedness of these genomes, with the exception of phylogroups A and B1 ($r = 0.02\text{-}0.11$, $p < 2.2\times10^{-16}$), as indicated by the mash distance analysis (**Figure S1C**). These observations indicate that, although HGT is an important route of defence system acquisition, vertical inheritance plays a major role in the evolution of the immune repertoires, at least at short phylogenetic distances. The low correlation in phylogroups A and B1 likely reflects ecological differences between subclades in which case HGT apparently becomes a defining factor.

In summary, our analysis of the distribution of defence systems across *E. coli* genomes demonstrates associations between specific defence systems and individual phylogroups, likely driven by selection for sets of defence mechanisms capable of efficiently protecting the bacteria against the specific repertoires of phages and other MGEs that they encounter in their respective environments.

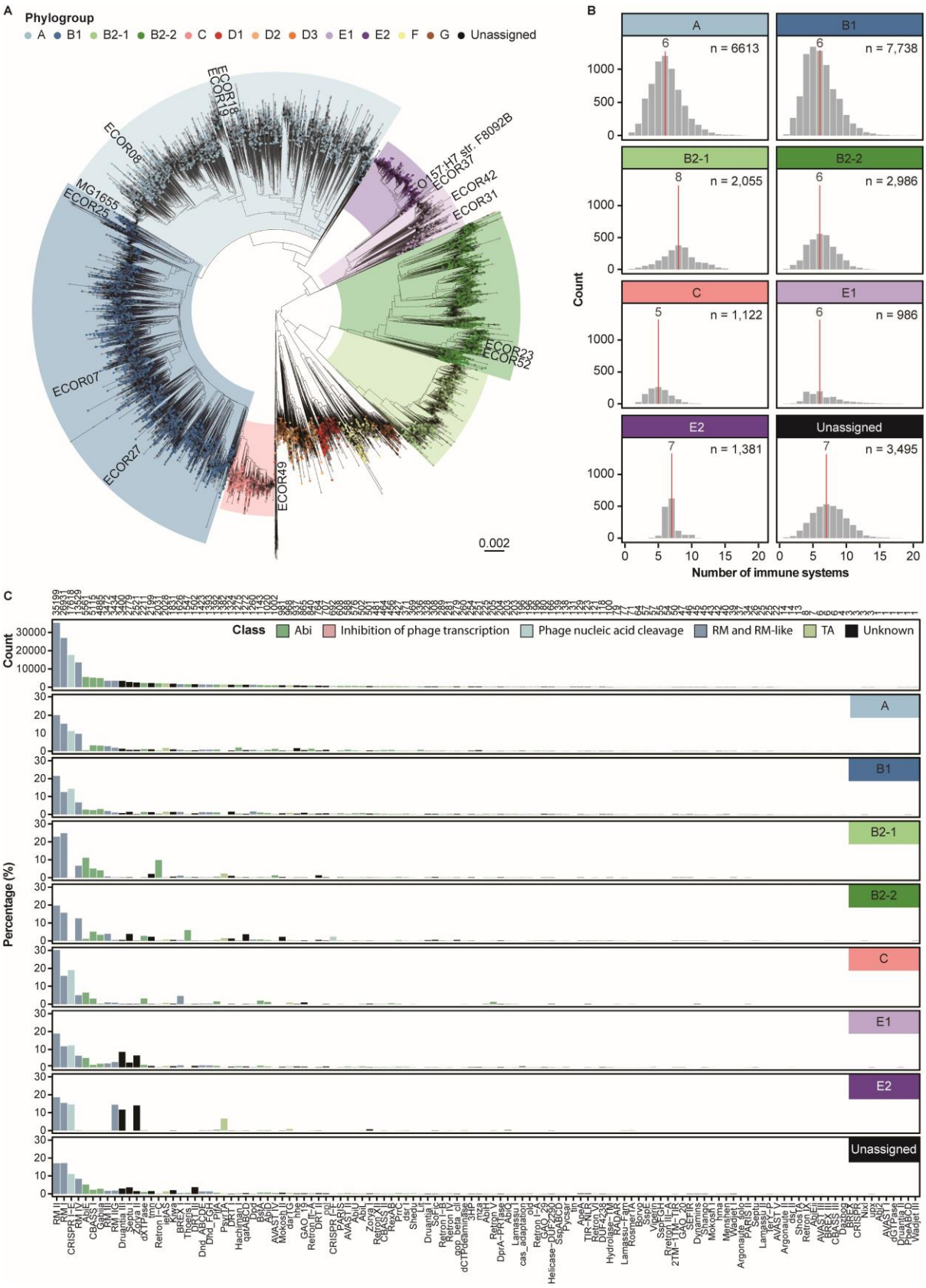


Figure 1 Distribution of defence systems across *E. coli* phylogroups.

(A) A phylogenetic tree displaying 26,362 *E. coli* genomes obtained from the RefSeq database. Phylogroups are assigned as described previously[45], and each phylogroup is colour-coded according to the key.

(B) Number of defence systems found per *E. coli* strain in each phylogroup. The mean number of defence systems is indicated by a red line.

(C) Prevalence of defence systems in the *E. coli* genomes. The defence systems are organised from most prevalent (left) to least prevalent (right) and their total count is shown in the top bar graph. The bars are colour-coded according to the mechanism of the defence system. The remaining bar graphs show the prevalence, in percentage, of the defence systems per phylogroup.

118 pairs of defence systems co-occur in *E. coli*

In previous studies, some defence systems have been shown to interact, resulting in enhanced or expanded protection against phages[29,31-34]. Here, we sought to determine if specific defence systems co-occurred more frequently than expected in bacterial genomes, potentially indicating interactions between different defence mechanisms enhancing protection against phages. To this end, we explored correlations between the occurrences of pairs of defence systems in *E. coli* genomes corrected for phylogenetic bias; we deemed such a correction to be essential because, as shown above, vertical inheritance of defence systems is common (see Methods). This analysis allowed us to identify pairs of defence systems that appear together in the same genome significantly more (co-occurring) or less (mutually exclusive) frequently than expected based on their individual prevalence (**Table S3**, **Figure 2A**).

Our analysis revealed that 183 interacting pairs of defence systems (7.4% of all the analysed pairs) were significantly correlated, positively or negatively (Pagel test for binary traits, with Bonferroni correction for multiple tests). Of these, 118 pairs were co-occurrences (64.5% of the correlated pairs), and the rest were cases of mutual exclusion. With the more permissive Benjamini-Hochberg correction, 379 pairs of defence systems (15.3% of the analysed pairs) were significantly correlated, of which 219 (57.8%) were co-occurrences (**Figure 2B**). Notably,

although the network of co-occurrences and especially of exclusions between the *E. coli* defence systems was sparse (**Figure 2**, **Figure S2A**), each system significantly co-occurred with at least one other system, and typically, with two or more under the permissive correction, and for most, at least one significant co-occurrence was detected under the strict correction, too (**Table S3**). Thus, co-occurrence between defence systems is a widespread phenomenon that involves (nearly) all such systems identified in *E. coli*. The greatest number of significant co-occurrences was observed for the CRISPR I-E system that is found in the majority of the *E. coli* genomes, but several less common systems, such as AbiE, PsyrTA and Mokosh II, also appeared to be particularly prone to co-occurrence with other systems (**Table S3**). Some of the co-occurring pairs appeared striking in that the great majority of the instantiations of the rarer system in the pair were found in genomes that also carried the more common system, suggestive of a functional dependence and indeed their interaction was better described by the dependent model (**Table S3**). For example, 882 of the 975 instances of Mokosh II co-occurred with RM IV, and 2151 of the 2518 instances of Zorya II co-occurred with Druantia III (**Figure 2A** and **Table S3**).

Notably, different subtypes of the same defence system displayed distinct co-occurrence patterns with other systems (**Figure 2B**, **Figure S2B**). For example, while Druantia III co-occurred with 8 other defence systems, no co-occurrences were found for Druantia I (Bonferroni correction). CBASS I, composed of cyclase and effector proteins, and CBASS II, characterised by the presence of cGasylation proteins cap2 (E1-E2 fusion) and cap3 (JAB)[19,46], co-occurred with distinct defence systems. CBASS I co-occurred with DndABCD and DndFGH (which also co-occurred with each other), Mokosh II, qatABCD, and RM I and IV, whereas CBASS II co-occurred with CRISPR-Cas type I-E, Hachiman I, and tmn. These specific co-occurrences might reflect distinct cooperative interactions between the respective defence system subtypes.

Conversely, we observed 65 (2.6%) pairs of defence systems that were mutually exclusive, such as CRISPR-Cas types I-E and I-F1, consistent with previous observations in

Proteobacteria[31] (**Figure 2B**). Like the number of co-occurrences, the number of significant exclusions notably varied across defence systems, with some, for example, PsyrTA and RM IV appearing particularly prone to avoiding other systems (**Figure S2B**). Interestingly, RM IV and BREX I were mutually exclusive in *E. coli*, even though previously shown to co-occur on a plasmid-encoded defence island and provide complementary protections against modified (RM IV) and non-modified (BREX I) invading DNA in *Escherichia fergusonii*[29]. However, here we consolidated all subtypes of RM IV together, and the majority of occurrences in our dataset were in the chromosome, showing that the observed pattern in *E. fergusonii* was specific and differed from the typical behaviour of RM IV. Additionally, location of one of the systems within an integrated element, such as prophage, can lead to mutual exclusion with systems active against that MGE, as it potentially happens in the case of ietAS and Zorya II (**Figure 2A**). The RM-like Dpd system, which acts by inserting 7-deazaguanine derivatives into the host DNA to distinguish it from the non-modified invading DNA[47], was found to be mutually exclusive with several RM and RM-like systems, such as RM I, RM III, RM IV, BREX I[48-50], Druantia III[2], and DndABCD and DndFGH[9].

Like the co-occurrence patterns, the patterns of mutual exclusion showed substantial differences among subtypes of the same defence system (**Figure 2B**, **Figure S2B**). In most cases, however, the exclusivity between defence systems was not strict, that is, the respective pairs were observed together in some genomes (**Table S3**). This observation implies that the exclusivity is not caused by incompatibility between the respective systems resulting in negative epistasis[31], but rather by genetic drift due to functional redundancy or by selection against such redundancy.

Overall, our results indicate that both non-random co-occurrence and (partial) mutual exclusion among defence systems are common in *E. coli*.

Figure 2 Co-occurrence and mutual exclusion among defence systems in *E. coli*.

(A) Graphical representation of the co-occurrence analysis, depicting one pair of defence systems that co-occur (Gabija and tmn), and one pair that is mutually exclusive (Zorya II and ietAS). The nodes of the *E. coli* phylogenetic tree are coloured according to the presence or absence of the defence system in each strain. Their location in chromosome, plasmid, or prophage regions is indicated in the middle. For visualization purposes, only leaves that carry at least one system from the pair are shown.

(B) Co-occurrence of defence system pairs in *E. coli*. Co-occurring systems are shown in orange, and mutually exclusive systems are shown in green. The correlation between pairs was calculated with Pagel test for binary traits. Asterisks show correlations that were significant after the less stringent Benjamini-Hochberg correction (*) or the most stringent Bonferroni correction (**) for multiple testing. In the main text, the results after Bonferroni correction are considered. The defence systems are colour-coded according to their broadly defined mechanism of defence.

(C,D) Distance histograms of (B) all and (C) co-occurring defence system pairs in 2,164 complete *E. coli* genomes. The median distance between genes encoding defence systems is shown by a red line. The analysis considered only those pairs that significantly co-occurred after Bonferroni correction.

Co-occurrence of defence systems is not tightly linked to physical proximity

Defence systems often cluster together in defence islands[2,4,7,24,26,51], which has been hypothetically attributed to fitness benefits conferred by such clustering on bacteria living in environments with high phage loads [27]. In particular, clustering of defence systems in defence islands, especially within integrated MGEs, increases the likelihood of horizontal co-transfer[52]. Therefore, if the defence system repertoire is predominantly shaped by HGT, and not by any functional benefits, the co-localizing systems will also be the ones that we observed as co-occurring in *E. coli*. However, our analysis showed that co-occurring defence systems were generally not located significantly closer to each other in the genome compared to the average distance between defence systems (**Figure 2C,D**). Thus, in general, the physical proximity of defence systems within the genome, although likely facilitating their concomitant horizontal transfer, does not appear to play a defining role in their co-occurrence.

Nonetheless, there are some exceptions where co-occurring defence systems did indeed non-randomly co-localize. These include Mokosh II and RM IV, Druantia III and RM I, Druantia III and RM IV, RM I and RM IIG, RM IIG and Zorya II, Druantia III and Zorya II, Druantia III and RM IIG, Gabija and tmn, DndABCDE and DndFGH, CBASS I and qatABCD, RM III and ietAS, and RM IV and SoFic (**Figure S2C**). Of these pairs, only DndABCDE and DndFGH were previously reported to co-localize and functionally interact[53,54].

Co-occurring defence systems act synergistically to counter phage infection

Next, we sought to explore whether co-occurrence of defence systems is driven by their complementary activities[30] or synergistic molecular cooperation[28,29]. To this end, we selected three pairs of significantly co-occurring defence systems found in *E. coli* strains from our collection: Gabija and tmn that co-localise frequently in plasmids, Druantia III and Zorya II that co-localise in the chromosome, and ietAS and Kiwa that do not co-localise (**Figure S2C**). Additionally, we tested the mutually exclusive combination of Zorya II and ietAS.

By cloning these systems into *E. coli* individually or jointly (**Figure 3A**), we assessed their effects on phage resistance using efficiency of plating (EOP) assays with a panel of 29 phages. This experiment demonstrated limited anti-phage activity for all single systems, with the exception of Gabija, which provided strong protection against multiple phages (**Figure S3A**). Combinations of the defence systems substantially increased the protection levels against specific phages. To further quantify these effects, we calculated the epistatic coefficients by comparing the combined effect of the defence system pair to the sum of the individual effects of the two partners. The results showed that all tested combinations of defence systems, with the exception of ietAS and Kiwa, displayed significant synergistic effects against at least some phages in our panel, with even the mutually exclusive pair Zorya II and ietAS showing unexpected synergy (**Figure 3B**, **Figure S3A**).

To further validate the findings from the EOP assays, we performed time post-infection assays using phages T1 and T3 to assess the impact of the defence system combinations on phage propagation in liquid cultures. The results from these assays consistently confirmed the synergy as the combinations of defence systems led to a reduction in phage propagation that significantly exceeded the sum of the individual effects (**Figure 3C**, **Figure S3B**). This synergistic trend was observed for phage T1 with Gabija and tmn, for T1 and T3 with Druantia III and Zorya II, and for T3 with Zorya II and ietAS. Even the combination of ietAS and Kiwa, which showed no significant synergy in the EOP assays (**Figure 3B**), displayed a minimal but detectable synergy against T3 propagation (**Figure 3C**). Moreover, the synergy was not restricted to phages targeted by both individual systems. For example, T1 was not affected by Gabija alone, but the combination of Gabija and tmn resulted in an increased protection compared to tmn alone (**Figure 3C**). For the mutually exclusive pair Zorya II and ietAS, only Zorya showed substantial activity against phage T3, but the combination with ietAS resulted in an improved reduction in phage propagation.

Additionally, we assessed synergy between defence systems in terms of bacterial survival by measuring the absorbance of bacterial cultures over time when infected with phages at different multiplicities of infection (MOIs) (**Figure 3D**). To quantify the synergy in these assays, we compared the areas under the curve (AUC) above the OD at the experiment start for individual systems and their combinations. These additional results from liquid cultures were consistent with the findings from the EOP and phage propagation assays, providing further evidence of synergistic interactions between the defence systems. We observed synergy occurring at low MOIs, resulting in increased bacterial survival. However, at higher MOIs, all bacterial cultures collapsed, likely due to overwhelming of the defence systems or abortive infection. In the case of Gabija and tmn, increased bacterial survival was not observed at low MOIs for the combination of systems because tmn alone was sufficient to restore normal growth of T1-infected bacteria (**Figure 3D**, **Figure S3C**).

Taken together, our assays provided robust evidence that some of the significantly co-occurring defence systems display synergistic activity against specific phages. This bolsters the notion that co-occurrence is maintained by environmental selection favouring combinations that are beneficial for the bacteria given the particular virome composition in their niche. Additionally, the observation that mutually exclusive defence systems in some cases also display synergistic activity suggests that inherent mechanistic incompatibilities between defence systems are unlikely to be a major driver of their mutual avoidance.

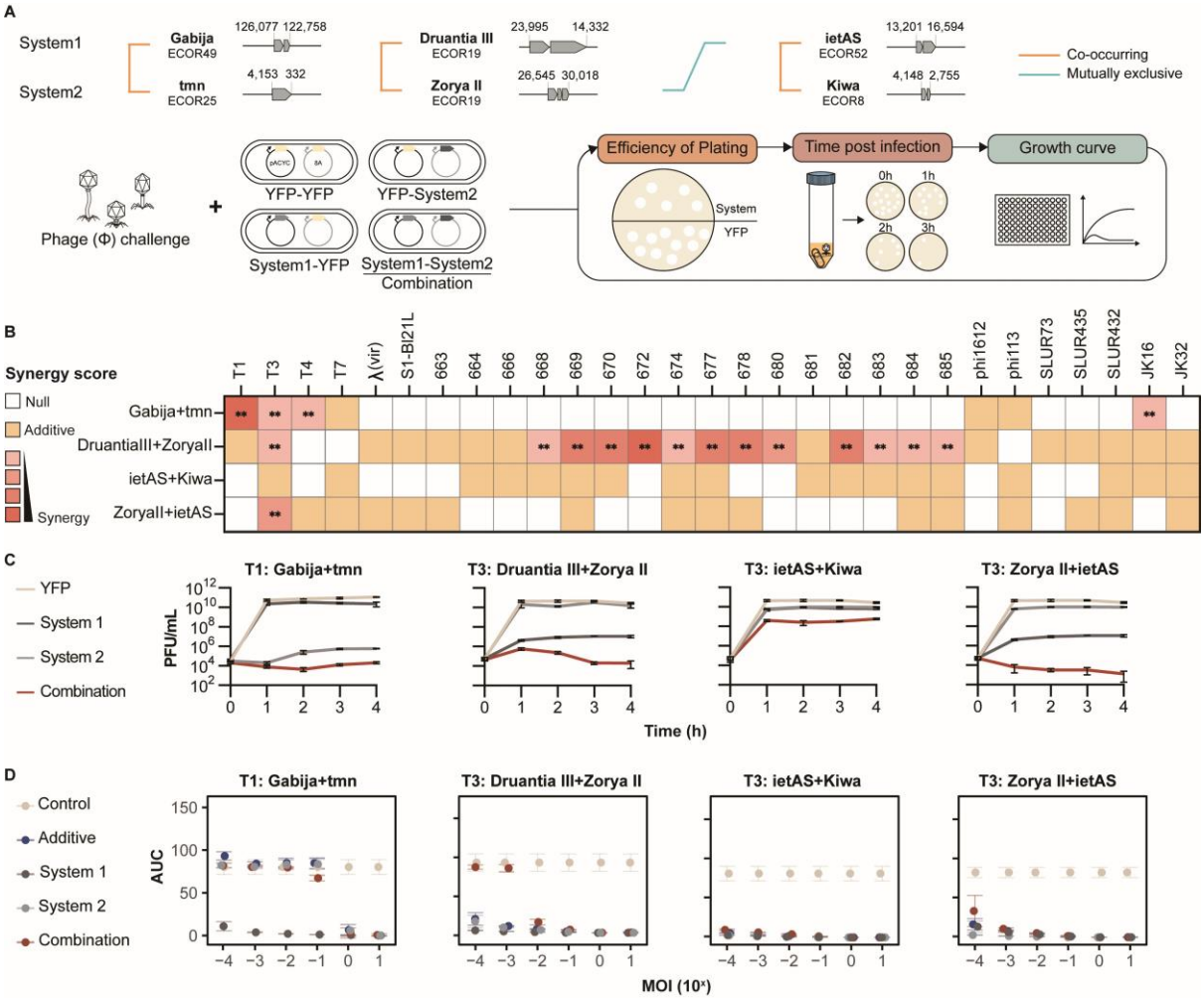


Figure 3 Defence system pairs provide synergistic anti-phage activity**.**

(A) Experimental set-up for the assessment of the anti-phage activity of individual defence systems and their combinations. YFP, yellow fluorescent protein.

(B) Heatmap of synergy score of protection provided by selected defence system pairs against a panel of 29 phages. The synergy score is the epistatic coefficient for pairs of defence systems (see Methods).

Null, EOP equivalent to the defence provide by one system; Additive, EOP corresponds to the combined defence of the two individual systems; Synergy, EOP exceeds the collective defence of the two systems. Data is shown as the average of three biological replicates. ** Statistically significant ($p < 0.01$).

(C) Time post infection assays, measuring T1 or T3 titres over the course of four hours in liquid cultures of *E. coli* containing individual or combined defence systems. Data is shown as the average and standard deviation of three biological replicates.

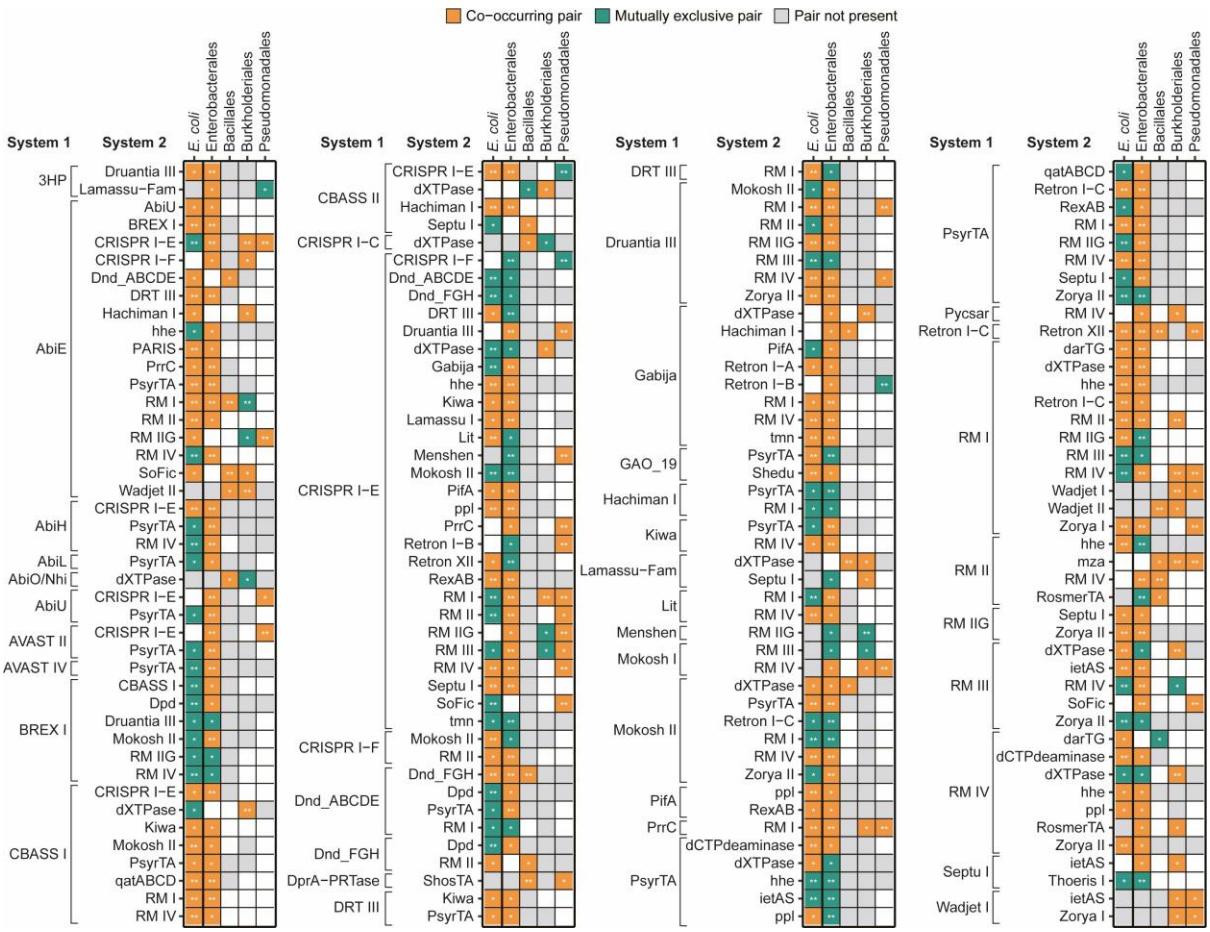
(**D**) Bacterial growth under phage predation at different multiplicities of infection (MOIs), represented as area under the curve (AUC) in OD·h. An defence system pair acts synergetic when its dot (red) is above the expected additive effect (blue). Data is shown as the confidence interval of three biological replicates. The raw data and growth curves used to calculate the AUCs are available on the associated Github and Zenodo databases.

Defence system co-occurrence is not conserved across bacterial taxa

Genomic analyses show that increasing phylogenetic distance between bacterial species, decreases the rate of HGT[55-58]. In particular, the host ranges of most phages are narrow so that transduction occurs mostly within the host species boundaries[59]. Similarly, only a small fraction of plasmids has been shown to cross the interspecies barrier[60-62]. Considering these limitations of HGT along with (largely) non-overlapping viromes[22], it could be expected that the co-occurrence of defence systems is poorly conserved among bacteria. To test this prediction, we extended our analysis to four bacterial orders, Bacillales, Burkholderiales, Enterobacterales, and Pseudomonadales. We first examined the sets of defence systems present in these bacteria. The different bacterial orders displayed variations in the type and abundance of defence systems (**Figure S4A**). For example, RM systems were the most prevalent in Bacillales, Enterobacterales (including *E. coli*), and Pseudomonadales, whereas Burkholderiales were characterised by a higher abundance of dXTPase, Zorya III, and Mokosh I. Additionally, while CRISPR I-E was abundant (50.8%) in Enterobacterales (including *E. coli*), its prevalence was markedly lower (<6%) in the other orders. Nonetheless, when considering

all four orders and one extensively characterized species (*E. coli*), they collectively shared 87 defence systems in common, with only a few systems unique to each of them (3 in *E. coli*, 18 in Bacillales, 2 in Enterobacterales, 2 in Pseudomonadales, and 1 in Burkholderiales) (**Figure S4B**).

As anticipated, the specific pairs of co-occurring and mutually exclusive defence systems differed across the taxa. Moreover, multiple pairs of defence systems that co-occurred in one order were found to be mutually exclusive in another (**Figure 4**). For instance, AbiE and RM I co-occur in Bacillales and Enterobacterales but are mutually exclusive in Burkholderiales. Similarly, CBASS type II and CRISPR-Cas type I-E co-occur in Enterobacterales but are mutually exclusive in Pseudomonadales. These findings indicate that mutually exclusive systems are not inherently functionally incompatible or redundant. Rather, the interactions between these systems likely depend on environmental and genetic factors that select for a particular anti-phage immunity strategy. Overall, our results indicate that defence systems are generally mechanistically compatible, allowing bacteria to adopt diverse, flexible strategies for anti-phage defence based on their unique environmental and genetic contexts.



354

355 **Figure 4.** Patterns of defence system co-occurrences across bacterial taxa.

356 Heatmap of defence system co-occurrence patterns in *E. coli* (n = 26,362) and in four bacterial orders:
357 Enterobacterales including *E. coli* (n = 9,124), Bacillales (n = 3,952), Burkholderiales (n = 2,199), and
358 Pseudomonales (n = 1,288). Grey squares indicate that at least one system in the pair is not present in
359 the taxonomic group. * Co-occurrence significant after Benjamini-Hochberg correction; ** Co-
360 occurrence significant after Bonferroni correction.

361

362 **Synergistic immunity provides an evolutionary advantage to bacterial populations**

363 Due to our observations, we reasoned that co-occurring and synergistic defence systems
364 could provide advantages at the population level. To assess the evolutionary and ecological
365 impact of synergistic interactions between defence systems, we performed a short-term
366 evolution experiment using chromogenic reporter plasmids expressing engineered coral

chromoproteins[63,64]. We mixed populations of strains containing either an individual defence system and a second, “empty” chromogenic plasmid, or carrying two defence systems on separate plasmids. These populations were then infected with either a phage shown to trigger a synergistic defence, or a phage eliciting no obvious synergy between the respective defence systems. Over a period of three days, we monitored the proportion of the population carrying one or both defence systems by counting colony forming units of different colours (**Figure 5A**). We confirmed that all plasmids were stably maintained in the populations throughout the experiment using plasmid loss assays, ruling out any influence of plasmid loss due to toxicity on the outcomes (**Figure S5A,B**). Further, given that resistant receptor mutants tend to spread in bacterial populations shortly after phage infection[65], we investigated whether this factor influenced the outcome of our short-term evolution experiments by evaluating the capacity of the phage to infect bacterial colonies retrieved from the experiment. Since the defence systems under examination in this study do not offer full protection against phage infection, the absence of infection indicates the emergence of receptor mutants leading to complete phage resistance. We observed a limited effect of receptor mutants on phage resistance, with greater relevance at high phage MOI (**Figure 5B**).

The results showed that in the absence of phage, populations containing either individual or combined defence systems remained relatively stable, and thus these systems were minimally toxic to the cells, if at all (**Figure 5C**). However, after exposure to phage infection at high or low MOI, a clear shift in the population composition towards cells harbouring both defence systems was observed starting from day 1. This shift was pronounced in cultures infected with the phage that elicited synergistic defence (T1 for Gabija and tmn, T3 for Druantia III and Zorya II, and T3 for ietAS and Zorya II). In cases the defence was not synergistic (Lambda for Gabija and tmn, phi113 for Druantia III and Zorya II, and T7 for ietAS and Zorya II), the outcomes varied. For Druantia III and Zorya II, as well as Zorya II and ietAS combinations, cells carrying both the combination of systems and the individual system active against the phage (Druantia III and Zorya II, respectively **Figure S3A**) became predominant in the

population. However, in the case of the Gabija and tmn combination, the system that was not active against the phage (tmn) dominated (**Figure S3A**). This discrepancy can be attributed to the considerably higher efficiency of Gabija against phage Lambda (10^5-fold) compared to that of Druantia III against phi113 (10^1-fold) and Zorya II against T7 (10^1-fold) (**Figure S3A**). This result suggests that the phage population was more effectively reduced when Gabija was active, allowing other cells in the population, even those lacking active defences against the phage, to survive. In the case of the Gabija-tmn combination, the reduction in the population of cells containing the active system, Gabija, was likely due to the abortive infection phenotype of this defence system[66].

These results underscore the pivotal role played by the degree of protection provided by specific, active sets of defence systems in the survival of phage-sensitive cells within a heterogeneous population, thereby shaping the dynamics of coexisting defence systems. Overall, the competition experiments validate the evolutionary advantage of synergistic defence system combinations against specific phages. Moreover, these findings also emphasise the substantial impact of factors such as the type and abundance of the encountered phages, as well as the effectiveness of the defence systems, on the resulting dynamics of the bacterial population.

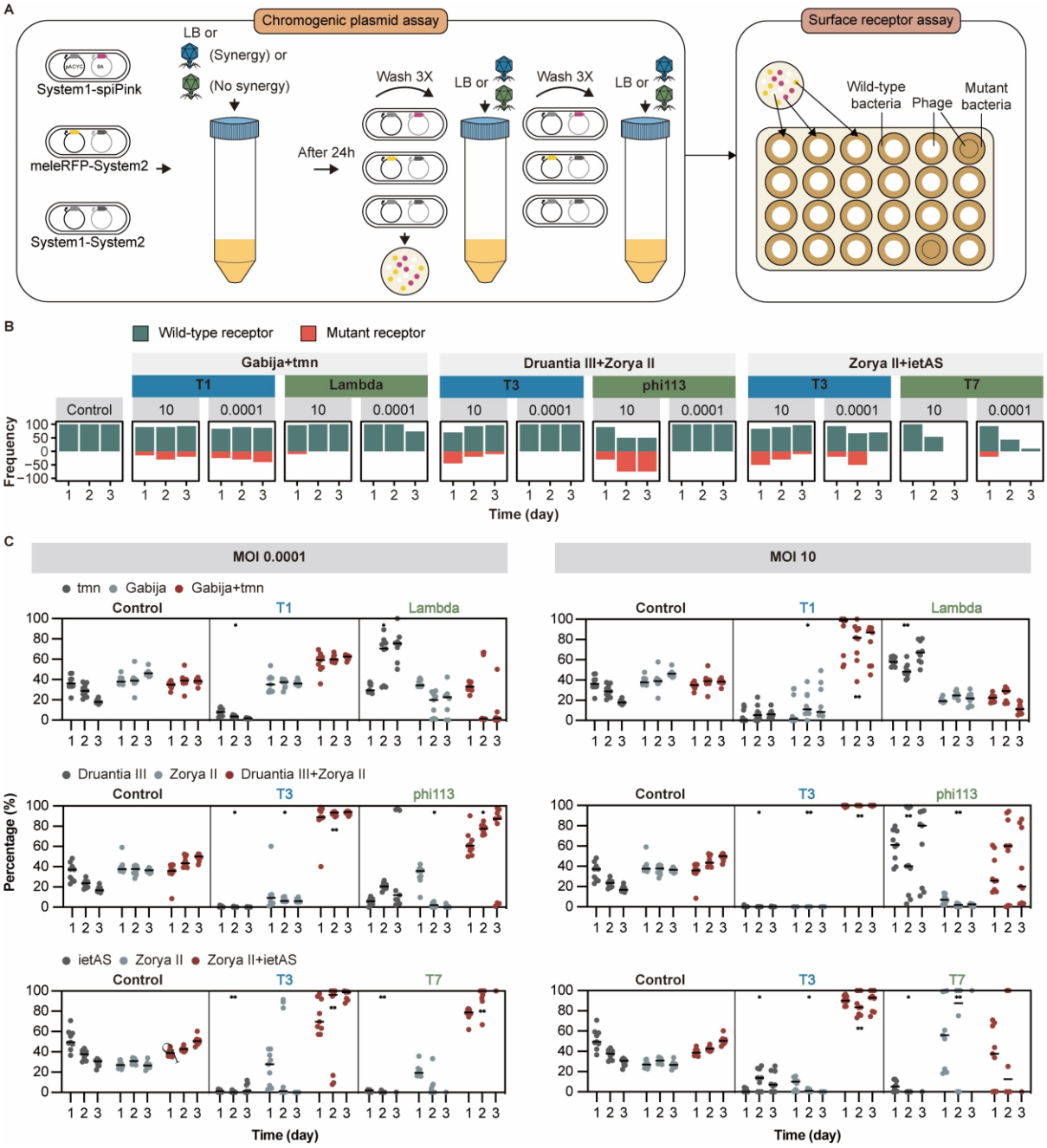


Figure 5 Synergistic defence system pairs provide an evolutionary advantage to bacteria**.**

(A) Set-up of experimental evolution assay of defence systems using chromogenic plasmids (left). Cells containing a single defence system carry a second plasmid that expresses a chromogenic reporter. Cells with defence system combinations were mixed in equal proportions with cells with single systems, and infected with phage. After 1, 2, and 3 days, cells were platted and the colonies of different colours were enumerated. A surface receptor assay (right) assessed the influence of receptor mutants on the outcome of the evolution assays, by subjecting colonies of different colours to phage infection in plaque assays.

(B) Prevalence of receptor mutants in the bacterial population during the evolution assay. The proportion of colonies with the wild-type receptor is represented as positive in the vertical axis, while the proportion of colonies with mutated receptor is shown as negative. Colonies with mutated receptor were identified by their complete resistance to phage infection in spot assays.

(C) Percentage of colonies from the evolution assay that carry each individual defence system or their combinations at 1, 2 and 3 days post infection with phage at low or high multiplicity of infection (MOI), compared to a non-infected control. Data is shown as the average of three biological and three technical replicates with individual counts shown. * p value < 0.01, ** p value < 0.001 determined by multiple comparison for nested one-way ANOVA, comparing to the corresponding uninfected control.

Tmn co-opts the ATPase domain of Gabija for synergy

To better understand the molecular mechanisms underlying the observed synergistic effect between defence systems, we focused on the combination of Gabija and tmn. This choice was motivated by several factors. Gabija has been well characterised previously, in contrast to all other tested defence systems, providing molecular detail[2,66-68] that helps understanding the synergy with the less thoroughly characterized tmn[3]. Furthermore, Gabija and tmn tend to physically co-localise on plasmids (**Figure S2D**). Comparison of plasmids carrying both tmn and Gabija showed pronounced genetic variability, except for regions that encompass conjugation-related genes and these particular defence systems (**Figure 6A**). Gabija and tmn, along with type II TA systems such as VapBC, are specifically located in the leading region of the plasmid. This region is crucial for maintaining plasmid stability during conjugation[69,70], and is enriched in anti-defence genes[71], which protect conjugative plasmids from host defence systems during the initial stages of plasmid invasion. The conserved location of Gabija and tmn within the leading region of plasmids suggests that they play a critical role in plasmid maintenance. By allowing the plasmid to fend off other, competing MGEs, these systems likely ensure the plasmid maintenance in the cell, and with it, the evolutionary success of this defence system combination in the face of ongoing inter-MGE conflicts[28].

To explore the mechanism of synergy between tmn and Gabija, we focused on determining the specific contributions of the functional domains of the individual defence systems to the synergistic anti-phage activity. We introduced point mutations into these functional domains and assessed their effects on the individual defence systems and their combination using EOP assays.

In tmn, mutations of conserved residues of the P-loop NTPase domain (G66A/K67A, and R276A/R279A) abolished protection by tmn and the synergy with Gabija against phage T1 (**Figure 6B**). In Gabija, the ABC ATPase domain of GajA senses the depletion of cellular nucleotides during phage infection, activating the DNA binding and nicking activity of its TOPRIM nuclease domain[68]. Introduction of nicks into the DNA activates nucleotide hydrolysis by the UvrD-like helicase domain of GajB, depleting essential nucleotides and leading to abortive infection[66]. Surprisingly, we observed that only the nucleotide-sensing ATPase domain of GajA appeared to be critical for the synergy with tmn (**Figure 6B**) because mutations of conserved residues in the respective catalytic sites[66] of the TOPRIM and UvrD-like helicase domains abolished protection by Gabija on its own but had no effect on the synergy. The individual K35A and H317A mutations introduced into the ATPase domain of GajA abolished Gabija activity but only H317A abolished the synergy with tmn. The histidine residue H317 has been shown previously to play a role in inhibiting the nicking activity of GajA in the presence of ATP, suggesting that this histidine is critical for sensing the nucleotide levels in the cell[66] and for the synergy with tmn.

Gabija forms a supramolecular complex composed of a GajA tetramer with two sets of GajB dimers docked on opposite sides[67]. While the helicase function of GajB and the TOPRIM activity of GajA do not appear to be required for the synergy with tmn, we sought to determine whether the intact complex was a critical factor for the molecular interactions leading to synergy. We introduced stop codons into each Gabija protein-coding sequence and assessed the impact of halted protein translation on the synergy with tmn. Our findings indicate that

GajA alone could not synergise with tmn (**Figure 6C**), highlighting the necessity of forming an intact GajAB supramolecular complex for the synergy.

These findings demonstrate the critical role of the ATPase domains of both GajA and tmn in driving the synergy between these defence systems. More specifically, the ABC ATPase domain of GajA enhances the activity of tmn, emphasizing the central role of tmn in the synergistic combination. The molecular mechanism of tmn enhancement remains to be elucidated. Tmn is a KAP NTPase[3], characterised by the presence of two transmembrane helices inserted into the P-loop NTPase domain, which anchor tmn in the membrane such that the P-loop domain is located on the intracellular side[72]. Because most P-loop ATPases are multimers (most commonly, hexamers[73]), we modelled the tmn dimer using Alphafold2 (pLDDT score 89.1) (**Figure S6A**), and observed that the residues situated between the two transmembrane helices (P177-P198) are likely located in the periplasmic region. This periplasmic loop might be involved in the recognition of phage components resulting in the activation of the NTPase. It has been proposed that the function of the KAP NTPase domain is the regulation of assembly/disassembly of other protein complexes that interact with the extended surfaces provided by the α-superhelical structural domains of tmn, in an NTP-dependent manner[72]. Hence, the ABC ATPase of Gabija might assist the NTPase domain of tmn in this regulation, increasing the downstream response, by a mechanism that remains to be elucidated.

Overall, our exploration of the synergistic interaction between tmn and Gabija revealed a specific molecular interplay between these systems, highlighting the pivotal role of the distinct ATPase domains of each of these defence systems.

Synergy between tmn and defence systems containing sensory switch ATPases

Because tmn appears to be the main driver of the synergy with Gabija, we examined the domain architectures of the other defence systems that significantly co-occurred with tmn,

namely, AbiL, PrrC, PsyrTA, Septu I, and CBASS II (**Figure 2A**). Except for CBASS II, all these systems, including Gabija, contain ATPase domains associated with effectors that likely cause DNA or RNA damage[2,74-76], suggesting a potential shared mechanism underlying synergistic interactions with tmn. To test this hypothesis, we analysed the anti-phage defence provided by tmn when combined with the co-occurring systems Septu I and PrrC. As observed for tmn and Gabija, the combination of tmn with Septu I or PrrC demonstrated synergistic effects against specific phages (**Figure 6D**). The synergy between tmn and Septu I was consistently observed when using Septu I gene clusters from different *E. coli* strains (**Figure S6B**).

We further characterised the synergy between tmn and Septu I by introducing mutations into critical residues of Septu I. Mutations in the Walker A (K43A), Walker B (D316A/E317A), and D-loop (H321) regions of the ATPase domain of PtuA, as well as in the predicted Mg^{2+} binding site (K6A) and active site (H75A) of the HNHc nuclease domain of PtuB, resulted in the loss of synergy against phage T1, indicating the essential role of both proteins in the synergy (**Figure 6E**). Both PtuA and PtuB proteins act as toxins of the retron Ec78 defence system, and it was proposed that PtuA, similarly to GajA in Gabija, is activated by NTP depletion during phage infection[77]. This feature is also observed in the PrrC defence system, where the ATPase domain of PrrC is inhibited by ATP and GTP, alongside negative regulation by the RM system PrrI. The release of this inhibition, triggered by phages deploying anti-restriction peptides that inhibit the PrrI restriction enzyme, leads to GTP hydrolysis and activates the C-terminal anti-codon nuclease (ACNase) HEPN domain[75,78,79]. The PsyrTA[76], also known as RqlHI[80], and AbiL[74] defence systems, which co-occur with tmn, likely function in a similar manner, given that the ATPase containing proteins of these systems have been shown to inhibit the toxic activity of the other protein.

In summary, our results show that tmn synergises with various defence systems containing ATPase domains that function as sensory switches unmasking the associated effector domains, as observed here with Gabija, PrrC, and Septu I. These domains likely aid the

NTPase domain of tmn in controlling its downstream response. Further exploration of the mechanisms of these defence systems is expected to provide deeper insights into the synergistic interactions.

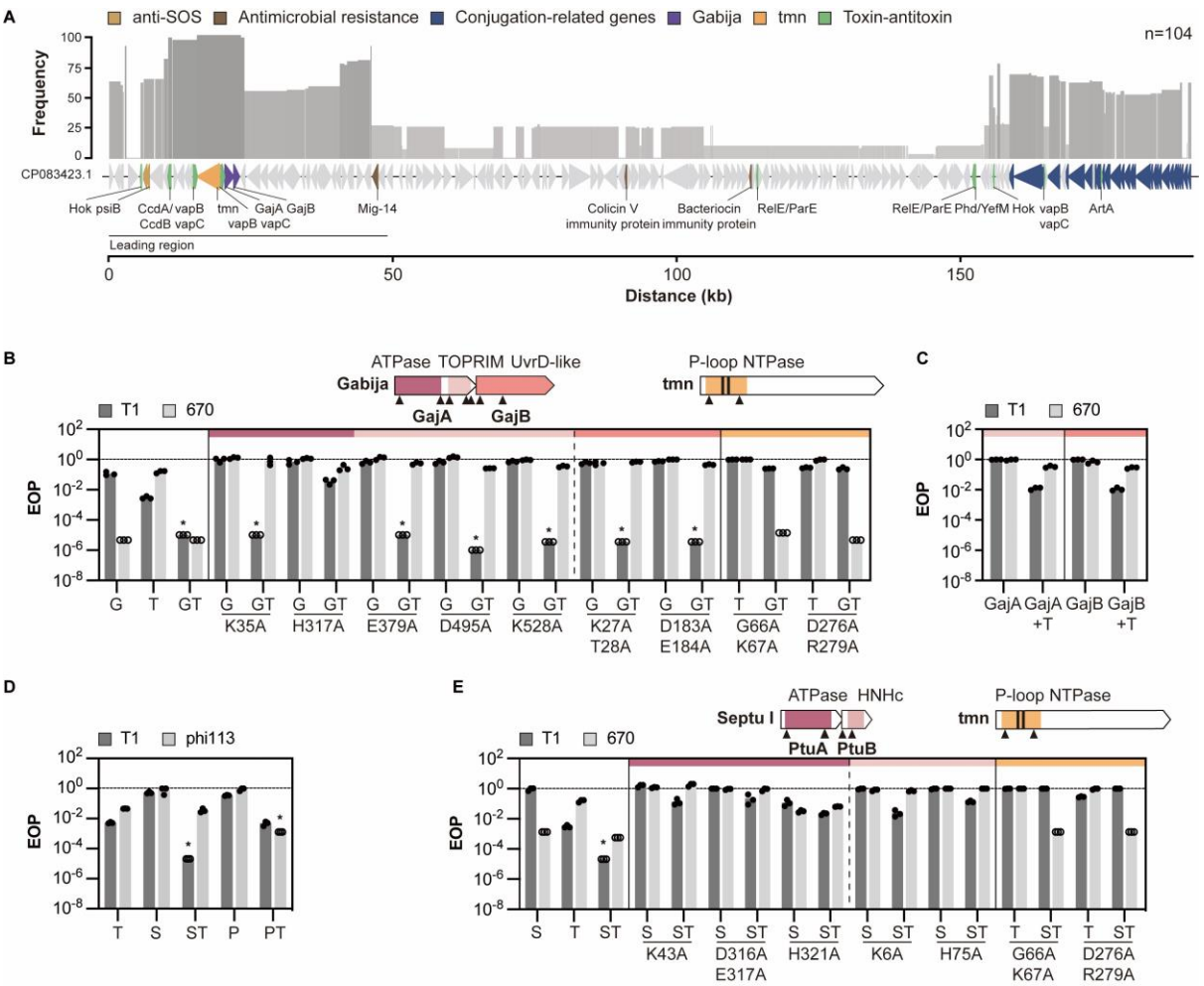


Figure 6 Mechanistic insight into the synergistic interaction between tmn and Gabija or Septu I.

(A) Whole-plasmid alignment of 104 plasmids containing tmn and Gabija from complete *E. coli* genomes with plasmid CP083423.1 as a reference (see Methods). The histogram shows the percentage of plasmids where the corresponding block from CP083423.1 was found. Annotated genes are coloured by function.

(B) Efficiency of plating (EOP) of phages T1 and 670 on cells expressing Gabija (G), tmn (T), Gabija and tmn (GT), and alanine mutants of specific functional domains. The mutations are organised by functional domains of Gabija (ATPase, TOPRIM, and UvrD-like) and tmn (P-loop NTPase). Unfilled

circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution. Asterisk (*) indicates cases of synergy.

(C) EOP of phages T1 and 670 on cells expressing tmn (T), Gabija (G), and tmn with either GajA or GajB. Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution.

(D) EOP of phages T1 and phi113 on cells expressing tmn (T), Septu I (S), PrrC (P), Septu I and tmn (ST), or PrrC and tmn (PT). Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution. Asterisk (*) indicates cases of synergy.

(E) EOP of phages T1 and 670 on cells expressing tmn (T), Septu (S), and Septu and tmn (ST), and variants with point mutations in specific functional domains. The mutations are organised by functional domains of Septu I (ATPase and HNHc) and tmn (P-loop NTPase). Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution. Asterisk (*) indicates cases of synergy.

DISCUSSION

In this work, we aimed to gain insights into the interactions between defence systems and the impacts of such interactions on bacterial immunity against phage predation both at cellular and population level. By comprehensive analysis of thousands *E. coli* genomes, we identified patterns of co-occurrence as well as mutual exclusion among defence systems. We showed that the co-occurrences are not conserved in more distant bacteria, suggesting the existence of many distinct defence strategies. Perhaps unexpectedly, the co-occurrence among defence systems was found not be strongly linked to their co-localisation in bacterial genomes, although we did identify several instances in which co-occurring systems co-localised. In several specific cases we explored, co-occurrence of both co-localising and non-co-localising pairs of systems was associated with synergistic interactions which led to a significantly greater protective effect against specific phages than expected from the sum of the effects of

individual systems. Notably, we observed a synergistic protective effect against certain phages even between some mutually exclusive defence systems, such as Zorya II and ietAS. Furthermore, we found that defence systems that are mutually exclusive in one bacterial order can co-occur in others, suggesting that the mutual exclusion between these systems is not caused by mechanistic incompatibility leading to negative epistasis. We also assessed the ecological implications of defence system synergy in short-term competition experiments, finding that bacterial populations carrying synergistic systems gained an evolutionary advantage over populations carrying any of the individual systems when targeted by specific phages that activated the synergistic defence.

Collectively, our results strongly suggest that interactions between defence systems are common and that non-random co-occurrence of defence systems in bacteria is an adaptive phenomenon driven by selection for enhanced immunity against specific phages. The defence strategies appear to differ substantially across bacterial taxa, most likely, driven by the species-specific viromes. Given the extensive horizontal transfer of defence systems, the evolution of bacterial defence strategies appear to follow the venerated general principle of adaptive evolution of microbes “everything is everywhere but the environment selects”[81].

We further investigated the molecular basis of the synergistic interactions between tmn and its co-occurring defence systems Gabija, Septu I, and PrrC, and uncovered a common mechanism of synergy, whereby tmn co-opted the sensory switch ATPase domains of the companion defence system, enhancing the anti-phage activity. These synergistic interactions and co-opting mechanisms likely play an important role in the evolution of defence systems and the emergence of multiple defence system variants, as observed for CRISPR-Cas[82], CBASS[19], Lamassu[6], Shield[83], and others. The modularity of anti-phage immunity is further evident in sharing of proteins and functional domains across distinct defence systems[84], such as HNH-endonuclease in Septu, Zorya II and type II CRISPR-Cas [2], NucC in CBASS III[85] and CRISPR-Cas type III[86], TOPRIM domain in Gabija[3], Wadjet[2], and PARIS[4], and P-loop NTPases, including helicases, in a broad diversity of defence systems including CRISPR-Cas

type I, Gabija, PrrC, tmn and many others[2,3,75]. Expanding our understanding of the mechanisms underlying the modularity of defence systems is expected to provide insights into their adaptive potential and evolutionary dynamics in the perennial arms race between bacteria and phages.

The results of this work underscore the importance of considering the interplay among defence system beyond their cumulative effect, taking into account the environmental context and the influence of phage pressure, for understanding prokaryotic immunity in its increasingly apparent complexity. Adaptation of prokaryotes to specific environments is driven by specific selective forces imposed by the virome, resulting in unique fitness landscapes in each niche. Thus, future research should aim to explore the broader patterns of defence system combinations in diverse prokaryotes in their specific ecological niches.

Limitations of the study

In this work, we investigated in detail the co-occurrence and mutual exclusion among the known defence systems only among isolates of a single bacterial species, *E. coli*. Our preliminary analysis of defence system co-occurrence in other bacterial orders revealed substantially different patterns emphasizing the importance of a broad exploration of bacterial immunity. Our mechanistic investigation of synergistic defence systems was obviously even more limited so that much further work is required to determine how general the domain co-option observed here might be and what other mechanisms contribute to the synergy.

ACKNOWLEDGEMENTS

S.K.G and E.V.K. are funded through the Intramural Research Program of the National Institutes of Health of the USA (National Library of Medicine). The work in F.L.N.'s group is supported by Wessex Medical Trust (AB03). We thank Dr Jennifer Mahony (University College Cork) for kindly providing phages JK16 and JK32. We also thank Fagenbank (Netherlands)

for providing phages T1, T3, T4, T7, and λ(vir), and Morgen Hedges (University of Southampton) for the phage drawings. We acknowledge the use of the IRIDIS High-Performance Computing Facility, and associated support services at the University of Southampton. We thank Victor Tobiasson for help with Alphafold, and Yuri Wolf and Sanasar Babajanyan for useful discussions.

AUTHOR CONTRIBUTIONS

Conceptualization, F.L.N.; Methodology, Y.W., S.K.G., Y.E.G., E.V.K., and F.L.N.; Formal Analysis, Y.W., S.K.G., A.v.d.H., and Y.O; Investigation, Y.W., S.K.G., A.v.d.H., C.A.M., S.K., Y.O., C.M.K., and T.T.; Resources, M.C., A.M., Y.E.G., E.V.K., and F.L.N.; Data Curation, Y.W., S.K.G., and F.L.N.; Writing – Original Draft, Y.W. S.K.G., E.V.K., and F.L.N.; Writing – Review & Editing, all authors; Visualization, Y.W., S.K.G., and F.L.N.; Supervision, Y.E.G., E.V.K., and F.L.N.; Funding Acquisition, E.V.K. and F.L.N.

DECLARATION OF INTERESTS

Y.E.G. is a full-time employee of SNIPR Biome.

FIGURE TITLES AND LEGENDS

Figure 1 Distribution of defence systems across *E. coli* phylogroups.

(A) A phylogenetic tree displaying 26,362 *E. coli* genomes obtained from the RefSeq database. Phylogroups are assigned as described previously[45], and each phylogroup is colour-coded according to the key.

(B) Number of defence systems found per *E. coli* strain in each phylogroup. The mean number of defence systems is indicated by a red line.

(C) Prevalence of defence systems in the *E. coli* genomes. The defence systems are organised from most prevalent (left) to least prevalent (right) and their total count is shown in the top bar graph. The bars are colour-coded according to the mechanism of the defence system. The remaining bar graphs show the prevalence, in percentage, of the defence systems per phylogroup.

Figure 2 Co-occurrence and mutual exclusion among defence systems in *E. coli*.

(A) Graphical representation of the co-occurrence analysis, depicting one pair of defence systems that co-occur (Gabija and tmn), and one pair that is mutually exclusive (Zorya II and ietAS). The nodes of the *E. coli* phylogenetic tree are coloured according to the presence or absence of the defence system in each strain. Their location in chromosome, plasmid, or prophage regions is indicated in the middle. For visualization purposes, only leaves that carry at least one system from the pair are shown.

(B) Co-occurrence of defence system pairs in *E. coli*. Co-occurring systems are shown in orange, and mutually exclusive systems are shown in green. The correlation between pairs was calculated with Pagel test for binary traits. Asterisks show correlations that were significant after the less stringent Benjamini-Hochberg correction (*) or the most stringent Bonferroni correction (**) for multiple testing. In the main text, the results after Bonferroni

correction are considered. The defence systems are colour-coded according to their broadly defined mechanism of defence.

(C,D) Distance histograms of (B) all and (C) co-occurring defence system pairs in 2,164 complete *E. coli* genomes. The median distance between genes encoding defence systems is shown by a red line. The analysis considered only those pairs that significantly co-occurred after Bonferroni correction.

Figure 3 Defence system pairs provide synergistic anti-phage activity**.**

(A) Experimental set-up for the assessment of the anti-phage activity of individual defence systems and their combinations. YFP, yellow fluorescent protein.

(B) Heatmap of synergy score of protection provided by selected defence system pairs against a panel of 29 phages. The synergy score is the epistatic coefficient for pairs of defence systems (see Methods). Null, EOP equivalent to the defence provide by one system; Additive, EOP corresponds to the combined defence of the two individual systems; Synergy, EOP exceeds the collective defence of the two systems. Data is shown as the average of three biological replicates. ** Statistically significant ($p < 0.01$).

(C) Time post infection assays, measuring T1 or T3 titres over the course of four hours in liquid cultures of *E. coli* containing individual or combined defence systems. Data is shown as the average and standard deviation of three biological replicates.

(**D**) Bacterial growth under phage predation at different multiplicities of infection (MOIs), represented as area under the curve (AUC) in OD·h. An defence system pair acts synergetic when its dot (red) is above the expected additive effect (blue). Data is shown as the confidence interval of three biological replicates. The raw data and growth curves used to calculate the AUCs are available on the associated Github and Zenodo databases.

Figure 4. Patterns of defence system co-occurrences across bacterial taxa.

Heatmap of defence system co-occurrence patterns in *E. coli* (n = 26,362) and in four bacterial orders: Enterobacterales including *E. coli* (n = 9,124), Bacillales (n = 3,952), Burkholderiales (n = 2,199), and Pseudomonales (n = 1,288). Grey squares indicate that at least one system in the pair is not present in the taxonomic group. * Co-occurrence significant after Benjamini-Hochberg correction; ** Co-occurrence significant after Bonferroni correction.

Figure 5 Synergistic defence system pairs provide an evolutionary advantage to bacteria**.**

(A) Set-up of experimental evolution assay of defence systems using chromogenic plasmids (left). Cells containing a single defence system carry a second plasmid that expresses a chromogenic reporter. Cells with defence system combinations were mixed in equal proportions with cells with single systems, and infected with phage. After 1, 2, and 3 days, cells were platted and the colonies of different colours were enumerated. A surface receptor assay (right) assessed the influence of receptor mutants on the outcome of the evolution assays, by subjecting colonies of different colours to phage infection in plaque assays.

(B) Prevalence of receptor mutants in the bacterial population during the evolution assay. The proportion of colonies with the wild-type receptor is represented as positive in the vertical axis, while the proportion of colonies with mutated receptor is shown as negative. Colonies with mutated receptor were identified by their complete resistance to phage infection in spot assays.

(C) Percentage of colonies from the evolution assay that carry each individual defence system or their combinations at 1, 2 and 3 days post infection with phage at low or high multiplicity of infection (MOI), compared to a non-infected control. Data is shown as the average of three biological and three technical replicates with individual counts shown. * p value < 0.01, ** p value < 0.001 determined by multiple comparison for nested one-way ANOVA, comparing to the corresponding uninfected control.

Figure 6 Mechanistic insight into the synergistic interaction between tmn and Gabija or Septu I.

(A) Whole-plasmid alignment of 104 plasmids containing tmn and Gabija from complete *E. coli* genomes with plasmid CP083423.1 as a reference (see Methods). The histogram shows the percentage of plasmids where the corresponding block from CP083423.1 was found. Annotated genes are coloured by function.

(B) Efficiency of plating (EOP) of phages T1 and 670 on cells expressing Gabija (G), tmn (T), Gabija and tmn (GT), and alanine mutants of specific functional domains. The mutations are organised by functional domains of Gabija (ATPase, TOPRIM, and UvrD-like) and tmn (P-loop NTPase). Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution. Asterisk (*) indicates cases of synergy.

(C) EOP of phages T1 and 670 on cells expressing tmn (T), Gabija (G), and tmn with either GajA or GajB. Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution.

(D) EOP of phages T1 and phi113 on cells expressing tmn (T), Septu I (S), PrrC (P), Septu I and tmn (ST), or PrrC and tmn (PT). Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution. Asterisk (*) indicates cases of synergy.

(E) EOP of phages T1 and 670 on cells expressing tmn (T), Septu (S), and Septu and tmn (ST), and variants with point mutations in specific functional domains. The mutations are organised by functional domains of Septu I (ATPase and HNHc) and tmn (P-loop NTPase). Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 was assumed at the corresponding dilution. Asterisk (*) indicates cases of synergy.

METHODS

Bacteria and phages

E. coli strain Dh5α was used to clone plasmids pACYCDuet-1 or 8A with individual defence systems. *E. coli* BL21-AI cells containing plasmid(s) with the defence systems were used for phage assays. All bacterial strains were grown at 37 °C in Lysogeny Broth (LB) with 180 rpm shaking for liquid cultures, or in LB agar (LBA) plates for solid cultures. Strains containing plasmid pACYCDuet-1 or 8A were grown in media supplemented with 25 µg/ml of chloramphenicol or 100 µg/ml of ampicillin, respectively. Phages used in this study and their origins are described in **Table S4**. All phages were produced in LB with their host strain, centrifuged, filter-sterilized, and stored as phage lysates at 4 °C.

Defence system detection

The FASTA amino acid (FAA), FAST Nucleic Acid (FNA), and Generic Feature Format (GFF) files of 26,384 *Escherichia coli* were downloaded from the NCBI Reference Sequence (RefSeq) Database. Complete genomes of 9,124 Enterobacterales, 1,288 Pseudomonadales, 3,952 Bacillales, and 2,199 Burkholderiales genomes were downloaded from Genbank in October 2021. Defence systems were detected in these genomes using PADLOC version 1.1.0 with database version 1.4.0[87] and DefenseFinder version 1.0.8[22]. Defence systems outputted by PADLOC in categories *other* and *adaptation* were removed from the analysis. Some of the *E. coli* genomes had extensive fragmentation, which negatively influenced the number of defence systems detected (two-sided Spearman, $p < 0.001$), but the effect was small enough ($r_s = -0.06$) that we opted to keep the fragmented genomes in our analysis. Defence systems found in all strains can be seen on **Table S1**.

Contig characterization and prophage detection

Platon version 1.6 [88] on accuracy mode was used to categorize bacterial contigs as plasmid or chromosome. Phigaro version 2.2.6[89] was used with default settings to detect prophages in

the bacterial genomes. Contigs shorter than 20 kbp were excluded from this analysis. Defence systems were considered to be located in a prophage region when at least one defence gene was fully within the prophage limits.

Phylogenetic analysis

Mash v2.3[90] distances were used to reconstruct phylogenetic trees for each dataset. Pairwise mash distances were calculated, employing a sketch size of 1,000 for the *E. coli* dataset and 100,000 for the order-level datasets. These pairwise distances were transformed into a distance matrix in phylip format, serving as input for the reconstruct of Neighbour-Joining phylogenetic trees with RapidNJ[91].

To remove potentially contaminated genomes from the phylogenetic trees, we implemented a three-step filtration process. First, we applied TreeShrink v 1.3.9[92] with centroid rerooting, using a quantile threshold of 0.1 for the *E. coli* dataset and 0.05 for the order-level datasets to remove leaves located on excessively long branches. Second, genomes with CheckM[93] contamination greater than 5 in the BV-BRC database (https://www.bv-brc.org/) were excluded. Lastly, leaves without metadata in RefSeq or Genbank (as of January 25, 2023), were removed from consideration due to potential errors in their genomic data. After filtration 26,362 genomes of *E. coli* were retained.

For further validation of the phylogenetic trees of the order level datasets, we colour-coded the leaves according to their genus level classification and visually confirmed their agreement with general taxonomy.

To additionally validate the phylogenetic tree for the *E. coli* dataset, we examined the proximity of samples from the same phylogroup the tree. We employed phylogroup assignments from a dataset of 10,667 *E. coli* genomes[45], and observed that all major clades contained genomes belonging to a single phylogroup. To extract clades containing samples from individual phylogroups, we first employed TreeCluster v 1.0.3[94] with a threshold parameter of 0.3, resulting in the division of the *E. coli* tree into 17 clusters. Subsequently, we further divided

phylogroups E1 and E2, and B1 and C using a custom R script. For each clade, we compiled the list of nodes belonging to that clade and performed phylogroup-specific analysis. All tree visualization and manipulations were performed using R v 4.1.2 with ggtree[95] and phytools[96].

Correlation between defence system content and phylogenetic distance

To investigate the relationship between defence system content and phylogenetic distance we performed a correlation analysis. The phylogenetic distance between pairs of genomes was determined using mash distances obtained as described above. To estimate the distance in defence system content between genomes, we employed Jaccard distance, which compares the presence and absence of defence systems in vectors. Defence systems present in less than 0.5% of genomes in the dataset were excluded from this analysis. For all unique pairs of genomes included in the analysis, we calculated Spearman correlation coefficients, and corresponding p-values, to quantify the correlation between phylogenetic distance and the dissimilarity in defence systems content.

Odds ratios for defence systems distribution between phylogroups

To test the hypothesis that defence systems are distributed unevenly among *E. coli* phylogroups, we performed a Chi-Squared test for homogeneity. The enrichment analysis aimed to assess whether the presence of a specific defence system in a particular phylogroup deviates from what would be expected by chance. This was quantified by calculating the odds ratio as the ratio of the observed number of genomes containing a particular system within a specific phylogroup to the expected number of genomes with that system.

Co-occurrence analysis

Given that the genomes under study exhibit phylogenetic relatedness, the co-occurrence analysis required a phylogenetically informed approach. We utilized the Pagel test for binary traits, as implemented in the R-package phytools with fitDiscrete model, to determine if pairs of defence systems exhibited non-random distribution across the phylogenetic tree, indicating potential interdependencies. To ensure robust results and avoid artefacts associated with

small sample sizes, systems that were present in less than 0.5% of genomes in the *E. coli* dataset and less than 1% in the order-level datasets were excluded from this analysis. Leaves carrying the defence system of interest were marked with ones, while leaves where the system was absent were marked with zeros. Prior to conducting the test, we standardized tree branch lengths to a mean branch length of 0.1, as recommended in the original implementation of the Pagel test outline in the BayesTraits manual.

We compared the results of the independent model with three alternative models: a) A model in which the distribution of two systems depends on each other; b) A model in which system A depends on system B; c) A model in which system B depends on system A. If model (a) produced a significant p (< 0.01), we further investigated which of the three models best fitted the observed data on the phylogenetic tree. For cases with significant p-value in model (a), we determined the directionality of interaction by using transition probabilities from the best fitted model. This involved two types of transition values: those assuming independent changes of states (e.g. transition from state (0,0) where no system is present, to state (1,0) where system B is present) and those assuming dependent changes of states (e.g. transition from state (0,1) to (1,1)). For each pair of transition values, we calculated the flux, e.g. for the transition from (1,0) to (1,1), by dividing the transition rate from (1,0) to (1,1) by the transition rate from (1,1) to (1,0). If the sum of fluxes into (1,1) exceeded the sum of fluxes from (0,0) to (1,0) and (0,0) to (0,1), we inferred that the systems co-occur; otherwise, they were considered mutually exclusive.

To correct for multiple testing, we used both the Bonferroni correction (the most stringent) and the Benjamini-Hochberg correction (less strict). Both sets of significant results after multiple testing correction were considered for downstream analysis, with a preference for those after Bonferoni correction due to their higher reliability.

Defence system cloning

The plasmids constructed in this work are listed in **Table S5**, and the primers used can be found in **Table S6**. Plasmid pACYCDuet-1 was modified to contain the pBAD promoter from plasmid 8A (MacroLab) by Gibson assembly. YFP, Druantia III (from *E. coli* ECOR19), ietAS (ECOR52), and tmn (ECOR25) were cloned into the modified pACYCDuet-1 by Gibson assembly. YFP, Kiwa (ECOR8), Gabija (ECOR49), and Zorya II (ECOR19) were cloned into plasmid 8A by Gibson assembly. Coral chromoproteins spisPink and meleRFP (Stanford Free Genes) were cloned into plasmids 8A and pACYC, respectively, by Gibson assembly. The plasmids were recovered in Dh5α cells, extracted using NucleoSpin Plasmid QuickPure Kit (Thermo Fisher Scientific) and confirmed by sequencing at Plasmidsaurus (USA). Mutations of the defence system operons were engineered by around-the-horn PCR, and confirmed by Sanger sequencing (Eurofins Genomics). Plasmids were transformed individually or in combinations into competent BL21-AI cells prepared using the Mix&Go! *E. coli* Transformation Kit (Zymo).

Efficiency of platting

Overnight cultures of the bacteria were diluted 1:50 in LB containing antibiotics, induced with 0.2% arabinose, and incubated for 5 hours before being used in double agar overlay assays. For this, bacterial cultures were mixed with 0.6% top agar and overlaid on LBA plates. Ten-fold serial dilutions of the phage stocks were spotted onto the bacterial lawn and the plates incubated overnight at 37 °C. The phage plaques were counted and used to calculate the EOP relative to the control. Epistatic coefficients, representing the interaction strength between defence systems, were determined as $|\text{Log10}\ (EOP_{system1+system2})| - |\text{Log10}\ (EOP_{system1})| - |\text{Log10}\ (EOP_{system2})|$. Synergy was considered when $|\text{Log10}\ (EOP_{system1+system2})| > |\text{Log10}\ (EOP_{system1})| + |\text{Log10}\ (EOP_{system2})| + 1$, additive effects were considered when $\text{Max}\ [|\text{Log10}\ (EOP_{system1})|, |\text{Log10}\ (EOP_{system2})|] < |\text{Log10}\ (EOP_{system1+system2})| < |\text{Log10}\ (EOP_{system1})| + |\text{Log10}\ (EOP_{system2})| + 1$, and antagonistic effects were considered when $|\text{Log10}\ (EOP_{system1+system2})| - \text{Max}\ (|\text{Log10}\ (EOP_{system1})|, |\text{Log10}\ (EOP_{system2})|] < -1$. Statistical significance was determined using the multiple comparison function from Two-way ANOVA with a p-value of <0.01.

Time post infection assay

Overnight bacterial cultures were diluted to an optical density at 600 nm of 0.1 in LB containing antibiotics and 0.2% arabinose. The cultures were infected with phage at an MOI of 0.0001. At 0, 1, 2, 3 and 4 hours post infection, a sample was taken and centrifuged at 12,000 × *g* for 2 minutes. The supernatant was serially diluted, and the phages were quantified by plaque assay on a bacterial lawn of cells with YFP. PFUs were counted after overnight incubation at 37 °C.

Liquid assay

Overnight bacterial cultures were diluted to an optical density at 584 nm of 0.25 in LB containing antibiotics and 0.2% arabinose. The bacterial suspension was distributed into wells of a 96-well plate to which phage dilutions or LB were added. The plates were incubated in a Fluostar Optima plate reader at 37 °C with shaking at 200 rpm, with optical density at 584 nm measured every 5 min for 24 h. To evaluate the impact of interactions between defence systems, we calculated the areas under the curve (AUC) for both individual systems and combination of systems. To calculate the AUC, the optical density at the start of the experiment was subtracted from each data point. If the AUC for a system combination exceeded the sum of the AUCs for the individual systems (the expected value), we considered those system as having a synergistic protective effect.

Short-term evolution experiment

Overnight bacterial cultures with single or double defence systems were diluted to an optical density at 600 nm of 0.1 in LB containing antibiotics and 0.2% arabinose, and mixed in equal proportions. The mixed cultures were infected with phage at an MOI of 0.0001, and incubated at 37 °C with shaking at 180 rpm for 24 h. A control without phage was used. The cultures were centrifuged at 8000 × *g* for 10 min, and the cell pellet was washed three times with LB at 12,000 × *g* for 2 min. Cells were resuspended in 1 ml of LB, serially diluted, and 100 µl of each dilution were spread onto LBA plates supplemented with antibiotics and 0.2% L-

arabinose. The plates were incubated overnight at 37 °C and the colonies of each colour counted. The experiment was repeated for 48h and 72h time points, using 50 µl of the previous cultures to inoculate fresh LB containing antibiotics and 0.2% arabinose, and challenging the cultures with phage at the same MOI.

Quantification of receptor mutants

Ten colonies were selected for each time point and condition of the short-term evolution experiment. The selected colonies were resuspended in 30 µl of sterile ddH_2O, and 5 µl of this cell suspension were spotted onto LBA plates. To assess the presence of receptor mutants, 2 µl of phage stock were spotted on top of the bacteria spots. The plates were left to incubate overnight at 37 °C. Colonies where no evidence of phage lysis was observed were identified and counted as receptor mutants.

Plasmid loss assay

Overnight bacterial cultures containing double defence systems were diluted to an optical density at 600 nm of 0.1 in LB containing 0.2% arabinose. The cultures were infected with phage at an MOI of 0.0001 and incubated at 37 °C with shaking at 180 rpm for 24h. A control without phage was used. Cultures were centrifuged at 8000 × *g* for 10 min, and the cell pellet was washed three times with LB at 12,000 × *g* for 2 min. Cells were resuspended in 1 ml of LB, serially diluted, and 100 µl of each dilution were spread onto LBA plates. The plates were incubated overnight at 37 °C and 96 colonies were picked, dissolved in LB, and streaked onto LBA plates with and without antibiotics, to determine the rate of plasmid loss. The experiment was repeated for 48h and 72h time points, using 50 µl of the previous cultures to inoculate fresh LB containing 0.2% arabinose, and challenging the cultures with phage at the same MOI.

Cactus plasmid analysis

E. coli plasmids containing both tmn and Gabija defence systems were extracted from the Enterobacterales dataset. The plasmid dataset comprised of a total of 104 plasmids from various *E. coli* strains. Next, we constructed a phylogenetic tree for these plasmids using the

Neighbour-Joining method as described above, using a mash sketch size of 1,000. Next, we generated whole length plasmid alignments using Cactus v 2.4.4[97] with default parameters and the phylogenetic tree described above as the guiding tree. The reference free alignment obtained was then transformed into multiple alignment format (MAF) using plasmid CP083423.1 as a reference. This MAF file was filtered and visualized in R. Alignment blocks smaller than 100 bp were excluded from the analysis. Alignment blocks were considered to be present in the plasmid if they exhibited a coverage above 50%.

DATA AVAILABILITY

All unique bacterial strains, phages, and plasmids generated in this study are available from the lead contact without restriction. Phages of SNIPR Biome are proprietary and can be shared with other non-competing parties upon written permission. Raw data have been deposited at Github (https://github.com/garushyants/synergy_bacterial_immune_systems/) and Zenodo (https://doi.org/10.5281/zenodo.10075784) and are publicly available as of the date of publication.

QUANTIFICATION AND STATISTICAL ANALYSIS

A two-sided binomial test was performed to determine if the observed co-occurrence of defence systems differed significantly from the expected co-occurrence, using R[98]. To correct for multiple testing we utilized both Benjamin-Hochberg and Bonferroni corrections, with a p-value < 0.001 considered significant. Unless stated otherwise, experimental data are presented as the mean of biological triplicates ± standard deviation. Statistical tests were performed using GraphPad Prism 9.2.0 and one sample t test or one-way ANOVA test. All statistical tests are described in detail in the corresponding chapters in Methods, and available as R code on Github and Zenodo.

SUPPLEMENTAL INFORMATION TITLES AND LEGENDS

Figure S1 Distribution of defence systems across *Escherichia coli* genomes.

(A) Phylogenetic tree of the *E. coli* genomes, with the count of defence systems per genome displayed and colour-coded based on their genomic locations. Each phylogroup is colour-coded according to the key.

(B) Distribution pattern of defence systems within the chromosome, plasmids, and prophage regions of *E. coli*.

(C) Density plot demonstrating the correlation between mash distance and similarity in defence systems content among the *E. coli* strains. A linear regression trend line is highlighted in red for reference.

Figure S2 Counts, genomic distance, and location of co-occurring defence systems in *E. coli*.

(A) Network of co-occurrence between defence systems. The width of the edge is proportional to the number of genomes in which the pair was observed, and the size of the node is proportional to the number of genomes where the specific defence system was found. Co-occurrences are shown in orange, while mutual exclusions are shown in green.

(B) Counts of co-occurrences and mutual exclusions for each defence system. Correlation between pairs was assessed using the Pagel test for binary traits. Significant correlations, adjusted with either the Benjamini-Hochberg or the Bonferroni method, are indicated by the colour scheme in the legend.

(C) Genomic distance between pairs of co-occurring defence systems.

(D) Distribution of co-occurring pairs across chromosome, plasmid, or prophage regions. Instances where each defence system resides in different locations are denoted as Plasmid-Chromosome, Prophage-Chromosome, and Prophage-Plasmid.

Figure S3 Synergistic anti-phage defence provided by combinations of defence systems.

(A) Efficiency of plating (EOP) of phages on strains carrying individual defence systems and their combinations. b.d., below limit of detection. Unfilled circles indicate instances where it was not possible to determine the number of phage plaques, hence a value of 1 at the respective dilution was assumed. Asterisks (*, $p< 0.01$; **, $p< 0.001$) indicate cases of synergy.

(B) Time post infection assays, measuring phage tires over the course of four hours in liquid cultures of *E. coli* containing individual or combined defence systems. Data is shown as the average and standard deviation of three biological replicates.

(C) Bacterial growth under phage predation at different multiplicities of infection (MOIs), represented as area under the curve (AUC) in OD·h. A defence system pair acts synergetic when its dot (red) is above the expected additive effect (blue). Data is shown as the confidence interval of three biological replicates. The raw data and growth curves used to calculate the AUCs are available on the associated Github and Zenodo databases.

Figure S4 Comparative analysis of defence systems across bacterial orders.

(A) Prevalence of defence systems in the bacterial orders Pseudomonadales, Enterobacterales, Burkholderiales, and Bacillales. The defence systems are colour-coded according to their mechanism of defence.

(B) Venn diagram illustrating the overlap in defence system presence among the four bacterial orders and *E. coli*.

Figure S5 Assessment of dual plasmid maintenance in *E. coli* using plasmid loss assays.

(A) Illustration outlining the plasmid loss assay procedure. Bacterial cultures harbouring two plasmids with one defence system each, were grown without antibiotic and assessed for plasmid persistence every 24h for three days. Plasmid maintenance was determined by spreading serial dilutions of the bacterial cultures onto LBA plates. Subsequently, 100 colonies were streaked on LBA with and without corresponding antibiotics. This assay was performed both in the absence and presence of phage.

(B) Proportion of colonies maintaining either one or both plasmids in the plasmid loss assay under normal conditions (control) and under phage predation.

Figure S6 Insights into tmn and its synergistic interaction with Gabija and Septu I.

(A) Tertiary structure of the tmn protein predicted by Alphafold2 (pLDDT 89.3). The figure depicts two chains of the potential hexamer complex. Light purple, P-loop NTPase; dark grey/beige, predicted transmembrane regions; pink, Walker A motif; purple, Walker B motif;

The P-loop NTPase domain is shown in light purple, the predicted transmembrane regions is shown in dark grey/beige. Residues mutated in the experiments are shown as spheres.

(B) Efficiency of plating (EOP) of phages T1, 670, and 678 on cells expressing tmn (T), Septu I (S), Gabija (G), Septu and tmn (ST), or Gabija and tmn (GT), derived from different strains. Unfilled circles represent cases where the number of phage plaques was indeterminable, with a value of 1 assumed at the respective dilution. Asterisks (**) denote cases of synergy.

Table S1 Defence systems found per strain of *E. coli*, Bacillales, Burkholderiales, Enterobacterales, and Pseudomonadales.

Table S2 Enrichment analysis of defence systems in *Escherichia coli* phylogroups, using the chi-squared test for homogeneity. Only systems exhibiting significant enrichment (p-value <0.001) are shown

Table S3 Co-occurrence analysis of defence systems in *E. coli*, Bacillales, Burkholderiales, Enterobacterales, and Pseudomonadales.

Table S4 List of phages used in this work.

Table S5 List of plasmids used and generated in this work.

Table S6 List of primers used in this work.

REFERENCES

1. Georjon, H., and Bernheim, A. (2023). The highly diverse antiphage defence systems of bacteria. Nat Rev Microbiol *21*, 686-700. 10.1038/s41579-023-00934-x.
2. Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., and Sorek, R. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. Science *359*, eaar4120. 10.1126/science.aar4120.
3. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E.V., and Zhang, F. (2020). Diverse enzymatic activities mediate antiviral immunity in prokaryotes. Science *369*, 1077-1084. 10.1126/science.aba0372.
4. Rousset, F., Depardieu, F., Miele, S., Dowding, J., Laval, A.-L., Lieberman, E., Garry, D., Rocha, E.P.C., Bernheim, A., and Bikard, D. (2022). Phages and their satellites encode hotspots of antiviral systems. Cell Host Microbe *30*, 740-753.e745. 10.1016/j.chom.2022.02.018.
5. Vassallo, C., Doering, C., Littlehale, M.L., Teodoro, G., and Laub, M.T. (2022). A functional selection reveals previously undetected anti-phage defence systems in the *E. coli* pangenome. Nat Rev Microbiology *7*, 1568-1579. 10.1038/s41564-022-01219-4.
6. Millman, A., Melamed, S., Leavitt, A., Doron, S., Bernheim, A., Hör, J., Lopatina, A., Ofir, G., Hochhauser, D., Stokar-Avihail, A., et al. (2022). An expanded arsenal of immune systems that protect bacteria from phages. Cell Host Microbe *30*, 1556-1569.e5. doi: 10.1016/j.chom.2022.09.017.
7. Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2013). Comparative genomics of defense systems in archaea and bacteria. Nucleic Acid Res *41*, 4360-4377. 10.1093/nar/gkt157.
8. Tock, M.R., and Dryden, D.T.F. (2005). The biology of restriction and anti-restriction. Cur Opin Microbiol *8*, 466-472. 10.1016/j.mib.2005.06.003.

9. Xiong, X., Wu, G., Wei, Y., Liu, L., Zhang, Y., Su, R., Jiang, X., Li, M., Gao, H., Tian, X., et al. (2020). SspABCD–SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities. Nat Microbiol *5*, 917-928. 10.1038/s41564-020-0700-6.
10. Hille, F., Richter, H., Wong, S.P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The Biology of CRISPR-Cas: Backward and Forward. Cell *172*, 1239-1259. 10.1016/j.cell.2017.11.032.
11. Bernheim, A., Millman, A., Ofir, G., Meitav, G., Avraham, C., Shomar, H., Rosenberg, M.M., Tal, N., Melamed, S., Amitai, G., and Sorek, R. (2021). Prokaryotic viperins produce diverse antiviral molecules. Nature *589*, 120-124. 10.1038/s41586-020-2762-2.
12. LeRoux, M., Srikant, S., Teodoro, G.I.C., Zhang, T., Littlehale, M.L., Doron, S., Badiee, M., Leung, A.K.L., Sorek, R., and Laub, M.T. (2022). The DarTG toxin-antitoxin system provides phage defence by ADP-ribosylating viral DNA. Nat Microbiol *7*, 1028-1040. 10.1038/s41564-022-01153-5.
13. Durmaz, E., and Klaenhammer, T.R. (2007). Abortive phage resistance mechanism AbiZ speeds the lysis clock to cause premature lysis of phage-infected Lactococcus lactis. J Bacteriol *189*, 1417-1425. 10.1128/jb.00904-06.
14. Parma, D.H., Snyder, M., Sobolevski, S., Nawroz, M., Brody, E., and Gold, L. (1992). The Rex system of bacteriophage lambda: tolerance and altruistic cell death. Genes Dev *6*, 497-510. 10.1101/gad.6.3.497.
15. Depardieu, F., Didier, J.-P., Bernheim, A., Sherlock, A., Molina, H., Duclos, B., and Bikard, D. (2016). A eukaryotic-like serine/threonine kinase protects Staphylococci against phages. Cell Host Microbe *20*, 471-481. 10.1016/j.chom.2016.08.010.
16. Zhang, T., Tamman, H., Coppieters ’t Wallant, K., Kurata, T., LeRoux, M., Srikant, S., Brodiazhenko, T., Cepauskas, A., Talavera, A., Martens, C., et al. (2022). Direct activation of a bacterial innate immune system by a viral capsid protein. Nature *612*, 132-140. 10.1038/s41586-022-05444-z.

17. Guegler, C.K., and Laub, M.T. (2021). Shutoff of host transcription triggers a toxin-antitoxin system to cleave phage RNA and abort infection. Mol Cell *81*, 2361-2373.e2369. 10.1016/j.molcel.2021.03.027.
18. Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., Oppenheimer-Shaanan, Y., and Sorek, R. (2020). Bacterial retrons function in anti-phage defense. Cell *183*, 1551-1561.e1512. 10.1016/j.cell.2020.09.065.
19. Millman, A., Melamed, S., Amitai, G., and Sorek, R. (2020). Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. Nat Microbiol *5*, 1608-1615. 10.1038/s41564-020-0777-y.
20. Tal, N., Millman, A., Stokar-Avihail, A., Fedorenko, T., Leavitt, A., Melamed, S., Yirmiya, E., Avraham, C., Brandis, A., Mehlman, T., et al. (2022). Bacteria deplete deoxynucleotides to defend against bacteriophage infection. Nat Microbiol *7*, 1200-1209. 10.1038/s41564-022-01158-0.
21. Ofir, G., Herbst, E., Baroz, M., Cohen, D., Millman, A., Doron, S., Tal, N., Malheiro, D.B.A., Malitsky, S., Amitai, G., and Sorek, R. (2021). Antiviral activity of bacterial TIR domains via immne signaling molecules. Nature *600*, 116-120. 10.1038/s41586-021-04098-7.
22. Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., and Bernheim, A. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. Nat Commun *13*, 2561. 10.1038/s41467-022-30269-9.
23. Piel, D., Bruto, M., Labreuche, Y., Blanquart, F., Goudenège, D., Barcia-Cruz, R., Chenivesse, S., Le Panse, S., James, A., Dubert, J., et al. (2022). Phage–host coevolution in natural populations. Nat Microbiol *7*, 1075-1086. 10.1038/s41564-022-01157-1.
24. Hochhauser, D., Millman, A., and Sorek, R. (2023). The defense island repertoire of the Escherichia coli pan-genome. PLoS Gen *19*, e1010694. 10.1371/journal.pgen.1010694.
25. Johnson, M.C., Laderman, E., Huiting, E., Zhang, C., Davidson, A., and Bondy-Denomy, J. (2023). Core defense hotspots within Pseudomonas aeruginosa are a consistent and

rich source of anti-phage defense systems. Nucleic Acids Res *51*, 4995-5005. 10.1093/nar/gkad317.

26. Makarova, K.S., Wolf, Y.I., Snir, S., and Koonin, E.V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. J Bacteriol *193*, 6039-6056. 10.1128/JB.05535-11.
27. Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. Nat Rev Microbiol *18*, 113-119. 10.1038/s41579-019-0278-2.
28. Rocha, E.P.C., and Bikard, D. (2022). Microbial defenses against mobile genetic elements and viruses: Who defends whom from what? PLoS Biol *20*, e3001514. 10.1371/journal.pbio.3001514.
29. Picton, D.M., Luyten, Y.A., Morgan, R.D., Nelson, A., Smith, Darren L., Dryden, D.T.F., Hinton, J.C.D., and Blower, Tim R. (2021). The phage defence island of a multidrug resistant plasmid uses both BREX and type IV restriction for complementary protection from viruses. Nucleic Acids Res *49*, 11257-11273. 10.1093/nar/gkab906.
30. Costa, A.R., van den Berg, D.F., Esser, J.Q., Muralidharan, A., van den Bossche, H., Estrada Bonilla, B., van der Steen, B.A., Haagsma, A.C., Nobrega, F.L., Haas, P.-J., and Brouns, S.J. (2022). Accumulation of defense systems in phage resistant strains of *Pseudomonas aeruginosa*. bioRxiv, 10.1101/2022.08.12.503731v2.
31. Bernheim, A., Bikard, D., Touchon, M., and Rocha, E.P.C. (2019). Atypical organizations and epistatic interactions of CRISPRs and cas clusters in genomes and their mobile genetic elements. Nucleic Acids Res *48*, 748-760. 10.1093/nar/gkz1091.
32. Li, M., Gong, L., Cheng, F., Yu, H., Zhao, D., Wang, R., Wang, T., Zhang, S., Zhou, J., Shmakov, S.A., et al. (2021). Toxin-antitoxin RNA pairs safeguard CRISPR-Cas systems. Science *372*, eabe5601. 10.1126/science.abe5601.
33. Shmakov, S.A., Barth, Z.K., Makarova, K.S., Wolf, Yuri I., Brover, V., Peters, Joseph E., and Koonin, Eugene V. (2023). Widespread CRISPR-derived RNA regulatory elements in CRISPR-Cas systems. Nucleic Acids Res. 10.1093/nar/gkad495.

34. Dupuis, M.-È., Villion, M., Magadán, A.H., and Moineau, S. (2013). CRISPR-Cas and restriction–modification systems are compatible and increase phage resistance. Nat Commun *4*, 2087. 10.1038/ncomms3087.

35. Clermont, O., Dixit, O.V.A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahmias, A., Denamur, E., and Gordon, D. (2019). Characterization and rapid identification of phylogroup G in Escherichia coli, a lineage with high virulence and antibiotic resistance potential. Environ Microbiol *21*, 3107-3117. 10.1111/1462-2920.14713.

36. Clermont, O., Christenson, J.K., Denamur, E., and Gordon, D.M. (2013). The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ Microbiol Rep *5*, 58-65. 10.1111/1758-2229.12019.

37. Yu, D., Banting, G., and Neumann, N.F. (2021). A review of the taxonomy, genetics, and biology of the genus Escherichia and the type species *Escherichia coli*. Can J Microbiol *67*, 553-571. 10.1139/cjm-2020-0508.

38. Keen, E.C. (2015). A century of phage research: Bacteriophages and the shaping of modern biology. BioEssays *37*, 6-9. 10.1002/bies.201400152.

39. Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010). The population genetics of commensal Escherichia coli. Nat Rev Microbiol *8*, 207-217. 10.1038/nrmicro2298.

40. Escobar-Páramo, P., Clermont, O., Blanc-Potard, A.-B.a., Bui, H., Le Bouguénec, C., and Denamur, E. (2004). A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. Mol Biol Evol *21*, 1085-1094. 10.1093/molbev/msh118.

41. Dadi, B.R., Abebe, T., Zhang, L., Mihret, A., Abebe, W., and Amogne, W. (2020). Distribution of virulence genes and phylogenetics of uropathogenic *Escherichia coli* among urinary tract infection patients in Addis Ababa, Ethiopia. BMC Infect Dis *20*, 108. 10.1186/s12879-020-4844-z.

42. Nowrouzian, F.L., Wold, A.E., and Adlerberth, I. (2005). *Escherichia coli* strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. J Infect Dis *191*, 1078-1083. 10.1086/427996.

43. Boroumand , M., Naghmachi, M., and Ghatee, M.A. (2021). Detection of phylogenetic groups and drug resistance genes of *Escherichia coli* causing urinary tract infection in southwest Iran. Jundishapur J Microbiol *14*, e112547. 10.5812/jjm.112547.
44. Touchon, M., Charpentier, S., Clermont, O., Rocha, E.P.C., Denamur, E., and Branger, C. (2011). CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. J Bacteriol *193*, 2460-2467. 10.1128/jb.01307-10.
45. Abram, K., Udaondo, Z., Bleker, C., Wanchai, V., Wassenaar, T.M., Robeson, M.S., 2nd, and Ussery, D.W. (2021). Mash-based analyses of *Escherichia coli* genomes reveal 14 distinct phylogroups. Commun Biol *4*, 117. 10.1038/s42003-020-01626-5.
46. Ledvina, H.E., Ye, Q., Gu, Y., Sullivan, A.E., Quan, Y., Lau, R.K., Zhou, H., Corbett, K.D., and Whiteley, A.T. (2023). An E1–E2 fusion protein primes antiviral immune signalling in bacteria. Nature *616*, 319-325. 10.1038/s41586-022-05647-4.
47. Thiaville, J.J., Kellner, S.M., Yuan, Y., Hutinet, G., Thiaville, P.C., Jumpathong, W., Mohapatra, S., Brochier-Armanet, C., Letarov, A.V., Hillebrand, R., et al. (2016). Novel genomic island modifies DNA with 7-deazaguanine derivatives. Proc Natl Acad Sci *113*, E1452-E1459. 10.1073/pnas.1518570113.
48. Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-Amikam, Y., Afik, S., Ofir, G., and Sorek, R. (2015). BREX is a novel phage resistance system widespread in microbial genomes. EMBO J *34*, 169-183. 10.15252/embj.201489455.
49. Isaev, A., Drobiazko, A., Sierro, N., Gordeeva, J., Yosef, I., Qimron, U., Ivanov, N.V., and Severinov, K. (2020). Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defence. Nucleic Acids Res *48*, 5397-5406. 10.1093/nar/gkaa290.
50. Gordeeva, J., Morozova, N., Sierro, N., Isaev, A., Sinkunas, T., Tsvetkova, K., Matlashov, M., Truncaitė, L., Morgan, R.D., Ivanov, N.V., et al. (2018). BREX system of Escherichia coli distinguishes self from non-self by methylation of a specific DNA site. Nucleic Acids Res *47*, 253-265. 10.1093/nar/gky1125.

51. Vassallo, C.N., Doering, C.R., Littlehale, M.L., Teodoro, G.I.C., and Laub, M.T. (2022). A functional selection reveals previously undetected anti-phage defence systems in the E. coli pangenome. Nat Microbiol *7*, 1568-1579. 10.1038/s41564-022-01219-4.
52. Pingoud, A., Wilson, G.G., and Wende, W. (2014). Type II restriction endonucleases--a historical perspective and more. Nucleic Acids Res *42*, 7489-7527. 10.1093/nar/gku447.
53. Xu, T., Yao, F., Zhou, X., Deng, Z., and You, D. (2010). A novel host-specific restriction system associated with DNA backbone S-modification in Salmonella. Nucleic Acids Research *38*, 7133-7141. 10.1093/nar/gkq610.
54. Wu, D., Tang, Y., Chen, S., He, Y., Chang, X., Zheng, W., Deng, Z., Li, Z., Wang, L., Wu, G., and Chen, S. (2022). The functional coupling between restriction and DNA phosphorothioate modification systems underlying the DndFGH restriction complex. Nat Catal *5*, 1131-1144. 10.1038/s41929-022-00884-2.
55. Popa, O., Hazkani-Covo, E., Landan, G., Martin, W., and Dagan, T. (2011). Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. Genome Res *21*, 599-609. 10.1101/gr.115592.110.
56. Andam, C.P., and Gogarten, J.P. (2011). Biased gene transfer in microbial evolution. Nat Rev Microbiol *9*, 543-555. 10.1038/nrmicro2593.
57. Soucy, S.M., Huang, J., and Gogarten, J.P. (2015). Horizontal gene transfer: building the web of life. Nat Rev Gen *16*, 472-482. 10.1038/nrg3962.
58. Puigbò, P., Wolf, Y.I., and Koonin, E.V. (2010). The Tree and Net Components of Prokaryote Evolution. Genome Biol Evol *2*, 745-756. 10.1093/gbe/evq062.
59. Popa, O., Landan, G., and Dagan, T. (2017). Phylogenomic networks reveal limited phylogenetic range of lateral gene transfer by transduction. ISME J *11*, 543-554. 10.1038/ismej.2016.116.
60. Iranzo, J., Puigbò, P., Lobkovsky, A.E., Wolf, Y.I., and Koonin, E.V. (2016). Inevitability of Genetic Parasites. Genome Biol Evol *8*, 2856-2869. 10.1093/gbe/evw193.
61. Koonin, E. (2016). Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. F1000Research *5*. 10.12688/f1000research.8737.1.

62. Redondo-Salvo, S., Fernández-López, R., Ruiz, R., Vielva, L., de Toro, M., Rocha, E.P.C., Garcillán-Barcia, M.P., and de la Cruz, F. (2020). Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. Nat Commun *11*, 3602. 10.1038/s41467-020-17278-2.

63. Liljeruhm, J., Funk, S.K., Tietscher, S., Edlund, A.D., Jamal, S., Wistrand-Yuen, P., Dyrhage, K., Gynnå, A., Ivermark, K., Lövgren, J., et al. (2018). Engineering a palette of eukaryotic chromoproteins for bacterial synthetic biology. J Biol Eng *12*, 8. 10.1186/s13036-018-0100-0.

64. Alieva, N.O., Konzen, K.A., Field, S.F., Meleshkevitch, E.A., Hunt, M.E., Beltran-Ramirez, V., Miller, D.J., Wiedenmann, J., Salih, A., and Matz, M.V. (2008). Diversity and evolution of coral fluorescent proteins. PLoS One *3*, e2680. 10.1371/journal.pone.0002680.

65. Gurney, J., Aldakak, L., Betts, A., Gougat-Barbera, C., Poisot, T., Kaltz, O., and Hochberg, M.E. (2017). Network structure and local adaptation in co-evolving bacteria–phage interactions. Mol Ecol *26*, 1764-1777. 10.1111/mec.14008.

66. Cheng, R., Huang, F., Lu, X., Yan, Y., Yu, B., Wang, X., and Zhu, B. (2023). Prokaryotic Gabija complex senses and executes nucleotide depletion and DNA cleavage for antiviral defense. Cell Host Microbe. 10.1016/j.chom.2023.06.014.

67. Antine, S.P., Johnson, A.G., Mooney, S.E., Leavitt, A., Mayer, M.L., Yirmiya, E., Amitai, G., Sorek, R., and Kranzusch, P.J. (2023). Structural basis of Gabija anti-phage defense and viral immune evasion. bioRxiv, 2023.2005.2001.538945. 10.1101/2023.05.01.538945.

68. Cheng, R., Huang, F., Wu, H., Lu, X., Yan, Y., Yu, B., Wang, X., and Zhu, B. (2021). A nucleotide-sensing endonuclease from the Gabija bacterial defense system. Nucleic Acids Res *49*, 5216-5229. 10.1093/nar/gkab277.

69. Takahashi, H., Shao, M., Furuya, N., and Komano, T. (2011). The genome sequence of the incompatibility group Iγ plasmid R621a: Evolution of IncI plasmids. Plasmid *66*, 112-121. 10.1016/j.plasmid.2011.06.004.

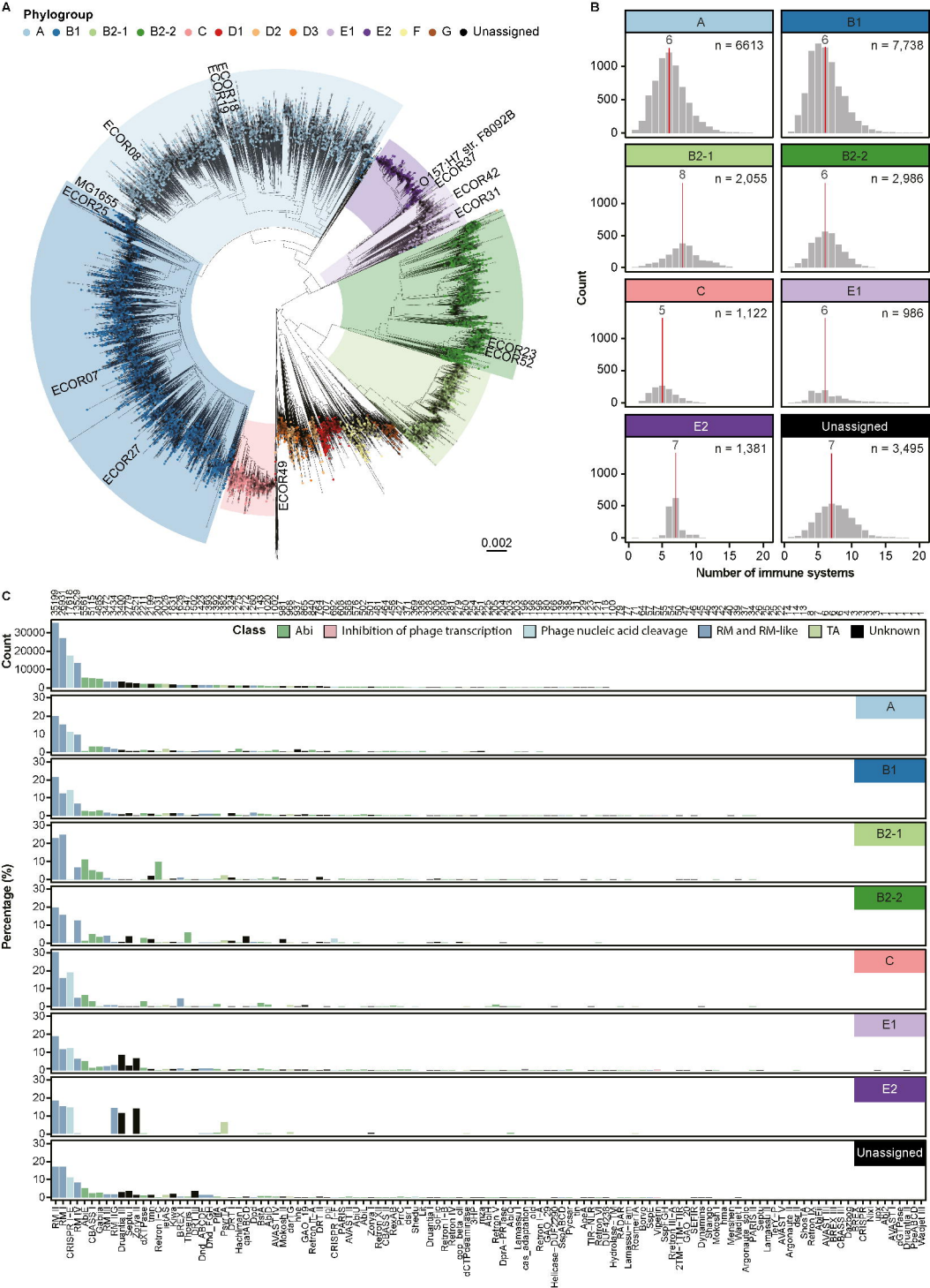
70. Venturini, C., Hassan, K.A., Roy Chowdhury, P., Paulsen, I.T., Walker, M.J., and Djordjevic, S.P. (2013). Sequences of two related multiple antibiotic resistance virulence plasmids sharing a unique IS26-related molecular signature isolated from different *Escherichia coli* pathotypes from different hosts. PLoS One *8*, e78862. 10.1371/journal.pone.0078862.
71. Samuel, B., and Burstein, D. (2023). A diverse repertoire of anti-defense systems is encoded in the leading region of plasmids. bioRxiv, 2023.2002.2015.528439. 10.1101/2023.02.15.528439.
72. Aravind, L., Iyer, L.M., Leipe, D.D., and Koonin, E.V. (2004). A novel family of P-loop NTPases with an unusual phyletic distribution and transmembrane segments inserted within the NTPase domain. Genome Biol *5*, R30. 10.1186/gb-2004-5-5-r30.
73. Zhang, G., Li, S., Cheng, K.-W., and Chou, T.-F. (2021). AAA ATPases as therapeutic targets: Structure, functions, and small-molecule inhibitors. Eur J Med Chem *219*, 113446. 10.1016/j.ejmech.2021.113446.
74. Deng, Y.-M., Liu, C.-Q., and W. Dunn, N. (1999). Genetic organization and functional analysis of a novel phage abortive infection system, AbiL, from Lactococcus lactis. J Biotechnol *67*, 135-149. 10.1016/S0168-1656(98)00175-8.
75. Uzan, M., and Miller, E.S. (2010). Post-transcriptional control by bacteriophage T4: mRNA decay and inhibition of translation initiation. Virol J *7*, 360. 10.1186/1743-422X-7-360.
76. Millman, A., Melamed, S., Leavitt, A., Doron, S., Bernheim, A., Hör, J., Garb, J., Bechon, N., Brandis, A., Lopatina, A., et al. (2022). An expanded arsenal of immune systems that protect bacteria from phages. Cell Host Microbe *30*, 1556-1569.e1555. 10.1016/j.chom.2022.09.017.
77. Azam, A.H., Chihara, K., Kondo, K., Nakamura, T., Ojima, S., Tamura, A., Yamashita, W., Cui, L., Takahashi, Y., Watashi, K., and Kiga, K. (2023). Viruses encode tRNA and anti-retron to evade bacterial immunity. bioRxiv, 2023.2003.2015.532788. 10.1101/2023.03.15.532788.

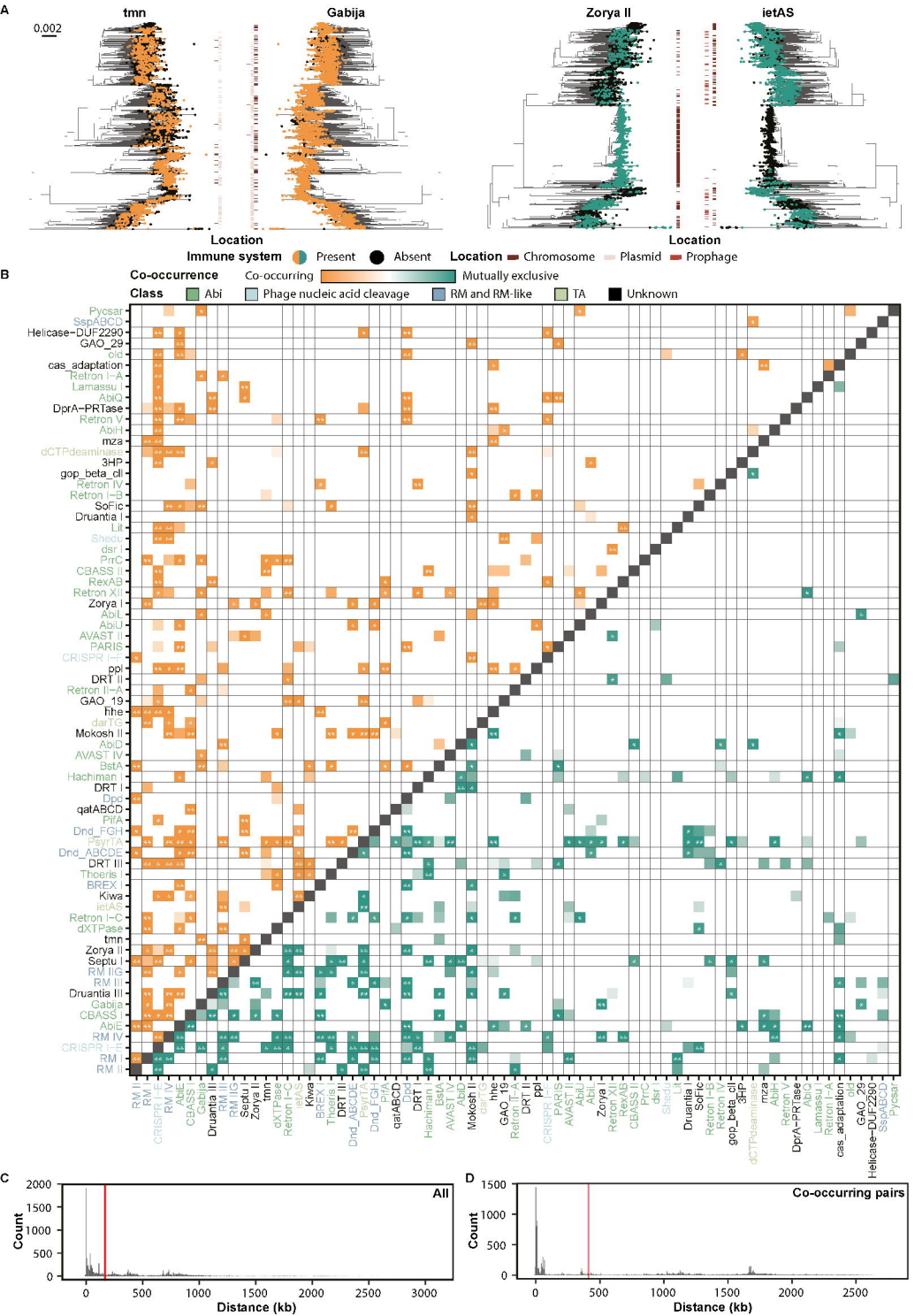
78. Blanga-Kanfi, S., Amitsur, M., Azem, A., and Kaufmann, G. (2006). PrrC-anticodon nuclease: functional organization of a prototypical bacterial restriction RNase. Nucleic Acids Res *34*, 3209-3219. 10.1093/nar/gkl415.
79. Krishnan, A., Burroughs, A.M., Iyer, L.M., and Aravind, L. (2020). Comprehensive classification of ABC ATPases and their functional radiation in nucleoprotein dynamics and biological conflict systems. Nucleic Acids Res *48*, 10045-10075. 10.1093/nar/gkaa726.
80. Russell, C.W., and Mulvey, M.A. (2015). The extraintestinal pathogenic *Escherichia coli* factor RqlI constrains the genotoxic effects of the RecQ-like helicase RqlH. PLoS Pathog *11*, e1005317. 10.1371/journal.ppat.1005317.
81. Becking, L.G.M.B. (1934). Geobiologie of inleiding tot de milieukunde (W.P. Van Stockum & Zoon).
82. Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., et al. (2020). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol *18*, 67-83. 10.1038/s41579-019-0299-x.
83. Macdonald, E., Wright, R., Connolly, J.P.R., Strahl, H., Brockhurst, M., van Houte, S., Blower, T.R., Palmer, T., and Mariano, G. (2023). The novel anti-phage system Shield co-opts an RmuC domain to mediate phage defense across Pseudomonas species. PLOS Gen *19*, e1010784. 10.1371/journal.pgen.1010784.
84. Mariano, G., and Blower, T.R. (2023). Conserved domains can be found across distinct phage defence systems. Mol Microbiol *120*, 45-53. 10.1111/mmi.15047.
85. Lau, R.K., Ye, Q., Birkholz, E.A., Berg, K.R., Patel, L., Mathews, I.T., Watrous, J.D., Ego, K., Whiteley, A.T., Lowey, B., et al. (2020). Structure and mechanism of a cyclic trinucleotide-activated bacterial endonuclease mediating bacteriophage immunity. Mol Cell *77*, 723-733.e726. 10.1016/j.molcel.2019.12.010.
86. Mayo-Muñoz, D., Smith, L.M., Garcia-Doval, C., Malone, L.M., Harding, K.R., Jackson, S.A., Hampton, H.G., Fagerlund, R.D., Gumy, L.F., and Fineran, P.C. (2022). Type III

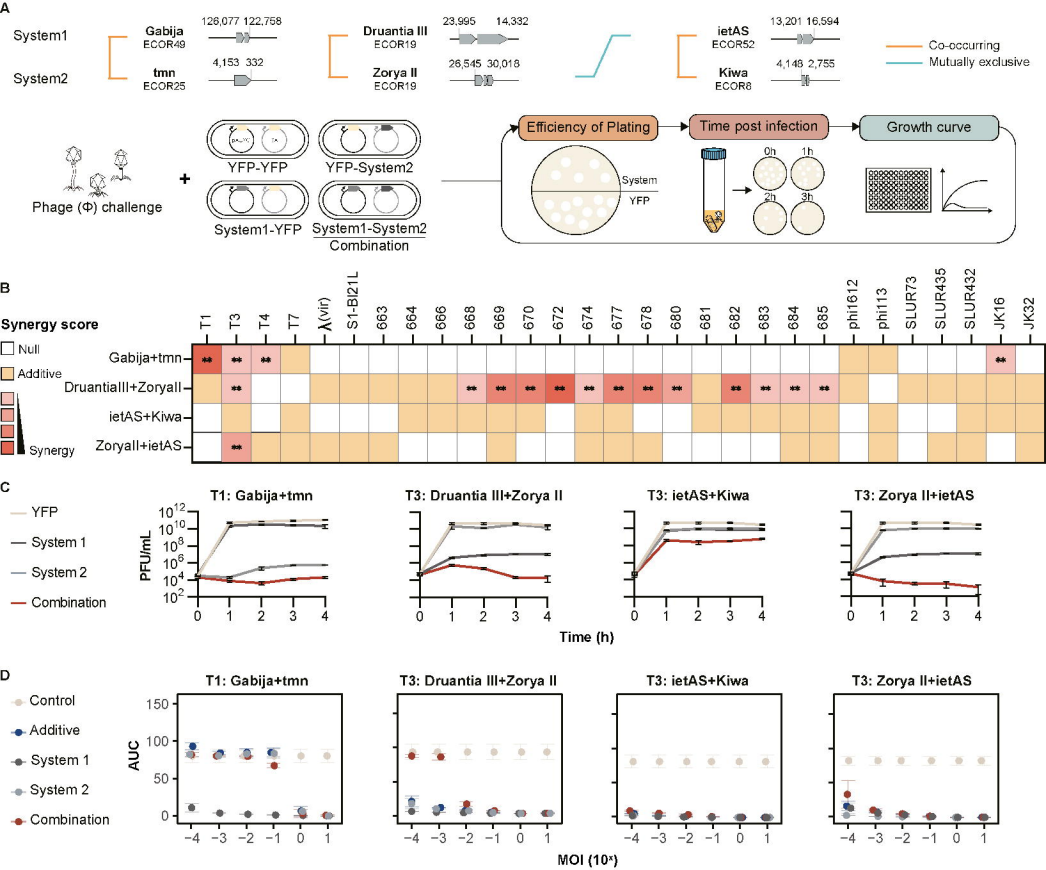
CRISPR-Cas provides resistance against nucleus-forming jumbo phages via abortive infection. Mol Cell *82*, 4471-4486.e4479. 10.1016/j.molcel.2022.10.028.

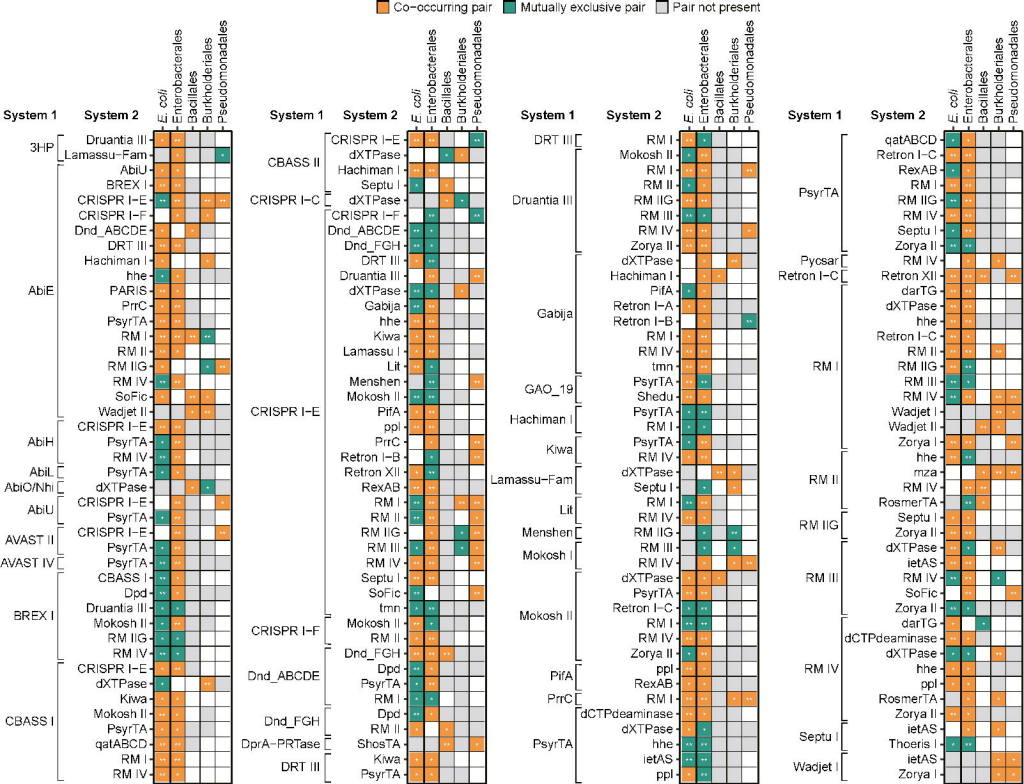
87. Payne, L.J., Todeschini, T.C., Wu, Y., Perry, B.J., Ronson, Clive W., Fineran, Peter C., Nobrega, Franklin L., and Jackson, Simon A. (2021). Identification and classification of antiviral defence systems in bacteria and archaea with PADLOC reveals new system types. Nucleic Acids Res *49*, 10868-10878. 10.1093/nar/gkab883.
88. Schwengers, O., Barth, P., Falgenhauer, L., Hain, T., Chakraborty, T., and Goesmann, A. (2020). Platon: identification and characterization of bacterial plasmid contigs in short-read draft assemblies exploiting protein sequence-based replicon distribution scores. Microb Genom *6*. 10.1099/mgen.0.000398.
89. Starikova, E.V., Tikhonova, P.O., Prianichnikov, N.A., Rands, C.M., Zdobnov, E.M., Ilina, E.N., and Govorun, V.M. (2020). Phigaro: high-throughput prophage sequence annotation. Bioinformatics *36*, 3882-3884. 10.1093/bioinformatics/btaa250.
90. Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., and Phillippy, A.M. (2016). Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol *17*, 132. 10.1186/s13059-016-0997-x.
91. Simonsen, M., Mailund, T., and Pedersen, C.N.S. (2008). Rapid Neighbour-Joining. Berlin, Heidelberg, (Springer Berlin Heidelberg), pp. 113-122.
92. Mai, U., and Mirarab, S. (2018). TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. BMC Genomics *19*, 272. 10.1186/s12864-018-4620-2.
93. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res *25*, 1043-1055. 10.1101/gr.186072.114.
94. Balaban, M., Moshiri, N., Mai, U., Jia, X., and Mirarab, S. (2019). TreeCluster: Clustering biological sequences using phylogenetic trees. PLoS One *14*, e0221068. 10.1371/journal.pone.0221068.

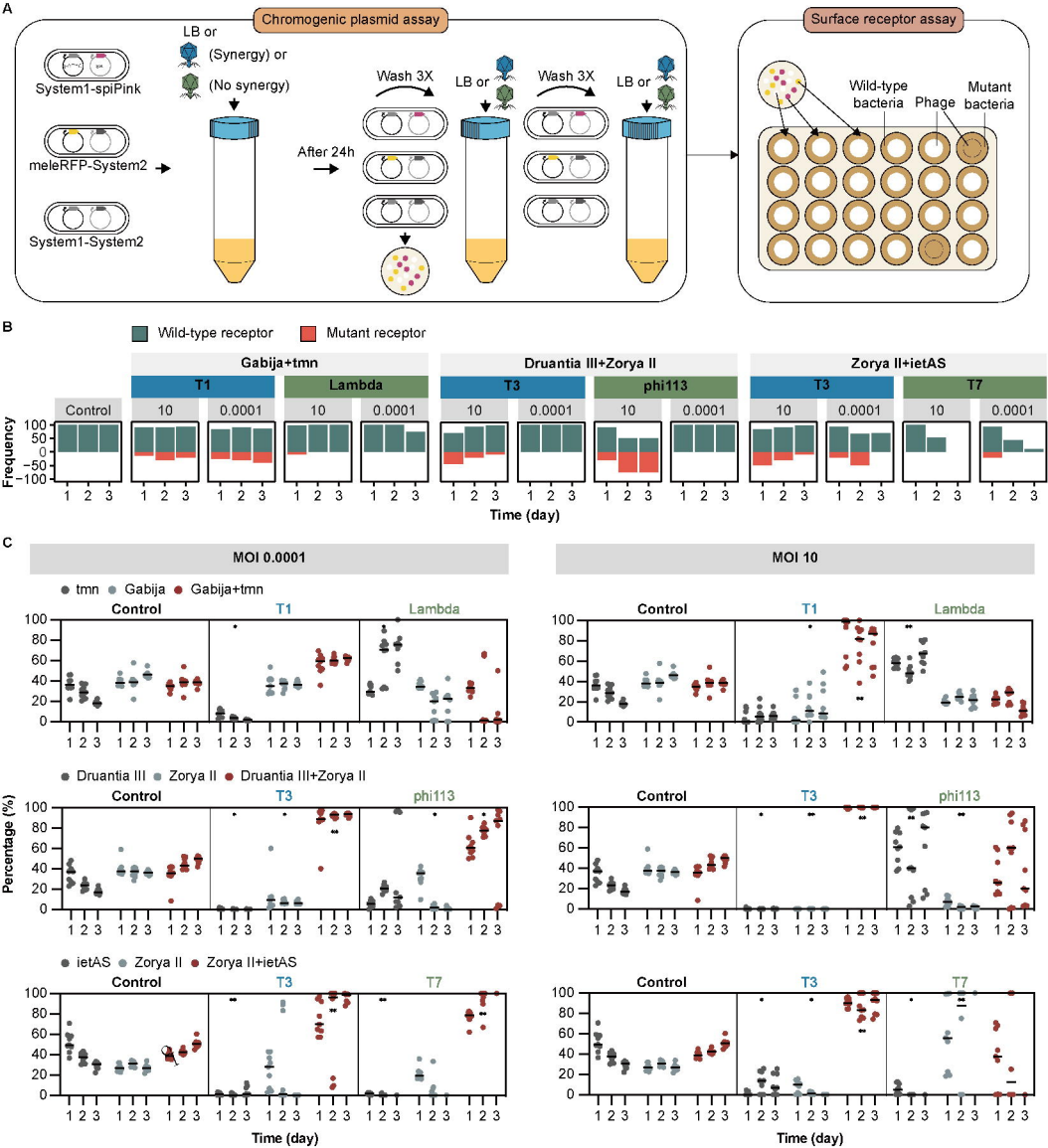
95. Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.-Y. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol *8*, 28-36. 10.1111/2041-210X.12628.
96. Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol *3*, 217-223. 10.1111/j.2041-210X.2011.00169.x.
97. Armstrong, J., Hickey, G., Diekhans, M., Fiddes, I.T., Novak, A.M., Deran, A., Fang, Q., Xie, D., Feng, S., Stiller, J., et al. (2020). Progressive Cactus is a multiple-genome aligner for the thousand-genome era. Nature *587*, 246-251. 10.1038/s41586-020-2871-y.
98. R Core Team (2018). R: A language and environment for statistical computing (R Foundation for Statistical Computing).

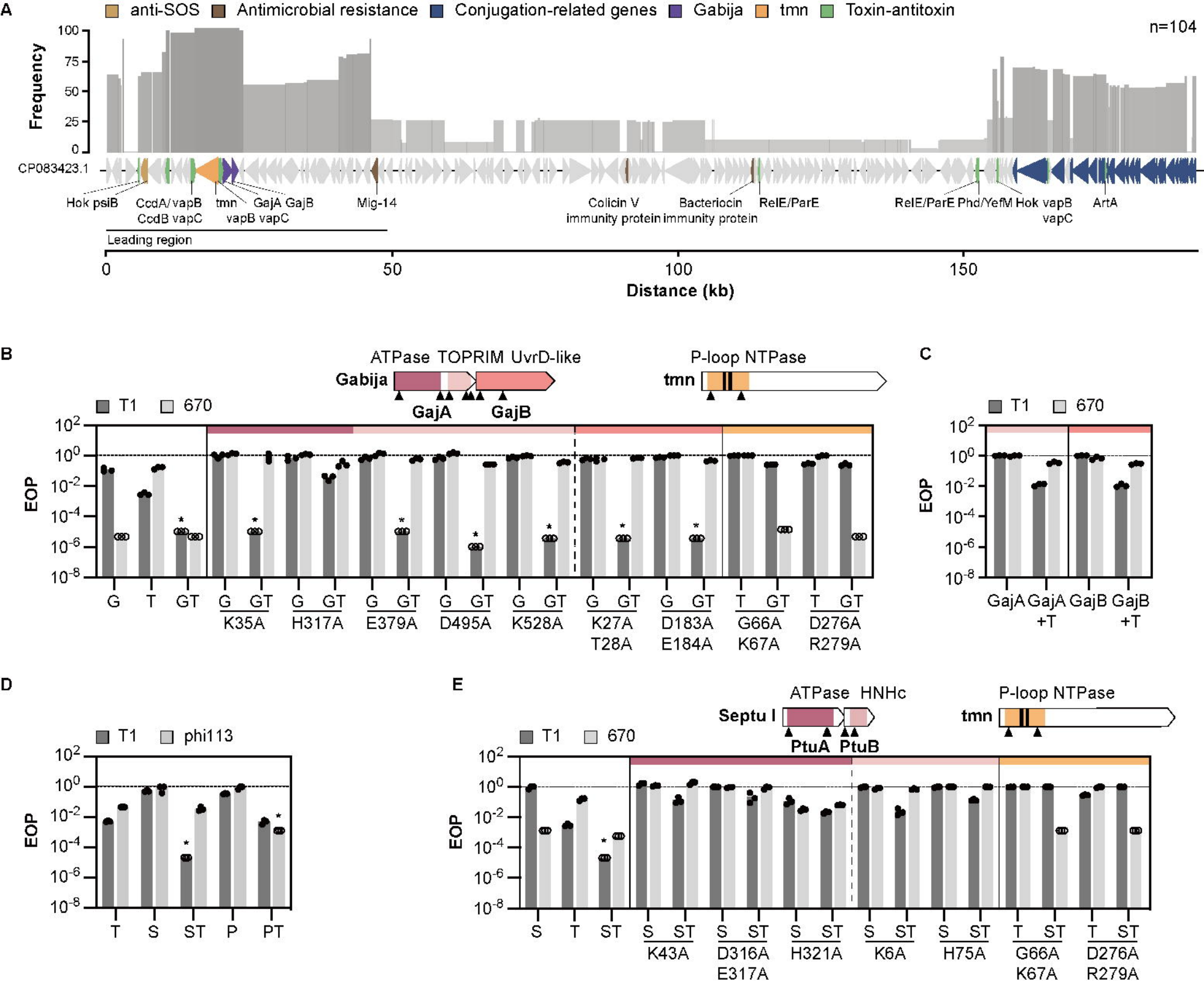

A
Phylogroup
A B1 B2-1 B2-2 C D1 D2 D3 E1 E2 F G Unassigned
ECOR18
ECOR19
ECOR08
MG1655
ECOR25
ECOR07
ECOR27
ECOR49
O157:H7 str. F8092B
ECOR37
ECOR42
ECOR31
ECOR23
ECOR52
0.002
B
A
n = 6613
B1
n = 7,738
B2-1
n = 2,055
B2-2
n = 2,986
C
n = 1,122
E1
n = 986
E2
n = 1,381
Unassigned
n = 3,495
Count
Number of immune systems
C
Class
Abi
Inhibition of phage transcription
Phage nucleic acid cleavage
RM and RM-like
TA
Unknown
Count
Percentage (%)
A
B1
B2-1
B2-2
C
E1
E2
Unassigned


A
tmn
Gabija
Zorya II
ietAS
0.002
Location
Location
Immune system
Present
Absent
Location
Chromosome
Plasmid
Prophage
B
Co-occurrence
Co-occurring
Mutually exclusive
Class
Abi
Phage nucleic acid cleavage
RM and RM-like
TA
Unknown
Pycsar
SspABCD
Helicase−DUF2290
GAO_29
old
cas_adaptation
Retron I−A
Lamassu I
AbiQ
DprA−PRTase
Retron V
AbiH
mza
dCTPdeaminase
3HP
gop_beta_cdi
Retron IV
Retron I−B
SoFic
Druantia I
Lit
Shedu
dsr I
PrrC
CBASS II
RexAB
Retron XII
Zorya I
AbiL
AbiU
AVAST II
PARIS
CRISPR I−C
ppl
DRT I
Retron II−A
GAO_19
hhe
darTG
Mokosh II
AbiD
AVAST IV
BstA
Hachiman
DRT I
Dpd
qatABCD
PifA
Dnd_FGH
PsyrTA
Dnd_ABCDE
DRT III
Thoeris I
BREX I
Kiwa
ietAS
Retron I−C
dXTPase
tmn
Zorya II
Septu I
RM IIG
RM III
Druantia III
Gabija
CBASS I
AbiE
RM IV
CRISPR I−E
RM I
RM II
RM II
RM
CRISPR I−E
RM IV
AbiE
CBASS
Gabija
Druantia II
RM III
RM IIG
Septu
Zorya II
tmn
dXTPase
Retron I−C
Kiwa
ietAS
BREX
Thoeris
DRT II
Dnd_ABCDE
PsyrTA
Dnd_FGH
PifA
qatABCD
DRT
Hachiman
BstA
AVAST IV
AbiD
Mokosh II
darTG
hhe
GAO_19
Retron II−A
DRT II
ppl
CRISPR I−E
PARIS
AVAST II
AbiU
AbiL
Zorya I
Retron XII
RexAB
CBASS I
PrrC
dsr I
Shedu
Lit
Druantia I
SoFic
Retron I−B
Retron IV
gop_beta_cdi
3HP
dCTPdeaminase
mza
AbiH
Retron V
DprA−PRTase
AbiQ
Lamassu I
Retron I−A
cas_adaptation
old
GAO_29
Helicase−DUF2290
SspABCD
Pycsar
C
All
Count
2000
1500
1000
500
0
0
500
1000
1500
2000
2500
3000
Distance (kb)
D
Co-occurring pairs
Count
1500
1000
500
0
0
500
1000
1500
2000
2500
Distance (kb)


A
System1
System2
Gabija
ECOR49
tmn
ECOR25
126,077 122,758
4,153 332
Druantia III
ECOR19
Zorya II
ECOR19
23,995 14,332
26,545 30,018
ietAS
ECOR52
Kiwa
ECOR8
13,201 16,594
4,148 2,755
Co-occurring
Mutually exclusive
Phage (Φ) challenge
YFP-YFP
YFP-System2
System1-YFP
System1-System2
Combination
Efficiency of Plating
Time post infection
Growth curve
System
YFP
0h
1h
2h
3h
B
Synergy score
Null
Additive
Synergy
T1
T3
T4
T7
λ(vir)
S1-BI21L
663
664
666
668
669
670
672
674
677
678
680
681
682
683
684
685
phi1612
phi113
SLUR73
SLUR435
SLUR432
JK16
JK32
Gabija+tmn
DruantiaIII+ZoryaII
ietAS+Kiwa
ZoryaII+ietAS
C
YFP
System 1
System 2
Combination
T1: Gabija+tmn
T3: Druantia III+Zorya II
T3: ietAS+Kiwa
T3: Zorya II+ietAS
PFU/mL
Time (h)
D
Control
Additive
System 1
System 2
Combination
AUC
MOI (10ˣ)


Co-occurring pair
Mutually exclusive pair
Pair not present
System 1
System 2
E. coli
Enterobacterales
Bacillales
Burkholderiales
Pseudomonadales
3HP
Druantia III
Lamassu-Fam
AbiU
BREX I
CRISPR I-E
CRISPR I-F
Dnd_ABCDE
DRT III
Hachiman I
hhe
AbiE
PARIS
PrrC
PsyrTA
RM I
RM II
RM IIG
RM IV
SoFic
Wadjet II
CRISPR I-E
AbiH
PsyrTA
RM IV
AbiL
PsyrTA
AbiO/Nhi
dXTPase
CRISPR I-E
AbiU
PsyrTA
CRISPR I-E
AVAST II
PsyrTA
AVAST IV
PsyrTA
CBASS I
Dpd
Druantia III
Mokosh II
RM IIG
RM IV
CRISPR I-E
dXTPase
Kiwa
Mokosh II
PsyrTA
qatABCD
RM I
RM IV
CBASS II
CRISPR I-C
CRISPR I-E
dXTPase
Hachiman I
Septu I
dXTPase
CRISPR I-F
Dnd_ABCDE
Dnd_FGH
DRT III
Druantia III
dXTPase
Gabija
hhe
Kiwa
Lamassu I
Lit
Menshen
Mokosh II
PifA
ppl
PrrC
Retron I-B
Retron XII
RexAB
RM I
RM II
RM IIG
RM III
RM IV
Septu I
SoFic
tmn
CRISPR I-F
Mokosh II
RM II
Dnd_FGH
Dnd_ABCDE
Dpd
PsyrTA
RM I
Dnd_FGH
Dpd
RM II
DprA-PRTase
ShosTA
DRT III
Kiwa
PsyrTA
DRT III
RM I
Druantia III
Mokosh II
RM I
RM II
RM IIG
RM III
RM IV
Zorya II
dXTPase
Hachiman I
Gabija
PifA
Retron I-A
Retron I-B
RM I
RM IV
GAO_19
tmn
PsyrTA
Shedu
Hachiman I
PsyrTA
RM I
Kiwa
PsyrTA
RM IV
Lamassu-Fam
dXTPase
Septu I
Lit
RM I
RM IV
Menshen
RM IIG
Mokosh I
RM III
RM IV
Mokosh II
dXTPase
PsyrTA
Retron I-C
RM I
RM IV
Zorya II
PifA
ppl
RexAB
PrrC
RM I
PsyrTA
dCTPdeaminase
dXTPase
hhe
ietAS
ppl
PsyrTA
qatABCD
Retron I-C
RexAB
RM I
RM IIG
RM IV
Septu I
Zorya II
Pycsar
RM IV
Retron I-C
Retron XII
darTG
dXTPase
hhe
RM I
Retron I-C
RM II
RM IIG
RM III
RM IV
Wadjet I
Wadjet II
Zorya I
RM II
hhe
mza
RM IV
RosmerTA
RM IIG
Septu I
Zorya II
dXTPase
RM III
ietAS
RM IV
SoFic
Zorya II
darTG
dCTPdeaminase
dXTPase
RM IV
hhe
ppl
RosmerTA
Zorya II
Septu I
ietAS
Thoeris I
Wadjet I
ietAS
Zorya I


A
Chromogenic plasmid assay
System1-spiPink
meleRFP-System2
System1-System2
LB or
(Synergy) or
(No synergy)
After 24h
Wash 3X
LB or
Wash 3X
LB or
Surface receptor assay
Wild-type bacteria
Phage
Mutant bacteria
B
Wild-type receptor
Mutant receptor
Gabija+tmn
Druantia III+Zorya II
Zorya II+ietAS
T1
Lambda
T3
phi113
T3
T7
Control
10
0.0001
Frequency
Time (day)
C
MOI 0.0001
MOI 10
tmn
Gabija
Gabija+tmn
Druantia III
Zorya II
Druantia III+Zorya II
ietAS
Zorya II
Zorya II+ietAS
Control
T1
Lambda
T3
phi113
T7
Percentage (%)
Time (day)


A
anti-SOS
Antimicrobial resistance
Conjugation-related genes
Gabija
tmn
Toxin-antitoxin
n=104
Frequency
100
75
50
25
0
CP083423.1
Hok psiB
CcdA/ vapB
CcdB vapC
tmn
vapB vapC
GajA GajB
Mig-14
Colicin V
immunity protein
Bacteriocin
immunity protein
RelE/ParE
RelE/ParE Phd/YefM Hok vapB
vapC
ArtA
Leading region
0
50
100
150
Distance (kb)
B
ATPase TOPRIM UvrD-like
Gabija
GajA
GajB
P-loop NTPase
tmn
T1
670
EOP
G
T
GT
K35A
H317A
E379A
D495A
K528A
K27A
T28A
D183A
E184A
G66A
K67A
D276A
R279A
C
GajA
GajA +T
GajB
GajB +T
D
phi113
T
S
ST
P
PT
E
ATPase HNHc
Septu I
PtuA
PtuB
K43A
D316A
E317A
H321A
K6A
H75A