Parallelized screening of virus accessory genes reveals diverse defense and counter-defense mechanisms

Sukrit Silas[1,10], Héloïse Carion[1], Kira S. Makarova[2], Eric Laderman[1], David Sanchez Godinez[3,4], Matthew Johnson[1], Andrea Fossati[5,6,7], Danielle Swaney[5,6,7], Michael Bocek[8], Eugene V. Koonin[2], Joseph Bondy-Denomy[1,6,9,10]

[1] Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA 94158, USA

[2] National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

[3] Cardiovascular Research Institute, University of California, San Francisco, San Francisco, CA 94158, USA

[4] Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA

[5] J. David Gladstone Institutes, San Francisco, 94158, California, USA

[6] Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, CA 94158, USA

[7] Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA

[8] Twist Biosciences, South San Francisco, CA 94080, USA

[9] Innovative Genomics Institute, Berkeley, CA 94720, USA

[10] Corresponding authors: Sukrit.Silas@ucsf.edu, Joseph.Bondy-Denomy@ucsf.edu

Abstract

Genomes of bacterial and archaeal viruses are replete with fast-evolving, uncharacterized accessory genes (AGs), most of which likely antagonize host defenses or other viruses. We developed a computational pipeline to find AGs in bacteriophage genomes and built a high-throughput screening platform to assay their functions. This approach is targeted towards identifying the most salient antiviral mechanisms in any bacterial niche. We tested 200 Enterobacteriophage AGs in 20 wild *Escherichia coli* strains challenged with 8 phages. The most prominent AG functions we observed were antagonism of O-antigen-based barrier defense and restriction-modification (R-M) systems. In response to phage-encoded anti-R-M strategies, some Type I and III R-M systems surprisingly act as programmed-cell-death modules. Moreover, several hyper-variable AGs trigger other abortive defense systems, demonstrating that phage-associated molecular patterns that activate immunity need not be well-conserved. Our approach yields insights into the multiple levels of virus-host competition and can be rapidly deployed in various bacteria.

Introduction

To a large extent, the evolution of life is a story of the virus-host arms race[1], during which the hosts evolved diverse, elaborate antiviral defense mechanisms, and viruses retorted by evolving equally versatile counter-defenses. Analysis of the rapidly growing genomic and especially metagenomic sequence databases uncovered a vast pool of proteins with mostly unknown functions encoded by viruses infecting bacteria and archaea[2]. The characterized examples of such proteins comprise only a small fraction of that pool, but possess diverse, exquisitely evolved capacities. These functionalities include creating novel interfaces in protein complexes[3], mimicking biomolecules to competitively inhibit enzymes[4], allosterically modifying cellular enzymes[5], interfering with host processes such as transcription[6], translation[7], adaptive immunity[8], motility[9], quorum sensing[10] and DNA transfer[11], inhibiting the reproduction of other viruses[12], and even participating in inter-kingdom competition[13]. The study of model phage and phage-defense systems over the past few decades has yielded key molecular biology tools, notably, recombinant DNA methodology that relies on restriction-modification (R-M) systems[14] and CRISPR genome editing[15]. A comprehensive investigation of the genes encoded by viruses can reveal an enormous variety of biological mechanisms involved in the virus-host arms race.

Bacterial and archaeal defense systems often cluster together in genomic loci known as defense islands[16]. Recently, a broad variety of distinct defense mechanisms have been discovered by investigating genes located in the same islands with known defense systems[17,18]. Counter-defense genes also often cluster together in virus genomes[19], but these “counter-defense islands” have not been studied as systematically as the defense islands in prokaryotic genomes. Because experiments with individual virus proteins can be slow and laborious, systematic approaches are essential to prioritize genes for study.

We developed an approach to identify bacteriophage genes that are likely to interfere with host biology. Core genes are present in (nearly) all members of a group of related phages and typically encode essential structural components of the virions as well as components of the replication and transcription machineries. In contrast, Accessory Genes (AGs) are present in some, but not all, closely-related genomes[20,21] and most are of unknown function, whereas those that have been characterized (such as numerous anti-CRISPR proteins) are often involved in counter-defense[22-24] or superinfection exclusion[11,12]. Our algorithm identifies “accessory regions” (ARs) that interrupt blocks of conserved core genes or are located at genome ends in closely related virus genomes (Figure 1a). AGs are typically non-essential, undergo frequent HGT, and are subject to enhanced evolutionary pressure due to their involvement in the virus-host arms race, resulting in extensive diversification.

To test AG functions, we developed the Phage Accessory Gene Exploration aNd Testing (PhAGENT) platform, which can experimentally evaluate thousands of phage genes at once. The platform utilizes genetic screening to rapidly identify AGs that affect host sensitivity to the phage or trigger programmed cell death (PCD). This yields host-phage-AG combinations that can then be interrogated using further screens, proteomics, and classical microbiology. AG libraries can be synthesized from any phage genomic dataset, and the PhAGENT platform can be deployed in any bacterial or archaeal strain collection where conjugation and transposon mutagenesis are supported. To validate the approach, we synthesized 200 phage AGs from diverse Enterobacteriophages and tested them in 20 wild strains of the *E. coli* Reference (ECOR) collection. We identified layered bacterial defenses and extensive co-evolution of defense and counter-defense mechanisms: a distinct mode of bacterial seroconversion by phage proteins that alter host metabolism, phage-encoded inhibitors of antiviral restriction mechanisms, an R-M system that can trigger PCD in addition to its canonical defense functions,

and decoy immunity in bacteria functionally similar to that found in plants[25] which can alert the host to viral attempts to interfere with cellular pathways.

Combined Results + Discussion

A high throughput screening platform for identifying and assaying accessory genes

AGs are usually defined within a cluster of closely-related viruses[20-23] and AG identification is highly sensitive to the method used to cluster virus genomes. We therefore sought to identify ARs without *a priori* sequence comparison or clustering of virus genomes (Supplementary File 1). First, all virus genes from the entire dataset were collected, and pairwise combinations of genes (the query pair) were evaluated sequentially (Figure 1b). Hence, every virus genome that contained both genes of the query pair was grouped into a possible virus cluster, and genes encoded between the genes of the query pair were considered members of a possible AR. Second, a series of filters were applied. Clusters containing few viruses (<6) or highly diverged ones were discarded. The longest ARs (>10 genes) or those that consisted of the same genes in every genome were also discarded. Third, all retained ARs were scored and ranked by diversity (see Methods). Crucially, we did not rely on any gene ontology to identify AGs (eg. genes similar to or encoded next to known counter-defense factors). A sample AR is shown in Figure 1c. Redundant, low scoring ARs that comprised mostly of genes which had already been found in higher scoring ARs were then discarded. Using an initial set of 1706 Enterobacteriophage genomes, 2014 non-redundant ARs containing 10888 putative AGs were identified in 1217 phages. A typical AR contained 11-30 unique AGs ~50-100 aa in length. A typical phage genome contained <5 ARs and <5 AGs (Figure 1d).

We manually inspected the highest-scoring 200 ARs and chose multiple (2-5) AGs from >60 of these (see Supplementary File 2 for detailed notes on each AR). A barcoded library of 200 AGs was constructed (with 65 AGs from ARs containing known counter-defense factors, and 135 AGs from ARs with all genes of unknown function) to enable multiplexed identification of AGs that target bacterial defenses. In the subset of 65 AGs, 11 phage and mobile genetic element (MGE)-encoded proteins with previously elucidated counter-defense or superinfection-exclusion functions were included as controls (Abc2, ArdA, Arn, Cor, Dmd, Imm, Ip1, Ocr, Ral, SieA, Stp), with the remaining 54 of unknown function (Supplementary File 3).

To select phage-host combinations, we used a subset of the *E. coli* Reference (ECOR) strains[26] and assayed their ability to resist infection by 8 model phages. We reasoned that we may find new antiviral mechanisms if the phage receptors were present on the host cell surface, but some host factor(s) were blocking infection. Hence, 45 phage-host combinations with attenuated infection were chosen, that is, those where a phage was able to produce clearings on the bacterial lawn, but only with significantly reduced titer (Supplementary Figure 1a).

To identify AG function, we developed a scalable, high-throughput screen to simultaneously assess how all AGs affect host immunity to infection (See Methods). Barcoded, chromosomally integrated, IPTG-inducible AGs were expressed in each host (one AG per cell) in pools, and fitness effects were measured in triplicate upon AG-induction and phage infection by next-generation sequencing (Figure 1e) (Supplementary File 4). Importantly, although a set of 8 model phages was used in our experiments, the AGs can be taken from any available phage genome. AG induction was limited to 1-1.5 hrs, and subsequent infections were only allowed to proceed for ~1 phage replication cycle, to prevent prolonged induction of various AGs from biasing the results. This setup also avoids biases resulting from epigenetic modification of phages by widespread restriction-modification (R-M) defenses in *E. coli*.

AGs that drastically reduced the fitness of some (but not all) hosts either upon expression, or upon phage infection, were selected for further study.

AGs target conserved host pathways and barrier defenses

To identify AGs with counter-defense phenotypes, we compared infected samples with uninfected controls and found that several host-phage-AG combinations were depleted from the infected pools, presumably because the respective AGs facilitate phage replication (Figure 2a). Each AG that sensitized any wild strain to infection was then individually validated in that strain against the entire phage set. In these validation experiments, some phage-host combinations that were omitted in the PhAGENT screen also exhibited counter-defense phenotypes of AGs (Supplementary Figure 1b). In all, increased plaquing of several phages was observed upon expression of 3 known type I R-M inhibitors (Ocr, Ral, ArdA), 6 novel AGs, and the T4 internal protein 2 (IpII) in a total of 28 phage-host combinations. We hypothesized that these AGs antagonize host factors to help the phage(s) replicate.

Six novel AGs (orfs 48, 63, 92, 116, 126, 148), which have not been previously implicated in phage-host interactions (Supplementary File 5), all exhibited counter-defense phenotypes in multiple host strains. Because the broad-spectrum effects of these AGs did not correlate with the defense systems in the hosts (Supplementary Figure 1c), we sought to determine whether they could be targeting conserved host pathways that have antiviral functions.

To identify the antiviral mechanisms inhibited by AGs, we employed two unbiased approaches. (1) Transposon-knockout screens in the hosts helped pinpoint gene disruptions that phenocopy AG expression (i.e. increased phage infection) (Supplementary File 6). (2) Affinity-purification and mass spectrometry (AP-MS) of AGs helped identify their binding partners in the wild strains (Supplementary File 7). We performed these assays for every host found to be sensitized to infection by any AG.

Genome-wide knockout screens indicated that disrupting the host O-antigen or capsule genes phenocopied the phage-sensitization effects of counter-defense AGs with almost all tested phages (Table 1a). In particular, transposon-mediated O-antigen disruption (Figure 2b) or expression of AGs 48, 63, and 92 (Figure 2c) both converged on the same phenotype, namely, enhanced infection by phage T5 (which does not natively encode these AGs). We hypothesized that T5 adsorption was blocked by the wild-type O-antigen structures in these hosts, but heterologous AG expression enabled infection by modifying or removing the O-antigen. We therefore examined LPS preparations from every host where any AG exhibited a counter-defense phenotype. Hosts expressing AGs 48, 63, and 92 exhibited subtle downshifts in the O-antigen banding pattern (Figure 2d), as observed previously in some seroconverting lysogens as a means to exclude superinfecting phages[27]. Specifically, the following host-AG combinations showed visible downshifts in O-antigen bands: orfs 63 and 92 in ECOR1, orfs 48 and 92 in ECOR3, and orfs 48 and 63 in ECOR22. These downshifts correlate with the host-AG pairs where T5 infection was especially ($10^{5\text{-}6}$-fold) enhanced (Figure 2c). We reasoned that the AGs might interfere with O-antigen biosynthesis, and tentatively named orfs 48, 63 and 92, *gnarl1, 2,* and *3* respectively, after the skin-eating demon in the popular TV series *Buffy the Vampire Slayer*.

Next, we examined the distribution of *gnarl* genes in the context of O-antigen seroconversion. All known seroconverting prophages modify O-antigens using enzymes with distinct biochemical activities (such as glycosyltransferases and acetyltransferases)[27-31]. *Gnarl*-encoding phages in our genomic dataset do not encode any such seroconverting enzymes. Gnarl proteins (40aa – 90aa) are also much smaller than typical enzymes, suggesting that these AGs do not possess O-antigen-modifying enzymatic activity but instead affect O-antigen biosynthesis. This potentially constitutes a distinct mode of host seroconversion. *Gnarl* genes are not restricted to either lytic or lysogenic phages; *gnarl1* and *gnarl3* are present in lytic phages (Supplementary File 5), whereas *gnarl2* is present in the classical temperate phage Mu (gene E6/Mup07). Seroconversion could occur in lysogenic infections to exclude superinfecting

phages, or in lytic infections as receptor-masking to prevent newly synthesized virions from unproductively binding lysed cell fragments[32].

To identify the cellular targets of gnarl proteins, we analyzed AP-MS data for all counter-defense AGs. Each AG-encoded protein bound distinct host proteins although the profiles of AGs 116 and 126 were somewhat correlated (Figure 2e). We performed Gene Ontology pathway analysis for each co-IP to and found that UDP-glucose biosynthesis factors were significantly enriched in the *gnarl3* set. Among these, GalU catalyzes the formation of UDP-glucose[33] to build the LPS outer core[34]; GalF possibly regulates the levels of UDP-glucose[35]; and GalE reversibly converts UDP-glucose to UDP-galactose (also an LPS outer core component)[36]. If *gnarl3* inhibits UDP-glucose biosynthesis, we reasoned that overexpressing the target of *gnarl3* in the wild hosts might overcome its phage-sensitization phenotype. As a control, we also overexpressed *wecB*, which was enriched in *gnarl3* AP-MS but is not involved in UDP-glucose synthesis.

Overexpression of *galU* completely reversed the *gnarl3* T5-sensitization phenotype in all tested hosts, namely, ECOR1, ECOR3, ECOR21 (not shown), and ECOR22 (Figure 2f), whereas overexpression of *galE, galF,* and *wecB* had no effect. The same effect was observed with unrelated phages T4 (Supplementary Figure 2) and P1 (not shown), indicating the *gnarl3-galU* interaction was phage-agnostic. Overexpressing *galU* without *gnarl3* did not affect phage sensitivity except in ECOR22, where it partially potentiated infection for unknown reasons. No other genes we tested (*galE, galF,* wecB) affected phage-sensitization by *gnarl3*, suggesting a specific interaction between *galU* and *gnarl3*. Coupled with LPS electrophoresis and AP-MS data, this observation suggests a hypothetical model where *gnarl3* inhibits GalU, reducing the availability of UDP-glucose and resulting in a modification of the cell envelope that potentiates phage absorption. Although the possibility remains that the *gnarl3* phenotype is reversed indirectly by *galU* overexpression (eg. through its role in galactose metabolism), the inability of certain phages to adsorb to UDP-glucose pyrophosphorylase mutants of both *E. coli*[37] and *B. subtilis*[38] suggests that glycosylation of cell-surface biopolymers is an ancient phage-blocking strategy.

Barrier defenses and restriction systems provide layered immunity

The existence of broad-spectrum counter-defense AGs underscores the importance of barrier defenses. Consistent with this, our genome-wide transposon screens identified an O121-specific O-antigen[39] as a principal barrier to infection in ECOR21 (Figure 3a). While disrupting O-antigen allowed other phages to infect this strain, we could not obtain corresponding transposon insertions that sensitized ECOR21 to T2 (Figure 4a). Because the same ECOR21-knockout library was used for transposon screens with every phage, we hypothesized that an additional defense mechanism in ECOR21 restricted T2 infection (Figure 3b). In the PhAGENT screen, we recovered a phenotype for a T4 internal-protein, IpII (orf143), that partially sensitized ECOR21 to T2 (Figure 2a), suggesting that a T2-restricting defense system was disabled by IpII. If this hypothetical defense system indeed complements barrier defenses, this would explain why no single transposon insertion in ECOR21 seems to enable T2 infection.

To test this hypothesis, we eliminated O-antigen by deleting *wecA*[40] from ECOR21 (Figure 3c) and attempted the genome-wide knockout screen again. The screen now yielded several insertions in an atypical BREX system (Figure 3d) in which a fused type IV restriction-modification enzyme GmrSD (also known as BrxU) is embedded[41]. Thus, the O-antigen barrier indeed obscured this defense locus.

Type IV R-M systems have co-evolved extensively with T-even phages. T-even phage DNA contains hydroxymethylcytosines (hmC) that are further modified to glucosyl-hmC (g-hmC), presumably to escape cleavage by type I R-M systems. Type IV R-M systems (comprising two factors, GmrS and GmrD) can restrict g-hmC modified DNA, but T4 encodes an

accessory gene product, internal protein 1 (IpI), that is packaged into the phage capsid, delivered to the host with the virus DNA, and disables type IV R-M. However, IpI cannot block restriction by a modified type IV R-M system that combines both GmrS and GmrD activities in a single effector GmrSD[32].

We wondered whether BREX or GmrSD was the target of IpII in ECOR21. We cloned the entire operon along with its natural promoter and recapitulated its phenotype in *E. coli* K12. The operon robustly protected against all tested T-even phages (T2, T4, T6). To isolate the defense activity, we first deleted *pglX*, an essential component of BREX, and found that the phages were still blocked. Next, we tested an active-site mutant of GmrSD (GmrSD*) and found that infection was almost completely restored. Coupled with previous observations that type IV R-M systems digest glycosylated T-even phage DNA whereas BREX systems do not[42], our results indicate that GmrSD, and not BREX, restricted T-even phages (Figure 3e). Moreover, overexpressing GmrSD alone was sufficient to block all T-even phages whereas a catalytic mutant (GmrSD*) had no effect (Figure 3f).

Next, we tested whether various AGs (IpII, Ocr, IpI) might modify BREX/GmrSD function. Ocr (a canonical R-M inhibitor) was recently demonstrated to have anti-BREX activity[43], whereas IpI blocks the ancestral two-component type IV R-M system but not GmrSD. We found that Ocr and IpI did not affect any configurations of the BREX/GmrSD system we tested. However, expressing IpII restored infection by T-even phages despite the presence of BREX/GmrSD, or the Δ*pglX* construct (Figure 3e), indicating that IpII inhibited GmrSD.

Of the many T-even phage internal-proteins, IpII is now only the second (after IpI) with an identified target. T4 naturally encodes IpII, and was previously known to infect GmrSD-containing hosts via an undiscovered mechanism[32]. Our results suggest that IpII provides this activity. Consistent with this, over-expressed GmrSD targeted T2, T4, and T6, but the uninduced leaky GmrSD construct protected against only T2 and T6 (Figure 3f). When IpI or IpII were supplied, IpII restored T2 and T6 plaquing but did not further enhance T4 infection (Figure 3g). The observed redundancy between barrier defenses and restriction systems might obscure phenotypes of other internal-proteins of T-even phages.

Restriction-modification systems can provide abortive defense in response to their inhibitors

A second group of host-interacting AGs produced “conditional lethal” phenotypes, that is, several AGs imposed a severe infection-independent fitness burden in a minority of strains (Figure 4a). We hypothesized that these strains might harbor abortive defenses that detect specific phage proteins and induce PCD in response to a perceived infection. Notably, several of these trigger AGs were known anti-restriction factors (Supplementary Figure 3a). Putative anti-restriction-induced abortive (ARIA) mechanisms might serve as a backup to thwart phage/MGE-encoded counter-defense strategies.

We transformed the respective hosts with trigger AGs and confirmed their abortive phenotypes (Figure 4b). We then performed genome-wide transposon suppressor screens to find the genetic locus responsible for PCD. Transposon insertions in the abortive system would be expected to inactivate PCD and allow survival of the bacteria despite expression of the trigger AG. We identified putative ARIAs in all hosts except in ECOR66, where no suppressor mutants (other than disruptions of the trigger AG) could be obtained (Table 1b, Supplementary Figure 4, Supplementary File 6).

Wild strains encoding these ARIAs respond variously to R-M inhibitors, even discriminating between the two DNA mimics, Ocr and ArdA, suggesting a diversity of mechanisms (Supplementary Figure 3a). Notably, Ral, which operates by allosterically hyperactivating the methyltransferase (unlike Ocr and ArdA)[5] did not trigger PCD in any strains. Conversely, a novel AG (orf7) that triggers PCD by activating the same ARIAs triggered by Ocr

and ArdA (Supplementary Figure 3a) does not inhibit the *EcoKI* type I R-M system (Supplementary Figure 3b). Further, the predicted structure of orf7 – hereafter, *aoa1* for Activator Of ARIA – does not resemble a DNA mimic (Supplementary File 8), and it is also not as negatively charged (pI ~4 for Ocr/ArdA, ~8 for *aoa1*). It thus seems possible for a canonical R-M inhibitor to not trigger ARIAs, and for ARIAs to be triggered by R-M inhibitors as well as other phage factors.

We were surprised to find that a type III R-M system in ECOR22 (hereafter, EcoR22III) was required for cell-death in response to Ocr (Table 1b, Figure 4c), which blocks both type I and type III R-M[44]. While R-M defenses are ubiquitous in bacteria (in particular, *E. coli*), so are anti-restriction factors in phages. In response, bacteria evolved so-called anti-anti-restriction mechanisms (such as retrons[45] and PARIS[46]) to serve as backup for their frontline defenses. Our results suggest that in ECOR22, the frontline (R-M) and backup (PCD) defense functions have coalesced into the same system. To test this explicitly, we reconstituted the PCD and R-M activities of EcoR22III in a model *E. coli* strain.

In its native host, the *mod* and *res* genes of EcoR22III are associated with an SNF2-like helicase found in DISARM systems, *drmD*[47], and all 3 genes are encoded in a P4 satellite prophage hotspot[46]. However, transposon-suppressors of Ocr-toxicity in ECOR22 were found only in *mod* and *res* (Figure 4c, Supplementary Figure 4). When overexpressed in a lab strain, the 3-gene system did not produce a PCD response to Ocr (not shown). We then overexpressed EcoR22III alone (only the *mod* and *res* genes) and found that EcoR22III was now toxic upon Ocr induction but did not react to ArdA or Aoa1, consistent with the behavior of the wild host (Figure 4d). Ocr was not toxic in the lab strain on its own (see Figure 5a). An F54V Ocr mutant, previously reported to evade detection by an ARIA known as PARIS[46], also lost the ability to trigger EcoR22III PCD (Supplementary Figure 3c). Since there were no other R-M loci in the lab strain nor its native host ECOR22, EcoR22III likely detects Ocr directly rather than its effect on another R-M system.

Next, we tested whether EcoR22III possessed the canonical R-M activity. We constructed an active-site mutant of the PD-(D/E)xK restriction nuclease (res*: D1024A, D1043A double mutant) and compared immunity conferred by wildtype and res* EcoR22III against a panel of phages (not shown). EcoR22III could block λvir, but infection was not restricted by the putative restriction-dead mutant (Figure 4e). Furthermore, λvir passaged through the res* host (and thus modified by the presumably intact EcoR22III methyltransferase) regained the ability to infect cells carrying wildtype EcoR22III, indicating that EcoR22III indeed restricts only unmodified phage DNA.

We wondered if the targeted (R-M) and abortive (PCD) defense functions of EcoR22III could be uncoupled. EcoR22III genes *mod* and *res* individually did not induce cell death, indicating that PCD is a property of the complex (Figure 4f). The res* nuclease-dead mutation in EcoR22III that had completely abrogated its ability to target phage λvir (Figure 4e), also decreased PCD by 100-fold but did not fully abolish Ocr-induced toxicity (Figure 4f). Therefore, the EcoR22III PD-(D/E)xK nuclease motif appears largely responsible for its PCD and R-M activities.

Overall, EcoR22III is capable of mounting targeted antiviral defenses as well as an abortive response in the presence of certain phage proteins, possibly through different mechanisms mediated by the same effector (Figure 4g). Coupling restriction and PCD could allow cells to respond dynamically depending on the progress of the infection[48].

A distinct Type I R-M derived system deploys a small toxin to provide decoy immunity

In contrast, a second R-M-like ARIA we identified in ECOR55 from our transposon screens (Table 1b, Supplementary Figure 4) has acquired a similar PCD function but lost the restriction capability. This locus includes *hsdS* and *hsdM* genes typical of type I R-M but lacks

the nuclease (*hsdR*), and instead encodes a tightly linked 61aa protein. We named this system *ronin*, for the master-less samurai in the 1962 Japanese film, *Harakiri*.

Ronin induces cell suicide when Ocr, ArdA, or *aoa1* are expressed in ECOR55 (Figure 4b), but only responds to ArdA and *aoa1* (and not Ocr) when cloned into a lab strain (Figure 5a). The same results were obtained with a nearly identical system found bioinformatically in ECOR66 (not shown).

Several lines of evidence suggest that the 61aa protein (hereafter, *ronA*) in Ronin is the PCD effector. A highly hydrophobic transmembrane helix suggests that RonA might interfere with membrane integrity (Figure 5b). Over-expression of *ronA* alone is lethal to the host, but this effect is suppressed by HsdMS co-expression (Figure 5c). Ronin activity is independent of the methyltransferase function of HsdMS as PCD was not abrogated by the inactivating F287G mutation in motif IV (NPPF) of the *hsdM* catalytic domain[49] (Figure 5d). However, removal of *ronA* inactivates Ronin, such that anti-restriction proteins no longer elicit a PCD response (Figure 5d). Live cell imaging revealed that Ronin activation arrested growth but did not lyse cells, consistent with RonA-mediated inner-membrane disruption as a hypothetical PCD mechanism (Figure 5e). As *ronA* is not lethal when co-expressed with other Ronin components, we suggest that its activity is sequestered by the putative complex formed by HsdM and HsdS (hereafter, HsdMS). This hypothetical model is reminiscent of PrrC activation upon blocking of EcoPrrI restriction by T4-encoded peptide Stp[50]: similarly to how Stp binds the EcoPrrI HsdRMS complex to release the abortive anticodon nuclease PrrC, an anti-restriction protein might bind Ronin HsdMS and release the lethal effector.

Next, we examined the gene neighborhoods of *ronA* homologs in all complete bacterial genomes. We found that they were always linked to *hsdS* and *hsdM* genes, and the vast majority were arranged in one of two genetic architectures (Figure 5f). The more prevalent architecture (~70%) contained an uninterrupted (>400 aa) integrase upstream, and an atypical R-M system downstream (with two DUF262 proteins often found in type IV R-M in addition to the type I R-M *hsdRMS* genes) transcribed from the opposite strand. This arrangement was never found in *E. coli* genomes. The less prevalent architecture (~30%) was exclusive to *E. coli* and often contained a downstream BREX system, also encoded in the opposite strand. Importantly, most (77%) *ronA*-encoding *E. coli* genomes did not contain any type I R-M systems anywhere in the genome, and therefore lacked *hsdR* genes (Supplementary File 9). Although it is certainly possible that an unlinked HsdR is recruited *in trans* in other species, the predominantly nuclease-lacking configuration in *E. coli* suggests that ARIA activity is Ronin's primary function. This is reminiscent of decoy immunity (especially common in plants), where a host factor mimics the target of a pathogen effector. When the pathogen effector binds a decoy instead of its actual target, it alerts the host to the presence of the pathogen[25].

Thus, Ronin appears to be a type I R-M derivative that has gained the ability to scout for anti-restriction effectors but lost its prototypical defense function. Ronin encodes the canonical target of Ocr/ArdA (the type I R-M HsdMS complex), and its activation possibly depends on the ability of these anti-restriction proteins to bind and inactivate HsdMS (Figure 5g). Consistent with this, the Ocr F54V mutant, which evades PARIS and EcoR22III but does not lose anti-restriction activity[46], could not evade Ronin (Supplementary Figure 3c). Notably, Ronin almost always appears linked to a type I R-M or BREX system, both of which can be blocked by Ocr[43]. Hence, Ronin might drive evolution of Ocr-encoding phages to a BREX/R-M susceptible state, suggesting that decoy immunity could be a powerful host strategy to constrict the evolution of virus escape mutants.

Prophage-encoded abortive infection mechanisms can be triggered by specific AGs

Three additional abortive loci triggered by counter-defense-associated AGs were identified through whole-genome suppressor screens (Table 1b, Figure 6). These were all

prophage-linked, implying that prophages can activate superinfection exclusion mechanisms in response to individual phage products.

First, we found that transposon-insertions in the P4 satellite prophage (especially the ~100-bp intergenic region marked by a star in Figure 6a) alleviated the lethality due to orf116 expression in ECOR15 (Table 1b, Supplementary Figure 4). We cloned the satellite prophage onto a plasmid and reconstituted this effect in a lab strain of *E. coli*. While P4 often encodes a hotspot typically occupied by superinfection exclusion systems[46], the P4 construct we cloned does not contain this hotspot. Therefore, the P4 prophage itself kills the host in response to expression of orf116 (Figure 6a). Orf116 is the *abc1* gene from phage P22 and is adjacent to a known recBCD inhibitor *abc2* (Supplementary File 5), but its own function is less clear[51]. Abc1 contains an HTH domain suggesting that it could be a DNA-binding protein (Supplementary Files 3, 9). *Abc1* expression triggers the PCD function of the P4 prophage, but this toxicity is blocked when expression of P4 is also simultaneously driven with the strong pBAD promoter (Figure 6a). Because the toxicity of Abc1 in this strain can be alleviated by interfering with P4 prophage regulation, this result confirms the specificity of the P4-Abc1 interaction. The P4 satellite prophage cannot be induced without a co-resident P2 prophage (which was absent in our lab strain), ruling out a prophage-induction-based mechanism of action. We speculate that *abc1* interferes with some aspect of transcriptional regulation of the prophage, perhaps, with the repression of the P4 *kil* gene known to have a cytotoxic effect[52]. This hypothesis seems to be supported by the observation that the intergenic region that likely contains the *kil* promoter was heavily targeted by transposon insertions in our suppressor screens (Supplementary Figure 4).

The gop-beta-cII abortive system has been previously identified as a toxin-antitoxin-like locus[53], and suggested to exert anti-phage activity against coliphages λ and P1[46]. Our suppressor screens showed that gop-beta-cII toxicity is activated by orf184 (Table 1b, Supplementary Figure 4), which is widely distributed among coliphages and often present in the same AR with the anti-restriction gene *ral* (Supplementary File 5). We cloned gop-beta-cII into a lab strain of *E. coli* and confirmed that orf184 triggered abortive defense (Figure 6b). However, orf184 is not present in λ or P1.

Finally, a novel prophage-encoded system in ECOR61 inhibits growth upon expression of orf24 (Table 1b, Supplementary Figure 4), which is extremely narrowly distributed and is associated with the anti-restriction gene *ocr* in phage genomes (Supplementary File 5). We reconstituted this effect in a lab strain of *E. coli* and found that the effect is more muted (Figure 6c) than in the native host (Figure 4b). An Icd-like cell division inhibitor protein is the likely effector, as the system loses function when it is removed (Figure 6c). Our results highlight parallels between guardianship of defense systems by ARIAs and prophage-mediated superinfection exclusion (Figure 6d).

Conclusion

The PhAGENT platform harnesses the prediction that AGs manipulate host processes to provide a snapshot of the phage-host interface. This approach can systematically measure the fitness effects of hundreds of AGs in parallel, as well as their ability to trigger or block bacterial responses to infection. Here, we focused PhAGENT towards the discovery of counter-defense AGs.

In addition to higher throughput, the PhAGENT screen offers several advantages over inferential experiments with individual strains and phages. First, wild bacterial isolates are often not sensitive to many phages. Even when extensive host-phage collections can be curated, infection experiments have binary outcomes (i,e. survival or death) which can present challenges for deconvoluting the multiple mechanisms at play. A gene-delivery approach

bypasses this barrier and allows for the discovery of both the interacting defense and counter-defense factors at once.

Second, phages that might encode modulators of specific defense systems frequently prove impossible to procure and grow. Gene-synthesis allows AGs to be derived from any sequenced phage. In particular, assaying AGs individually in many wild bacterial strains can directly identify thus-far elusive triggers of defense systems. Abortive defense systems are generally thought to detect conserved phage factors (nucleic acids, capsid, terminase, portal, or a general effect such as host transcriptional arrest)[54-56], but our results show that poorly conserved phage genes can also trigger suicide mechanisms in the host and its prophages. Broadly understanding the principles of protein detection in antiviral immunity could yield a generalizable protein-targeting toolkit.

Third, we identified exaptations of well-studied R-M systems for mounting PCD responses to anti-restriction proteins that were not identified by sequence comparisons with known defense/counter-defense factors, which can miss subtle differences or miss divergent genes entirely. For instance, EcoR22III does not contain known domains that might suggest its PCD activity. Similarly, *ronA*-like toxins have not previously been linked with type I R-M genes (missed likely due to their small size).

Fourth, the PhAGENT screen offers more specificity than overexpressing putative defense systems in model strains to determine whether they are involved in antiviral defense in nature because the phage-host interaction is (a) revealed principally by a phage countermeasure to thwart the defense system, and (b) discovered in wild host strains where the defense systems are natively regulated. While redundant defense pathways may lead to missed interactions, co-expressing colinear configurations of multiple AGs in the same hosts could further increase the sensitivity of the approach.

Finally, any genetically tractable bacteria or archaea can be queried by PhAGENT. Our approach is especially suitable for discovering rare interactions that may be present only in a particular evolutionary niche. Deploying this platform more broadly in various pathogenic bacteria will help determine how phage-therapy approaches will need to be tailored to overcome species-specific antiviral mechanisms.

While we synthesized a library of 200 AGs to test the PhAGENT platform, the same workflow can scale to substantially larger libraries, since we routinely perform transposon mutagenesis screens with ~10^5 elements. The throughput could further be expanded by a double-barcoding strategy that allows a single read to identify both the host strain and AG. The platform is therefore only limited by our ability to synthesize phage genes. Stepwise screens could help explore the phage pangenome more efficiently; for instance, we sampled >60 accessory regions here, but could focus subsequent libraries on ARs with confirmed host-interacting AGs. Expanding the scope beyond AGs could also highlight secondary host-modifying functions of conserved phage genes such as capsids, polymerases, tail proteins, etc.

Overall, PhAGENT uncovers the fundamentally layered nature of anti-phage immunity in bacteria, and can tease the layers apart, providing a holistic view of bacterial defenses. We discovered phage effectors that interact with this multilayered defense architecture at every step: virus proteins that inhibit metabolism to modify the cell surface and block viral entry, counter-defense effectors that block the activity of antiviral restriction mechanisms inside the cell, cellular abortive immunity that guards against counter-defense strategies deployed by viruses, and phage-encoded triggers of superinfection exclusion mechanisms in prophages and other mobile elements in the host genome. With more detailed maps of phage-host interactions in many species, we hope to generalize these themes in phage-host co-evolution.

Figure 1. Phage Accessory Gene Exploration aNd Testing (PhAGENT) platform. (A) Schematic of a genomic region showing core (shown in grey) and accessory (in various colors) genes in a hypothetical phage genome cluster. **(B)** Schematic of algorithm to exhaustively identify accessory regions using pairwise combinations of genes, without *a priori* genome clustering. **(C)** A sample high-scoring AR in a family of *Salmonella* phages. Purple shading denotes >80% and green denotes >90% nucleotide identity. **(D)** Statistical features of ARs and AGs. 2D chart depicts saturation curves (cumulative distribution functions) of unique AGs (left axis) and phage genomes (right axis) represented in ~2000 non-redundant ARs ranked by score on the X-axis. 3D bar graphs show various distributions (length of AGs, number of AGs per AR, number of AGs per phage, and number of ARs per phage). Color-coded histograms in each bar graph are cumulative i.e., the first histogram (dark blue) shows the top 20% highest-scoring ARs, the next (light blue) represents top 40% and so on. **(E)** Schematic of the PhAGENT screen. A barcoded library of IPTG-inducible AGs is transformed into each host in triplicate via Tn7 transposition resulting in single-copy chromosomal insertions. Each replicate host library is then challenged with each phage, and samples are collected at time 0, and 30-60 min post infection. Uninfected but induced controls are maintained for every experiment. Another round of plating and scraping depletes barcodes from cells killed by phage infection or AG expression, and high-throughput sequencing is used to quantify remaining barcodes. >>10X as many colonies as the number of AGs are scraped at each step to avoid bottlenecking.

Figure 2. AGs that sensitize hosts to infection by modifying barrier defenses. (A) Heatmap of log2-transformed fitness scores comparing infected and uninfected samples. Of 200 AGs tested, only AGs with a host-sensitizing phenotype are shown on the X-axis. Red boxes/bold names highlight counter-defense AGs selected for further study. Named AGs are orf1:*ocr*, orf7:*aoa1*, orf48:*gnarl1*, orf63:*gnarl2*, orf87:*ral*, orf92:*gnarl3*, orf116:*abc1*, orf143:*IpII* and orf169:*ardA*. **(B)** Schematic of follow-up transposon knockout screens, performed in the native host without AGs. Graphs show fitness of transposon-mediated knockouts in infected (Y-axis) and uninfected (X-axis) Tn5-libraries from three representative hosts challenged by T5. Red dots are gene-disruptions that lower bacterial fitness upon infection (e.g., disruptions in O-antigen biosynthesis). **(C)** Plaque assays with phage T5 and hosts from (D) with *gnarl* genes repressed (top) or induced (bottom). **(D)** Lipopolysaccharide (LPS) visualized by electrophoresis from host-AG combinations in (E) where AGs produced a phage-sensitizing effect. **(E)** Log2-transformed protein representation scores (MaxLFQ) for the top 30 enriched host proteins upon expression of AGs. **(F)** log10-transformed T5 infection scores in wild hosts ECOR1 (green), ECOR3 (orange), ECOR22 (blue) with or without UDP-glucose biosynthesis pathway genes *galU, galE, galF,* and ECA precursor *wecB* cloned onto a plasmid and overexpressed. (See also Figure S2)

Figure 3. AGs that block canonical antiviral defenses. (A) Fitness of transposon-mediated knockouts in infected and uninfected Tn5-libraries from ECOR21 challenged by various phages. **(B)** Schematic of ECOR21 defense against T2 through both barrier defenses and a defense system which is putatively blocked by orf143/IpII. **(C)** LPS visualized from ECOR21 with the O-antigen biosynthesis initiator *wecA* deleted by allelic exchange. **(D)** Whole-genome transposon-knockout screens as in (A) but with the ECOR21::Δ*wecA*. **(E)** Variants of the BREX/GmrSD operon cloned onto a plasmid along with their native promoters and tested in a lab strain of *E. coli* (TOP10) against T-even phages (T2, T4, T6). Removal of the PglX gene inactivates BREX. Active-site mutations that disable GmrSD are shown as two parallel red bars (GmrSD* double mutant; D474A, H475A). Ocr, IpI and IpII are expressed from single-copy chromosomal insertions. Bar graphs show log10 infection scores. **(F)** GmrSD and its inactive variant GmrSD* cloned onto a plasmid and tested in TOP10 for restriction activity against T-even phages. "pBAD" indicates the defense construct was overexpressed from the plasmid. **(G)** Plaque

assays with TOP10 cells with uninduced GmrSD (i.e., not overexpressed) challenged by T-even phages.

Figure 4. AGs that trigger abortive defense in the host. (A) Heatmap of log2-transformed fitness scores comparing induced and uninduced samples without phage infection. Of 200 AGs tested, only AGs that exhibited lethality in some (but not all) hosts are shown. Red boxes/bold names highlight conditional-lethal AGs selected for further study. Named AGs are orf7:*aoa1*, orf87:*ral*, orf63:*gnarl2*, orf1:*ocr*, orf92:*gnarl3*, orf116:*abc1*, orf169:*ardA* and orf2:*arn*. **(B)** Host-AG combinations in red boxes from (A) tested individually for AG-induced lethality (serial tenfold dilutions of saturated culture; AGs repressed: left, induced: right). **(C)** Schematic depicts hypothetical strain-specific abortive defense system being triggered by AGs, and inactivated in follow-up transposon knockout screens (inset). Data from a representative screen with a Tn5 library constructed from ECOR22 with chromosomally encoded *ocr*. Proportion of transposon insertions per gene in surviving cells with *ocr* repressed is plotted on the X-axis, and with *ocr* induced on the Y-axis. Y-axis interrupted to resolve genes other than type III R-M *mod*, *res*. (See also Figure S4) **(D)** Abortive phenotype of EcoR22III in its native host, and reconstituted when cloned onto a plasmid and expressed in TOP10 (AGs repressed: left, induced: right)**. (E)** Plaque assay showing R-M phenotype of EcoR22III against methylated and unmethylated λvir in TOP10. λvir* was methylated by passaging in TOP10 strain with nuclease-dead EcoR22III (res* double mutant; D1024A, D1043A) and tested for its ability to evade restriction by wildtype or res* EcoR22III. **(F)** As in (D), PCD activity of EcoR22III variants in TOP10 (wildtype, nuclease dead (res*), and *mod*, *res* genes expressed individually). **(G)** Hypothetical model of EcoR22III R-M and PCD activities. EcoR22III kills its host in the presence of Ocr, partially relying on its nuclease activity. “pBAD” indicates the construct was overexpressed from the plasmid in (F-H).

Figure 5. Decoy immunity in bacteria. (A) Schematic of Ronin system from ECOR55. Abortive phenotype reconstituted in TOP10 upon expression of Ronin from a plasmid (serial tenfold dilutions of saturated culture; AGs repressed: left, induced: right). **(B)** Prediction of transmembrane character along the 61 aa RonA sequence. **(C)** Lethality of Ronin components tested when overexpressed in TOP10 (indicated as “pBAD”) without AGs. **(D)** As in (A), genetic requirements for the abortive phenotype of Ronin. Red bar in *hsdM* (second row) denotes the F287G active-site mutation. (See also Figure S5) **(E)** Time-lapse imaging of TOP10 cells expressing Ronin with (bottom) or without (top) chromosomally-integrated ArdA. **(F)** Two prevalent architectures of *ronA*-linked type I R-M-like abortive system (ronin), based on the presence of an upstream DUF4102 integrase. Number of genomes with a full type I R-M system (with *hsdR, hsdM, hsdS* genes) anywhere in the genome is indicated. **(G)** Schematic depicting hypothetical RonA “uncaging” when ArdA is present.

Figure 6. AGs that trigger prophage-encoded superinfection-exclusion mechanisms. (A) Abortive phenotype of P4 satellite prophage from ECOR15 cloned onto a plasmid and expressed in TOP10. Orf116 is expressed from a single-copy chromosomal insertion (serial tenfold dilutions of saturated culture; AGs repressed: left, induced: right). Star marks an intergenic region that was strongly targeted in transposon suppressor screens. (See also Figure S4) **(B)** As in (A), abortive phenotype of gop-beta-cII system from ECOR59 upon expression of orf184. **(C)** As in (A), abortive phenotype of novel prophage-encoded superinfection exclusion system from ECOR61 upon expression of orf24. **(D)** Schematic representation of similar defense architecture of ARIAs and superinfection exclusion systems. “pBAD” indicates the ARIA was overexpressed from the plasmid in (A-C).

Table 1. Summary of transposon-mediated whole-genome knockout screens. Follow-up screens to find **(A)** phage exclusion mechanisms that could be the targets of counter-defense

AGs identified in PhAGENT screen and **(B)** abortive defense mechanisms that might trigger PCD in response to conditional-lethal AGs. A total of 26 whole-genome knockout screens were performed in (A), with 7 Tn5 libraries (one in each host) challenged by the indicated phages. In contrast, each host-AG combination in (B) was screened as a separate Tn5 library.

Acknowledgements

S.S. was supported by the Damon Runyon Fellowship Award (DRG 2352-19). Experiments were funded by awards to J.B.D by the Vallee, Searle, and Kleberg Foundations. We thank Jason Peters for technical discussions, Andy Millard for the Enterobacteriophage genome dataset, and Ry Young, Andrew Fire, Carol Gross and Jonathan Weissman for guidance.

Author Contributions

Conceptualization, S.S., J.B.D.; Methodology, S.S., E.L.; Resources, A.F., D.S.; Software, S.S.; Investigation, S.S., H.C., D.S.G.; Formal Analysis, S.S., H.C., M.J., K.S.M., E.V.K.; Visualization, S.S., H.C., M.J., K.S.M.; Writing – Original Draft, S.S., H.C.; Writing – Reviewing & Editing, S.S., H.C., M.B., E.V.K., J.B.D.; Supervision, S.S., J.B.D.; Funding Acquisition, S.S., J.B.D.

Declaration of Interests

J.B.D. is a scientific advisory board member of SNIPR Biome, Excision Biotherapeutics, and LeapFrog Bio, consults for BiomX, and is a scientific advisory board member and co-founder of Acrigen Biosciences. The Bondy-Denomy lab received research support from Felix Biotechnology.

Inclusion and Diversity

We support inclusive, diverse, and equitable conduct of research.

Supplementary Figure 1. Phage-host combinations included in PhAGENT screen (A) Phage susceptibility of select wild ECOR strains. Numbers in boxes are log10-transformed infection scores (7 = 10^7-fold higher plaquing than 0). Borders indicate that individual plaques were visible in that phage-host combination. Combinations with attenuated phage infection that were selected for the PhAGENT screen are in yellow. Combinations with attenuated infection that were attempted but yielded no survivors in liquid infection experiments (perhaps due to abortive defenses) are in red. **(B)** Phage-host combinations where counter-defense phenotypes were confirmed for any AGs. Red denotes combinations where the screen could not be completed (as in (A)). Yellow: combinations that were tested where none of the 200 AGs produced counter-defense effects. Green: combinations where phage-sensitizing effects could not be verified by plaque assays. Blue: combinations where counter-defense phenotypes were confirmed by plaque assay (at least 100X higher EOP upon AG expression). **(C)** Defense repertoire of selected ECOR strains. Computational prediction of the presence or absence of various defense systems in ECOR strains used in the PhAGENT screen. Predictions were generated using the ISLAND software suite[57].

Supplementary Figure 2. Over-expression of *galU* reverses T4 sensitization by *gnarl3*. log10-transformed T4 infection scores in wild hosts ECOR1 (green), ECOR3 (orange), ECOR22 (blue) with or without UDP-glucose biosynthesis pathway genes *galU, galE, galF,* and ECA precursor *wecB* cloned onto a plasmid and overexpressed.

Supplementary Figure 3. Abortive effects of known R-M inhibitors in wild *E. coli.* (A) Schematic representation of conditional-lethality data from Figure 5b, showing only known R-M inhibitors and *aoa1*. **(B)** AGs from (A) tested for their ability to block EcoKI restriction of phage λvir in plaque assays with restriction-competent *E. coli* MG1655. **(C)** F54V mutant of Ocr (known to evade a previously described ARIA, PARIS) still triggers ARIA activity in ECOR55 but not ECOR22.

Supplementary Figure 4. Transposon insertions that suppress AG-induced conditional-lethal phenotypes. Number of unique transposon insertions recovered from survivors from Tn5 libraries, mapped to putative abortive loci. Tn5 libraries were prepared in the indicated wild hosts and survivors obtained upon AG-induction. E.g., in ECOR15, >500 insertions were recovered from the ~100 bp intergenic region marked by a star, and at the very 3'end of the DNA primase CDS. Scale is indicated for each locus.

Supplementary Figure 5. Upstream non-coding region is required for Ronin PCD activity. (A) The first 135-bp of the 235-bp upstream non-coding region (marked by a thick red line) are required for activity as truncation down to 100-bp causes inactivation. ArdA is chromosomally integrated under pLlacO-1 control and strains in the left (ArdA repressed) and right (ArdA induced) panels are isogenic. **(B)** Over-expression of 235-bp upstream region is not toxic in TOP10 indicating that it is not the PCD effector. **(C)** As in (A), abortive phenotype of truncated Ronin system when overexpressed (indicated as "pBAD"). Over-expression of truncated Ronin did not rescue the PCD defect indicating that the upstream region contributes something in addition to the promoter.

Supplementary File 1. Worked example of PhAGENT algorithm. Readme.txt describes how to run the code. Sample database of Salmonella phage genomes is included. All ARs from this phage dataset are reported in the output file SPFMphages_accRegions_nr_all.txt. Any AR can be extracted and visualized using PhamLite[58]. The AR shown in Figure 1d is extracted by default.

Supplementary File 2. Manually inspected accessory regions from Enterobacteriophage genomes. Visualizations of all ~200 accessory regions (ARs) that were manually inspected. The first slide in the file shows the layout for the rest, with the AR score calculated according to the formula in Methods, followed by a unique identifier for each hypothetical phage group ("phamily"). Each AR was also given a short, memorable name (top right). Slides with green backgrounds indicate ARs from which genes were selected for PhAGENT screen. Authors' impressions upon AR inspection are included in the Notes section.

Supplementary File 3. Phage accessory genes selected for PhAGENT screen. 200 accessory genes (AGs) selected for screening. Genomic locations, protein and phage genome identifiers, and protein descriptions are indicated. IDs denoted by "UDP" correspond to the orf/AG numbers used throughout the manuscript.

Supplementary File 4. Experimental record of PhAGENT screen. A detailed record of all PhAGENT experiments is presented. The first sheet contains phage-host combinations, dates of the experiment, and preliminary results. The next three sheets each correspond to one replicate of PhAGENT screening. The timeline of each experiment indicates library growth characteristics. Empirically measured multiplicity at time of infection, and dilutions used before plating are also recorded. Screens that had to be repeated for technical reasons are color-coded.

Supplementary File 5. Accessory regions of AGs selected for further study. Visualizations of ARs where AGs that produced counter-defense or conditional-lethal phenotypes in PhAGENT screen (see Table 1) are located.

Supplementary File 6. Experimental record of follow-up transposon whole-genome knockout screens. A detailed record of all follow-up transposon screens is presented. The first sheet contains details of screens designed to phenocopy counter-defense phenotypes of AGs identified in PhAGENT screen. These screens are performed with RB-TnSeq libraries prepared with naïve ECOR hosts (no AGs present). Each library is challenged by various phages one at a time. The timeline of each experiment indicates library growth characteristics. Empirically measured multiplicity at time of infection, and dilutions used before plating are also recorded. Screens were performed in two stages: first, a pilot with one phage for each RB-TnSeq library, and then a larger experiment with all other phages that showed enhanced infection upon expression of a counter-defense AG. The second sheet contains details of screens designed to inactivate putative abortive defense mechanisms responsible for AG-triggered conditionally-lethal phenotypes. Tn5 libraries are constructed in ECOR strains already transformed with Tn7 transposons carrying trigger AGs (maintained under strict glucose suppression). No phages are involved. Survivors are selected and sequenced after trigger AG induction upon plating on IPTG. Screens are color coded to indicate whether they were ultimately attempted.

Supplementary File 7. Host-AG combinations for affinity-purification/mass-spectrometry. Matrix of all host-AG combinations where potential binding partners of AGs were affinity-purified from the native host expressing counter-defense AGs, and identified by mass-spectrometry. Experiments were performed in triplicate. Orf74 was inert in PhAGENT screening and was used as a control for all experiments in the relevant hosts.

Supplementary File 8. AlphaFold predictions for all AGs that produced phenotypes in PhAGENT screens. PDB files containing predicted structures of orfs 7 (aoa1; triggers PCD by Ronin), 24 (triggers PCD in ECOR61), 48, 63, 92, (*gnarl*1-3; putative O-antigen modifiers), 116 (*abc1*; triggers PCD by P4 prophage), 126 (broad-spectrum counter-defense), 143 (IpII; inhibitor

of GmrSD), 145 (triggers PCD in MG1655), 148 (broad-spectrum counter-defense), and 184 (triggers PCD by gop-beta-cII).

Supplementary File 9. Genomic architectures of all ECOR55-like Ronin systems. Spreadsheet contains accession numbers and genomic locations of all *ronA* homologs. Grouping variables indicate whether a type I R-M system (containing *hsdR, hsdM,* and *hsdS*) was found anywhere in the genome, and also whether a full (>400 aa) DUF4102 integrase was present immediately upstream of Ronin. HTML files split Ronin systems into two groups based on the presence of the integrase, and gene annotations can be viewed by mouse-over.

Methods

Algorithm to identify phage accessory regions

1706 non-redundant Enterobacteriophage genomes were obtained (with help from Dr. Andrew Millard, Leicester) from public databases. These phages were both lytic and lysogenic, but were sequenced and assembled from cultured isolates (i.e. prophages and metagenomic samples were not included). All annotated protein records from all the phages were assigned unique numeric IDs and clustered by MMseqs2[59]. ORF IDs in phage genomes were then replaced by their MMseqs2 cluster representative IDs in each genome. Thus, each phage genome can be represented as an ordered list of IDs of clustered proteins. The gene-neighborhood of every ORF (up to 10 genes upstream and downstream) was extracted from every virus. Given that phage genomic termini can be difficult to identify, every genome was considered potentially circularly permuted. Allowing for circularization also ensures all gene-neighborhoods were the same size and putative ARs towards ends of genomes would not be missed. A 2D matrix of all possible pairwise combinations of genes was constructed, with every matrix position containing lists of (a) the phages that each gene appears in, (b) the 10 upstream, and (c) 10 downstream neighboring ORFs of each gene in every genome. We then traversed the matrix and assessed every gene pair to identify all possible ARs, defined as the region between a pair of genes that are both present in at least 6 different phage genomes but are no more than 10 genes apart in any genome. Regions trivially containing the same genes in each participating genome were discarded. Genomes involved in a putative AR were then compared pairwise at the nucleotide level using FastANI[60], and ARs with divergent genomes were removed. In practice, this was done by pre-computing an all-vs-all matrix of genome ANI and then traversing subsets of the matrix containing all the genomes participating in the putative AR to ensure no cells were empty (FastANI produces no output when very divergent genomes are compared). Various parameters were extracted from surviving ARs. *N* is the number of unique genes in the AR. The vector *c* contains conservation numbers (the number of different genomes that each gene appears in; also indicated on top of each gene) for all the unique genes in this AR. The vector *l* contains the lengths (in numbers of genes) of the AR in each genome. *CV* denotes coefficient of variation. *GCdev* is a vector containing one minus the absolute deviation in GC content of each accessory gene from its resident genome. A scaling factor is also included to normalize the score by the number of phage genomes in the group that have at least one gene in the AR. ARs were scored according to the formula below.

$$ARscore = \frac{N * median(GCdev) * CV(c)}{CV(l) * mean(c)} * \frac{\#non\_empty}{\#genomes}$$

To filter out ARs that were subsets of other ARs, we successively removed redundant, low scoring ARs with at least 90% of genes that had already been encountered in higher scoring ARs. The top 200 ARs were examined manually and 1-15 AGs were selected from 62 ARs (Supplementary Files 2,3). A worked example of the code is provided in Supplementary File 1.

Strains and Vectors

The *E. Coli* Reference (ECOR) collection of wild *E. coli* strains was obtained from the STEC Center (Michigan State University). Lab strains were obtained as follows: MG1655 (from Dr. Carol Gross, UCSF), C990 (from Dr. Ry Young, Texas A&M), DH5α (NEB), DH10ß/TOP10 (Life Technologies), BW25113 and BW25141 (from Dr. Vivek Mutalik, LBNL), and WM6026 (from Dr. Jason Peters, UW-Madison). All strains were cultured at 37°C with shaking (180-225 rpm) in lysogeny broth (LB) (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl with 15 g/l agar for plates) supplemented with 100 µg/ml of ampicillin/carbenicillin, 10-20 µg/ml gentamicin (except

WM6026, which yielded more consistent colony sizes with 50 µg/ml), 50 µg/ml kanamycin, and 0.5% w/v glucose, 0.1% w/v L-arabinose, and 1 mM IPTG (isopropyl-b-D-thiogalactopyranoside) as needed. The WM6026 strain required diaminopimelic acid (DAP) for growth, which was supplied at 300 µM. All strains were stored in 20% glycerol at -80°C for long term storage.

R6K-origin pTn7 plasmids (from Dr. Jason Peters, UW-Madison) – the pJMP1039/pTn7C1 helper plasmid (carbenicillin) and the pJMP1360/pTn7C185 transposon vector (gentamicin, carbenicillin) – were used for integration of phage AGs into the *E. coli* genome. All host defense systems identified in the various screens were placed under the araBAD promoter (arabinose inducible) in the pBAD Myc/His A vector (Invitrogen). Conjugative allelic exchange was carried out using R6K-origin pKEK2201 (kanamycin) (from Karl E. Klose, UT San Antonio). All plasmids were constructed by Gibson Assembly.

R6K-origin plasmids were maintained in BW25141 (routine cloning) and WM6026 (conjugative donor strain). All other plasmids were maintained in DH5α or TOP10. All plasmid sequences were verified by Sanger or whole-plasmid sequencing (Plasmidsaurus, Primordium) and are available upon request. Frequent plasmid dimerization during the Gibson Assembly procedure (especially when using the NEBuilder reagent) was ameliorated by using 10-fold less of one of the insert DNA fragments in a reaction involving at least 2 insert fragments (not including the backbone).

AG conjugation into host strains

AGs were delivered to *E. coli* hosts by Tn7 transposition via conjugation performed by triparental mating as described previously[61]. Briefly, the transposon and helper plasmids were delivered to the recipient from WM6026 donor strains by overnight incubation on LB agar plates supplemented with 300 µM DAP and 0.5% glucose to repress AG expression. The next day, cells were scraped and resuspended in 1 ml LB, and 1000X-20000X dilutions of this conjugation mixture were plated on LB plates (lacking DAP) supplemented with the appropriate antibiotics and 0.5% glucose. All strains were maintained under strict glucose repression of the pLlacO-1 promoter at every step until screens were performed. Thus, all AGs were maintained as single-copy chromosomal inserts at the same attTn7 site in the genomic region between *glmS* and *pstS* genes.

Efficient Tn7 transposition is necessary to construct a AG library in any given *E. coli* host. Tn7 conjugation-transposition efficiencies were determined in 44 ECOR stains containing an intact *glmS-pstS* region (we note that other strains can support Tn7 integration as well). Tn7 transposons carried only a barcode sequence and the gentamicin-resistance gene to minimize payload-dependent variation in conjugation efficiency between strains. Efficiency of conjugation was determined by streaking the conjugation mixture on two different gentamicin concentrations (25 µg/ml and 10 µg/ml) to account for possible differences in natural antibiotic resistance between the wild strains. The barcode sequence was verified by Sanger sequencing and both flanks were checked by PCR to ensure proper integration of the Tn7 transposon at the expected attTn7 site. Strains that supported Tn7 transposition at high levels and were amenable to agar-overlay plaque assays were selected for further experiments.

DNA extraction + Nextera Genomic Sequencing + Analysis

Genomic DNA was extracted using a modified SDS–proteinase K method: Briefly, cells pelleted from 50 µl of saturated culture (or 5 µl of the cell mixture scraped from selective plates during AG library construction) were re-suspended in 200 µl of lysis buffer (10 mM tris, 10 mM EDTA, 400 mg/ml proteinase K, 0.5% SDS) and incubated at 50-55°C for 1 hour. Subsequently, the temperature was lowered to 37°C and RNase A (Thermo Scientific) was added to a concentration of 1 mg/ml. After 30-60 min of incubation, the digests subsequently purified using the Genomic DNA Clean & Concentrator Kit (Zymo Research). DNA was prepared for whole-genome sequencing using the Nextera Flex Library Prep Kit (now, the Illumina DNA Prep Kit)

according to manufacturer's instructions, except that we lowered the volumes of all tagmentation reactions 5-fold, and used a custom dual-indexed primer set JSW-SS-22:33 (CAAGCAGAAGACGGCATACGAGAT NNNNNNNN GTCTCGTGGGCTCGG) and JSW-SS-34:41 (AATGATACGGCGACCACCGAGATCTACAC NNNNNNNN TCGTCGGCAGCGTC) to amplify libraries for 11-13 cycles using Phusion High Fidelity PCR Master Mix (NEB) instead of the Illumina-supplied PCR reagents. The N_8 sequences correspond to reverse-complemented Nextera DNA indexes N701 to N712 and N501 to N508, respectively (Illumina). Libraries were resolved by agarose gel electrophoresis and DNA was excised in the 300-400 bp range. Gel slices were purified using the Zymo Gel DNA Recovery Kit and sequenced on the Illumina MiSeq using 150-cycle v3 or 600-cycle v2 reagent kits.

Sequencing adaptors were trimmed from reads using cutadapt, mapped to reference genomes using bowtie2, and visualized using IGV. Reference genomic contigs of ECOR strains were obtained from the USDA[62]. SNPs, indels, and large structural variants were analysed using a mapping-free variant finder[63].

Phage propagation and plaque assays

High-titer phage stocks were generated by growth in liquid culture. All phage infections were performed at 37°C in LB supplemented with 10 mM $MgSO_4$ and 5 mM $CaCl_2$. *E. coli* BW25113 or C990 were grown to an optical density (at 600 nm; OD_{600}) ~ 0.5 (1-5 X 10^8 cfu/ml) and infected with the desired phage at MOI ~ 5. Infections were allowed to proceed at 37°C with agitation for 5-6 hours until the cultures were clear. Lysates were clarified by centrifugation (8000 X g for 2 min) and sterile-filtered using a 0.45 µm SFCA syringe filter (Millipore). Phage lysates were stored at 4°C.

Bacteriophages T2, T3, T4, T5, T6, T7, P1vir were obtained from Dr. Vivek Mutalik (LBNL), and λvir was obtained from Dr. Ry Young (Texas A&M). To the extent possible, all phages were grown on *E. coli* C990 which lacks restriction-modification. T4 and P1vir could not be grown on C990 and were instead grown on *E. coli* BW25113. To ensure that our phage lysates were free from ancestral type I R-M methylation marks or other host-passage derived DNA modifications, we serially passaged all phages in the desired hosts 3 times. All phage genomes were subsequently verified by Nextera high-throughput sequencing.

Phage titers were determined against various hosts by the agar-overlay method with LB top agar containing 0.7% agar, 10 mM $MgSO_4$ and 5 mM $CaCl_2$. Phage spot-titration assays were performed with 10-fold serial dilutions of phage lysates (3-5 µl of each dilution pipetted onto bacteria immobilized in top-agar overlay using a multichannel pipette).

AG Library Preparation + NGS verification + testing post-conjugation evenness

200 AGs with unique 10 bp DNA barcodes were synthesized (Twist and IDT) and cloned individually under pLlacO-1 control (IPTG inducible, repressed by glucose) into a modified pTn7C185 vector (with mobile CRISPRi components sgRNA and dCas9 removed). Plasmid assembly reactions were performed in 96-well format using the 2X Gibson assembly master mix (NEB) and transformed into chemically-competent BW25141 cells in 96-well format. Transformations were individually plated on LB agar supplemented with 20 µg/ml gentamicin and 0.5% glucose and one colony was initially checked by Sanger sequencing from each plate. Where a successful transformant could not be obtained on the first attempt, more colonies were sequenced in successive rounds until a plasmid with no mutations in the region spanning the pLlacO-1 promoter, the AG ORF, and the barcode sequence was obtained. Verified vectors were miniprepped using the ZR Plasmid miniprep – Classic kit (Zymo) and transformed into chemically-competent WM6026 cells in 96-well format, and each transformation was plated individually on LB agar supplemented with 50 µg/ml gentamicin, 300 µM DAP, and 0.5% glucose.

One colony from each of the 200 WM6026 transformants was grown in 2 ml square-well 96-well plates with shaking overnight, and 100 ul of each culture was then pooled together to yield ~20 ml of the donor mixture. This mixture was concentrated down to 10 ml by centrifugation at 6000 g for 10 min, and 5 ml of 60% glycerol (20% final) was added to yield the donor library. The 15 ml library was divided into single-use 50 ul aliquots (each containing >50M cells, i.e. >300K of each donor strain) and stored at -80°C.

Genomic DNA was extracted from one of the library aliquots and all constructs in the final pool were verified by Nextera high-throughput sequencing. This allowed us to assess library evenness as well as verify the fidelity of our cloning, colony picking, and strain pooling process. Overall, 195 AGs were cloned successfully. One AG (orf88) never yielded any colonies in multiple cloning attempts, possibly due to its cytotoxicity. Three AGs (orfs 36, 52, 94) were dropouts during the WM6026 growth step, and another two (orfs 25, 84) were discovered to contain mutations upon high-throughput sequencing that had been missed by Sanger sequencing. Orf25 had previously presented difficulties in cloning, and did not yield a mutation-free construct even after checking >12 colonies (whereas most transformations yielded a correct construct on the first attempt). The orf84 construct contained a mixed base at the last position of the second repeat of the Lac-operator sequence upstream of the ORF, which would not be expected to affect any downstream assays significantly.

The donor library was conjugated into MG1655 *E. coli* according to the Tn7 transposition procedure (above) and a 1000X dilution of the conjugation mixture was plated on LB agar supplemented with 20 µg/ml gentamicin and 0.5% glucose. We routinely use multiple 150 mm X 15 mm bacteriological petri dishes (Corning) to obtain sufficient numbers of transconjugant colonies to assemble libraries. The next day, >>2,000-20,000 transconjugant colonies (10-100X the number of AGs in the library) were scraped and combined thoroughly, and the thick cell mixture was directly frozen at -80°C in 20% glycerol. Genomic DNA was extracted from this mixture and the distribution of DNA barcodes corresponding to individual AGs was measured in both this post-conjugation strain mixture as well as the genomic DNA extracted from the WM6026 donor library aliquot.

To amplify barcodes for high-throughput sequencing, genomic DNA was subjected to two successive rounds of PCR. The DNA barcodes were amplified from genomic DNA using universal primers JSW-SS-170 (CGACGCTCTTCCGATCTNNNNN TGATGTCGTTGTTGCCATCG) and JSW-SS-171 (ACTGACGCTAGTGCATCA CTTTCTGAGCCAGTGTTGCT) and the Q5 Hot Start High-Fidelity 2X PCR master mix (NEB) according to manufacturer's instructions. Sequencing adaptors were attached in a second round of PCR using amplicons from the first PCR round as the templates, with dual-indexed primer sets JSW-SS-42:53 (CAAGCAGAAGACGGCATACGAGAT NNNNNNNN GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT ACTGACGCTAGTGCATCA) and JSW-SS-54:61 (AATGATACGGCGACCACCGAGATCTACAC NNNNNNNN ACACTCTTTCCCTACACGACGCTCTTCCGATCT), where the N_8 sequences correspond to reverse-complemented TruSeq HT indexes D701 to D712 and D501 to D508, respectively (Illumina). Template-matching regions in the primers are underlined. Cycling conditions for round 1 were as follows: one cycle at 98°C for 1 min; two cycles at 98°C for 10s, 66°C for 30s, and 72°C for 10s; 22 cycles at 98°C for 10s, and 72°C for 20s; and one cycle at 72°C for 2 min. Conditions for round 2 were one cycle at 98°C for 1 min; two cycles at 98°C for 10 s, 64°C for 30 s, and 72°C for 10 s; 4 cycles at 98°C for 10 s, and 72°C for 20 s; and one cycle at 72°C for 2 min. 0.5-1 ul of unpurified 1st round reaction product was used as template for the 2nd round of PCR. Amplicons from the 2nd round were gel-purified by electrophoresis (3% agarose gel, 4.2 V/cm, 2 hours) and quantified using the 1X dsDNA HS kit with the Qubit 4.0 Fluorometer (Invitrogen). Amplicons were sequenced on a MiSeq using the 150-cycle v3 reagent kit (Illumina) in single-end format (100 bp) with two (8 bp) index reads. DNA barcodes were

trimmed from the reads and the proportions of various AGs in the mixtures were determined using the prevalence of the corresponding barcode sequence.

Concordance between the Nextera whole-genome sequencing and the barcode sequencing of the WM6026 donor mixture showed that barcode representation was a suitable proxy for the true distribution of AGs in the mixture and that amplification biases during library preparation were not a significant source of noise. Concordance between barcode sequencing in the WM6026 donor and MG1655 post-conjugation recipient mixtures indicated that the Tn7 transposition and colony scraping process does not meaningfully bottleneck AG distribution.

The AG library was then conjugated (in triplicate) into ECOR strains that were previously confirmed to support robust Tn7 transposition. Strain mixtures were prepared in a procedure analogous to library construction in MG1655, except higher dilutions of the ECOR+AG conjugation mixtures (10,000-20,000X) were plated and the gentamicin concentration on selective plates was lowered to 10 µg/ml to account for higher steady state cell densities and higher gentamicin sensitivity of the wild *E. coli* strains. As with the MG1655 library, thick cell mixtures comprised of >>2,000-20,000 colonies scraped from each replicate conjugation were directly stored at -80°C in 20% glycerol and used to seed cultures for the PhAGENT screens (below). Genomic DNA was extracted from each mixture and barcode sequencing was performed to measure the AG distribution in every replicate of all ECOR strain derived libraries. The replicates were very highly correlated, and AG distributions in all libraries were confirmed to be free of significant bottlenecking. These datasets are the reference libraries that serve as the baseline for AG induction and phage infection experiments.

MOI calculation for liquid killing assays

Phage adsorption to host cells may be inefficient in very dilute cultures ($10^{6\text{-}7}$ cfu/ml). Therefore, the apparent multiplicity of infection (MOI-a) might be lower than the expected MOI (MOI-e). We added 2x serial dilutions of a λvir lysate to a 1:100 dilution of a saturated culture of *E. coli* BW25113 that had been allowed to acclimate at 37°C with shaking for 15 min. The proportion of surviving cells was determined (relative to uninfected controls) after a 60-75 min incubation (one phage replication cycle) by plating on LB agar. An apparent MOI value was calculated using the Poisson distribution function with zero occurrences $MOIa = -ln\ (MOIe)$ for each phage dilution, assuming that all infected cells would be killed by the virus (i.e. the proportion of surviving cells would be the zero class in the Poisson distribution). The apparent and expected MOI values were plotted against each other and the inverse of the slope (1/m) of the line (MOI-a = m*MOI-e + 0) fit through the data was determined as the fold-change reduction in MOI-a as compared to MOI-e for these experimental conditions. This fold-change reduction was determined to be ~26x. Hence, all phage infection screens were targeted to an MOI-a value of ~100 to ensure efficient phage adsorption and infection within the timeframe of a single phage replication cycle.

AG Screen + Plaque assay confirmation + Data analysis

Thick cell mixtures from each of the three replicates of the various ECOR+AG libraries were thawed on ice and diluted into LB supplemented with Mg^{2+}, Ca^{2+}, and 10 µg/ml gentamicin. Typically, 1-5 µl of the library mixture was diluted in 2 ml broth to yield an initial OD_{600} of 0.1-0.15. The cultures were immediately split x2, with one tube used to monitor OD_{600} values as the experiments progressed, and the other left untouched for the screen. AG expression was induced by the addition of 1 mM IPTG after acclimatization for 15 min at 37°C in a shaking incubator and OD_{600} values were measured again for each ECOR+AG library at time of induction. Libraries were then grown for ~80-120 min depending on the relative growth rates of the various ECOR strains, to OD_{600} ~ 1.2-1.5. A 10^{-5} dilution of each late-log phase culture was then plated on LB agar (with 10 µg/ml gentamicin and 0.5% glucose to repress AGs and avoid fitness costs) to obtain a count of viable colony forming units (cfu). Cultures were immediately

diluted ~1:20 in fresh LB (with IPTG, Mg^{2+}, Ca^{2+}, and 10 µg/ml gentamicin) and 1 ml aliquots were gently mixed with phage lysates at an MOI-e of ~30-100 (the actual MOI-e values were recorded the next day when cfu counts were obtained). Parallel uninfected controls were maintained for each experiment, even when the same ECOR+AG library was being tested against different phages and the uninfected controls would be interchangeable. Infections were allowed to proceed for 30-35 min for T-phages and 60-70 min for P1vir and λvir, i.e. one replication cycle. We note that allowing sufficient time for the phages to complete cell lysis was necessary as fast-growing cells were sometimes able to form colonies on LB agar despite having been infected. These colonies had an uneven size distribution and a "splotchy" appearance, perhaps due to ongoing infection on the plate that was nevertheless not able to outpace colony formation. After the infection period was complete, the cultures were vortexed, diluted 1:100, vortexed again, and 50 µl of the dilution was plated on 2 large (150 mm X 15 mm) LB agar plates (with 10 µg/ml gentamicin and 0.5% glucose to repress AGs and avoid fitness costs). The next day >>2,000-20,000 colonies were scraped from each infected and uninfected set of plates. 5 µl of the thick resulting cell mixtures was reserved for genomic DNA extraction, and the rest was frozen at -80°C in 20% glycerol as a backup. A full technical record of our screens is provided in Supplementary File 4.

AG barcodes were amplified and sequenced as with the WM6026 donor library, with a slight modification. The three biological replicate sets of ECOR+AG experiments were amplified using JSW-SS-171 and different versions of primer JSW-SS-170 containing extra N's at the start of the sequencing read. The first replicate was amplified with JSW-SS-170 (containing 5 N's), the second replicate with JBD-SS-254 (CGACGCTCTTCCGATCTNNNNNN TGATGTCGTTGTTGCCATCG) with 6 N's, and the third replicate with JBD-SS-255 (CGACGCTCTTCCGATCTNNNNNNN TGATGTCGTTGTTGCCATCG) with 7 N's. This ensured that reads from the three sets would not cross contaminate analyses of replicate experiments. Amplicons were sequenced to the same depth as the baseline libraries on Illumina MiSeq and NextSeq instruments in single-end format (100 bp) with two (8 bp) index reads.

We required exact matches for the 20 bp primer-derived invariant sequences flanking the barcodes. These sequences were trimmed and the proportions of various AG-encoding strains in the mixtures were determined using the prevalence of the corresponding 10 bp barcode sequences. Barcode counts per AG were normalized by dividing each count by the median of each experiment, which would represent neutral AG fitness. A log2 transformation was subsequently applied to obtain fitness scores. Care was taken to sequence each library to similar depth such that fitness scores across experiments were distributed in the same range. Poorly represented AGs were identified in the baseline libraries as well as the infected and uninfected experimental samples using a low read cutoff of 100, or if the log2 normalized average read counts across the three replicates for any AG were >3x the median absolute deviation below the median for that experiment. Reproducibility was determined using T-tests across the three replicates, and we required a P-value of at least 0.2. The magnitude of the AG fitness effect was determined by (a) subtracting the AG fitness scores in the uninfected library from the infected libraries to assess whether AGs changed host sensitivity to phage infection, and (b) subtracting the AG fitness scores in the baseline library from the uninfected libraries to assess whether AGs change host fitness directly. AGs were selected for further study if they were not poorly represented to begin with, produced reproducible effects across replicates, and changed host fitness by at least 4-fold. Although we had targeted our experiments to phage-host combinations where the host was mostly resistant to the phage in order to find counter-defense AGs, a few hosts turned out to be more sensitive to certain phages in liquid infection than suggested by agar-overlay plaque assays (perhaps due to abortive antiviral defenses). In these cases, our platform was also able to robustly identify AGs that conferred a protective advantage to the host (e.g. Imm and Cor, which both block phage DNA entry [11,12]). In practice

however, these effects can be very variable and the reproducibility filters typically needed to be relaxed in order to find such superinfection exclusion AGs.

All counter-defense AGs were subsequently conjugated into the wild ECOR hosts where they demonstrated an effect, and tested against all 8 phages (not just the ones that were used in the screen with that host). In several hosts, the AG was found to additionally sensitize the host to phages that were omitted in the screen. Not all the screen phenotypes could be confirmed in agar-overlay plaque assays, possibly due to variations in cell growth between liquid and solid media, or due to the requirement for successive infection cycles for plaque formation that would render plaque assays less sensitive than single-infection cycle dynamics in our screen. Nevertheless, further experiments were performed only in phage-host-AG combinations that could be confirmed by plaque assay due to ease of follow up testing.

Conditional lethal AGs were identified directly from the heatmap of fitness scores, with a subset selected for further experimentation using an arbitrary fitness cutoff of ~6-7 (~100-fold).

Tn5 Dropout Screen + Data analysis to find phage exclusion systems

Tn5 whole-genome knockout strain libraries were constructed using the RB-TnSeq donor library (pKMW7-derived strain APA766 from Dr. Adam Deutschbauer) conjugated with the recipient ECOR strain in a biparental mating on LB plates supplements with 300 µM DAP. The next day, cell mixtures were scraped and resuspended in 1 ml LB and 1000X dilutions of this conjugation mixture were plated on 10 large (150 mm X 15 mm) LB plates (lacking DAP) supplemented with 50 µg/ml kanamycin. Approximately 0.5-1 million colonies were scraped taking care to break colonies apart on the surface of the plates, mixed very thoroughly, aliquoted and stored at -80°C with 20% glycerol, with ~5-10 µl reserved for genomic DNA extraction and reference library preparation.

RB-TnSeq allows the same mutant library to be challenged under various conditions such as challenge by various phages, and simplifies the downstream sequencing-library preparation by substituting barcode sequencing for all experimental conditions after initial characterization of the reference library. To characterize the Tn5-knockout reference libraries in the various ECOR strains, genomic DNA was extracted from 5-10 µl of the cell scrape and prepared for sequencing using the NEBNext Ultra II FS DNA Library Prep Kit (NEB) with the following modifications. 500 ng DNA was fragmented for 25 min according to manufacturer's instructions. The adaptor ligation step was performed as directed albeit with a custom pre-annealed oligo adaptor comprising of YL001 (/5Phos/GATCGGAAGAG/3ddC/) annealed to one of YL002:005 (AGCGGCAATTTCACACAGGACAAGCAGAAGACGGCATACGAGATNNNNNNNNAGTGGTG ACTGGAGTTCAGACGTGTGCTCTTCCGATC*T) where the 4-base underlined sequence serves as a variable barcode unique to each oligo. The USER Enzyme steps post-ligation were skipped and the bead-cleanup was performed using 40 µl and 20 µl of Sample Purification Beads successively as instructed for library size selection in the 150-250 bp range. The PCR amplification of transposon-genome junctions was performed using the cycling parameters described in the kit for 11 cycles with Q5 Ultra II FS Master Mix (NEB) using primers YL006 (AGCGGCAATTTCACACAGGA) and JBD-SS-272/oAD493 (ACACTGGCAGAGCATTACGCCCT) instead of the kit-supplied primers. A second nested-PCR enrichment step was performed for 11 cycles using the same cycling parameters as before, using the entire bead-purified reaction product from the first step as a template with primers YL009 (CAAGCAGAAGACGGCATACGAG) and JBD-SS-273:280 (AATGATACGGCGACCACCGAGATCTACAC TATAGCCT ACACTCTTTCCCTACACGACGCTCTTCCGATCT NNNNNNN GCAGGGATGTCCACGAGGTCTCT) where N_8 barcodes correspond to reverse-complemented Illumina TruSeq HT D501-508 indexes. Bead-purified amplicons were quantified using the 1X

dsDNA HS kit with the Qubit 4.0 Fluorometer (Invitrogen) and sequenced on Illumina MiSeq and NovaSeq instruments in single-end format (150 bp) with two (8 bp) index reads.

The first 100 bp of these reads contain transposon derived sequences and a DNA barcode that uniquely identifies a particular Tn5 insertion. The last 50 bp of these reads contain host-genome derived sequence that help determine the location of the transposon insertion and hence construct a map between the Tn5 barcode and the insertion site (IS) in the genome. Briefly, 20 bp barcodes and 16-50 bp IS sequences were trimmed from the sequencing reads. To isolate the barcode and IS, we allowed 3 mismatches in the 23 bp invariant sequence upstream and up to 6 mismatches in the 50 bp downstream of the barcode using a Hamming distance function to compare the expected and actual sequences. Any reads failing these sequence-fidelity requirements were discarded. Reads containing an IS that matched the Tn5 transposon vector in the donor strain were also discarded. The 4 Ns adjacent to the underlined index sequences in YL002:005 were also sequenced as part of one 8 bp index read and served as a tag to remove PCR duplicates. This allowed us to distinguish between sequencing “reads” and DNA “fragments” present in the sample at the adaptor ligation step. Thus, reads containing unique combinations of barcode, PCR tag, and IS were entered into the barcode/IS map for each Tn5 library, each representing a distinct DNA fragment (but also machine errors in sequencing, chimeric PCR amplicons, 3’ truncations etc.). The filtered ISs were then mapped to genes within their respective host genomes. As described previously[64], it is possible for transposon barcodes to map to the genomes ambiguously. Barcodes associated with multiple distinct ISs were retained only if their most frequently represented IS accounted for over 90% of the fragments containing that barcode. We then counted the number of DNA fragments recovered per gene and used this as a reference library for subsequent BarSeq experiments to assay depletion of strains with mutations in specific genes upon phage infection.

Tn5 mutant strain libraries of the various ECOR strains were then challenged by various phages in a procedure analogous to the ECOR+AG screens. Cultures of Tn5 libraries were initiated with 100 µl of thick cell scrape mixture in 100 ml LB supplemented with 50 µg/ml kanamycin, Mg^{2+}, and Ca^{2+}, to an initial OD_{600} value of 0.1-0.15. Cultures were grown at 37°C with shaking for 80-90 min to an OD_{600} of 0.7-1. A 10^{-5} dilution of each late-log phase culture was then plated on LB agar (with 50 µg/ml kanamycin) to obtain a count of viable colony forming units (cfu). Cultures were immediately diluted ~1:10 in fresh LB (with Mg^{2+}, Ca^{2+}, and 50 µg/ml kanamycin) and 1 ml aliquots were gently mixed with lysates of the appropriate phages at an MOI-e of ~30-100 (the actual MOI-e values were recorded the next day when cfu counts were obtained). Since the parental ECOR strains were typically quite resistant to phage infection to begin with, only mutants with transposon-inactivated defense systems would be expected to drop out upon infection. In practice however, some strains exhibited more pronounced mortality upon prolonged phage infection in the Tn5 screens. In particular, in experiments with T2 and the Tn5 library constructed in the Δ*wecA* strain of ECOR21 (which was generally more susceptible to infection by T2 and other phages), MOI was titrated in three parallel infection screens (with varying amounts of phage lysate added: 350 µl, 100 µl, and 25 µl) to empirically ensure specific depletion of only the phage-sensitive mutants. Parallel uninfected controls were maintained for each experiment, and each set of experiments was performed in biological duplicates. Infections were allowed to proceed for 5 min at the benchtop without shaking, and 75 min at 37°C with shaking. Since an interaction between the host and phage had already been identified in the earlier AG screen, infections were not restriction to a single phage replication cycle, but allowed to proceed further to achieve full depletion of the phage-sensitive Tn5-knockout mutants in the pool. After the infection period was complete, the cultures were vortexed, diluted 1:10, vortexed again, and 100 µl of the dilution was plated on 2 large (150 mm X 15 mm) LB agar plates (with 50 µg/ml kanamycin). We note that this dilution yields bacterial lawns and in our experience this preserves diversity of surviving colonies without degrading the signal of dropouts by phage-killing in the experiment. The next day, bacterial lawns were scraped from each infected and

uninfected set of plates. 5 µl of the thick resulting cell mixtures was reserved for genomic DNA extraction, and the rest was frozen at -80°C in 20% glycerol as a backup. A full technical record of our screens is provided in Supplementary File 6.

To amplify barcodes from the post-selection infected and uninfected Tn5 libraries, we followed a modified BarSeq98 method[64], with two successive rounds of PCR. The transposon barcode sequence was amplified by PCR using Q5 Ultra II FS Master Mix (NEB) with primers JBD-SS-310 (CGACGCTCTTCCGATCTNNNNNNN GATGTCCACGAGGTCTCT) and JBD-SS-311 (ACTGACGCTAGTGCATCA GTCGACCTGCAGCGTACG). Template-matching regions in the primers are underlined. Sequencing adaptors were attached in a second round of PCR using amplicons from the first PCR round as the templates, with dual-indexed primer sets JSW-SS-42:53 and JSW-SS-54:61 (sequences already provided in section describing AG screening). Cycling conditions for round 1 were as follows: one cycle at 98°C for 4 min; two cycles at 98°C for 30s, 63°C for 30s, and 72°C for 30s; 22 cycles at 98°C for 30s, and 72°C for 30s; and one cycle at 72°C for 4 min. Conditions for round 2 were one cycle at 98°C for 1 min; two cycles at 98°C for 10 s, 64°C for 30 s, and 72°C for 10 s; 4 cycles at 98°C for 10 s, and 72°C for 20 s; and one cycle at 72°C for 2 min. 0.5-1 ul of unpurified 1st round reaction product was used as template for the 2nd round of PCR. Amplicons from the 2nd round were gel-purified by electrophoresis (3% agarose gel, 4.2 V/cm, 2 hours) and quantified using the 1X dsDNA HS kit with the Qubit 4.0 Fluorometer (Invitrogen). Amplicons were sequenced on Illumina MiSeq and NovaSeq platforms in single-end format (150 bp) with two (8 bp) index reads.

DNA barcodes were trimmed from the reads and the proportions of various Tn5-knockout mutants in the mixtures were determined using the prevalence of the corresponding barcode sequence. Briefly, we allowed 3 mismatches in the 18 bp invariant sequence upstream, and 4 mismatches in the 18 bp downstream of the transposon barcodes using a Hamming distance function to compare the expected and actual sequences. The frequency of occurrence of each barcode (representation number) was counted. Transposon barcodes were matched to those found in the reference libraries, inferred to represent knockouts in the genes containing the corresponding ISs, and the various barcode representation numbers were added together for each gene. Cumulative transposon insertions per gene were then compared between the infected and uninfected samples (with the reference library as a baseline to filter out poorly represented, presumably essential genes) using the same data analysis procedure as for the PhAGENT screen, albeit with two biological replicates instead of three.

Tn5 Suppressor Screen + Data analysis to find abortive loci

Tn5 whole-genome knockout strain libraries were constructed using the RB-TnSeq donor conjugated with the recipient ECOR strain as before, with a slight modification. The recipient ECOR strain was equipped with the AG (as a single copy chromosomal Tn7 insertion as usual) which triggers PCD, presumably by activating an abortive defense system. The trigger AG was inserted prior to Tn5 library construction in order to minimize passages of the library prior to screening and avoid potential bottlenecking due to fitness effects of the various Tn5 knockouts. Libraries were constructed as above, except that conjugation was carried out under strict glucose repression to prevent PCD activation at this stage, and the conjugation mixture scraped from the LB+DAP+glucose plates was then plated directly on LB plates containing 10 µg/ml gentamicin and 1 mM IPTG in addition to 50 µg/ml kanamycin. We typically needed to plate the entire conjugation mixture scraped from five 100 mm LB+DAP+glucose plate onto five 150 mm LB+IPTG+Kan+Gm plate in order to obtain sufficient numbers of suppressor colonies for characterization by high-throughput sequencing. Control samples were obtained by plating 1000-2000X dilutions on LB+glucose+Kan+Gm plates. Suitable dilutions were empirically determined for the various ECOR hosts. The next day, suppressor colonies were scraped from the large plates and stored at -80°C with 20% glycerol, with ~5-10 µl reserved for genomic DNA extraction and sequencing. Genomic DNA was extracted and libraries were prepared and

sequenced akin to the treatment of the RB-TnSeq/NEBNext reference libraries detailed above. There was no need for barcode sequencing since the screen identifies survivors instead of dropouts, and each library was only characterized once after IPTG induction/selection for suppressor mutants. Tn5 reference libraries were constructed from the data using the same procedures for the baseline libraries constructed in the various naïve (without AGs) ECOR strains above. The number of transposon insertions (i.e. the number of uniquely mapping ISs from distinct DNA fragments) recovered per gene were used directly as a measure of survival of mutants despite expression of PCD-triggering AGs. In practice, no statistical analysis was necessary to identify the genetic loci responsible for PCD, as the signal-to-noise ratio was sufficiently high to allow visual identification of hits when transposon insertion frequency was plotted in control vs experimental samples (i.e. when the trigger AG was repressed by plating on glucose-containing media, vs induced with IPTG).

LPS gel electrophoresis

Lipopolysaccharide was prepared from liquid overnight cultures of wild *E. coli* strains expressing the relevant AGs (LB + IPTG). Briefly, 50-100 µl of saturated culture was spun down and cells were thoroughly resuspended in 100-200 µl of lysing buffer (62.5 mM Tris-HCl, pH6.8, 10% glycerol, 2% SDS, 4% ß-mercaptoethanol). We routinely supplement 1X Laemmli Buffer (BioRad) with SDS and ßME to make lysing buffer. Samples were boiled at 100°C for 15 min, and allowed to rest at room temperature for 15 min. 100 µg of Protease K were added to each sample and protein was digested at 60°C for at least 1 hr. 15 µl of each sample were run onto a 12% polyacrylamide TGX Mini-Protean Precast gel (BioRad) at 200V for 35 min, and visualized using the Pro-Q Emerald 300 LPS stain kit (Thermo Fisher) precisely according to manufacturer's instructions.

Allelic Exchange for *wecA* deletion

The R6K-origin pKEK2201 vector (from Prof. Karl E. Klose) carrying the desired genetic modification (eg. a *wecA* deletion) flanked by 1000 bp homology arms was prepared by Gibson assembly, and delivered as a suicide vector into wild *E. coli* strains via conjugation from the WM6026 auxotrophic donor. Conjugation was performed using the same method used to deliver AGs, albeit with biparental instead of triparental mating. Successful transconjugants (containing a chromosomally integrated plasmid after the first homologous recombination event HR1) were selected by incubation on LB plates supplemented with kanamycin. The second recombination event HR2 to complete the allelic exchange was selected via *sacB*-mediated counterselection on LB plates supplemented with 300 mM sucrose after overnight growth of HR1 transconjugants. Successful allele exchange was verified by PCR, and by LPS electrophoresis to assay for the absence of O-antigen in the case of the *wecA* deletion.

Live cell Microscopy

Cells were prepared for microscopy, immobilized under an LB agarose pad, and imaged as described previously[65]. Where necessary, AGs were pre-induced with 1 mM IPTG for 30 min immediately before harvesting for microscopy.

Verification of abortive loci activity

Abortive/PCD systems identified by transposon screening of conditionally lethal AGs were cloned by Gibson assembly using the NEBuilder HiFi DNA assembly master mix (NEB) onto pBAD Myc/His A vectors along with their native promoters wherever possible. AGs that trigger PCD systems (or appropriate controls) were conjugated into DH10ß/TOP10 strains carrying pBAD-PCD plasmids using the Tn7 transposition method (above). Where the restriction/PCD systems prevented conjugation of AGs, the order of transformations was reversed – the AG was conjugated into naïve DH10ß/TOP10 cells first, and subsequently these

strains were transformed with the pBAD-PCD vectors. Successful transconjugants were grown overnight with glucose to repress trigger AGs, and viability was tested the next day by spotting 10-fold serial dilutions of saturated cultures on LB agar plates with four additives: 1) Glucose repression to assess background toxicity, (2) Arabinose induction of PCD systems alone to assess fitness costs of over-expression of abortive components without the trigger AGs, (3) IPTG to assess ability of AGs to trigger PCD without over-expression of abortive components (i.e. under leaky expression conditions, or when expressed from the native promoters where present), and (4) IPTG and Arabinose to assess ability of AGs to trigger PCD when abortive components were also over-expressed.

Affinity purification of AGs from wild *E. coli* hosts

FLAG-tagged AGs were conjugated into ECOR strains by Tn7-transposition (above) according to Supplementary File 7. Each strain was grown overnight at 37°C in 2 mL of LB supplemented with .5% glucose and 10 mg/mL gentamicin, then diluted 1:100 into 100 mL LB supplemented with 10 mg/mL gentamicin and grown at 37°C for 30 minutes. The media was then supplemented with 1 mM IPTG and grown for another 2.5 hours. “Popcorn” pellets were collected by centrifugation for 30 minutes at 4,000 X g, resuspension in 100 µL of ice-cooled lysis buffer [50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM $MgCl_2$, 0.5% NP40 (Research Products International), 125 U/mL Benzonase (Millipore), 1x protease inhibitor cocktail (Roche, cOmplete ULTRA tablets, mini, EDTA-free), 0.5 mg/ml lysozyme (Fisher Scientific)], followed by dropwise addition into liquid nitrogen and stored at -80°C. Popcorn pellets were lysed by 10 cycles of cryomilling for 2 minutes at 12 counts per second in a SPEX SamplePrep 6870D Freezer/Mill and stored at -80°C.

For each phage-host combination, 3 replicate lysates were reconstituted using 200 mg of the cryomilled powder resuspended in 750 µL of lysis buffer. Lysates were cleared by centrifugation at 16,000 X g at 4°C for 15 minutes. Anti-FLAG magnetic agarose beads (Pierce) were washed on a magnet with 1 mL of ice-cooled IP buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1mM $MgCl_2$) twice and resuspended to their original volume in IP buffer. Cleared lysates were incubated with 30 µL of FLAG beads on an end-over-end tube rotator at 10 rpm at 4°C overnight. The supernatant was separated from the beads on a magnet. Beads were washed once with 700 µL of .05% NP-40 and 1x protease inhibitor cocktail in IP, once with 700 µL of .05% NP-40 in IP, and 3 times with 700 µL of IP. Proteins were eluted by incubation with 25 µL of 100 µg/ml 3xFLAG peptide (Sigma) in 0.05% RapiGest (Waters Corp) in IP buffer on an end-over-end tube rotator at 10 rpm at 4°C overnight. The eluate was separated from the beads on a magnet. The elution step was then repeated with the same beads and fresh elution buffer at room temperature for 1 hour. The eluates were combined and stored at -20°C.

Mass spectrometry sample preparation and data acquisition

Samples were prepared for Mass Spectrometry (MS) as described previously in the Filter-Aided Sample Preparation Protocol[66] with the following modifications. Before transferring to an AcroPrep Advance 10k Omega plate, each eluate was resuspended in 100 µL of 8M Urea (Thermo Scientific) dissolved in 50 mM ammonium bicarbonate (MP Biomedicals) and 5 mM TCEP (Aldrich), followed by centrifugation to dryness. During alkylation of cysteine residues, 10 mM chloroacetamide (Thermo Scientific) was used instead of 10 mM iodoacetamide. Wash steps prior to proteolysis were carried out with 200 µL of 20 mM ammonium bicarbonate (instead of 100 µL). Proteolysis was performed without the addition of lysyl endopeptidase. Proteolysis was quenched by addition of 10% trifluoroacetic acid (Thermo Scientific) to reach ≈ pH 2. Peptides were then purified over a BioPureSPE Mini 96-Well Plate (Nest Group). Briefly, C18 columns were activated by washing once with 400 µL of acetonitrile (Fisher Scientific) and twice with 400 µL of 0.1% formic acid (Thermo Scientific) in water solution, LC–MS grade (Thermo Scientific). Samples were loaded and centrifuged at 1500 X g for 2 minutes. The bound

peptides were washed 3 times with 400 µL of 1% formic acid before elution with 150 µL of 50% acetonitrile in 0.1% trifluoroacetic acid. Peptides were vacuum dried and stored at -80°C.

Samples were resuspended in 15 µl of 0.1% formic acid prior to loading for MS. The MS acquisition was performed as described[67] with same MS parameters and LC configuration.

Mass Spectrometry data analysis and statistical analysis

Raw files were processed in FragPipe[68] using the LFQ-MBR workflow with minor modifications. The files were searched against a combined database of protein sequences from ECOR1, ECOR3, ECOR15, ECOR21, ECOR22, ECOR66[62], and MG1655 with duplicate entries removed and common contaminants added. Decoys were generated by pseudo-inversion. Protein and peptide FDR was fixed to 1% and MBR was disabled. Cysteine carbamylation was set as fixed modification while N-term acetylation, methionine oxidation and pyro-glu formation were set as variable modification with a maximum of 3 modifications per peptide.

MaxLFQ intensities were used for statistical analysis[69]. Each triplicate set of host-AG samples was quantile normalized against a triplicate set of samples from the same host expressing orf74. For quantile normalization, all proteins in each sample were ranked according to their MaxLFQ intensity. Next, the mean value of the intensities of proteins occupying a given rank was calculated, then each protein's value was replaced with the mean value according to its rank. Fold changes were calculated by taking the log2 ratio of average normalized values between the test set and the control set. For each host-AG sample, all positively enriched proteins were ranked according to their fold change. For each AG, proteins were assigned "summed rank" values by calculating the sum of the ranks of each protein across all hosts where that AG was tested. P-values were calculated by student's t-test on the normalized sets of replicates from the test and control samples. Top enriched candidates for each AG were selected based on their summed ranks and p-values.

Ronin systems phylogenetic analysis

The *ronA* nucleotide sequence in the ECOR55 Ronin system was used to identify similar genes by BLASTN. *RonA*-encoding bacterial genomes were downloaded using the entrez direct e-utilities command line tools. Of 100 complete genomes retrieved, *ronA* homologs had been annotated in 84. *hsdR* genes and other type I R-M components were identified using phmmer and Defense Finder[70] respectively. Ronin-adjacent genetic architectures were visualized using the PhamLite software suite[58]. All ronin-adjacent architectures are shown in HTML files with mouse-over annotations for genes.

References

1. Koonin, E.V., and Wolf, Y.I. (2012). Evolution of microbes and viruses: a paradigm shift in evolutionary biology? Front Cell Infect Microbiol *2*, 119. 10.3389/fcimb.2012.00119.

2. Kristensen, D.M., Mushegian, A.R., Dolja, V.V., and Koonin, E.V. (2010). New dimensions of the virus world discovered through metagenomics. Trends Microbiol *18*, 11-19. 10.1016/j.tim.2009.11.003.

3. Bingham, R., Ekunwe, S.I., Falk, S., Snyder, L., and Kleanthous, C. (2000). The major head protein of bacteriophage T4 binds specifically to elongation factor Tu. J Biol Chem *275*, 23219-23226. 10.1074/jbc.M002546200.

4. Wang, H.C., Ho, C.H., Hsu, K.C., Yang, J.M., and Wang, A.H. (2014). DNA mimic proteins: functions, structures, and bioinformatic analysis. Biochemistry *53*, 2865-2874. 10.1021/bi5002689.

5. Loenen, W.A., and Murray, N.E. (1986). Modification enhancement by the restriction alleviation protein (Ral) of bacteriophage lambda. Journal of molecular biology *190*, 11-22. 10.1016/0022-2836(86)90071-9.

6. Nechaev, S., and Severinov, K. (2003). Bacteriophage-induced modifications of host RNA polymerase. Annual review of microbiology *57*, 301-322. 10.1146/annurev.micro.57.030502.090942.

7. Mizuno, C.M., Guyomar, C., Roux, S., Lavigne, R., Rodriguez-Valera, F., Sullivan, M.B., Gillet, R., Forterre, P., and Krupovic, M. (2019). Numerous cultivated and uncultivated viruses encode ribosomal proteins. Nat Commun *10*, 752. 10.1038/s41467-019-08672-6.

8. Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., and Davidson, A.R. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature *493*, 429-432. 10.1038/nature11723.

9. Chung, I.Y., Jang, H.J., Bae, H.W., and Cho, Y.H. (2014). A phage protein that inhibits the bacterial ATPase required for type IV pilus assembly. Proc Natl Acad Sci U S A *111*, 11503-11508. 10.1073/pnas.1403537111.

10. Shah, M., Taylor, V.L., Bona, D., Tsao, Y., Stanley, S.Y., Pimentel-Elardo, S.M., McCallum, M., Bondy-Denomy, J., Howell, P.L., Nodwell, J.R., et al. (2021). A phage-encoded anti-activator inhibits quorum sensing in Pseudomonas aeruginosa. Molecular cell *81*, 571-583 e576. 10.1016/j.molcel.2020.12.011.

11. Lu, M.J., and Henning, U. (1994). Superinfection exclusion by T-even-type coliphages. Trends Microbiol *2*, 137-139. 10.1016/0966-842x(94)90601-7.

12. Uc-Mass, A., Loeza, E.J., de la Garza, M., Guarneros, G., Hernandez-Sanchez, J., and Kameyama, L. (2004). An orthologue of the cor gene is involved in the exclusion of

temperate lambdoid phages. Evidence that Cor inactivates FhuA receptor functions. Virology *329*, 425-433. 10.1016/j.virol.2004.09.005.

13. Brussow, H., Canchaya, C., and Hardt, W.D. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev *68*, 560-602, table of contents. 10.1128/MMBR.68.3.560-602.2004.

14. Berg, P., and Mertz, J.E. (2010). Personal reflections on the origins and emergence of recombinant DNA technology. Genetics *184*, 9-17. 10.1534/genetics.109.112144.

15. Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096. 10.1126/science.1258096.

16. 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. Journal of bacteriology *193*, 6039-6056. 10.1128/JB.05535-11.

17. 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*. 10.1126/science.aar4120.

18. Gao, L., Altae-Tran, H., Bohning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E.V., and Zhang, F. (2020). Diverse enzymatic activities mediate antiviral immunity in prokaryotes. Science *369*, 1077-1084. 10.1126/science.aba0372.

19. Pinilla-Redondo, R., Shehreen, S., Marino, N.D., Fagerlund, R.D., Brown, C.M., Sorensen, S.J., Fineran, P.C., and Bondy-Denomy, J. (2020). Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. Nat Commun *11*, 5652. 10.1038/s41467-020-19415-3.

20. Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F., and Hendrix, R.W. (2000). Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. Journal of molecular biology *299*, 27-51. 10.1006/jmbi.2000.3729.

21. Comeau, A.M., Bertrand, C., Letarov, A., Tetart, F., and Krisch, H.M. (2007). Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. Virology *362*, 384-396. 10.1016/j.virol.2006.12.031.

22. Bondy-Denomy, J., Qian, J., Westra, E.R., Buckling, A., Guttman, D.S., Davidson, A.R., and Maxwell, K.L. (2016). Prophages mediate defense against phage infection through diverse mechanisms. The ISME journal *10*, 2854-2866. 10.1038/ismej.2016.79.

23. Tsao, Y.F., Taylor, V.L., Kala, S., Bondy-Denomy, J., Khan, A.N., Bona, D., Cattoir, V., Lory, S., Davidson, A.R., and Maxwell, K.L. (2018). Phage Morons Play an Important Role in Pseudomonas aeruginosa Phenotypes. Journal of bacteriology *200*. 10.1128/JB.00189-18.

24. Cumby, N., Davidson, A.R., and Maxwell, K.L. (2012). The moron comes of age. Bacteriophage *2*, 225-228. 10.4161/bact.23146.

25. Moffett, P. (2016). Using Decoys to Detect Pathogens: An Integrated Approach. Trends Plant Sci *21*, 369-370. 10.1016/j.tplants.2016.04.002.

26. Ochman, H., and Selander, R.K. (1984). Standard reference strains of Escherichia coli from natural populations. Journal of bacteriology *157*, 690-693. 10.1128/JB.157.2.690-693.1984.

27. Kulikov, E.E., Golomidova, A.K., Efimov, A.D., Belalov, I.S., Letarova, M.A., Zdorovenko, E.L., Knirel, Y.A., Dmitrenok, A.S., and Letarov, A.V. (2021). Equine Intestinal O-Seroconverting Temperate Coliphage Hf4s: Genomic and Biological Characterization. Applied and environmental microbiology *87*, e0112421. 10.1128/AEM.01124-21.

28. Allison, G.E., and Verma, N.K. (2000). Serotype-converting bacteriophages and O-antigen modification in Shigella flexneri. Trends Microbiol *8*, 17-23. 10.1016/s0966-842x(99)01646-7.

29. Vander Byl, C., and Kropinski, A.M. (2000). Sequence of the genome of Salmonella bacteriophage P22. Journal of bacteriology *182*, 6472-6481. 10.1128/JB.182.22.6472-6481.2000.

30. Newton, G.J., Daniels, C., Burrows, L.L., Kropinski, A.M., Clarke, A.J., and Lam, J.S. (2001). Three-component-mediated serotype conversion in Pseudomonas aeruginosa by bacteriophage D3. Mol Microbiol *39*, 1237-1247. 10.1111/j.1365-2958.2001.02311.x.

31. Losick, R. (1969). Isolation of a trypsin-sensitive inhibitor of O-antigen synthesis involved in lysogenic conversion by bacteriophage epsilon-15. Journal of molecular biology *42*, 237-246. 10.1016/0022-2836(69)90040-0.

32. Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nature reviews. Microbiology *8*, 317-327. 10.1038/nrmicro2315.

33. Weissborn, A.C., Liu, Q., Rumley, M.K., and Kennedy, E.P. (1994). UTP: alpha-D-glucose-1-phosphate uridylyltransferase of Escherichia coli: isolation and DNA sequence of the galU gene and purification of the enzyme. Journal of bacteriology *176*, 2611-2618. 10.1128/jb.176.9.2611-2618.1994.

34. Genevaux, P., Bauda, P., DuBow, M.S., and Oudega, B. (1999). Identification of Tn10 insertions in the rfaG, rfaP, and galU genes involved in lipopolysaccharide core biosynthesis that affect Escherichia coli adhesion. Arch Microbiol *172*, 1-8. 10.1007/s002030050732.

35. Marolda, C.L., and Valvano, M.A. (1996). The GalF protein of Escherichia coli is not a UDP-glucose pyrophosphorylase but interacts with the GalU protein possibly to regulate

cellular levels of UDP-glucose. Mol Microbiol *22*, 827-840. 10.1046/j.1365-2958.1996.01531.x.

36. Pierson, D.E., and Carlson, S. (1996). Identification of the galE gene and a galE homolog and characterization of their roles in the biosynthesis of lipopolysaccharide in a serotype O:8 strain of Yersinia enterocolitica. Journal of bacteriology *178*, 5916-5924. 10.1128/jb.178.20.5916-5924.1996.

37. Franklin, N.C. (1969). Mutation in gal U gene of E. coli blocks phage P1 infection. Virology *38*, 189-191. 10.1016/0042-6822(69)90144-5.

38. Young, F.E. (1967). Requirement of glucosylated teichoic acid for adsorption of phage in Bacillus subtilis 168. Proc Natl Acad Sci U S A *58*, 2377-2384. 10.1073/pnas.58.6.2377.

39. Fratamico, P.M., Briggs, C.E., Needle, D., Chen, C.Y., and DebRoy, C. (2003). Sequence of the Escherichia coli O121 O-antigen gene cluster and detection of enterohemorrhagic E. coli O121 by PCR amplification of the wzx and wzy genes. J Clin Microbiol *41*, 3379-3383. 10.1128/JCM.41.7.3379-3383.2003.

40. Lehrer, J., Vigeant, K.A., Tatar, L.D., and Valvano, M.A. (2007). Functional characterization and membrane topology of Escherichia coli WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. Journal of bacteriology *189*, 2618-2628. 10.1128/JB.01905-06.

41. 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 research *49*, 11257-11273. 10.1093/nar/gkab906.

42. Gordeeva, J., Morozova, N., Sierro, N., Isaev, A., Sinkunas, T., Tsvetkova, K., Matlashov, M., Truncaite, 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 research *47*, 253-265. 10.1093/nar/gky1125.

43. 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 research *48*, 5397-5406. 10.1093/nar/gkaa290.

44. Kruger, D.H., Reuter, M., Hansen, S., and Schroeder, C. (1982). Influence of phage T3 and T7 gene functions on a type III(EcoP1) DNA restriction-modification system in vivo. Mol Gen Genet *185*, 457-461. 10.1007/BF00334140.

45. Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., Oppenheimer-Shaanan, Y., and Sorek, R. (2020). Bacterial Retrons Function In Anti-Phage Defense. Cell *183*, 1551-1561 e1512. 10.1016/j.cell.2020.09.065.

46. Rousset, F., Depardieu, F., Miele, S., Dowding, J., Laval, A.L., Lieberman, E., Garry, D., Rocha, E.P.C., Bernheim, A., and Bikard, D. (2022). Phages and their satellites encode

hotspots of antiviral systems. Cell host & microbe *30*, 740-753 e745. 10.1016/j.chom.2022.02.018.

47. Ofir, G., Melamed, S., Sberro, H., Mukamel, Z., Silverman, S., Yaakov, G., Doron, S., and Sorek, R. (2018). DISARM is a widespread bacterial defence system with broad anti-phage activities. Nat Microbiol *3*, 90-98. 10.1038/s41564-017-0051-0.

48. Koonin, E.V., and Zhang, F. (2017). Coupling immunity and programmed cell suicide in prokaryotes: Life-or-death choices. Bioessays *39*, 1-9. 10.1002/bies.201600186.

49. Murray, N.E. (2000). Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol Mol Biol Rev *64*, 412-434. 10.1128/MMBR.64.2.412-434.2000.

50. Penner, M., Morad, I., Snyder, L., and Kaufmann, G. (1995). Phage T4-coded Stp: double-edged effector of coupled DNA and tRNA-restriction systems. Journal of molecular biology *249*, 857-868. 10.1006/jmbi.1995.0343.

51. Murphy, K.C. (2000). Bacteriophage P22 Abc2 protein binds to RecC increases the 5' strand nicking activity of RecBCD and together with lambda bet, promotes Chi-independent recombination. Journal of molecular biology *296*, 385-401. 10.1006/jmbi.1999.3486.

52. Forti, F., Polo, S., Lane, K.B., Six, E.W., Sironi, G., Deho, G., and Ghisotti, D. (1999). Translation of two nested genes in bacteriophage P4 controls immunity-specific transcription termination. Journal of bacteriology *181*, 5225-5233. 10.1128/JB.181.17.5225-5233.1999.

53. Ghisotti, D., Finkel, S., Halling, C., Deho, G., Sironi, G., and Calendar, R. (1990). Nonessential region of bacteriophage P4: DNA sequence, transcription, gene products, and functions. Journal of virology *64*, 24-36. 10.1128/JVI.64.1.24-36.1990.

54. Lopatina, A., Tal, N., and Sorek, R. (2020). Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. Annu Rev Virol *7*, 371-384. 10.1146/annurev-virology-011620-040628.

55. Gao, L.A., Wilkinson, M.E., Strecker, J., Makarova, K.S., Macrae, R.K., Koonin, E.V., and Zhang, F. (2022). Prokaryotic innate immunity through pattern recognition of conserved viral proteins. Science *377*, eabm4096. 10.1126/science.abm4096.

56. Zhang, T., Tamman, H., Coppieters 't Wallant, K., Kurata, T., LeRoux, M., Srikant, S., Brodiazhenko, T., Cepauskas, A., Talavera, A., Martens, C., et al. (2022). Direct activation of a bacterial innate immune system by a viral capsid protein. Nature *612*, 132-140. 10.1038/s41586-022-05444-z.

57. Johnson, M.C., Laderman, E., Huiting, E., Zhang, C., Davidson, A., and Bondy-Denomy, J. (2022). Core Defense Hotspots within Pseudomonas aeruginosa are a

consistent and rich source of anti-phage defense systems. bioRxiv, 2022.2011.2011.516204. 10.1101/2022.11.11.516204.

58. Johnson, M.C. (2023). PhamLite. https://console.latch.bio/explore/61249/info.

59. Steinegger, M., and Soding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol *35*, 1026-1028. 10.1038/nbt.3988.

60. Jain, C., Rodriguez, R.L., Phillippy, A.M., Konstantinidis, K.T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun *9*, 5114. 10.1038/s41467-018-07641-9.

61. Peters, J.M., Koo, B.M., Patino, R., Heussler, G.E., Hearne, C.C., Qu, J., Inclan, Y.F., Hawkins, J.S., Lu, C.H.S., Silvis, M.R., et al. (2019). Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi. Nat Microbiol *4*, 244-250. 10.1038/s41564-018-0327-z.

62. Patel, I.R., Gangiredla, J., Mammel, M.K., Lampel, K.A., Elkins, C.A., and Lacher, D.W. (2018). Draft Genome Sequences of the Escherichia coli Reference (ECOR) Collection. Microbiol Resour Announc *7*. 10.1128/MRA.01133-18.

63. Hansen, L. (2021). SeqDiff. https://github.com/hansenlo/SeqDiff.

64. Wetmore, K.M., Price, M.N., Waters, R.J., Lamson, J.S., He, J., Hoover, C.A., Blow, M.J., Bristow, J., Butland, G., Arkin, A.P., and Deutschbauer, A. (2015). Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. mBio *6*, e00306-00315. 10.1128/mBio.00306-15.

65. Li, Y., Guan, J., Hareendranath, S., Crawford, E., Agard, D.A., Makarova, K.S., Koonin, E.V., and Bondy-Denomy, J. (2022). A family of novel immune systems targets early infection of nucleus-forming jumbo phages. bioRxiv, 2022.2009.2017.508391. 10.1101/2022.09.17.508391.

66. Fossati, A., Frommelt, F., Uliana, F., Martelli, C., Vizovisek, M., Gillet, L., Collins, B., Gstaiger, M., and Aebersold, R. (2021). System-Wide Profiling of Protein Complexes Via Size Exclusion Chromatography-Mass Spectrometry (SEC-MS). Methods in molecular biology *2259*, 269-294. 10.1007/978-1-0716-1178-4_18.

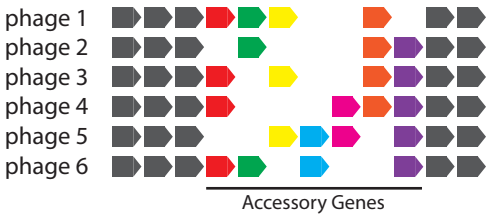
67. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature *583*, 459-468. 10.1038/s41586-020-2286-9.

68. Kong, A.T., Leprevost, F.V., Avtonomov, D.M., Mellacheruvu, D., and Nesvizhskii, A.I. (2017). MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. Nature methods *14*, 513-520. 10.1038/nmeth.4256.

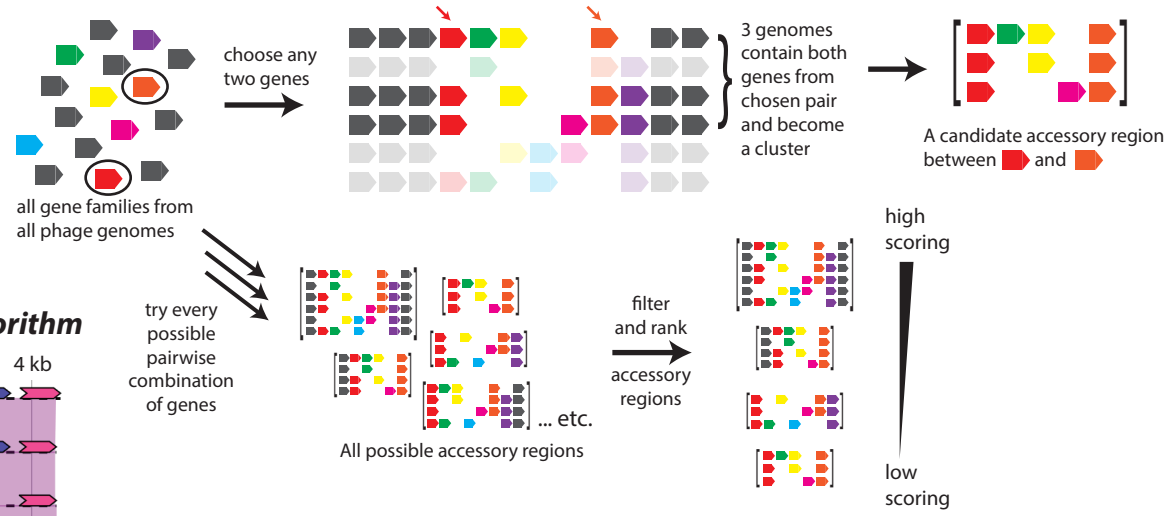
69. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics *13*, 2513-2526. 10.1074/mcp.M113.031591.

70. Tesson, F., Herve, A., Mordret, E., Touchon, M., d'Humieres, C., Cury, J., and Bernheim, A. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. Nat Commun *13*, 2561. 10.1038/s41467-022-30269-9.

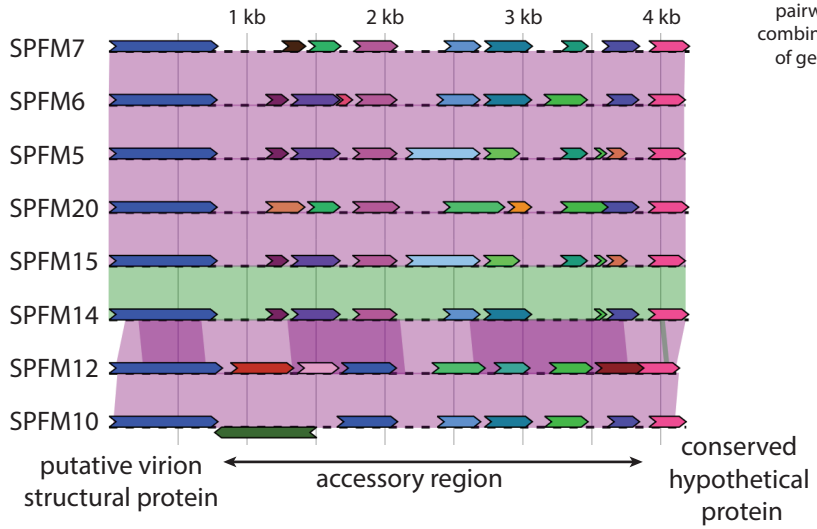
A. Core vs. Accessory genes in a hypothetical phage cluster



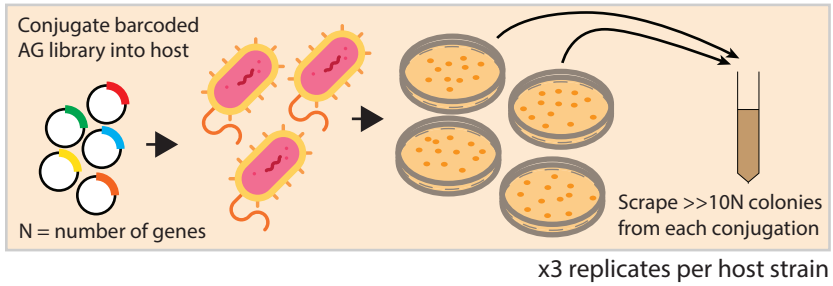
B. Algorithm to find accessory regions without a priori genome clustering



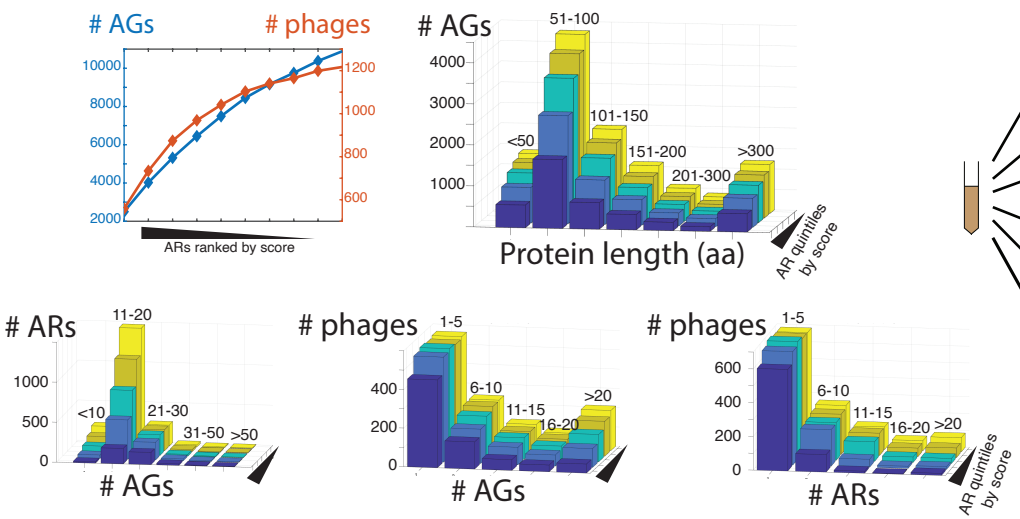
C. Sample accessory region found by the algorithm



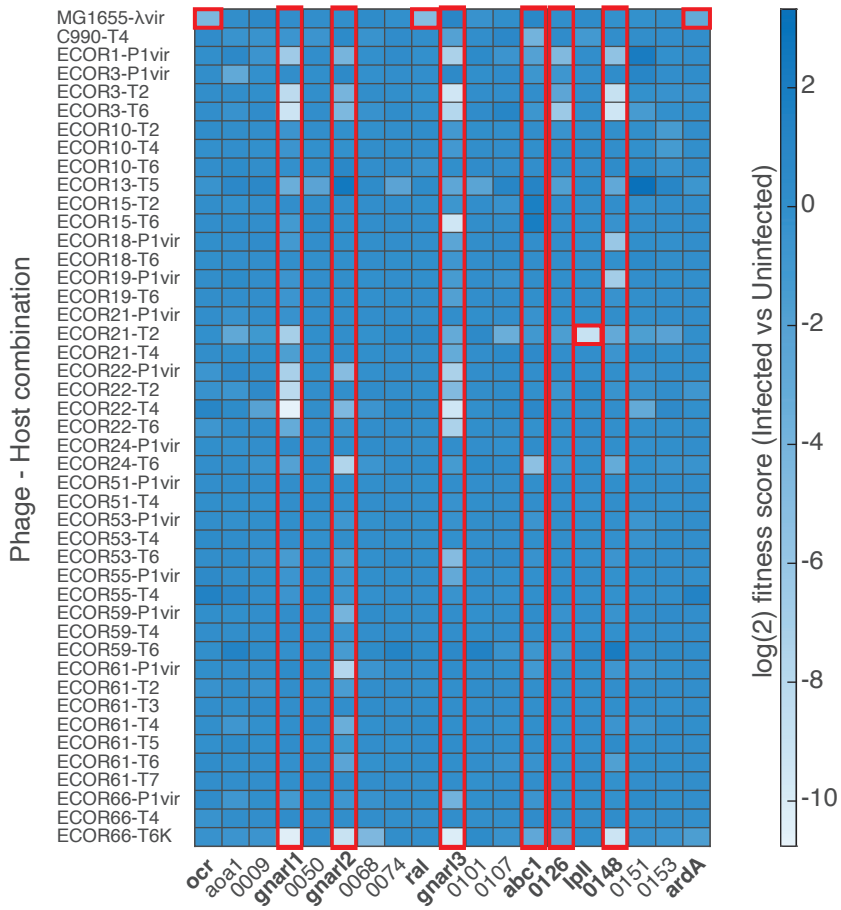
E. PhAGENT screen to assay AG phenotypes



D. Characteristics of ARs and AGs

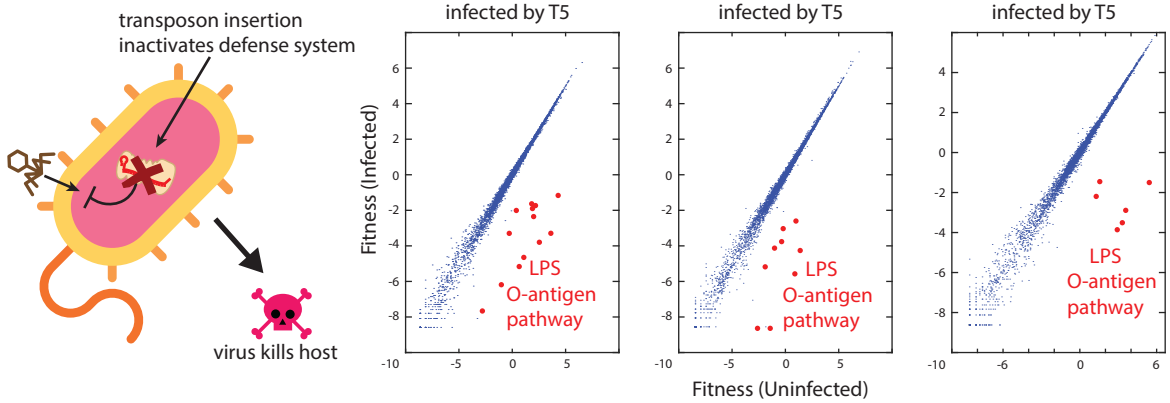


A. Accessory genes with counter-defense phenotypes in any host

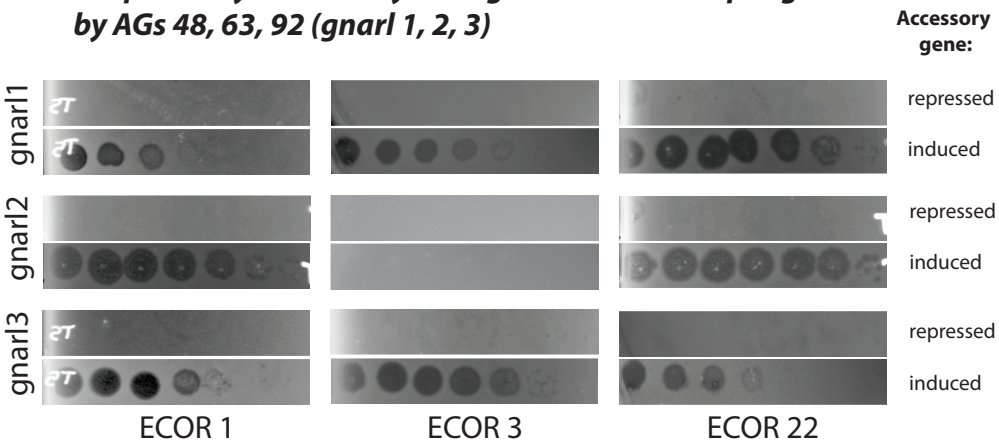


AGs that had a counter-defense effect in at least one host strain

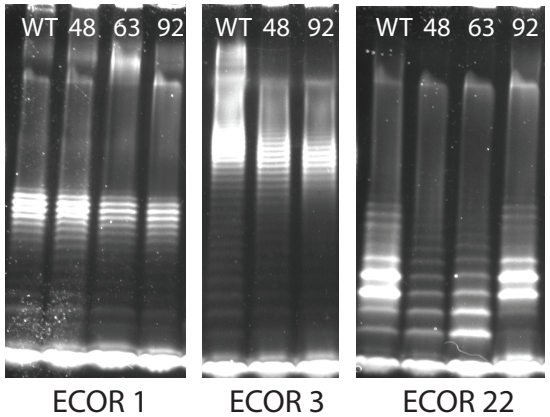
B. Whole genome Tn5-knockout screening to find phage exclusion mechanisms



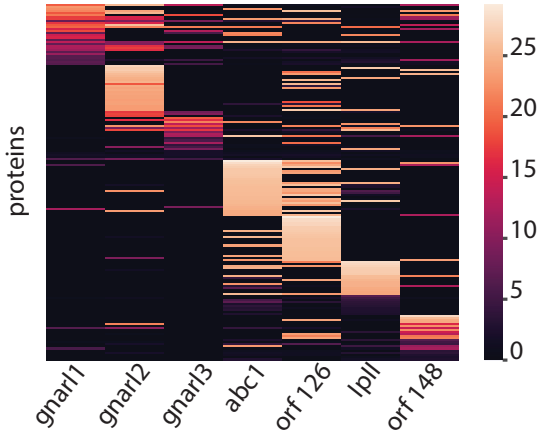
C. Plaque assays show very strong sensitization to phage T5 by AGs 48, 63, 92 (gnarl 1, 2, 3)



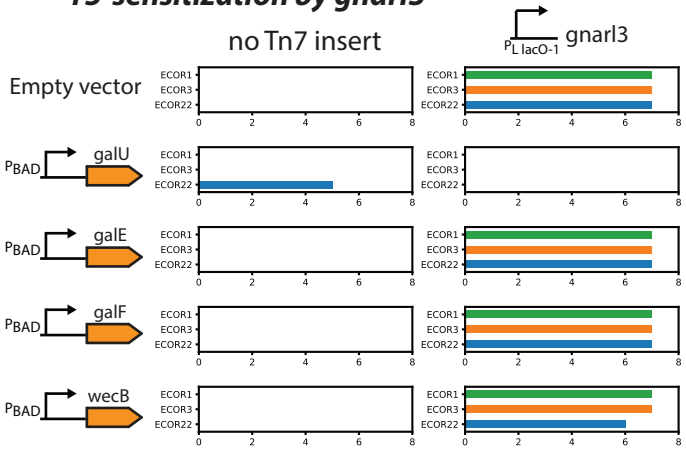
D. Gel electrophoresis of LPS from wild E. coli expressing AGs 48, 63, 92 (gnarl 1, 2, 3)



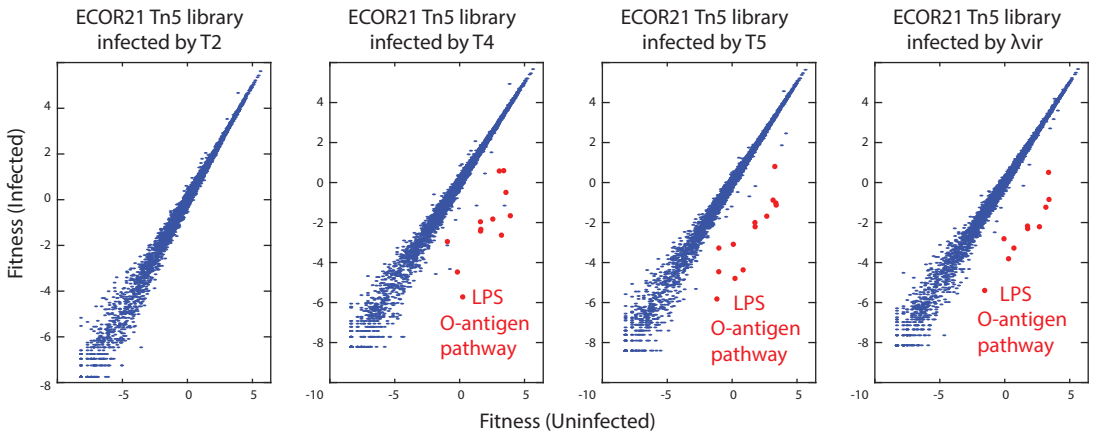
E. Heatmap of enriched host-proteins upon AP-MS of broad-spectrum phage AGs



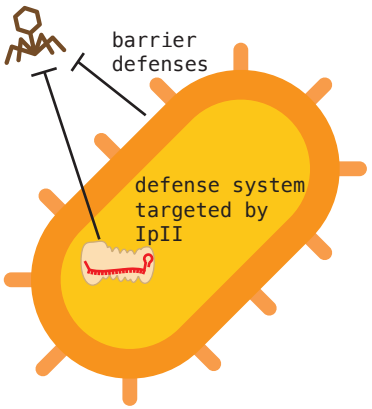
F. Over-expression of galU reverses T5-sensitization by gnarl3



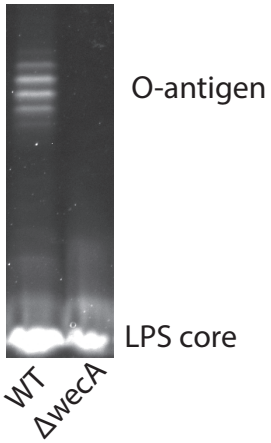
A. Whole genome Tn5-knockout screening in ECOR21 with various phages



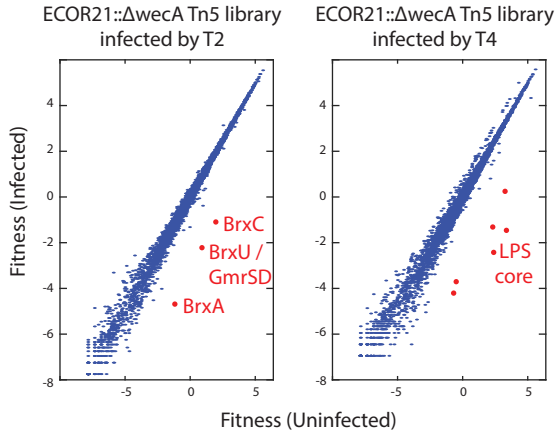
B. ECOR21 might block T2 at two levels simultaneously



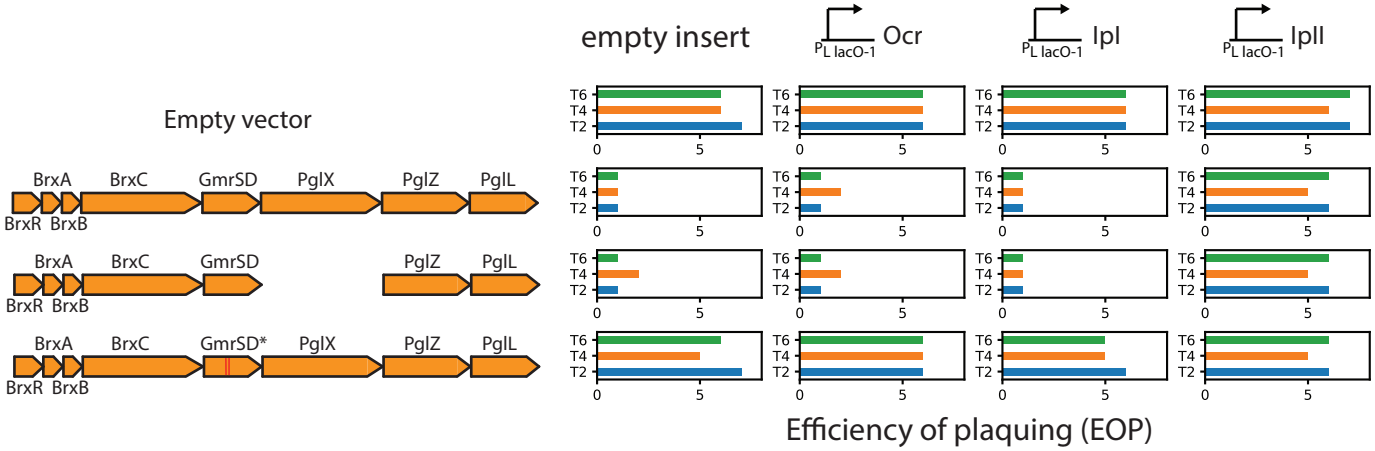
C. O-antigen removed in ECOR21::ΔwecA



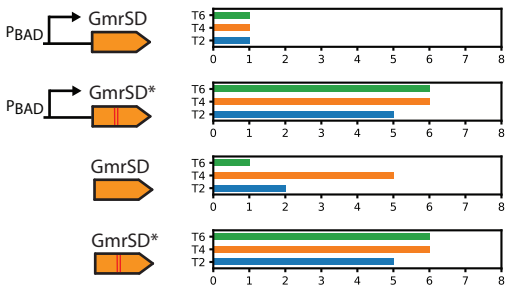
D. Tn5 screens after removing O-antigen layer



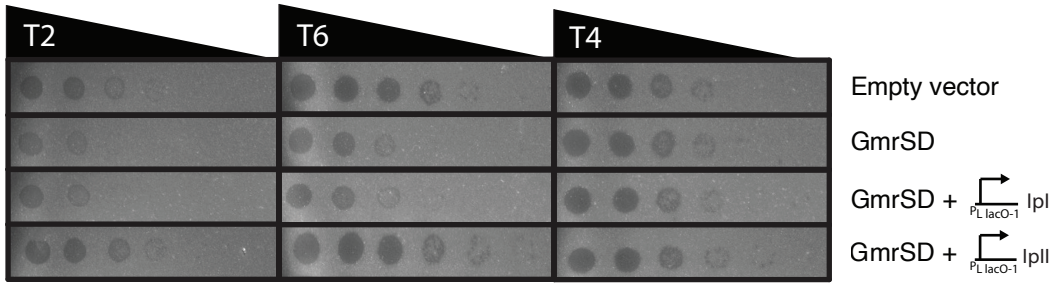
E. BREX-operon configurations tested against T-even phages with various inhibitors

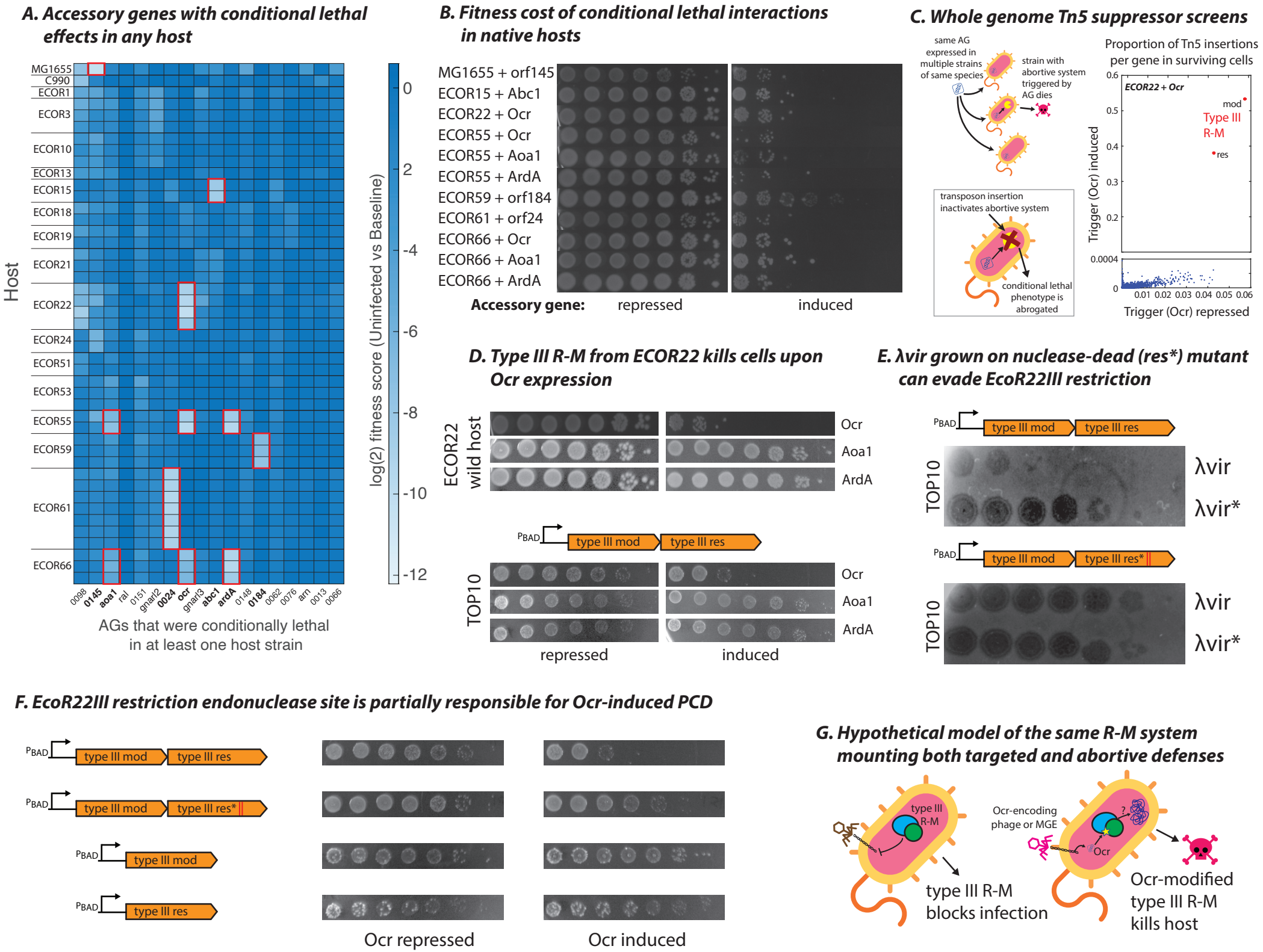


F. GmrSD is sufficient to confer resistance against T-even phages

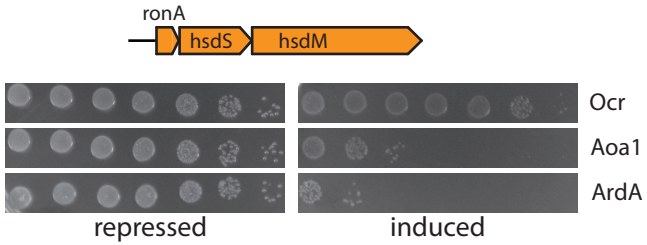


G. T4 naturally encodes IpII and does not require additional IpII to overcome basal levels of GmrSD restriction

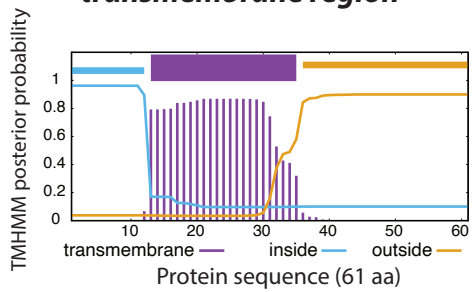



A. Accessory genes with conditional lethal effects in any host
Host
MG1655
C990
ECOR1
ECOR3
ECOR10
ECOR13
ECOR15
ECOR18
ECOR19
ECOR21
ECOR22
ECOR24
ECOR51
ECOR53
ECOR55
ECOR59
ECOR61
ECOR66
0098 0145 aoa1 ral 0151 gnarl2 0024 ocr gnarl3 abc1 ardA 0148 0184 0062 0076 arn 0013 0066
AGs that were conditionally lethal in at least one host strain
log(2) fitness score (Uninfected vs Baseline)
0 -2 -4 -6 -8 -10 -12
B. Fitness cost of conditional lethal interactions in native hosts
MG1655 + orf145
ECOR15 + Abc1
ECOR22 + Ocr
ECOR55 + Ocr
ECOR55 + Aoa1
ECOR55 + ArdA
ECOR59 + orf184
ECOR61 + orf24
ECOR66 + Ocr
ECOR66 + Aoa1
ECOR66 + ArdA
Accessory gene: repressed induced
C. Whole genome Tn5 suppressor screens
same AG expressed in multiple strains of same species
strain with abortive system triggered by AG dies
transposon insertion inactivates abortive system
conditional lethal phenotype is abrogated
Proportion of Tn5 insertions per gene in surviving cells
ECOR22 + Ocr
mod
Type III R-M
res
Trigger (Ocr) induced
Trigger (Ocr) repressed
D. Type III R-M from ECOR22 kills cells upon Ocr expression
ECOR22 wild host
Ocr
Aoa1
ArdA
PBAD
type III mod
type III res
TOP10
repressed induced
E. λvir grown on nuclease-dead (res*) mutant can evade EcoR22III restriction
type III res*
TOP10
λvir
λvir*
F. EcoR22III restriction endonuclease site is partially responsible for Ocr-induced PCD
Ocr repressed
Ocr induced
G. Hypothetical model of the same R-M system mounting both targeted and abortive defenses
type III R-M
Ocr-encoding phage or MGE
Ocr
type III R-M blocks infection
Ocr-modified type III R-M kills host

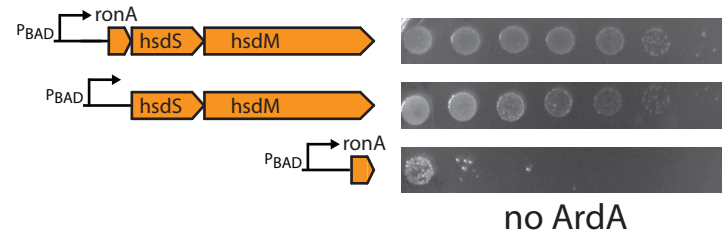
A. Ronin system from ECOR55 kills cells upon aoa1 or ardA induction



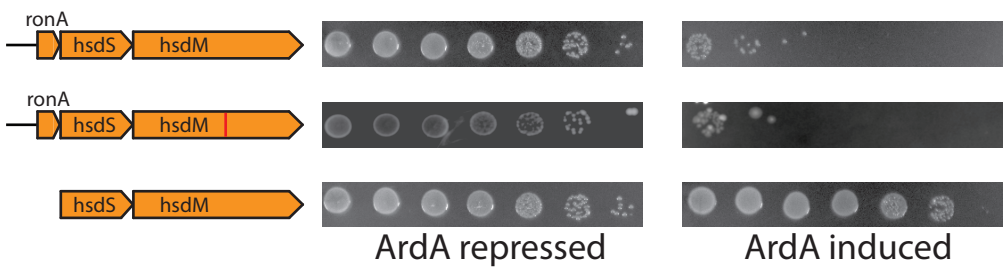
B. RonA contains a predicted transmembrane region



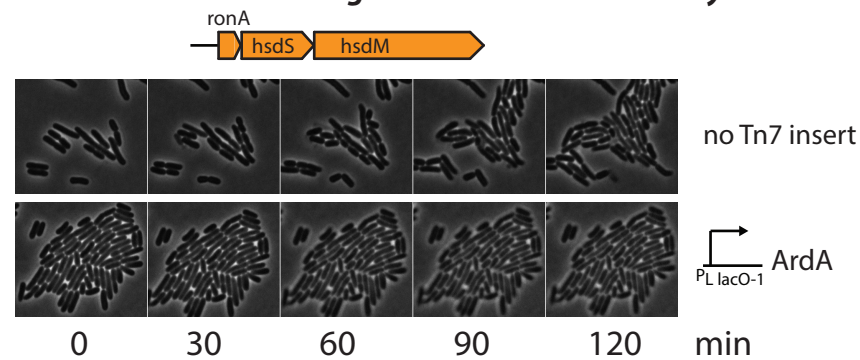
C. RonA is toxic upon overexpression but not with hsdMS present



D. RonA is necessary for ArdA-induced PCD activity



E. Ronin activation causes growth arrest but not cell lysis



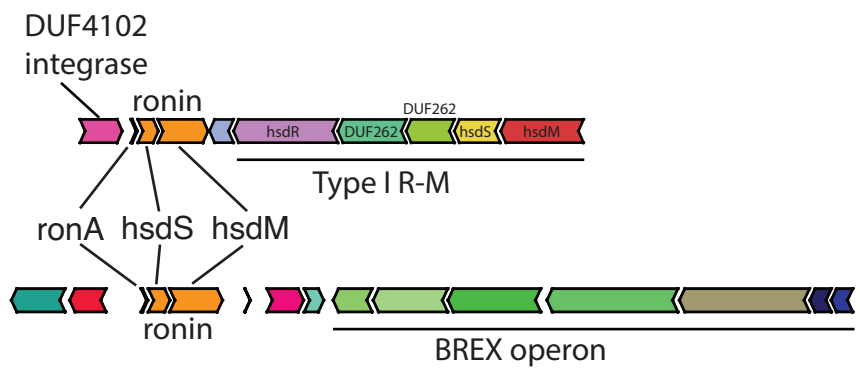
F. RonA-associated type I R-M systems (ronin) lack the restriction endonuclease

genomes with *ronA* (84)

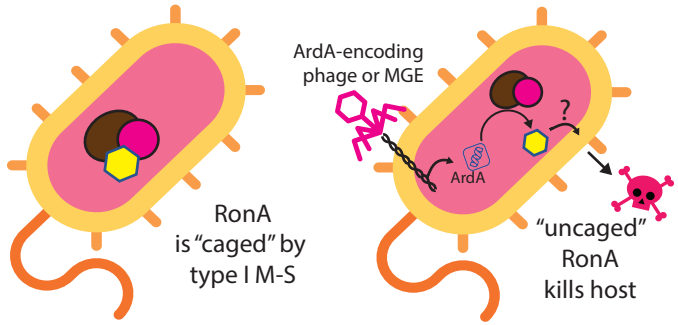
↓

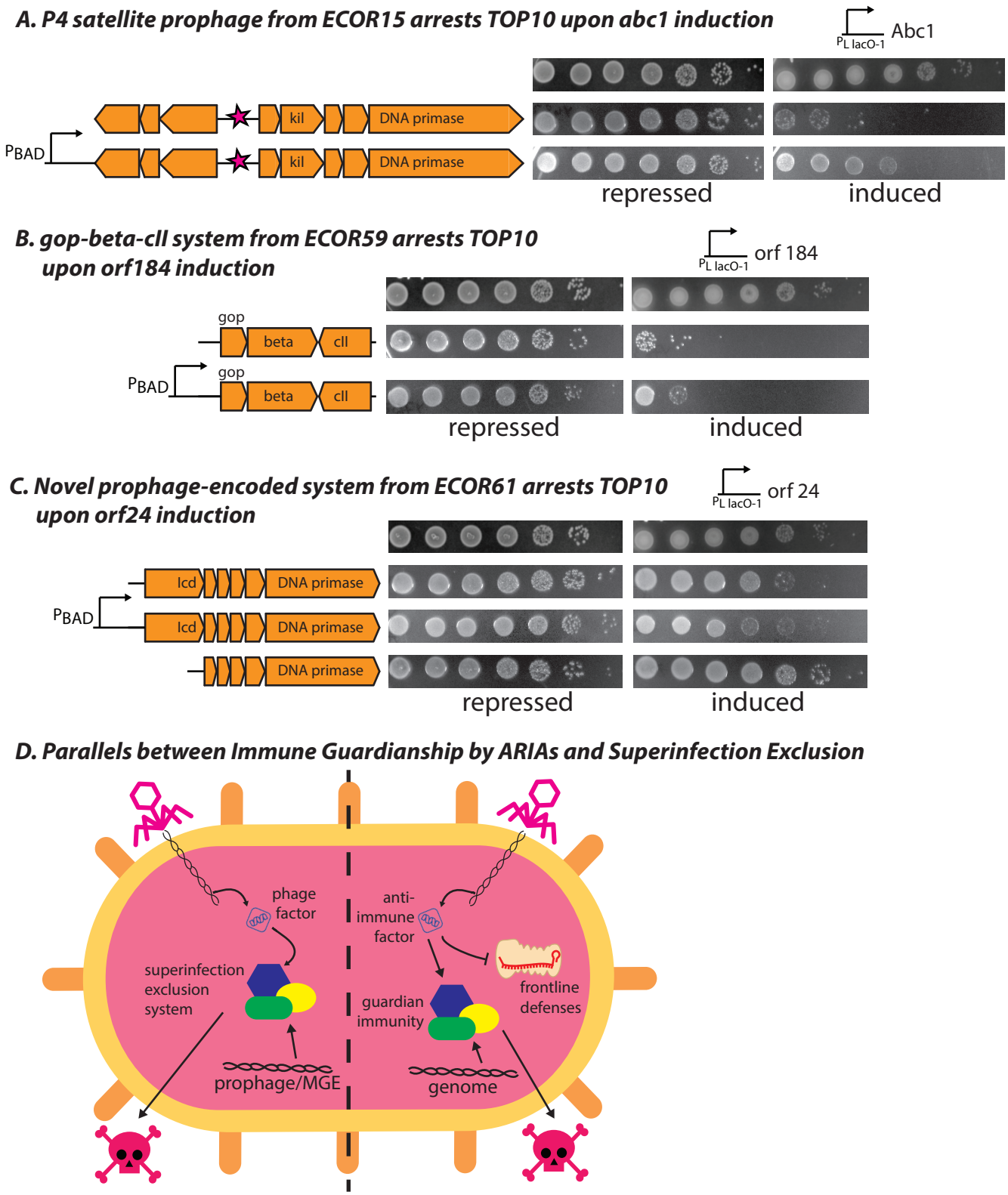
ronA linked to *hsdMS* (84)

	present in *E. coli*	full type I R-M anywhere
DUF4102 integrase upstream (58)	0	55
lacking full integrase upstream (26)	26	6



G. Hypothetical model of RonA "uncaging"




A. P4 satellite prophage from ECOR15 arrests TOP10 upon abc1 induction
PL lacO-1
Abc1
PBAD
kil
DNA primase
kil
DNA primase
repressed
induced
B. gop-beta-cII system from ECOR59 arrests TOP10 upon orf184 induction
PL lacO-1
orf 184
gop
beta
cII
PBAD
gop
beta
cII
repressed
induced
C. Novel prophage-encoded system from ECOR61 arrests TOP10 upon orf24 induction
PL lacO-1
orf 24
PBAD
lcd
DNA primase
lcd
DNA primase
DNA primase
repressed
induced
D. Parallels between Immune Guardianship by ARIAs and Superinfection Exclusion
phage factor
superinfection exclusion system
prophage/MGE
anti-immune factor
guardian immunity
frontline defenses
genome

A. Phage susceptibility of ECOR strains

	T2	T3	T4	T5	T6	λvir	P1vir	T7
C990	6	6	3	7	6	7	6	7
MG1655	6	4	6	7	6	3	6	7
ECOR1							3	
ECOR3	2				3		2	
ECOR4	4	4		7	6			7
ECOR6								
ECOR10	0		0		0			
ECOR13	6	2	6	4	7	7	3	7
ECOR15	0				1			
ECOR18					1		1	
ECOR19					1		1	
ECOR21	1		2				1	
ECOR22	0		3		0		2	
ECOR24					0		3	
ECOR25								
ECOR51			4		6		2	
ECOR53			1		2		2	
ECOR55			1		5		2	
ECOR59			0		1		1	
ECOR61	2	0	1	3	1		2	0
ECOR66			3		5		3	

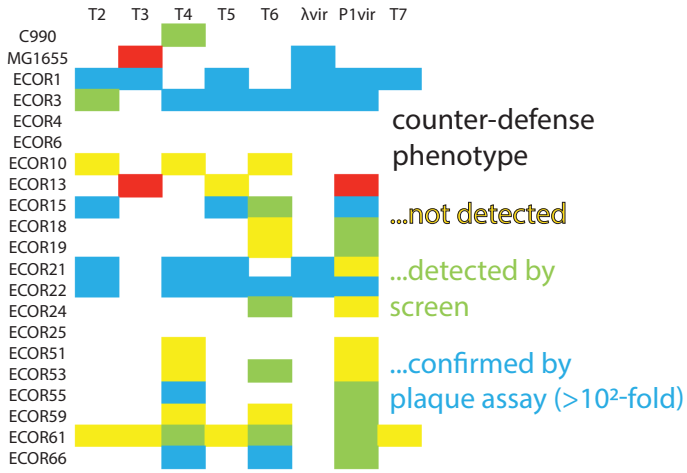
n EOP = n

n EOP = n, and single plaques visible

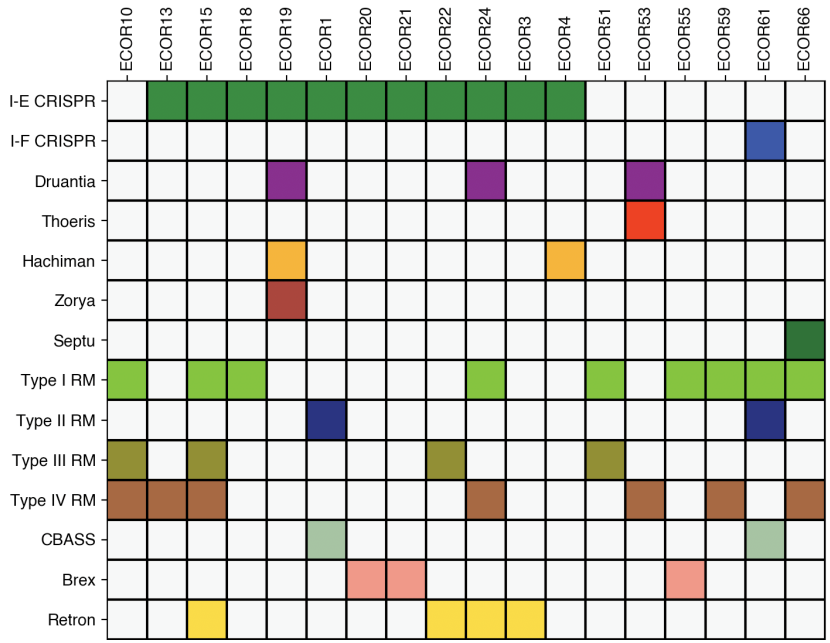
selected for PhAGENT screen

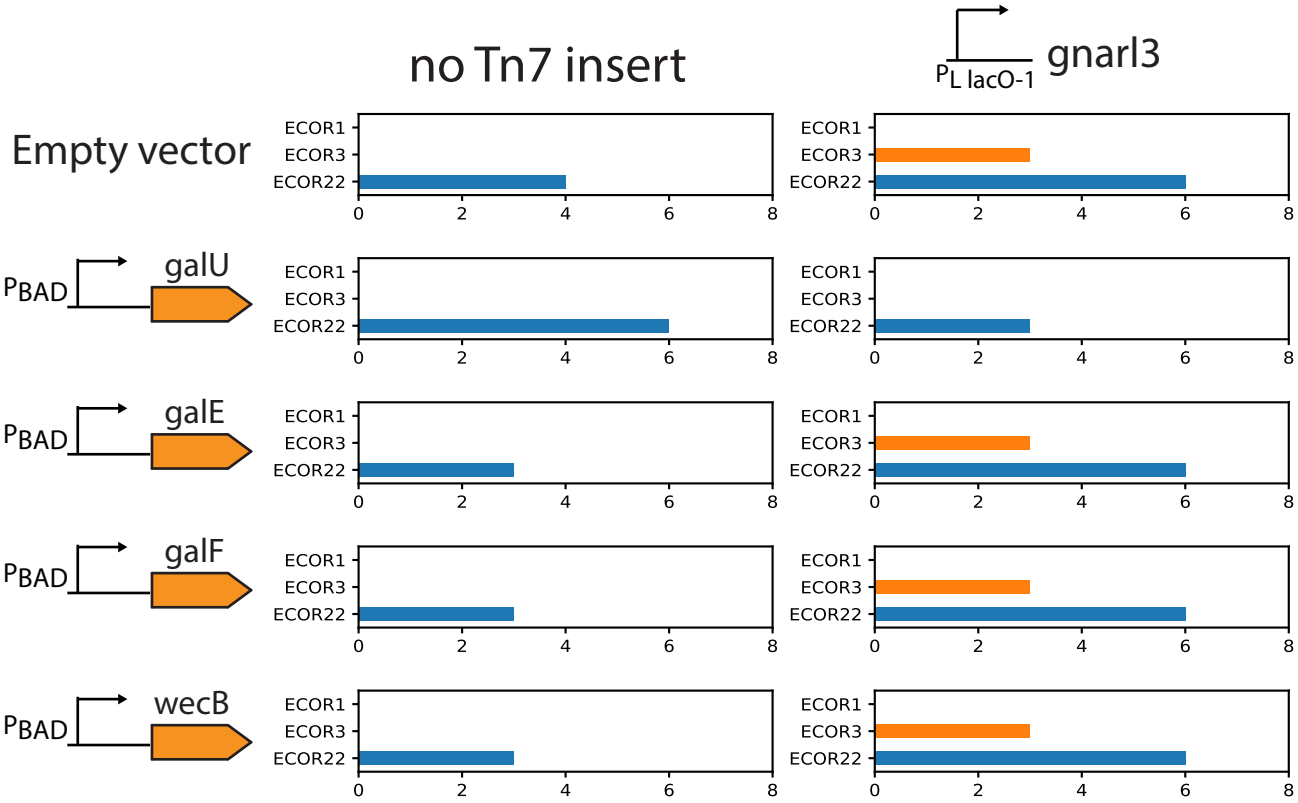
attempted but screen could not be performed

B. Phage-host combinations with confirmed counter-defense phenotypes

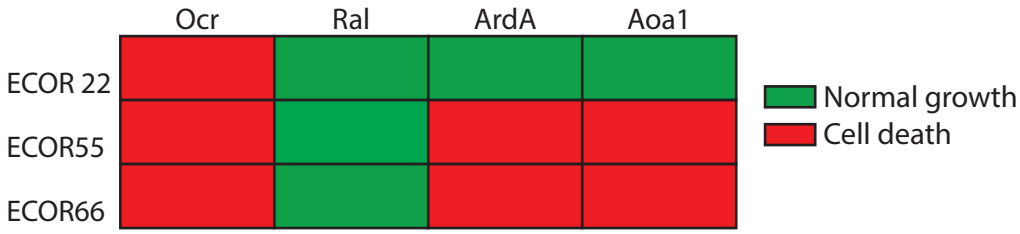


C. Defense systems in ECOR strains

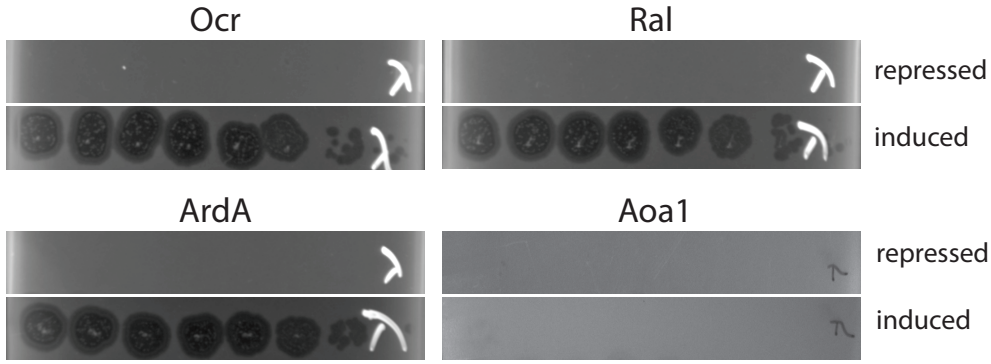



no Tn7 insert
PL lacO-1
gnarl3
Empty vector
PBAD
galU
galE
galF
wecB
ECOR1
ECOR3
ECOR22
0
2
4
6
8

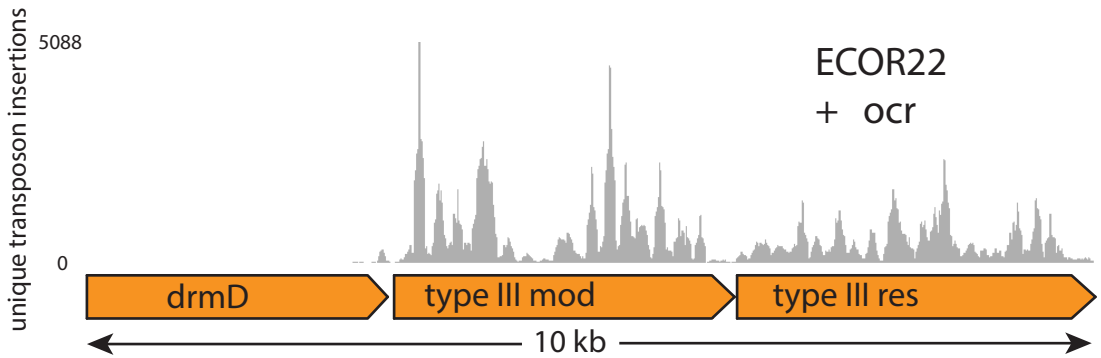
A. R-M inhibitors trigger abortive defenses in wild E. coli

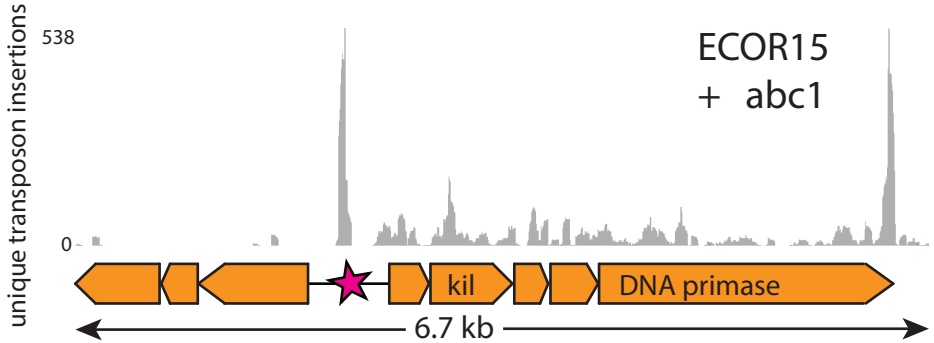


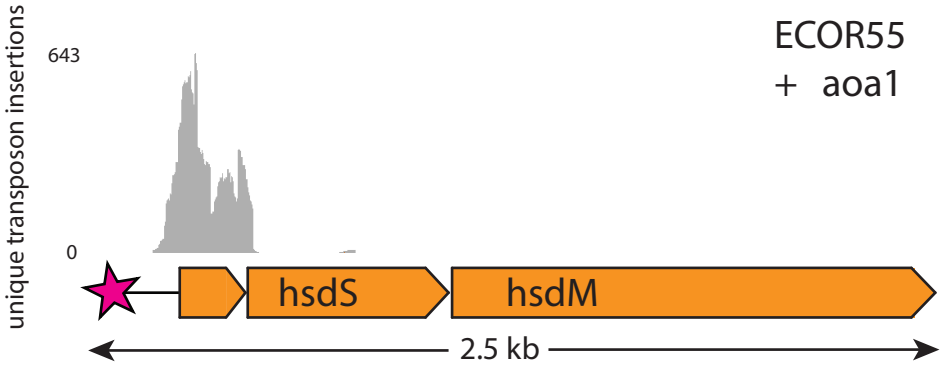
B. R-M inhibition by AGs in MG1655

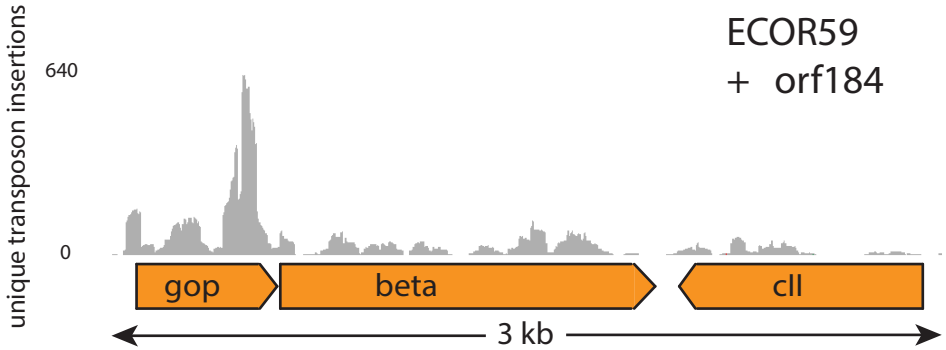


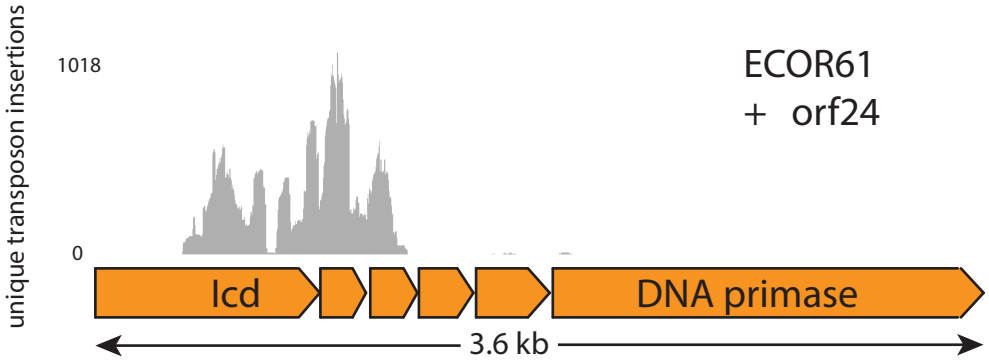
C. Effect of Ocr F54V mutation on ARIA activity


unique transposon insertions
5088
0
ECOR22
+ ocr
drmD
type III mod
type III res
10 kb

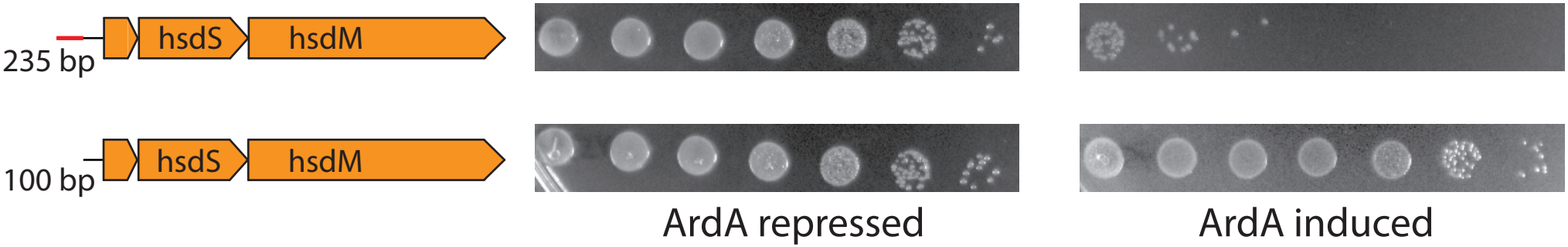

unique transposon insertions
538
0
ECOR15
+ abc1
kil
DNA primase
6.7 kb


unique transposon insertions
643
0
ECOR55
+ aoa1
hsdS
hsdM
2.5 kb

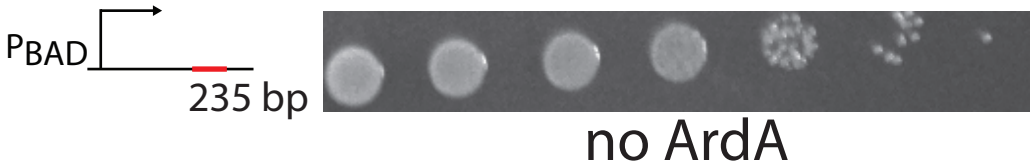
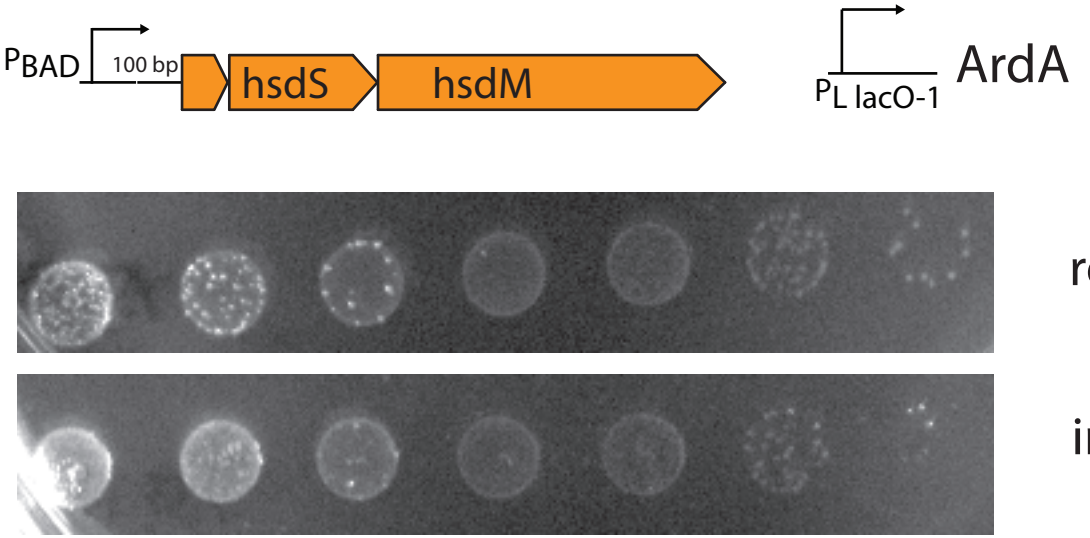

unique transposon insertions
640
0
ECOR59
+ orf184
gop
beta
cII
3 kb


unique transposon insertions
1018
0
ECOR61
+ orf24
lcd
DNA primase
3.6 kb

A. Ronin upstream region is required for PCD activity



B. Ronin upstream region is not toxic when overexpressed on its own



C. Over-expression of Ronin with truncated upstream region