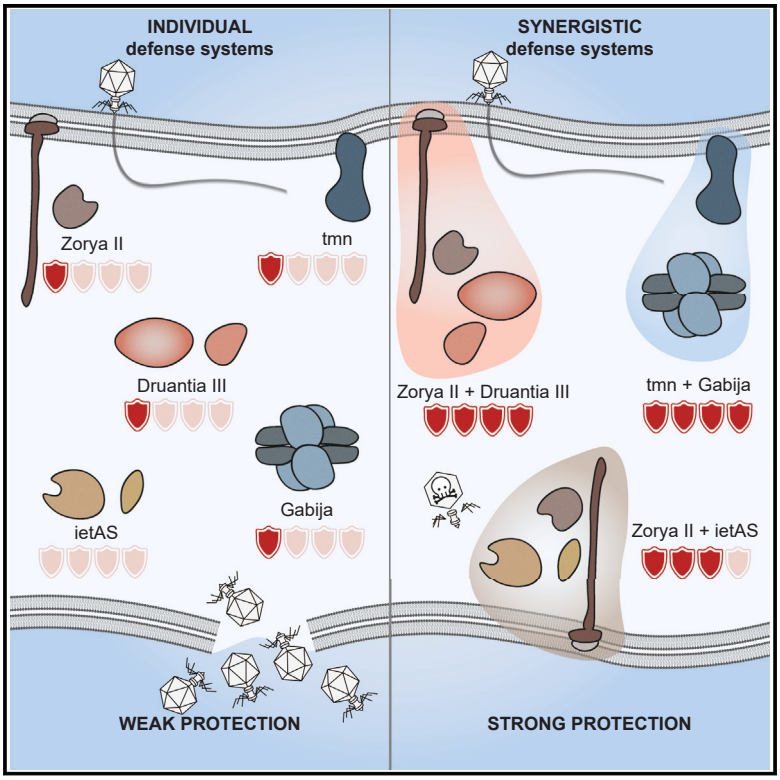
Article

Cell Host & Microbe

Bacterial defense systems exhibit synergistic anti-phage activity

Graphical abstract



Highlights

- Co-occurring bacterial defense systems display synergistic anti-phage activity
- Zorya II synergizes with Druantia III and ietAS, and tmn synergizes with Gabija and Septu I
- Tmn synergizes with defense systems containing sensory switch ATPase domains
- Active systems recruit functional domains of inactive systems for enhanced efficacy

Authors

Yi Wu, Sofya K. Garushyants, Anne van den Hurk, ..., Yilmaz Emre Gençay, Eugene V. Koonin, Franklin L. Nobrega

Correspondence

f.nobrega@soton.ac.uk

In brief

Bacteria defend against phages using a variety of defense systems, yet their interactions are poorly understood. Wu and Garushyants et al. reveal that these defense systems are generally compatible and, in some instances, interact resulting in synergistic anti-phage effects, conferring an evolutionary advantage on bacteria under specific environmental conditions.

Wu et al., 2024, Cell Host & Microbe *32*, 557–572
April 10, 2024
https://doi.org/10.1016/j.chom.2024.01.015

CellPress

Cell Host & Microbe

CellPress
OPEN ACCESS

Article

Bacterial defense systems exhibit synergistic anti-phage activity

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] and Franklin L. Nobrega[1,8,*]

[1]School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, UK
[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 and Science (BITS), Pilani, Rajasthan, India
[4]Wellcome Centre for Cell-Matrix Research, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
[5]Department of Genetics and Genome Biology, University of Leicester, Leicester, UK
[6]SNIPR biome, Copenhagen, Denmark
[7]These authors contributed equally
[8]Lead contact
*Correspondence: f.nobrega@soton.ac.uk
https://doi.org/10.1016/j.chom.2024.01.015

SUMMARY

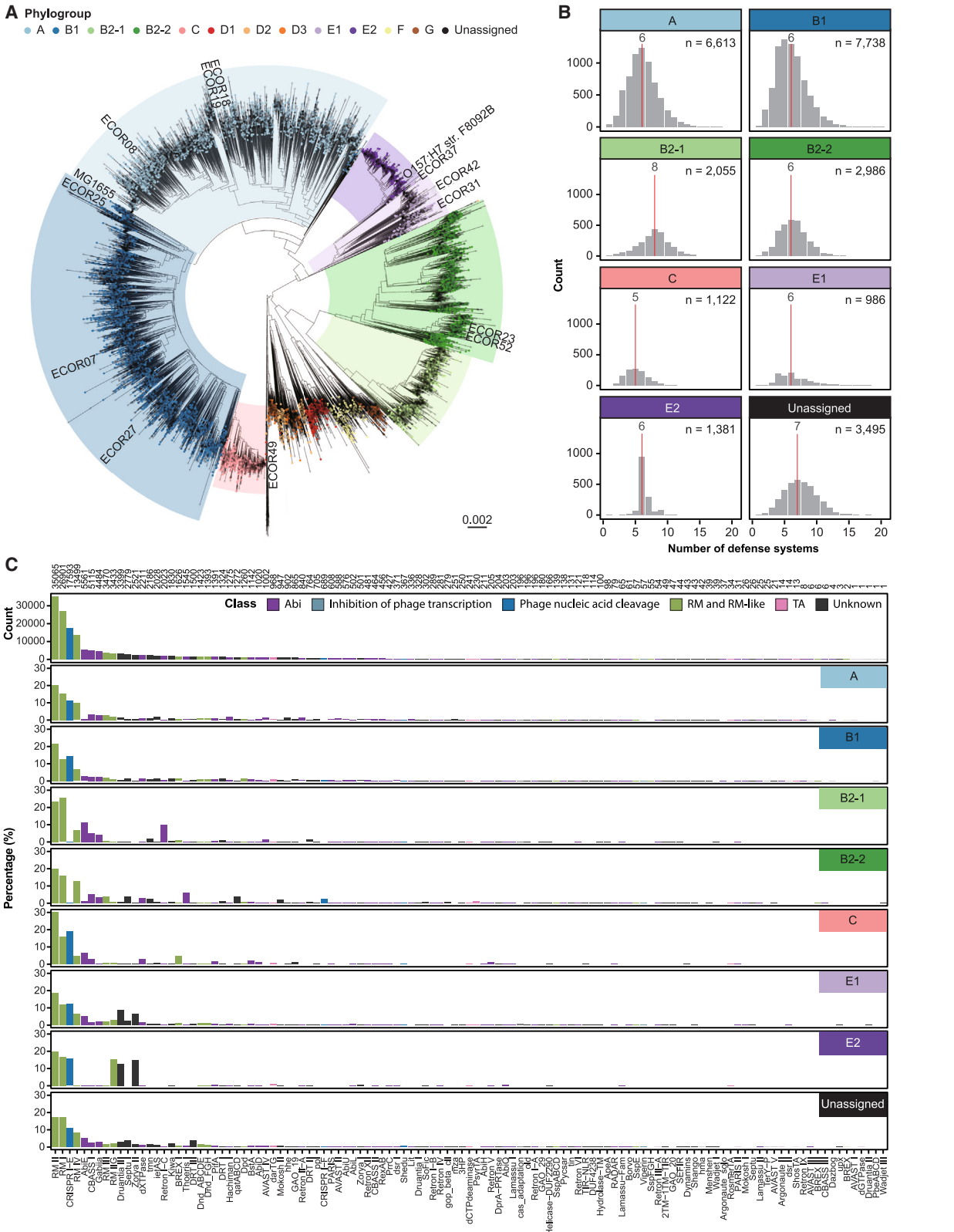
Bacterial defense against phage predation involves diverse defense systems acting individually and concurrently, yet their interactions remain poorly understood. We investigated >100 defense systems in 42,925 bacterial genomes and identified numerous instances of their non-random co-occurrence and negative association. For several pairs of defense systems significantly co-occurring in *Escherichia coli* strains, we demonstrate synergistic anti-phage activity. Notably, Zorya II synergizes with Druantia III and ietAS defense systems, while tmn exhibits synergy with co-occurring systems Gabija, Septu I, and PrrC. For Gabija, tmn co-opts the sensory switch ATPase domain, enhancing anti-phage activity. Some defense system pairs that are negatively associated in *E. coli* show synergy and significantly co-occur in other taxa, demonstrating that bacterial immune repertoires are largely shaped by selection for resistance against host-specific phages rather than negative epistasis. Collectively, these findings demonstrate compatibility and synergy between defense systems, allowing bacteria to adopt flexible strategies for phage defense.

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 defense systems that widely differ in protein composition and modes of action.[2–7] The mechanisms employed by bacterial defense 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 defense systems, the mechanism of action remains unknown.

A bacterial genome carries, on average, about five distinct (currently identifiable) defense 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 defense islands, sometimes associated with mobile genetic elements (MGEs) integrated into distinct hotspots in the bacterial genome.[24–26] Defense systems are believed to undergo frequent horizontal transfer between bacteria, and close proximity of the respective genes could facilitate simultaneous transfer of multiple systems.[27]

Despite the recent burst of bacterial defense system discovery, the causes of their clustering in defense islands remain poorly understood. It has been argued that co-localization of defense systems in MGEs and the resulting joint horizontal gene transfer (HGT) could provide fitness advantages to recipient bacteria, especially in phage-rich environments.[28] Additionally, it has been suggested that synergistic interactions between defense systems could drive their co-localization and favor their joint transfer,[29,30] as supported by the conservation of certain sets of defense systems.[31] For example, CRISPR-Cas systems of different subtypes often co-occur and the CRISPR arrays interact with Cas proteins across different systems.[32] Furthermore, toxin-antitoxin (TA) RNA pairs[33] and possibly other TA modules[34] safeguard CRISPR immunity by making cells dependent on CRISPR-Cas for survival. CRISPR-Cas and restriction-modification (RM) systems,[35] as well as BREX and the restriction enzyme BrxU,[30] co-occur resulting in expanded phage protection. However, these



(legend on next page)

examples of interaction between bacterial defense systems notwithstanding, their co-occurrence in bacteria and the connections between co-occurrence and co-localization in bacterial genomes have not been analyzed 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 defense islands evolve neutrally through a preferential attachment process whereby multiple defense 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 defense 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 defense systems in anti-phage immunity. Our findings show that defense system co-occurrence varies substantially across *E. coli* phylogroups and taxa and is not directly related to their co-localization in the genome. For several pairs of non-randomly co-occurring and negatively associated defense systems in *E. coli*, we experimentally demonstrated synergistic interactions that provided an evolutionary advantage to the bacterial population. Moreover, some of the defense systems that are negatively associated 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 defense system repertoire in bacteria.

RESULTS

Distinct defense system repertoires across *E. coli* phylogroups

To explore the variation in the immune repertoires among closely related bacteria, we analyzed the defense system content in a comprehensive dataset of 26,362 *E. coli* genomes from the NCBI Reference Sequence (RefSeq) database.[36,37] *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-characterized, complete genomes as well as a large panel of well-studied phages.[38,39] We found that, in agreement with previous observations, on average, *E. coli* genomes carry 5–7 defense systems, but some clades, such as those in phylogroup B2-1, harbor a greater diversity of such systems (Figure S1A; Table S1). The majority of the defense systems are encoded in the chromosomes, but there are some clades where additional systems, especially Gabija, tmn, PifA, ppl, and AbiQ 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 and B2) or domestic/wild animal microbiomes (B1)[40] (Figure 1A). Our analysis revealed significant differences in the number (Figure 1B, Wilcoxon two-sided test, $p < 2e{-}16$) and types (Figure 1C, chi-squared test for homogeneity, $p < 0.001$) of defense systems among the phylogroups. Phylogroup B2-1, which includes extra-intestinal pathogenic (ExPEC) strains,[41] was particularly noteworthy, with the highest average number of defense systems (8) among the examined phylogroups (Figure 1B). The genomes in phylogroup B2 accumulate virulence factors[42,43] as well as antibiotic resistance genes,[44] 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[45] (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 = 5.11), Thoeris I (OR = 6.17), Septu I (OR = 3.17), PsyrTA (OR=7.19), 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 = 4.87 and 6.92, respectively) compared with the other phylogroups. Phylogroup E2 additionally showed 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 ($r = 0.33–0.65$, $p < 2.2 \times 10^{-16}$) between the types of defense systems found in *E. coli* genomes within phylogroups and the genetic relatedness of these genomes, with the exception of phylogroups A and B1 ($r = 0.02–0.12$, $p < 2.22 \times 10^{-16}$), as indicated by the mash distance analysis (Figure S1C). These observations indicate that, although HGT is an important route of defense 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 defense systems across *E. coli* genomes demonstrates associations between specific defense systems and individual phylogroups, likely driven by selection for sets of defense mechanisms capable of efficiently protecting the bacteria against the specific repertoires of phages and other MGEs that they encounter in their respective environments.

117 pairs of defense systems co-occur in *E. coli*

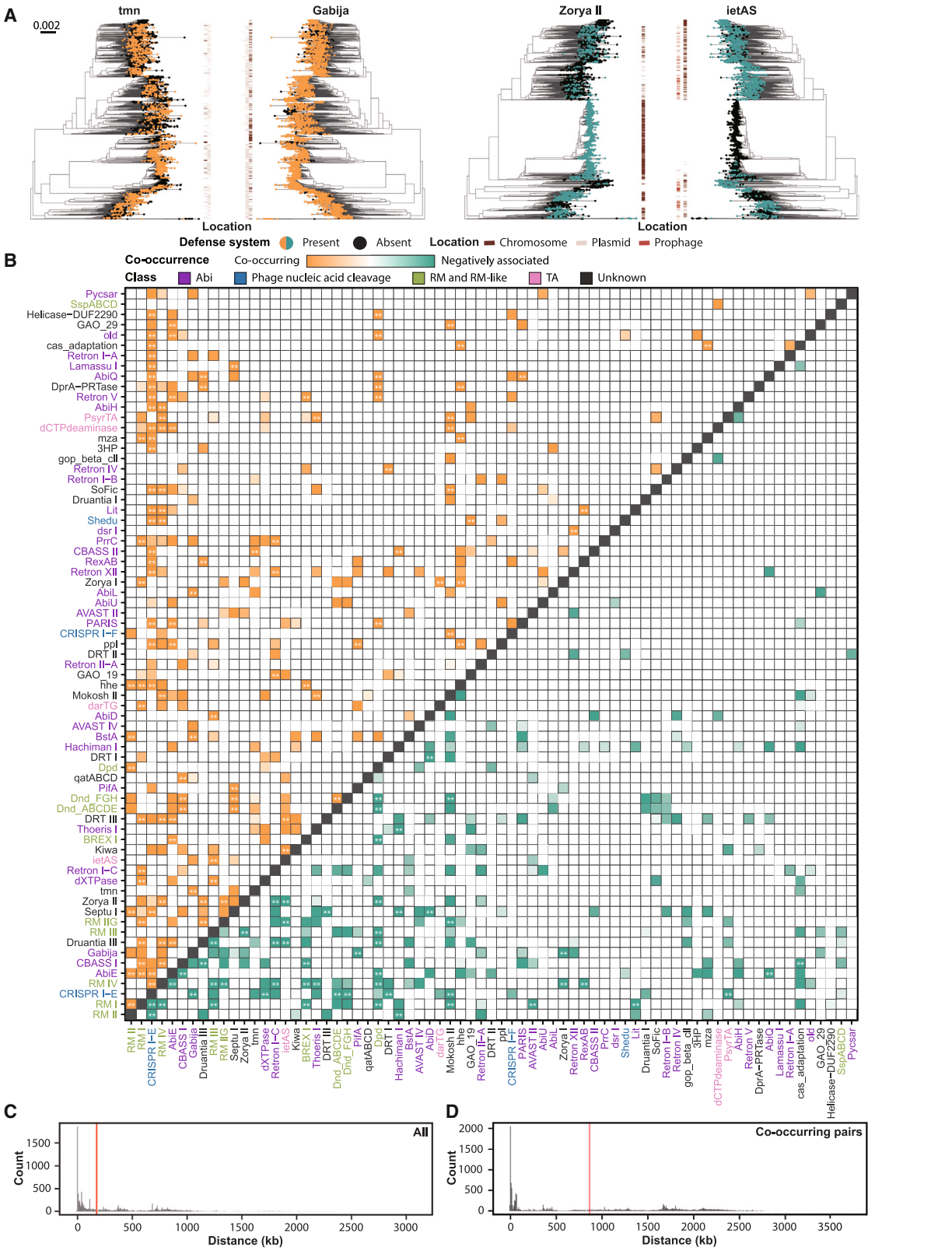
In previous studies, some defense systems have been shown to interact, resulting in enhanced or expanded protection against

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

(A) A phylogenetic tree displaying 26,362 *E. coli* genomes obtained from the RefSeq database. Phylogroups are color-coded according to the key.

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

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



(legend on next page)

phages.[30,32–35] Here, we sought to determine whether specific defense systems co-occurred more frequently than expected in bacterial genomes, potentially indicating interactions between different defense mechanisms enhancing protection against phages. To this end, we explored correlations between the occurrences of pairs of defense systems in *E. coli* genomes corrected for phylogenetic bias; we deemed such a correction to be essential because, as shown above, vertical inheritance of defense systems is common (see STAR Methods). This analysis allowed us to identify pairs of defense systems that appear together in the same genome significantly more (co-occurring) or less (negatively associated) frequently than expected based on their individual prevalence (Table S3; Figure 2A).

Our analysis revealed that 171 interacting pairs of defense systems (6.8% of all the analyzed pairs) were significantly correlated, positively or negatively (Pagel test for binary traits, with Bonferroni correction for multiple tests). Of these, 117 pairs were co-occurrences (68.4% of the correlated pairs), and the rest were cases of negative association. With the more permissive Benjamini-Hochberg correction, 265 pairs of defense systems (10.7% of the analyzed pairs) were significantly correlated, of which 211 (79.6%) were co-occurrences (Figure 2B). Notably, although the network of co-occurrences and especially of exclusions between the *E. coli* defense systems was sparse (Figures 2 and 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 defense 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 Septu I, 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, 856 of the 947 instances of Mokosh II co-occurred with RM IV, and 2,148 of the 2,518 instances of Zorya II co-occurred with Druantia III (Figure 2A; Table S3).

Notably, different subtypes of the same defense system displayed distinct co-occurrence patterns with other systems (Figures 2B and S2B). For example, while Druantia III co-occurred with 8 other defense systems, no co-occurrences were found for Druantia I (Bonferroni correction). CBASS I, composed of cyclase and effector proteins, and CBASS II, characterized by the presence of cGasylation proteins cap2 (E1-E2 fusion) and cap3 (JAB),[19,46] co-occurred with distinct defense systems. CBASS I co-occurred with DndABCD and DndFGH (which also co-occurred with each other), 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 defense system subtypes.

Conversely, we observed 54 (2.2%) pairs of defense systems that were negatively associated, such as CRISPR-Cas type I-E and RM I (Figure 2B). Similar to the number of co-occurrences, the number of significant exclusions notably varied across defense systems, with some, for example, RM IV, CRISPR I-E, and Dpd appearing particularly prone to avoiding other systems (Figure S2B). Interestingly, RM IV and BREX I were negatively associated in *E. coli*, even though previously shown to co-occur on a plasmid-encoded defense island and provide complementary protections against modified (RM IV) and non-modified (BREX I) invading DNA in *Escherichia fergusonii*.[30] 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 behavior of RM IV. Additionally, location of one of the systems within an integrated element, such as prophage, can lead to negative association 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 negatively associated 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]

Similar to the co-occurrence patterns, the patterns of negative association showed substantial differences among subtypes of the same defense system (Figures 2B and S2B). In most cases, however, the exclusivity between defense 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,[32] 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) negative association among defense systems are common in *E. coli*.

Figure 2. Co-occurrence and negative association among defense systems in *E. coli*

(A) Graphical representation of the co-occurrence analysis, depicting one pair of defense systems that co-occur (Gabija and tmn) and one pair that is negatively associated (Zorya II and ietAS). The nodes of the *E. coli* phylogenetic tree are colored according to the presence or absence of the defense 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 defense system pairs in *E. coli*. Co-occurring systems are shown in orange, and negatively associated 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 defense systems are color-coded according to their broadly defined mechanism of defense.

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

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

Defense systems often cluster together in defense islands,[2,4,5,7,24,26] which has been hypothetically attributed to fitness benefits conferred by such clustering on bacteria living in environments with high phage loads.[28] In particular, clustering of defense systems in defense islands, especially within integrated MGEs, increases the likelihood of horizontal co-transfer.[51] Therefore, if the defense 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*. To investigate the potential connection between the co-occurrence and co-localization of defense systems, we analyzed the genomic distance between defense systems in all complete *E. coli* genomes (2,164) and showed that co-occurring defense systems were generally not located significantly closer to each other in the genome compared with the average distance between defense systems (Figures 2C and 2D). Thus, in general, the physical proximity of defense 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 defense 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.[52,53]

Co-occurring defense systems act synergistically to counter phage infection

Next, we sought to explore whether the co-occurrence of defense systems is driven by their complementary activities[31] or synergistic molecular cooperation.[29,30] To this end, we selected three pairs of significantly co-occurring defense systems found in *E. coli* strains from our collection: Gabija and tmn that co-localize frequently in plasmids, Druantia III and Zorya II that co-localize in the chromosome, and ietAS and Kiwa that do not co-localize (Figure S2C). Additionally, we tested the negatively associated combination of Zorya II and ietAS. These systems were individually or jointly cloned into the *E. coli* strain BL21-AI (Figure 3A), which harbors the defense systems Retron II-A, Mokosh II, RM I, and RM IV. Among these, RM I co-occurs with Druantia III, and RM IV co-occurs with Zorya II and Druantia III (Figure 2B, considering the most stringent Bonferroni correction). Nevertheless, using the same genetic background across all experiments and assessing the isolated effects of defense systems within this background ensures that observed effects result from the interaction between the tested defense systems. We assessed the effects of the defense systems 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 defense 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 defense system pair with the sum of the individual effects of the two partners. The results showed that all tested combinations of defense systems, with the exception of ietAS and Kiwa, displayed significant synergistic effects against at least some phages in our panel, with even the negatively associated pair Zorya II and ietAS showing unexpected synergy (Figures 3B and 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 defense system combinations on phage propagation in liquid cultures. The results from these assays consistently confirmed the synergy as the combinations of defense systems led to a reduction in phage propagation that significantly exceeded the sum of the individual effects (Figures 3C and 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 with tmn alone (Figure 3C). For the negatively associated 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 defense 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 (AUCs) 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 defense 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 defense 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 (Figures 3D and S3C).

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

CellPress
OPEN ACCESS

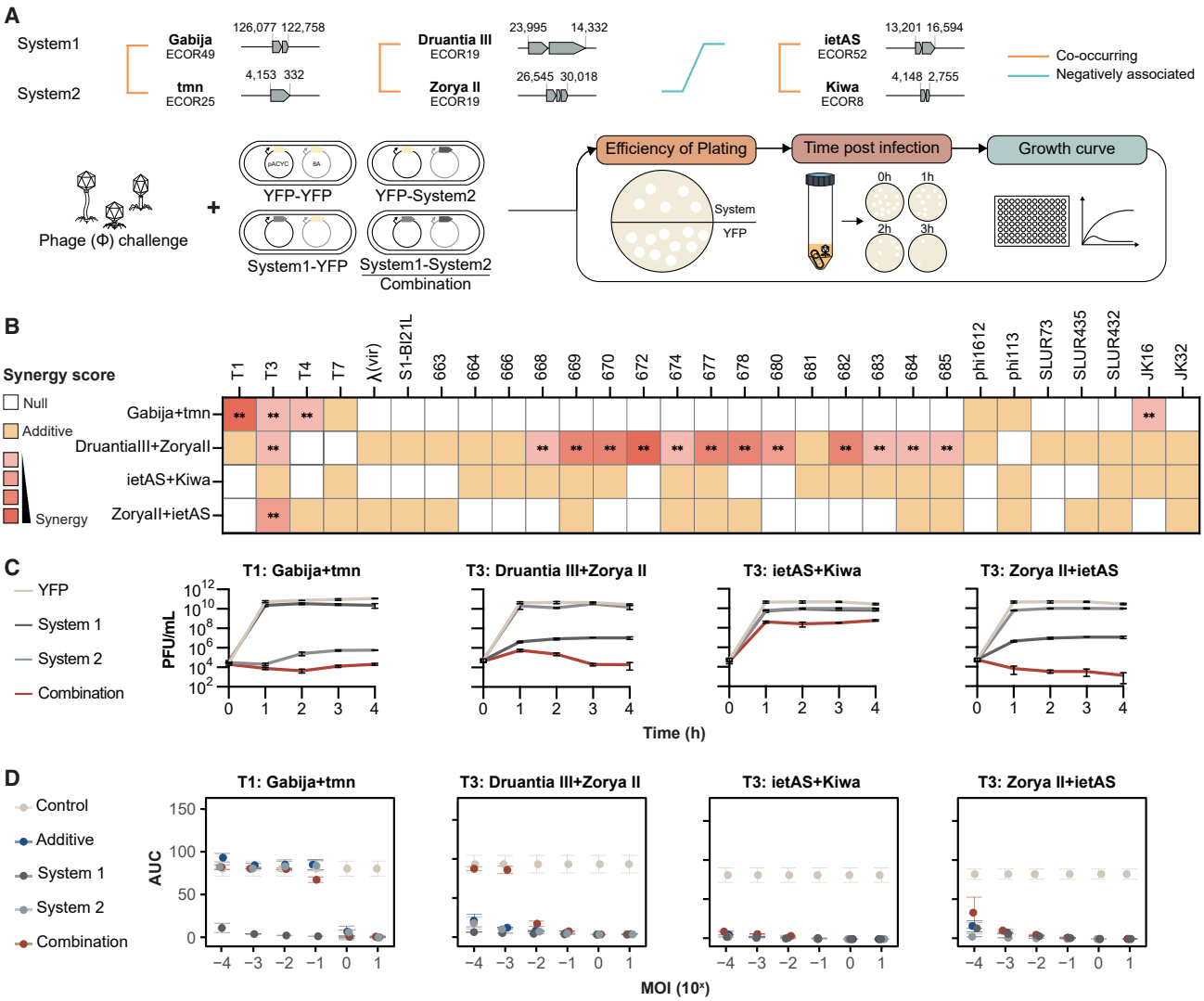


Figure 3. Defense system pairs provide synergistic anti-phage activity

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

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

(C) Time post-infection assays, measuring T1 or T3 titers over the course of four hours in liquid cultures of *E. coli* containing individual or combined defense systems. Data are 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. A defense system pair acts synergetic when its dot (red) is above the expected additive effect (blue). Data are 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: https://github.com/garushyants/synergy_bacterial_immune_systems/ and Zenodo: https://zenodo.org/doi/10.5281/zenodo.10075783 databases.

Defense system co-occurrence is not conserved across bacterial taxa

Genomic analyses show that increasing phylogenetic distance between bacterial species, decreases the rate of HGT.[54–57] In particular, the host ranges of most phages are narrow so that transduction occurs mostly within the host species boundaries.[58] Similarly, only a small fraction of plasmids has been shown to cross the interspecies barrier.[59–61] Considering these limitations of HGT along with (largely) non-overlapping viromes,[22] it could be expected that the co-occurrence of defense 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 defense systems present in these bacteria. The different bacterial orders displayed variations in the type and abundance of defense systems (Figure S4A). For example, RM systems were the most prevalent in Bacillales, Enterobacterales (including *E. coli*), and Pseudomonadales,

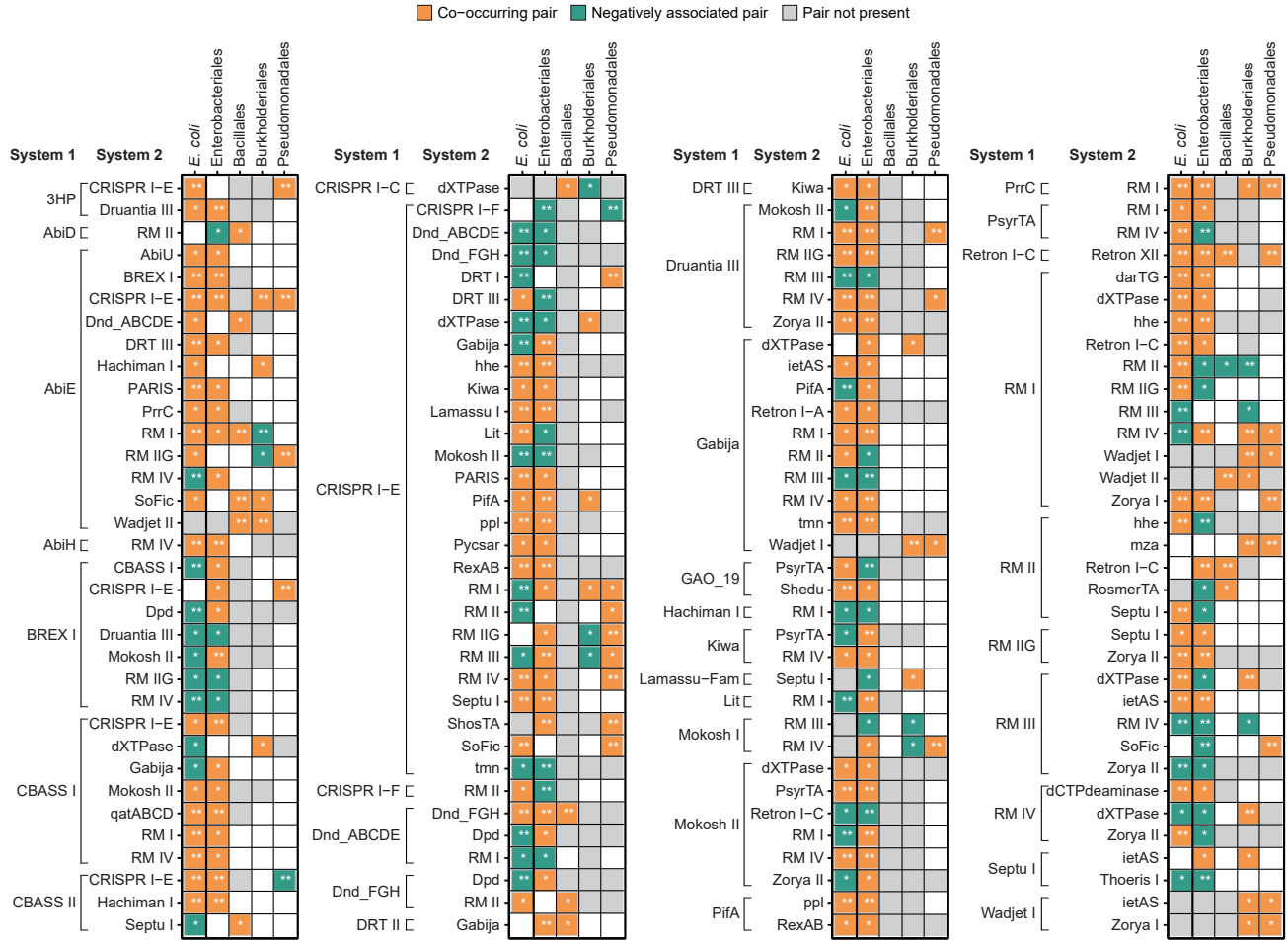


Figure 4. Patterns of defense system co-occurrences across bacterial taxa

Heatmap of defense 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). Gray 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.

whereas Burkholderiales were characterized by a higher abundance of dXTPase, Zorya III, and Mokosh I. Additionally, while CRISPR I-E was abundant (50.6%) in Enterobacterales (including *E. coli*), its prevalence was markedly lower (<6.1%) in the other orders. Nonetheless, when considering all four orders and one extensively characterized species (*E. coli*), they collectively shared 86 defense systems in common, with only a few systems unique to each of them (2 in *E. coli*, 17 in Bacillales, 4 in Enterobacterales, 2 in Pseudomonadales, and 1 in Burkholderiales) (Figure S4B).

As anticipated, the specific pairs of co-occurring and negatively associated defense systems differed across the taxa. Moreover, multiple pairs of defense systems that co-occurred in one order were found to be negatively associated in another (Figure 4). For instance, AbiE and RM I co-occur in Bacillales and Enterobacterales but are negatively associated in Burkholderiales. Similarly, CBASS type II and CRISPR-Cas type I-E co-occur in Enterobacterales but are negatively associated in Pseudomonadales. These findings indicate that negatively associated 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 defense systems are generally mechanistically compatible, allowing bacteria to adopt diverse, flexible strategies for anti-phage defense based on their unique environmental and genetic contexts.

Synergistic immunity provides an evolutionary advantage to bacterial populations

Due to our observations, we reasoned that co-occurring and synergistic defense systems could provide advantages at the population level. To assess the evolutionary and ecological impact of synergistic interactions between defense systems, we performed a short-term evolution experiment using chromogenic reporter plasmids expressing engineered coral chromoproteins.[62,63] We mixed populations of strains containing either an individual defense system and a second, "empty" chromogenic plasmid, or carrying two defense systems on separate plasmids. These populations were then infected with either a

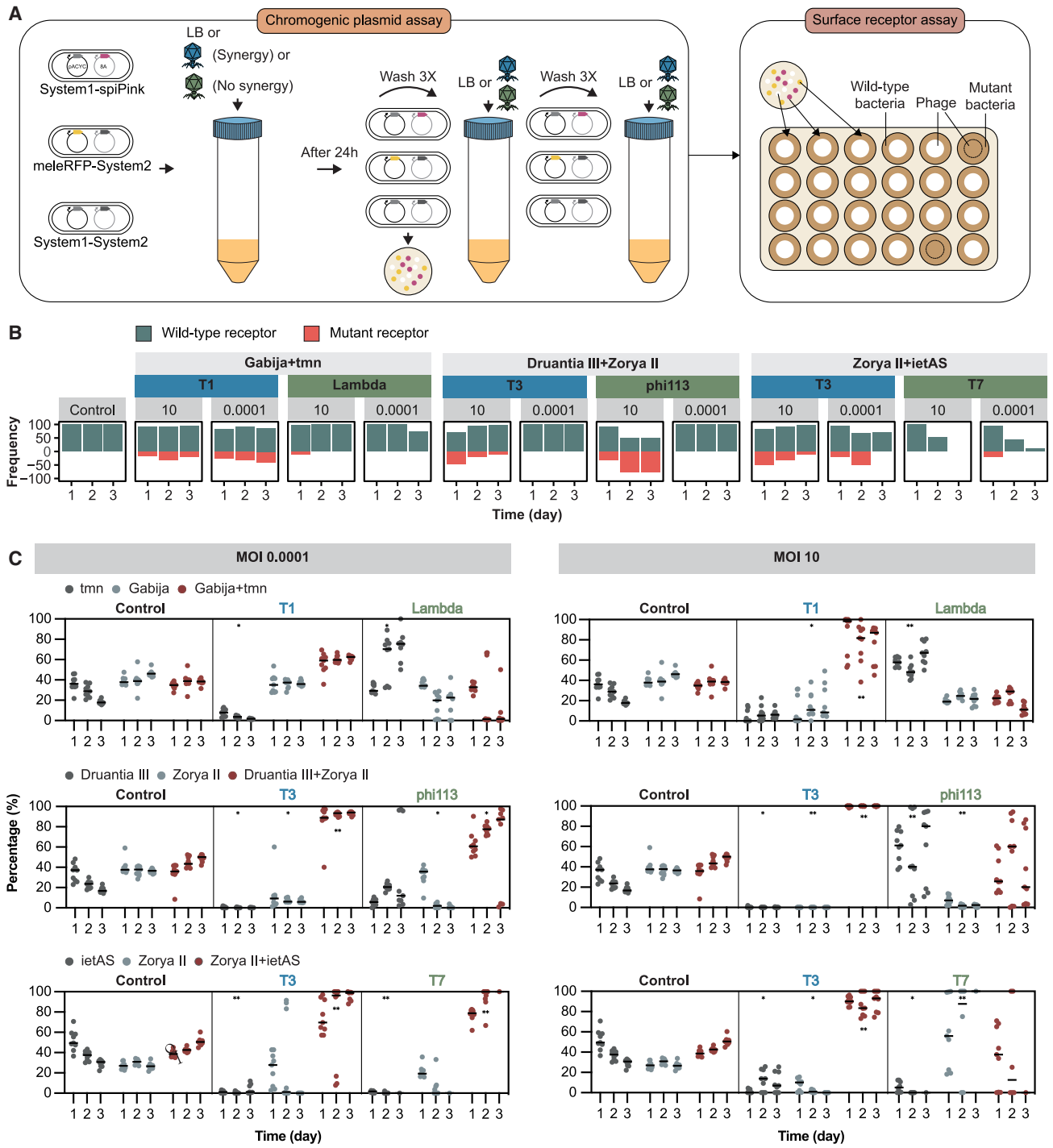


Figure 5. Synergistic defense system pairs provide an evolutionary advantage to bacteria

(A) Setup of experimental evolution assay of defense systems using chromogenic plasmids (left). Cells containing a single defense system carry a second plasmid that expresses a chromogenic reporter. Cells with defense system combinations were mixed in equal proportions with cells with single systems and infected with phage. After 1, 2, and 3 days, cells were plated, and the colonies of different colors 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 colors 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.

(legend continued on next page)

phage shown to trigger a synergistic defense, or a phage eliciting no obvious synergy between the respective defense systems. Over a period of 3 days, we monitored the proportion of the population carrying one or both defense systems by counting colony-forming units of different colors (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 (Figures S5A and S5B). Further, given that resistant receptor mutants tend to spread in bacterial populations shortly after phage infection,[64] 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 defense 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 defense 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 toward cells harboring both defense systems was observed starting from day 1. This shift was pronounced in cultures infected with the phage that elicited synergistic defense (T1 for Gabija and tmn, T3 for Druantia III and Zorya II, and T3 for ietAS and Zorya II). In cases the defense 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 with 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 defenses 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 defense system.[65]

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

Tmn co-opts the ATPase domain of Gabija for synergy

To better understand the molecular mechanisms underlying the observed synergistic effect between defense systems, we focused on the combination of Gabija and tmn. This choice was motivated by several factors. Gabija has been well characterized previously, in contrast to all other tested defense systems, providing molecular detail[2,65–67] that helps understanding the synergy with the less thoroughly characterized tmn.[3] Furthermore, Gabija and tmn tend to physically co-localize 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 defense 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[68,69] and is enriched in anti-defense genes,[70] which protect conjugative plasmids from host defense 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 defense system combination in the face of ongoing inter-MGE conflicts.[29]

To explore the mechanism of synergy between tmn and Gabija, we focused on determining the specific contributions of the functional domains of the individual defense systems to the synergistic anti-phage activity. We introduced point mutations into these functional domains and assessed their effects on the individual defense 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.[67] 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.[65] 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[65] 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

(C) Percentage of colonies from the evolution assay that carry each individual defense system or their combinations at 1, 2, and 3 days post infection with phage at low or high multiplicity of infection (MOI), compared with a non-infected control. Data are 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 with the corresponding uninfected control.

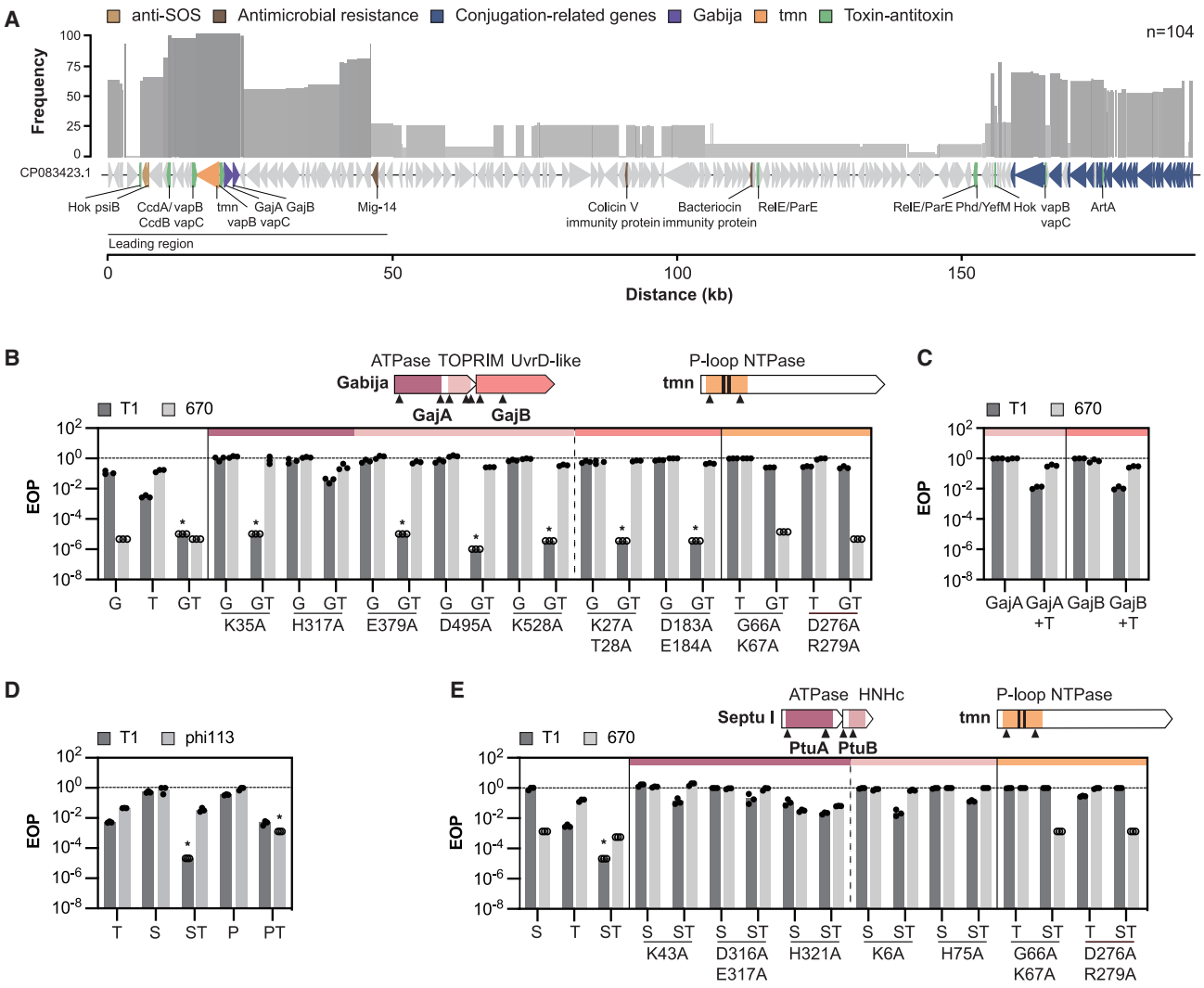


Figure 6. Mechanistic insight into the synergistic interaction between tmn and Gabija or Septu I

(A) Whole-plasmid alignment of 104 plasmids containing tmn and Gabija from complete *E. coli* genomes with plasmid CP083423.1 as a reference (see STAR Methods). The histogram shows the percentage of plasmids where the corresponding block from CP083423.1 was found. Annotated genes are colored 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 organized 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. The bars represent the average of three biological replicates.

(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. The bars represent the average of three biological replicates.

(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. The bars represent the average of three biological replicates.

(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 organized 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. The bars represent the average of three biological replicates.

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[65] 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.[66] 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 synergize 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 defense 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] characterized 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.[71] Because most P-loop ATPases are multimers (most commonly, hexamers[72]), we modeled the tmn dimer using Alphafold2 (pLDDT score 89.1) (Figure S6A) and observed that the residues (P177-P198) situated between the two transmembrane helices 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.[71] 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 defense systems.

Synergy between tmn and defense 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 defense 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,[2,73–75] suggesting a potential shared mechanism underlying synergistic interactions with tmn. To test this hypothesis, we analyzed the anti-phage defense 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 characterized 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 defense system, and it was proposed that PtuA, similarly to GajA in Gabija, is activated by NTP depletion during phage infection.[76] This feature is also observed in the PrrC defense system, where the ATPase domain of PrrC is inhibited by ATP and guanosine triphosphate (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.[74,77,78] The PsyrTA,[75] also known as RqlHI,[79] and AbiL[73] defense 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 synergizes with various defense 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 defense systems is expected to provide deeper insights into the synergistic interactions.

DISCUSSION

In this work, we aimed to gain insights into the interactions between defense 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 negative association among defense systems. We showed that the co-occurrences are not conserved in more distant bacteria, suggesting the existence of many distinct defense strategies. Perhaps unexpectedly, we found that the co-occurrence among defense systems was not strongly linked to their co-localization in bacterial genomes, although we did identify several instances in which co-occurring systems co-localized. In several specific cases we explored, co-occurrence of both co-localizing and non-co-localizing 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 negatively associated defense systems, such as Zorya II and ietAS. Furthermore, we found that defense systems that are negatively associated in one bacterial order can co-occur in others, suggesting that the negative association between these systems is not caused by mechanistic incompatibility leading to negative epistasis. We also assessed the ecological implications of defense 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 defense.

Consistent with other studies,[22,24,25,31] we observed an average of 5–7 defense systems per genome. This limited abundance of defense systems is likely due to the fitness cost

incurred by each of them such that the defense landscape is shaped by the trade-off between the benefits of protection against multiple viruses and the cumulative cost of multiple defense mechanisms. Our tests with negatively associated pairs of defense systems, although limited in scope, revealed no discernible fitness disadvantages, suggesting that the cost of defense systems is largely additive, without substantial negative epistasis.

Collectively, our results strongly suggest that interactions between defense systems are common and that non-random co-occurrence of defense systems in bacteria is an adaptive phenomenon driven by selection for enhanced immunity against specific phages. The defense strategies appear to differ substantially across bacterial taxa, most likely, driven by the species-specific viromes. Given the extensive horizontal transfer of defense systems, the evolution of bacterial defense strategies appear to follow the venerated general principle of adaptive evolution of microbes "everything is everywhere but the environment selects."[80]

We further investigated the molecular basis of the synergistic interactions between tmn and its co-occurring defense 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 defense system, enhancing the anti-phage activity. These synergistic interactions and co-opting mechanisms likely play an important role in the evolution of defense systems and the emergence of multiple defense system variants, as observed for CRISPR-Cas,[81] CBASS,[19] Lamassu,[6] Shield,[82] and others. The modularity of anti-phage immunity is further evident in sharing of proteins and functional domains across distinct defense systems,[83] such as HNH-endonuclease in Septu, Zorya II, and type II CRISPR-Cas[2]; NucC in CBASS III[84] and CRISPR-Cas type III,[85]; TOPRIM domain in Gabija,[3] Wadjet,[2] and PARIS,[4]; and P-loop NTPases, including helicases, in a broad diversity of defense systems including CRISPR-Cas type I, Gabija, PrrC, tmn, and many others.[2,3,74] Expanding our understanding of the mechanisms underlying the modularity of defense 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 defense 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 defense 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 negative association among the known defense systems only among isolates of a single bacterial species, *E. coli*. Our preliminary analysis of defense 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 defense 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. Moreover, the defense systems were expressed from plasmids, with gene dosage effects known to impact the protection range.[86] However, this is unlikely to confound our identified synergies, given the consistent expression levels of each defense system in both individual and combined configurations.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacteria
  - Phages
- METHOD DETAILS
  - Defense system detection
  - Contig characterization and prophage detection
  - Phylogenetic analysis
  - Correlation between defense system content and phylogenetic distance
  - Odds ratios for defense systems distribution between phylogroups
  - Analysis of co-occurrence of defense systems
  - Analysis of proximity between defense systems
  - Defense system cloning
  - Efficiency of plating
  - Time post infection assay
  - Liquid assay
  - Short-term evolution experiment
  - Quantification of receptor mutants
  - Plasmid loss assay
  - Cactus plasmid analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2024.01.015.

ACKNOWLEDGMENTS

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 (the 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.C.T.; resources, M.R.J.C., A.D.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.

Received: November 12, 2023
Revised: January 23, 2024
Accepted: January 30, 2024
Published: February 22, 2024

REFERENCES

1. Georjon, H., and Bernheim, A. (2023). The highly diverse antiphage defence systems of bacteria. Nat. Rev. Microbiol. *21*, 686–700.
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.
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.
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.e5.
5. 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 defense systems in the *E. coli* pangenome. Nat. Microbiol. *7*, 1568–1579.
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.
7. Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2013). Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Res. *41*, 4360–4377.
8. Tock, M.R., and Dryden, D.T.F. (2005). The biology of restriction and anti-restriction. Curr. Opin. Microbiol. *8*, 466–472.
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 defense system with broad anti-phage activities. Nat. Microbiol. *5*, 917–928.
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.
11. Bernheim, A., Millman, A., Ofir, G., Meitav, G., Avraham, C., Shomar, H., Rosenberg, M.M., Tal, N., Melamed, S., Amitai, G., et al. (2021). Prokaryotic viperins produce diverse antiviral molecules. Nature *589*, 120–124.
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.
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.
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.
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.
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.
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.e9.
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.e12.
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.
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.
21. Ofir, G., Herbst, E., Baroz, M., Cohen, D., Millman, A., Doron, S., Tal, N., Malheiro, D.B.A., Malitsky, S., Amitai, G., et al. (2021). Antiviral activity of bacterial TIR domains via immune signalling molecules. Nature *600*, 116–120.
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.
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.
24. Hochhauser, D., Millman, A., and Sorek, R. (2023). The defense island repertoire of the Escherichia coli pan-genome. PLOS Genet. *19*, e1010694.
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.
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.
27. Puigbò, P., Makarova, K.S., Kristensen, D.M., Wolf, Y.I., and Koonin, E.V. (2017). Reconstruction of the evolution of microbial defense systems. BMC Evol. Biol. *17*, 94.
28. Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defense as a community resource. Nat. Rev. Microbiol. *18*, 113–119.
29. 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.
30. Picton, D.M., Luyten, Y.A., Morgan, R.D., Nelson, A., Smith, D.L., Dryden, D.T.F., Hinton, J.C.D., and Blower, T.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.
31. 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., et al. (2022). Accumulation of defense systems in phage resistant strains of Pseudomonas aeruginosa. bioRxiv. https://doi.org/10.1101/2022.08.12.503731v2.

32. Bernheim, A., Bikard, D., Touchon, M., and Rocha, E.P.C. (2020). Atypical organizations and epistatic interactions of CRISPRs and cas clusters in genomes and their mobile genetic elements. Nucleic Acids Res. *48*, 748–760.

33. 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.

34. Shmakov, S.A., Barth, Z.K., Makarova, K.S., Wolf, Y.I., Brover, V., Peters, J.E., and Koonin, E.V. (2023). Widespread CRISPR-derived RNA regulatory elements in CRISPR-Cas systems. Nucleic Acids Res. *51*, 8150–8168.

35. 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.

36. 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.

37. 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.

38. 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.

39. Keen, E.C. (2015). A century of phage research: Bacteriophages and the shaping of modern biology. BioEssays *37*, 6–9.

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

41. Escobar-Páramo, P., Clermont, O., Blanc-Potard, A.B., 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.

42. 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.

43. 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.

44. 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.

45. 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.

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.

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. USA *113*, E1452–E1459.

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.

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.

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. (2019). BREX system of Escherichia coli distinguishes self from non-self by methylation of a specific DNA site. Nucleic Acids Res. *47*, 253–265.

51. Pingoud, A., Wilson, G.G., and Wende, W. (2014). Type II restriction endonucleases–a historical perspective and more. Nucleic Acids Res. *42*, 7489–7527.

52. 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 Res. *38*, 7133–7141.

53. Wu, D., Tang, Y., Chen, S., He, Y., Chang, X., Zheng, W., Deng, Z., Li, Z., Wang, L., Wu, G., et al. (2022). The functional coupling between restriction and DNA phosphorothioate modification systems underlying the DndFGH restriction complex. Nat. Catal. *5*, 1131–1144.

54. 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.

55. Andam, C.P., and Gogarten, J.P. (2011). Biased gene transfer in microbial evolution. Nat. Rev. Microbiol. *9*, 543–555.

56. Soucy, S.M., Huang, J., and Gogarten, J.P. (2015). Horizontal gene transfer: building the web of life. Nat. Rev. Genet. *16*, 472–482.

57. Puigbò, P., Wolf, Y.I., and Koonin, E.V. (2010). The Tree and Net Components of Prokaryote Evolution. Genome Biol. Evol. *2*, 745–756.

58. 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.

59. 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.

60. Koonin, E.V. (2016). Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. F1000Res *5*. F1000 Faculty Rev-1805.

61. 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.

62. 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.

63. 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.

64. 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.

65. 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 *31*, 1331–1344.e5.

66. Antine, S.P., Johnson, A.G., Mooney, S.E., Leavitt, A., Mayer, M.L., Yirmiya, E., Amitai, G., Sorek, R., and Kranzusch, P.J. (2024). Structural basis of Gabija anti-phage defence and viral immune evasion. Nature *625*, 360–365.

67. 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.

68. 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.

69. 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.

70. Samuel, B., and Burstein, D. (2023). A diverse repertoire of anti-defense systems is encoded in the leading region of plasmids. bioRxiv. https://doi.org/10.1101/2023.02.15.528439.

71. 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.

72. 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.

73. Deng, Y.M., Liu, C.Q., and Dunn, N.W. (1999). Genetic organization and functional analysis of a novel phage abortive infection system, AbiL, from Lactococcus lactis. J. Biotechnol. *67*, 135–149.

74. Uzan, M., and Miller, E.S. (2010). Post-transcriptional control by bacteriophage T4: mRNA decay and inhibition of translation initiation. Virol. J. *7*, 360.

75. 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.e5.

76. Azam, A.H., Chihara, K., Kondo, K., Nakamura, T., Ojima, S., Tamura, A., Yamashita, W., Cui, L., Takahashi, Y., Watashi, K., et al. (2023). Viruses encode tRNA and anti-retron to evade bacterial immunity. bioRxiv. https://doi.org/10.1101/2023.03.15.532788.

77. 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.

78. 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.

79. 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.

80. Becking, B., and Marinus, L.G. (1934). In Geobiologie of inleiding tot de milieukunde (W.P. Van Stockum & Zoon).

81. 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.

82. 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 Genet. *19*, e1010784.

83. Mariano, G., and Blower, T.R. (2023). Conserved domains can be found across distinct phage defense systems. Mol. Microbiol. *120*, 45–53.

84. 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.e6.

85. 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.e9.

86. Aframian, N., Bendori, S.O., Hen, T., Guler, P., and Eldar, A. (2023). High defense system expression broadens protection range at the cost of increased autoimmunity. bioRxiv. https://doi.org/10.1101/2023.11.30.569366.

87. 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.

88. 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.

89. Van der Auwera, GA, and O'Connor, BD (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra, 1st Edition (O'Reilly Media).

90. 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.

91. 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.

92. Payne, L.J., Todeschini, T.C., Wu, Y., Perry, B.J., Ronson, C.W., Fineran, P.C., Nobrega, F.L., and Jackson, S.A. (2021). Identification and classification of antiviral defense systems in bacteria and archaea with PADLOC reveals new system types. Nucleic Acids Res. *49*, 10868–10878.

93. Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. *3*, 217–223.

94. 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.

95. 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*.

96. R Core Team (2018). R: A language and environment for statistical computing (R Foundation for Statistical Computing).

97. Simonsen, M., Mailund, T., and Pedersen, C.N.S. (2008). Rapid Neighbour-Joining. In Algorithms in Bioinformatics: 8th International Workshop (Springer), pp. 113–122.

98. Mai, U., and Mirarab, S. (2018). TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. BMC Genom. *19* (*Supplement 5*), 272.

99. 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.

100. Balaban, M., Moshiri, N., Mai, U., Jia, X., and Mirarab, S. (2019). TreeCluster: clustering biological sequences using phylogenetic trees. PLOS One *14*, e0221068.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Dh5α	Thermo Fisher Scientific	Cat# 16512160
BL21-AI	Thermo Fisher Scientific	Cat# C607003
ECOR7, ECOR8, ECOR19, ECOR25, ECOR35, ECOR49, ECOR52, ECOR64, ECOR66, ECOR70	Fagenbank	N/A
All bacteriophages are listed and described in Table S4	N/A	N/A
Chemicals, peptides, and recombinant proteins		
Agar	Formedium	Cat# AGA04
Agarose	Melford	Cat# A20090
Ampicillin	Melford	Cat# A40040-10.0
Arabinose	Melford	Cat# A51000-100.0
Gibson assembly Master Mix	New England Biolabs	Cat# E2611L
Glycerol	Melford	Cat# G1345-5L
Chloramphenicol	Acros Organics	Cat# A0414716
Kanamycin	Gibco	Cat# 11815024
Lysogeny Broth (LB)	Formedium	Cat# LBX0103
Nuclease free water	New England Biolabs	Cat# B1500L
OneTaq 2X MasterMix with Standard Buffer	New England Biolabs	Cat# M0486S
Q5 DNA polymerase	New England Biolabs	Cat# M0491L
TAE 50X	Melford	Cat# T60015-1000.0
Critical commercial assays		
DNA Clean & Concentrator Kit	Zymo Research	Cat# D4029
GeneJET Genomic DNA Purification Kit	Thermo Fisher Scientific	Cat# K0722
NucleoSpin Plasmid QuickPure Kit	Thermo Fisher Scientific	Cat# 11902422
Mix&Go! *E. coli* Transformation Kit (Zymo)	Zymo Research	Cat# T3001
Qubit™ 1X dsDNA High Sensitivity (HS) Kit	Invitrogen	Cat# Q33231
Zymoclean Gel DNA Recovery Kit	Zymo Research	Cat# D4002
BacTiter-Glo™ Microbial Cell Viability Assay	Promega	Cat# G8230
Deposited data		
Code and raw experimental data	This study	GitHub: https://github.com/garushyants/synergy_bacterial_immune_systems/ and Zenodo: https://zenodo.org/doi/10.5281/zenodo.10075783
Oligonucleotides		
All DNA oligonucleotides are listed in Table S6	IDT	N/A
Recombinant DNA		
All plasmids are listed and described in Table S5	N/A	N/A
Software and algorithms		
Adobe Illustrator 24.2	Adobe	N/A
BioRender	BioRender	N/A
BWA v0.7.17	(Li 2013)	https://github.com/lh3/bwa

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
bcl-convert v3.9.3	N/A	https://support.illumina.com/downloads/bcl-convert-v4-0-3-installer.html
Cactus v2.4.4	Armstrong et al[87]	https://github.com/ComparativeGenomicsToolkit/cactus
CheckM	Parks et al[88]	https://github.com/Ecogenomics/CheckM
DefenseFinder version 1.0.8	Tesson et al[22]	https://github.com/mdmparis/defense-finder
gatk v4.2.6.1	(van der Auwera and O'Connor 2020)[89]	https://github.com/broadinstitute/gatk/releases
ggtree for R	Yu et al[90]	https://github.com/YuLab-SMU/ggtree
GraphPad Prism 9.2.0	GraphPad	N/A
Jupyter v4.6.3	N/A	https://github.com/jupyter/notebook/
Mash v2.3	Ondov et al[91]	https://github.com/marbl/Mash/releases/tag/v2.3
PADLOC version 1.1.0 with database version 1.4.0	Payne et al[92]	https://github.com/padlocbio/padloc
phytools for R	Revell[93]	https://github.com/liamrevell/phytools
Phigaro version 2.2.6	Starikova et al[94]	https://github.com/bobeobibo/phigaro
Platon version 1.6	Schwengers et al[95]	https://www.uni-giessen.de/fbz/fb08/Inst/bioinformatik/software/platon
Python v3.7	N/A	https://www.python.org/downloads/release/python-370/
R version 4.1.2	R Core Team [96]	https://cran.r-project.org
RapidNJ	Simonsen et al[97]	https://github.com/somme89/rapidNJ
TreeShrink v1.3.9	Mai and Mirarab[98]	https://github.com/uym2/TreeShrink
Other		
FLUOstar OPTIMA plate reader	BMG LABTECH	Cat# 0413B0001J
Nanodrop 2000	Thermo Fisher Scientific	Cat# ND2000
Qubit 2.0 Fluorometer	Invitrogen	Cat# Q32866

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Franklin L. Nobrega (F.Nobrega@soton.ac.uk).

Materials 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.

Data and code availability

- Raw data have been deposited at Github and Zenodo and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code has been deposited at Github and Zenodo and is publicly available as of the date of publication. Links are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacteria

E. coli strain Dh5α was used to clone plasmids pACYCDuet-1 or 8A with individual defense systems. *E. coli* BL21-AI cells containing plasmid(s) with the defense 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

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.

METHOD DETAILS

Defense 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. Defense systems were detected in these genomes using PADLOC version 1.1.0 with database version 1.4.0[92] and DefenseFinder version 1.0.8.[22] Defense systems outputted by PADLOC in categories *other* and *adaptation* were removed from the analysis. Defense systems predicted by DefenseFinder to be located on different contigs or longer than 30kb were discarded. Some of the *E. coli* genomes had extensive fragmentation, which negatively influenced the number of defense 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. Defense systems found in all strains can be seen on Table S1.

Contig characterization and prophage detection

Platon version 1.6[95] on accuracy mode was used to categorize bacterial contigs as plasmid or chromosome. Phigaro version 2.2.6[94] was used with default settings to detect prophages in the bacterial genomes. Contigs shorter than 20 kbp were excluded from this analysis. Defense systems were considered to be located in a prophage region when at least one defense gene was fully within the prophage limits.

Phylogenetic analysis

Mash v2.3[91] 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 reconstruction of Neighbor-Joining phylogenetic trees with RapidNJ.[97]

To remove potentially contaminated genomes from the phylogenetic trees, we implemented a three-step filtration process. First, we applied TreeShrink v 1.3.9[98] 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[88] 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 color 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 in the tree. We employed phylogroup assignments from a dataset of 10,667 *E. coli* genomes,[99] 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[100] 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[90] and phytools.[93]

Correlation between defense system content and phylogenetic distance

To investigate the relationship between defense 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 defense system content between genomes, we employed Jaccard distance, which compares the presence and absence of defense systems in vectors. Defense 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 defense systems content.

Odds ratios for defense systems distribution between phylogroups

To test the hypothesis that defense 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 defense 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.

Analysis of co-occurrence of defense systems

Given that the genomes under study are related, 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 defense systems were non-randomly distributed across the phylogenetic tree, indicating potential interdependencies. This analysis considered only the presence or absence of defense systems, not taking into account their localization in the genome. 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 defense 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-value (< 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 negatively associated.

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.

Analysis of proximity between defense systems

To assess whether co-occurring defense systems exhibited a tendency to co-localize in the genomes more frequently than random pairs of neighbor defense systems within a genome, we compared the shortest distances between pairs of the co-occurring systems with the mean of the background distribution of the distances between defense systems. The background distribution was generated by calculating distances between all neighbour pairs of defense systems across complete *E. coli* genomes. For each system, we calculated the shortest distance to the neighbor system. When a system was located at the edge of a defense island, we considered the distance to the next system within the same island rather than to the next island. Distances were calculated uniquely, ensuring that if system A's closest neighbor was system B and vice-versa, the distance between A and B was included in the analysis only once. The mean of all calculated distances served as the conservative measure of the proximity between neighbor defense systems. For additional details of the method employed for this analysis, see the positive_vs_negative_genome_distance.R script, available in the associated Github: https://github.com/garushyants/synergy_bacterial_immune_systems/ or Zenodo: https://zenodo.org/doi/10.5281/zenodo.10075783 repositories.

Defense 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 defense 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 plating

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 defense systems, were determined as $|\mathrm{Log}_{10}(\mathrm{EOP}_{\mathrm{system1+system2}})| - |\mathrm{Log10}(\mathrm{EOP}_{\mathrm{system1}})| - |\mathrm{Log10}$

($EOP_{system2}$)|. Synergy was considered when |Log_{10} ($EOP_{system1+system2}$)| > |Log_{10} ($EOP_{system1}$)| + |Log_{10} ($EOP_{system2}$)| +1, additive effects were considered when Max [|Log10 ($EOP_{system1}$)|, |Log10 ($EOP_{system2}$)|] < |Log10 ($EOP_{system1+system2}$)| < |Log_{10} ($EOP_{system1}$)| + |Log_{10} ($EOP_{system2}$)| + 1, and antagonistic effects were considered when |Log_{10} ($EOP_{system1+system2}$)| – Max (|Log10 ($EOP_{system1}$)|, |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 defense 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 defense 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 color counted. The experiment was repeated for 48 h and 72 h 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 defense 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 48 h and 72 h 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 defense systems were extracted from the Enterobacterales dataset. The plasmid dataset comprised of a total of 104 plasmids from various *E. coli* strains. We constructed a phylogenetic tree for these plasmids using the Neighbour-Joining method as described above, using a mash sketch size of 1,000. We then generated whole length plasmid alignments using Cactus v 2.4.4[87] 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%.

QUANTIFICATION AND STATISTICAL ANALYSIS

A two-sided binomial test was performed to determine if the observed co-occurrence of defense systems differed significantly from the expected co-occurrence, using R.[96] 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: https://github.com/garushyants/synergy_bacterial_immune_systems/ and Zenodo: https://zenodo.org/doi/10.5281/zenodo.10075783 .