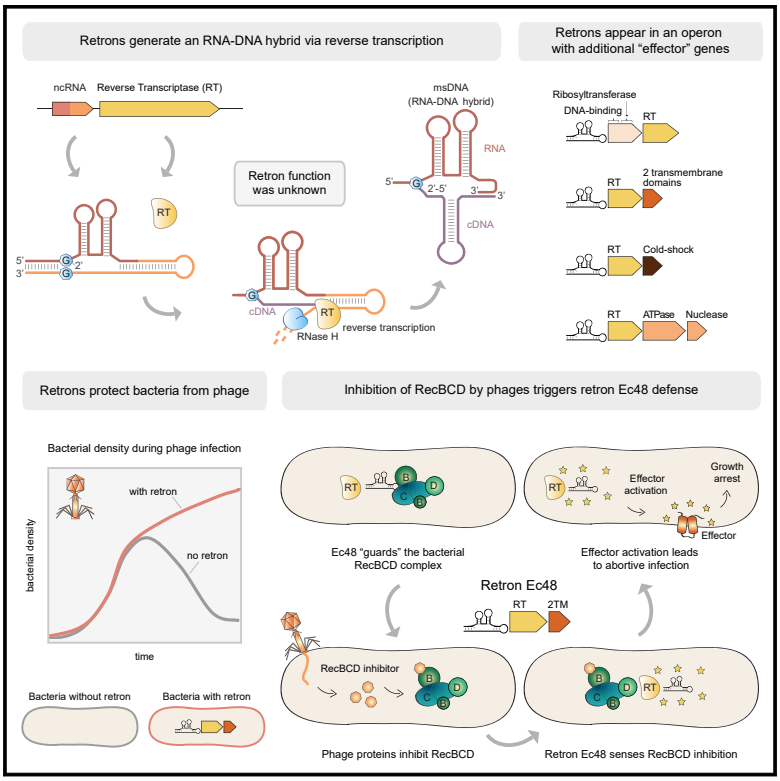
Cell

Article

Bacterial Retrons Function In Anti-Phage Defense

Graphical Abstract



Highlights

- Retrons are preferentially located in defense islands
- Retrons, together with their effector genes, protect bacteria from phages
- Protection from phage is mediated by abortive infection
- Retron Ec48 guards RecBCD. Inhibition of RecBCD by phages triggers retron defense

Authors

Adi Millman, Aude Bernheim, Avigail Stokar-Avihail, ..., Azita Leavitt, Yaara Oppenheimer-Shaanan, Rotem Sorek

Correspondence

rotem.sorek@weizmann.ac.il

In Brief

Retrons are part of a large family of anti-phage defense systems that are widespread in bacteria and confer resistance against a broad range of phages, mediated by abortive infection.

Millman et al., 2020, Cell *183*, 1551–1561
December 10, 2020
https://doi.org/10.1016/j.cell.2020.09.065

CellPress

Cell

CellPress

Article

Bacterial Retrons Function In Anti-Phage Defense

Adi Millman,[1,3] Aude Bernheim,[1,3] Avigail Stokar-Avihail,[1,3] Taya Fedorenko,[1] Maya Voichek,[1,2] Azita Leavitt,[1] Yaara Oppenheimer-Shaanan,[1] and Rotem Sorek[1,4,*]

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

[2]Present address: Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna Biocenter (VBC), Dr. Bohrgasse 3, 1030 Vienna, Austria

[3]These authors contributed equally

[4]Lead Contact

*Correspondence: rotem.sorek@weizmann.ac.il

https://doi.org/10.1016/j.cell.2020.09.065

SUMMARY

Retrons are bacterial genetic elements comprised of a reverse transcriptase (RT) and a non-coding RNA (ncRNA). The RT uses the ncRNA as template, generating a chimeric RNA/DNA molecule in which the RNA and DNA components are covalently linked. Although retrons were discovered three decades ago, their function remained unknown. We report that retrons function as anti-phage defense systems. The defensive unit is composed of three components: the RT, the ncRNA, and an effector protein. We examined multiple retron systems and show that they confer defense against a broad range of phages via abortive infection. Focusing on retron Ec48, we show evidence that it "guards" RecBCD, a complex with central anti-phage functions in bacteria. Inhibition of RecBCD by phage proteins activates the retron, leading to abortive infection and cell death. Thus, the Ec48 retron forms a second line of defense that is triggered if the first lines of defense have collapsed.

INTRODUCTION

Retrons are genetic elements composed of a non-coding RNA (ncRNA) and a specialized reverse transcriptase (RT). These elements typically generate a chimeric RNA-DNA molecule, in which the RNA and DNA components are covalently attached by a 2′-5′ phosphodiester bond (Figure S1). Retrons were originally discovered in 1984 in *Myxococcus xanthus*, when Yee et al. (1984) identified a short, multi-copy single-stranded DNA (msDNA) that is abundantly present in the bacterial cell. Further studies showed that this single-stranded DNA (ssDNA) is covalently linked to an RNA molecule (Dhundale et al., 1987) and later deciphered in detail the biochemical steps leading to the formation of the RNA-DNA hybrid (Lampson et al., 1989). It was found that the retron ncRNA is the precursor of the hybrid molecule and folds into a typical structure that is recognized by the RT (Hsu et al., 1989). The RT then reverse transcribes part of the ncRNA, starting from the 2′-end of a conserved guanosine residue found immediately after a double-stranded RNA structure within the ncRNA (Lampson et al., 1989). A portion of the ncRNA serves as a template for reverse transcription, which terminates at a defined position within the ncRNA (Lampson et al., 1989). During reverse transcription, cellular RNase H degrades the segment of the ncRNA that serves as template, but not other parts of the ncRNA, yielding the mature RNA-DNA hybrid (Lampson et al., 1989; Figure S1).

Dozens of retrons have been documented in a variety of microbial genomes, and 16 of them were studied experimentally in detail (Simon et al., 2019). The documented retrons were all named following a naming convention that includes the first letters of their genus and species names, as well as the length of reverse-transcribed DNA (e.g., Ec48 is a retron found in *Escherichia coli* whose reverse transcribed DNA segment is 48 nt long). All studied retrons contain an RT and a ncRNA, with the conserved guanosine from which reverse transcription is initiated (Lampson et al., 2001). However, the sequences and lengths of the reverse transcribed template significantly vary and frequently show no sequence similarity between retrons (Das et al., 2011). The ability of retrons to produce ssDNA *in situ* has been adapted for multiple applications of synthetic biology and genome engineering (Farzadfard and Lu, 2014; Sharon et al., 2018; Simon et al., 2018, 2019).

Although retrons have been studied for over 35 years, their biological function remained unknown. It has been suggested that retrons are a form of selfish genetic elements (Rice and Lampson, 1995) or have a function in coping with starvation (Herzer, 1996), pathogenesis (Elfenbein et al., 2015), and cell-specialization (Simon et al., 2019). However, evidence for these functions was circumstantial, and the mechanism by which retrons would exert these putative functions was not identified. In the current study, we show that retrons form a functional component in a large family of anti-phage defense systems that are widespread in bacteria and confer resistance against a broad range of phages.

RESULTS

A Retron-Containing Genetic System Protects against Phage Infection

We initiated the current study by searching for RT genes that may participate in defense against phages. This search was inspired by prior reports on the involvement of RTs in bacterial defense (Fortier et al., 2005; Silas et al., 2016; Wang et al., 2011) and phage counter-defense mechanisms (Doulatov et al., 2004). Because bacterial defense systems tend to cluster in "defense islands" in microbial genomes (Cohen et al., 2019; Doron et al., 2018; Makarova et al., 2011), we focused on RT genes of unknown function that are frequently encoded near known anti-phage systems such as restriction enzymes (STAR Methods).

One of these RT genes is presented in Figures 1A and 1B. Homologs of this gene appear in a diverse set of bacteria (102 homologs found in species belonging to the Proteobacteria and Firmicutes phyla) and show marked tendency to co-localize with known defense systems, with 60 (59%) located near known anti-phage operons (Figures 1A and 1B). The RT gene is always found next to a second gene with a predicted OLD-family endonuclease domain (Schiltz et al., 2020), and we therefore hypothesized that the RT together with the endonuclease form a two-gene phage resistance system. To test this hypothesis, we cloned this two-gene system from *E. coli* 200499, together with its flanking intergenic regions, into the laboratory strain *E. coli* MG1655, which naturally lacks the system. We then challenged the transformed strain with a set of 12 coliphages and found that the system conferred protection against phages from a variety of families: T7 (*Podoviridae*), T4 and T6 (*Myoviridae*), and SECphi4, SECphi6, and SECphi18 (*Siphoviridae*) (Figures 1C and S2).

The presence of an atypically large (>100 nt), conserved intergenic region between the predicted endonuclease and RT genes led us to hypothesize that this intergenic region might contain a non-coding RNA (Figure 1B). Indeed, examining RNA sequencing (RNA-seq) data from *Paenibacillus polymyxa* (Voichek et al., 2020), which naturally encodes this system, showed high levels of expression from the intergenic region (Figure 1D). Similar expression patterns were observed in RNA-seq data from the *E. coli* strain into which we cloned the defense system, consistent with the presence of a ncRNA in the intergenic region (Figure 1D).

Because the RT gene showed significant homology to retron-type RTs (Simon and Zimmerly, 2008), we hypothesized that the newly discovered defense system contains a retron, and the ncRNA we detected is the retron ncRNA precursor. In support of this, we found that the predicted secondary structure of the ncRNA conforms with the characteristics of known retron ncRNA precursors, including the conserved local dsRNA structure immediately followed by non-paired guanosine residues on both strands (Figures 1E and S1). These structural features were conserved among homologs of this system (Figure 1E). To check if the ncRNA indeed forms a precursor for ssDNA synthesis, we extracted the DNA from a strain into which we cloned the RT and ncRNA, and found a ssDNA species sized between 70–80 nt, which was absent from the control strain that contained a GFP gene instead (Figure 1F). These results confirm that the new defense system we discovered contains a previously unidentified retron.

To examine whether the retron features are involved in the anti-phage activity of the new defense system, we experimented with mutated versions of the system. Point mutations in the conserved YADD motif of the catalytic core of the RT (D200A and D201A) (Lampson et al., 2005) rendered the system inactive (Figure 1G). Similarly, a point mutation in the ncRNA, mutating the guanosine predicted as the branching residue priming the reverse transcription (Hsu et al., 1992) (G > C at position 17 of the ncRNA), or the second conserved guanosine that was shown in other retrons to be essential for initiation of reverse transcription (Hsu et al., 1992) (G > A at position 147), completely abolished defense against phages (Figure 1G). These results suggest that proper reverse transcription of the retron ncRNA is essential for its defensive function. We also found that a point mutation in the ATP-binding motif of the associated predicted OLD-family endonuclease gene (K36A) completely eliminated the defense phenotype, showing that the predicted endonuclease gene is an indispensable component of the retron-containing defense system. We therefore conclude that the new defense system consists of three components essential for its anti-phage activity: the RT and ncRNA (that together form an active retron) and an additional gene that contains a predicted endonuclease domain. Following a recently proposed revised nomenclature for retrons (Simon et al., 2019), we termed this defense system Retron-Eco8.

Known Retrons Are Part of Multi-Gene Systems and Are Located in Defense Islands

The identification of a novel defense system that contains a retron led us to ask whether retrons in general may have a role in defense against phages. If this is the case, we would expect to find retrons enriched in defense islands, near known anti-phage defense systems. To test this hypothesis, we searched for homologs of the RT proteins of previously characterized retrons in a set of 38,167 bacterial and archaeal genomes (STAR Methods). We detected 4,802 homologs of retron RTs in 4,446 of the genomes and used hierarchical clustering to divide these homologs into eight clades (Figure 2A; Table S1). We found that in six of the clades, the RT genes had a strong tendency to be genomically associated with other known defense genes in defense islands (STAR Methods). Between 38% and 47% of the genes in each of these clades were found to be located near known defense systems, implying that most retrons may participate in anti-phage defense.

Retrons have been previously described as two-component systems comprised of the RT and the precursor ncRNA (Simon et al., 2019). However, when examining the genomic environments of known retrons and their homologs, we observed that the vast majority are encoded as part of a gene cassette that includes one or two additional protein-coding genes (Figures 2B–2D; Table S1). For example, retrons Ec73, St85, and Vc81 (found in *E. coli*, *Salmonella enterica*, and *Vibrio cholerae*, respectively) all have an upstream gene containing a ribosyltransferase and a DNA-binding domain (Figure 2B); and retron Ec48 and its homologs are linked to a gene encoding a protein with two predicted transmembrane (2TM) helices (Figure 2C). Some retrons,

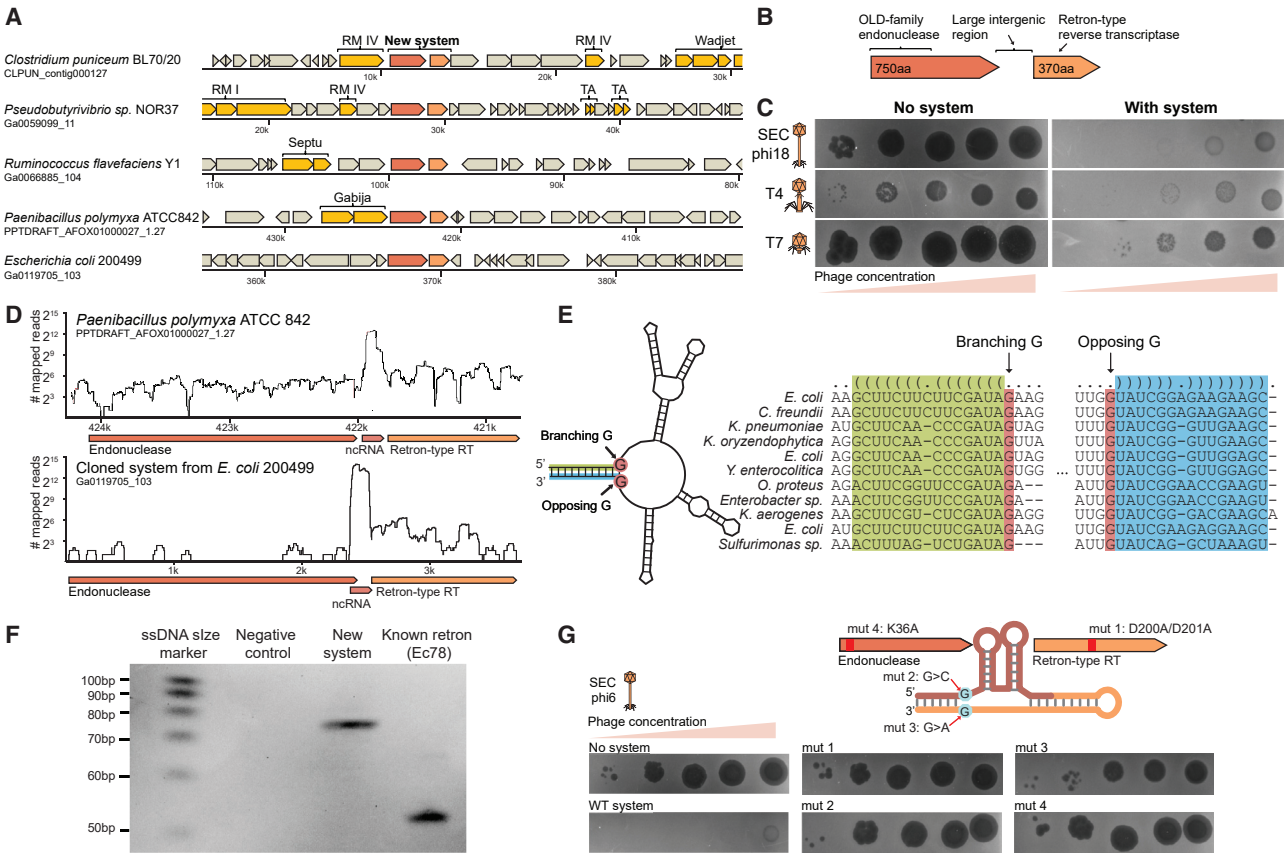


Figure 1. A Retron-Containing Genetic System Protects against Phage Infection

(A) A conserved two-gene cassette containing a reverse transcriptase is found in defense islands. Representative instances of the system, and their genomic environments, are presented. Genes known to be involved in defense are shown in yellow. RM, restriction modification; TA, toxin-antitoxin. Wadjet, Septu, and Gabija are recently described defense systems (Doron et al., 2018). The encoding strain and the respective DNA scaffold accession on the IMG database (Chen et al., 2019) are indicated to the left.

(B) Domain organization of the gene cassette.

(C) Serial dilution plaque assays shown for three phages on *E. coli* MG1655 strain transformed with the two-gene cassette ("with system") or with a control plasmid encoding an RFP ("no system"). Images are representative of three replicates. See also Figure S2.

(D) RNA-seq coverage of the systems' loci in *Paenibacillus polymyxa* ATCC 842 (natively encoding the system, top) and *E. coli* MG1655 into which the system from *E. coli* 200499 was transformed (bottom).

(E) Predicted structure of the ncRNA in the intergenic region of the tested system shows features of retron-type ncRNA (left). The structure corresponds to positions 369202–369366 in the *E. coli* 200499 genome assembly (scaffold Ga0119705_103). Alignment and structure prediction of the intergenic region from homologs in multiple bacteria are shown on the right. Areas of conserved base pairing are highlighted. Predicted structure is presented on the top, with each pair of parentheses representing base-pairing.

(F) Production of multicopy single-stranded DNA (msDNA). DNA was extracted from *E. coli* strains expressing the ncRNA and RT of the new defense system, a known retron Ec78, or a negative control expressing GFP instead (STAR Methods). Extracted msDNA was examined on a 10% denaturing polyacrylamide gel.

(G) Mutational analysis of elements within the new defense system. Shown are 10-fold serial dilution plaque assays with phage SECphi6, comparing the strain with the wild-type system to strains with mutated versions. Images are representative of two replicates.

See also Figure S1.

including Ec78, Yf79, and Vc95 are encoded as part of a cassette that includes two additional genes: one with an ATPase domain and another with an HNH endonuclease domain, a gene organization that was previously identified in the anti-phage system Septu (Doron et al., 2018; Figure 2D). In other cases (e.g., retron Ec67), the associated gene is fused to the RT gene (Figure 2E). The tight genetic linkage of retrons with these genes suggests that the functional unit that includes the retrons also includes the associated genes. We refer to these associated genes as retron "effectors," due to reasons explained below. Overall, we found 10 different types of such effector genes that are associated with the retron RTs (Figure 2E).

Retrons, Together with Their Effector Genes, Protect against Phages

To test whether retrons function as anti-phage defense systems, we experimentally examined six previously characterized retrons that were identified in *E. coli* strains, as well as five validated retrons encoded by *S. enterica* and *V. cholerae*. We cloned each retron, together with the predicted effector gene(s), into an

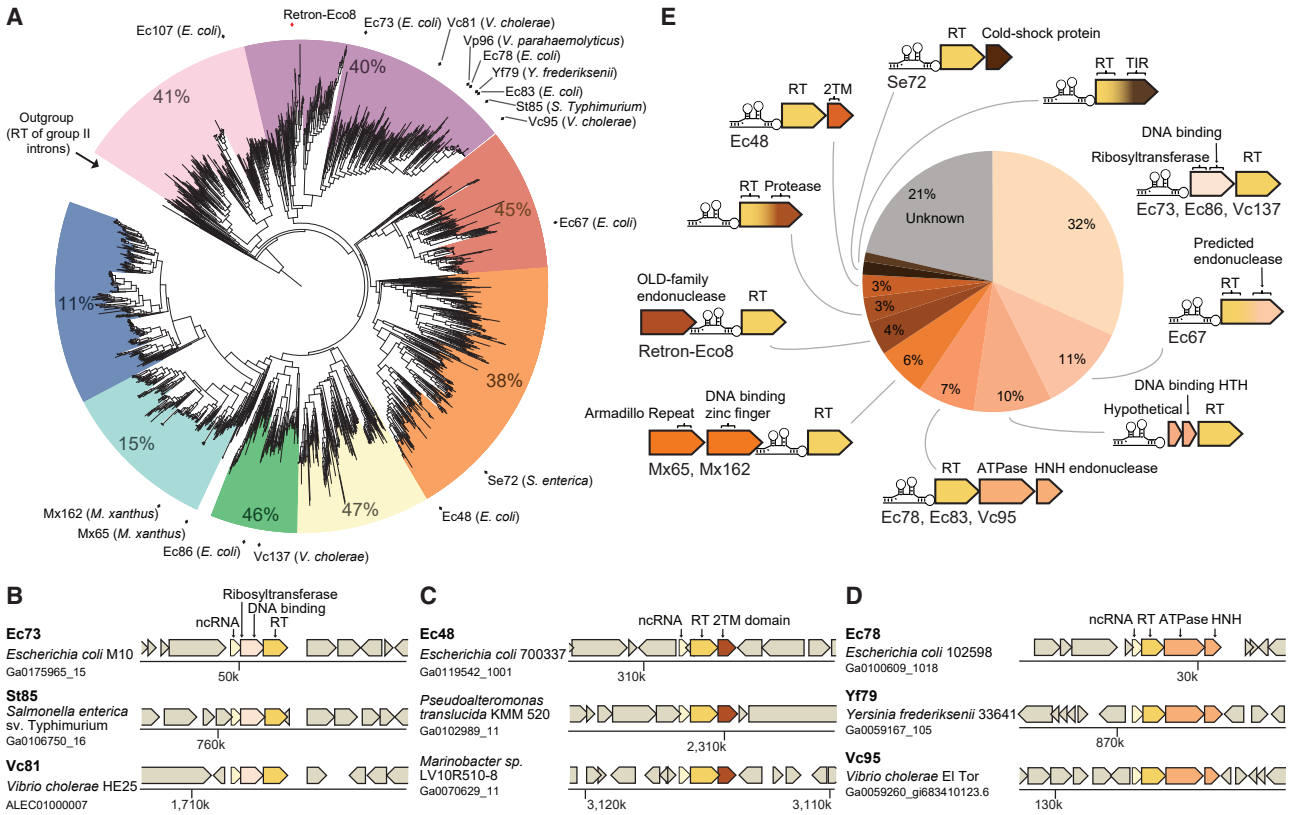


Figure 2. Known Retrons Are Part of Multi-Gene Systems and Are Located in Defense Islands

(A) Phylogenetic analysis of homologs of retron RTs. Known retrons are marked around the tree with the encoding species stated in parentheses. For each clade, the percentage of genes that appear next to known defense systems is presented (see STAR Methods).

(B–D) The genomic environments of known retrons and their homologs, divided into systems containing an associated ribosyltransferase (B), a two-transmembrane (2TM) domain gene (C), or genes with ATPase and HNH endonuclease domains (D). Names of known retrons are in bold. The encoding strain and the respective DNA scaffold accession in the IMG database (Chen et al., 2019) are indicated for each retron or retron homolog.

(E) Distribution of different types of genes associated with homologs of retron RTs (N = 4,802). Examples for known retrons are stated below the system type they belong to.

See also Figure S4 and Table S1.

E. coli MG1655 strain that is not known to encode retrons. We then challenged the retron-containing bacteria with an array of 12 phages that span several major phage families. For eight of the 11 systems, we observed marked anti-phage activity against at least one phage (Figures 3 and S3). For some strains (e.g., those harboring retrons Ec86 and Se72), anti-phage activity was restricted to one phage, while for others (Ec73 and Ec48) defense was broad, spanning phages from several different phage families (Figures 3 and S3).

To assess whether both the activity of the retron and its effector gene are necessary for anti-phage defense, we further experimented with the two retron systems that showed the broadest defense (Ec48 and Ec73). Point mutations predicted to inactivate the catalytic site of the RT (D216A/D217A in the RT of Ec48 and D189A/D190A in the RT of Ec73) completely abolished defense, indicating that reverse transcription of the retron is essential for defense. In addition, an E117Q point mutation predicted to inactivate the catalytic site of the ribosyltransferase domain of the Ec73 effector (Sikowitz et al., 2013), led to a non-functional system, and similarly, deletion of the transmembrane helices of the gene associated with the Ec48 retron also abolished defense (Figures 3 and S3). Together, these results show that retrons, functioning together with their associated effector genes, form anti-phage defense systems.

To assess the prevalence of retron defense systems in prokaryotes, we examined their distribution in the set of 38,167 bacterial and archaeal genomes that we analyzed, which belong to 14,566 species. Homologs of the retron RT were found in 1,928 species, belonging to over 20 different phyla. Retrons homologs were common in Proteobacteria and Cyanobacteria, and in these phyla, more than 20% of the species were found to encode retrons (Figure S4A). In contrast, retron homologs were rare in Archaea, with only 4 out of 347 species (1%) of the Euryarchaeota phylum harboring retrons (Figure S4A; Table S1). The most common retron effector type was a ribosyltransferase effector, appearing next to 1,525 (32%) of the retron RT homologs we found (Figures 2E and S4B). Retrons with this effector type are present in bacteria spanning a wide phylogenetic

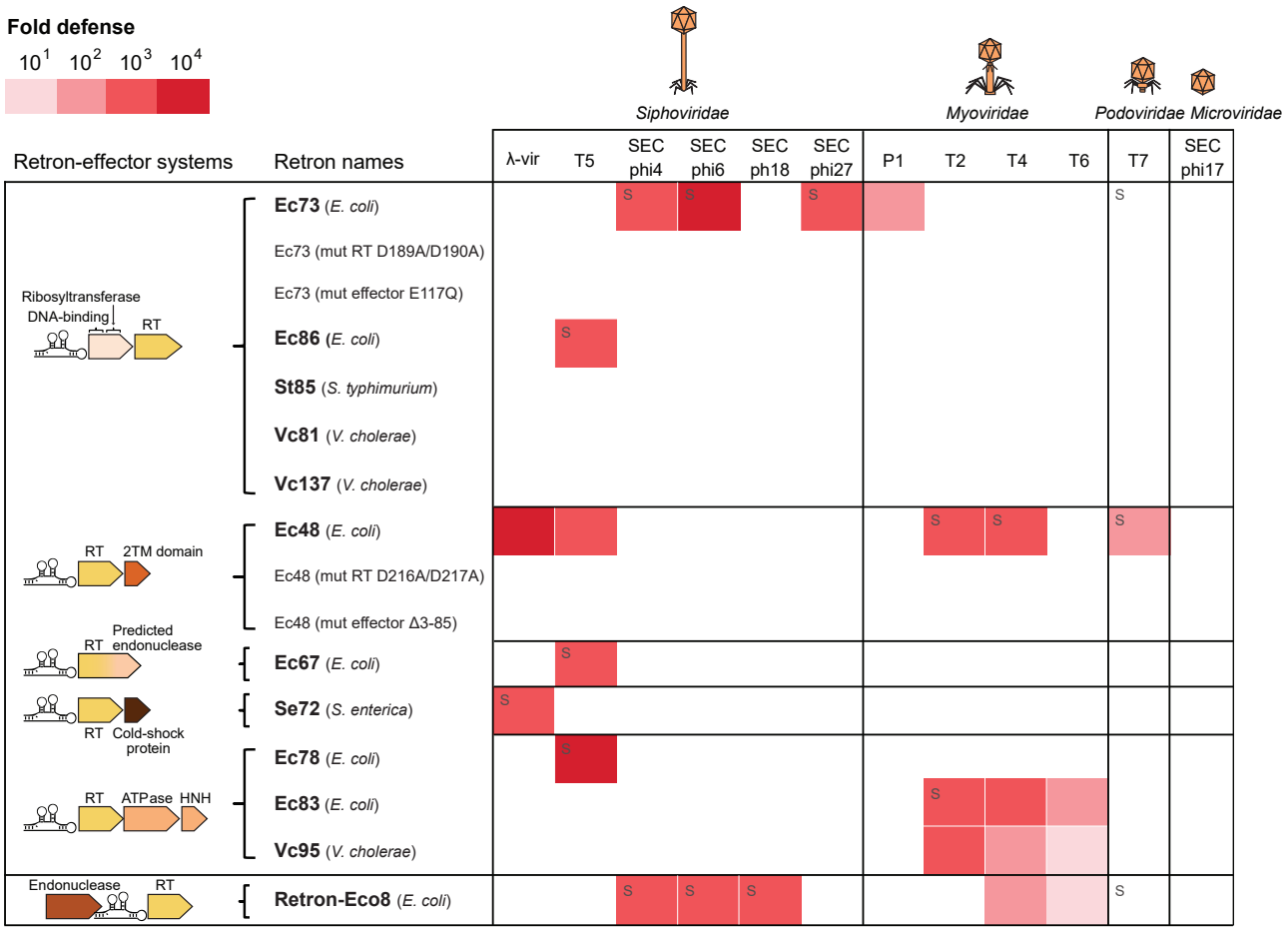


Figure 3. Retrons, Together with Their Effector Genes, Protect against Phages

Fold protection was measured using serial dilution plaque assays, comparing the efficiency of plating (EOP) of phages on the system-containing strain to the EOP on a control strain that lacks the system. Data represent an average of three replicates (Figure S3). A designation of 's' stands for a marked reduction in plaque size. On the left, gene organization of the defense systems, with identified domains indicated. For retrons Ec73 and Ec48, data for wild-type (WT) strains are presented as well as data for strains mutated in either the reverse transcriptase (RT) or the associated gene (effector). The Δ3-85 mutation in the effector gene of Ec48 reflects a deletion of the two predicted transmembrane helices.

See also Figure S3.

diversity, but other types of retrons appear only in a narrow range of bacteria. For example, retrons with predicted cold-shock protein, transmembrane domains (2TM), and ATPase + HNH nuclease effectors are only found in Proteobacteria (Figure S4B).

Retron Anti-Phage Systems Function through Abortive Infection

Some of the retron effector genes include protein domains that are also found in effector proteins of CBASS anti-phage defense systems, which cause the cell to commit suicide once phage infection has been sensed (Cohen et al., 2019). In CBASS, effector proteins that encode transmembrane, endonuclease, and TIR domains are responsible for the cell-killing effect after receiving a signal indicative of phage infection, causing the metabolism of the infected bacteria to arrest before the phage is able to complete its replication cycle, a defense strategy that is generally referred to as abortive infection (Bernheim and Sorek, 2020; Lopatina et al., 2020). This has led us to hypothesize that retrons also function via abortive infection, and the retron effector genes may be responsible for the toxic effect in response to phage infection. In support of this hypothesis, a recent study by Bobonis et al. (2020a) showed that in the *Salmonella* retron St85 the associated effector protein is a toxin, which is inhibited by the activity of the retron RT and msDNA.

If retrons function via abortive infection, it is predicted that infection with a high multiplicity of infection (MOI) (in which nearly all bacteria are infected in the first cycle) would cause cell death or growth arrest in all infected bacteria, even for cells that contain the defense system. To test this hypothesis, we infected retron-containing bacteria with varying MOIs and examined the infection dynamics in liquid culture (Figure 4A). At an MOI of 0.02, cell cultures that do not contain the defense system eventually collapsed due to phage propagation and eventual lysis, whereas retron-containing cultures did not collapse. However, at an MOI of 2, all retron-containing cultures either collapsed (Eco8) or entered a state of growth stasis. These results suggest that

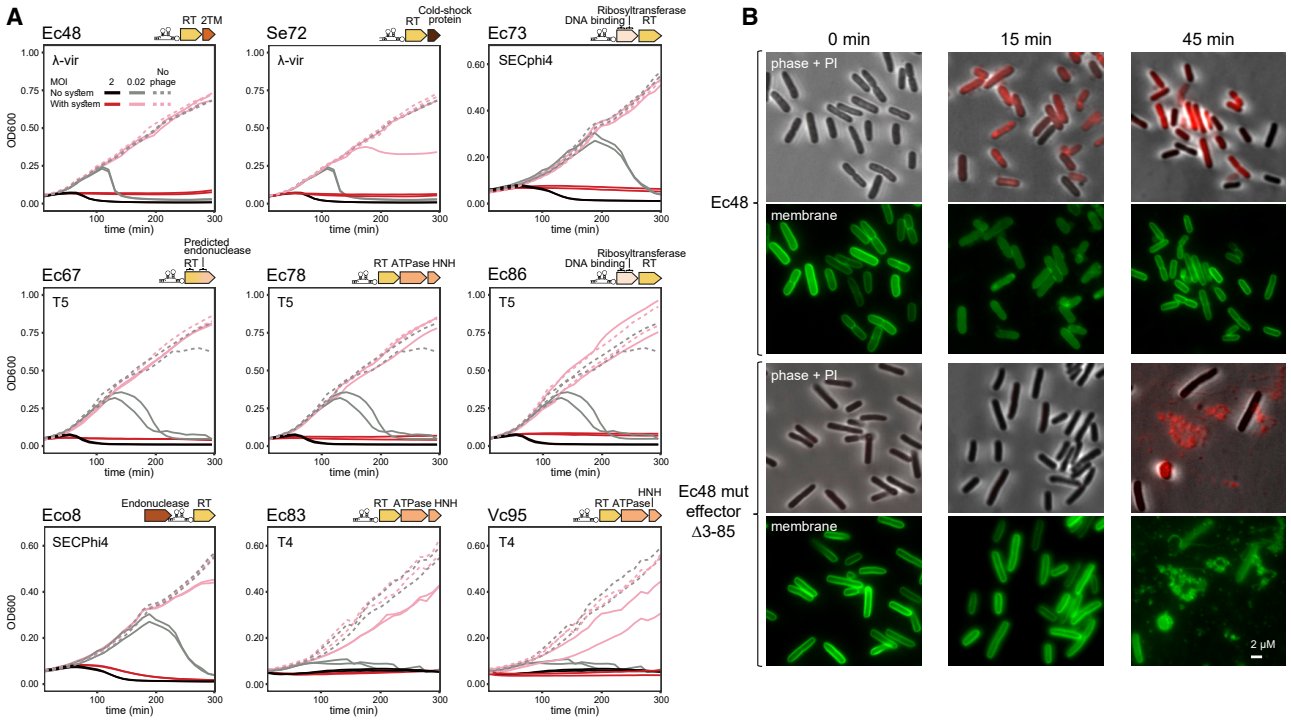


Figure 4. Retron Anti-Phage Systems Function through Abortive Infection

(A) Growth curves in liquid culture for *E. coli* MG1655 bacteria that contain the retron anti-phage system ("with system") and bacteria that lack the retron system ("no system"), infected by one of the following phages: λ-vir, SECphi4, T5, and T4. Bacteria were infected at time 0 at an MOI of 0.02 or 2, and the first measurement of the optical density (OD_{600}) was taken 10 min post infection. For each experiment, two biological replicates are presented as individual curves. All experiments with the same phage were performed simultaneously with the same negative controls, and thus negative control curves are duplicated for visualization purposes.

(B) Fluorescence microscopy images of phase contrast overlay with propidium iodide (PI) (red) and membrane stain (green) captured at 0 min, 15 min, and 45 min after infection with phage λ-vir at an MOI of 3. The two top rows show *E. coli* MG1655 cells containing the Ec48 retron system. The two bottom rows show *E.coli* MG1655 cells containing an inactive Ec48 retron system with deletion of the two predicted transmembrane helices of the effector gene (mut effector Δ3-85), rendering the defense system inactive. Membrane becomes permeable after 15 min in cells containing the Ec48 retron system, but not in cells in which the system is mutated. After 45 min, phage-mediated cell lysis is observed in cells in which the system is mutated. Representative images from a single replicate out of three independent replicates are shown.

retrons, in general, protect against phages via an abortive infection defense strategy.

To gain further insight into the abortive infection process, we focused on retron Ec48, as it provided strong defense against phages belonging to three different families (*Siphoviridae*, *Myoviridae*, and *Podoviridae*) (Figure 3). The Ec48 retron defense system contains an effector gene with two transmembrane-spanning helices (Figure 2C). Such a transmembrane-spanning domain organization is common in effector proteins of CBASS systems, in which it is predicted to impair the membrane integrity, causing the infected bacteria to die before the phage is able to complete its replication cycle (Cohen et al., 2019). We therefore hypothesized that in the process of the abortive infection inflicted by the Ec48 retron defense system, the transmembrane-spanning effector also exerts its toxicity by causing the cell membrane of infected cells to become permeable. To assess this possibility, we examined Ec48-containing cells under the microscope during infection with the λ-vir phage. Cells were stained with a membrane dye and with propidium iodide (PI), a fluorescent DNA-binding agent that penetrates cells only if they lost plasma membrane integrity, thus marking cells with impaired membranes (Figure 4B). Cells harboring a mutated, inactive Ec48 system, where the two predicted transmembrane helices were deleted in the effector gene, were not protected against the phage and exhibited phage-mediated cell lysis 45 min after initial infection. However, cells containing an intact Ec48 retron system became stained with PI already 15 min post infection, indicating that their membranes became permeable before the phage replication cycle has finished. These cells remained stained at 45 min post infection, but did not lyse, confirming that the phages did not complete their replication cycle within these cells.

Inhibition of RecBCD Triggers Ec48

To try and elucidate the mechanism by which the Ec48 retron recognizes and mitigates phage infection, we attempted to find phage mutants that escape retron defense. We were able to isolate six mutants of phage λ-vir, as well as two mutants of phage T7, which could overcome the defense conferred by Ec48 and its effector gene (Figure 5A). We then sequenced the full genome of each of the phage mutants and compared the resulting sequences to the sequence of the

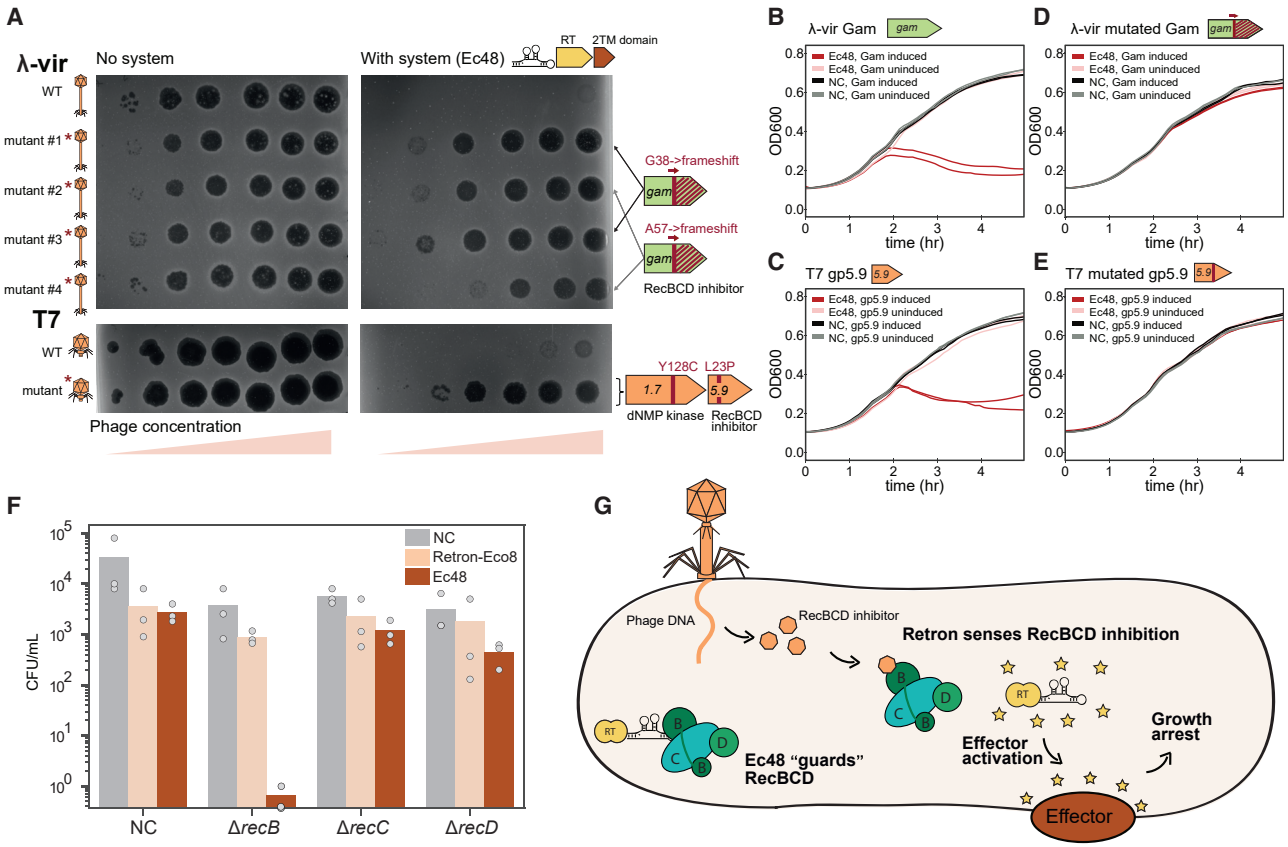


Figure 5. Inhibition of RecBCD Triggers Ec48

(A) Representative phage mutants capable of escaping Ec48 retron system defense. Shown are 10-fold serial dilution plaque assays, comparing the plating efficiency of WT and mutant phages on bacteria that contain the Ec48-retron anti-phage system and a control strain that lacks the system and contains an empty vector instead. Images are representative of three replicates. Genes mutated in mutant phages are presented on the right. λ-vir mutant #1 contains 15 additional mutations as detailed in Table S2.

(B–E) Expression of phage-encoded RecBCD inhibitors activates toxicity in Ec48-containing strains. Growth curves of *E. coli* expressing the Ec48 system or a negative control vector that lacks the system (NC). The expression of Gam protein of phage λ (B), gp5.9 protein of phage T7 (C), or their mutated versions (D and E) was induced by the addition of 0.3% arabinose to exponentially growing cells at OD_{600} 0.3. For each experiment, two biological replicates are presented as individual curves, each a mean of two technical replicates. See also Figure S5.

(F) A vector that contains the Ec48 anti-phage system cannot be transformed into cells lacking RecB. Transformation efficiency of a plasmid containing either the Ec48 retron system, the newly discovered anti-phage retron (Retron-Eco8), or RFP (as a negative control, NC), into *E. coli* strains with a deletion in one of the RecBCD components or a deletion of iscR (an unrelated bacterial gene as a negative control [NC]). The number of colony forming units following transformation is presented. Bar graph represents an average of three replicates, with individual data points overlaid.

(G) Model for the anti-phage activity of the Ec48 retron system.

wild-type (WT) phage. In five of the six retron-overcoming mutants of phage λ-vir, we found a single point mutation that distinguished the mutant phage from the WT phage. In all five cases, the point mutation (either a single base deletion or a single base insertion) resulted in a frameshift in a λ-vir gene called *gam*, and in the sixth mutant, we detected 16 mutations, one of which also causing a frameshift in *gam* (Figure 5A; Table S2). These results suggest that mutations that inactivate the Gam protein of phage λ-vir enable the phage to overcome the defensive activity of the retron Ec48 defense system. In the retron-escaping T7 mutants, we found two missense point mutations appearing in both mutants, one in gene *1.7* (tyrosine at position 128 mutated into cysteine) and the other in gene *5.9* (changing a leucine at position 23 of the protein to proline) (Figure 5A; Table S2).

Remarkably, both the Gam protein of phage λ and the gp5.9 protein of phage T7 have the same biological role: inhibition of a bacterial complex called RecBCD (Lin, 1992; Pacumbaba and Center, 1975; Unger et al., 1972). The RecBCD complex is known to have a central role in DNA repair as well as in anti-phage activity (Cheng et al., 2020; Dillingham and Kowalczykowski, 2008). RecBCD rapidly degrades linear dsDNA and hence phage DNA, which is usually injected into the infected cell in a linear form, is susceptible to RecBCD-mediated degradation (Dillingham and Kowalczykowski, 2008). RecBCD also plays an important role in acquisition of new CRISPR spacers from phage genomes (Levy et al., 2015) and in the formation of guide DNAs for prokaryotic argonaute proteins that defend against foreign nucleic acids (Kuzmenko et al., 2020). Because RecBCD is a central immunity hub in the bacterial cell, numerous phages

are known to encode anti-RecBCD proteins (including the λ Gam and T7 gp5.9 proteins), which bind the RecBCD complex and inactivate it (Bobay et al., 2013; Dillingham and Kowalczykowski, 2008). Our findings that mutations in RecBCD inhibitors of both phage λ-vir and phage T7 enabled the phages to escape defense suggest that the Ec48 retron senses these phage proteins, or their activities, as a signal for phage infection.

The above results suggest that the Ec48 retron defense system somehow recognizes the phage RecBCD inhibitors, and, once recognition occurs, the effector's toxicity becomes activated leading to growth arrest. To test this hypothesis, we expressed the RecBCD inhibitors from phages λ-vir and T7 (Gam and gp5.9, respectively) in cells that contain the Ec48 defense system and followed bacterial growth after induction of expression. We observed a marked decrease in bacterial cell density following the expression of either of the two RecBCD inhibitors in bacteria that contain the Ec48 retron defense system, but not in bacteria lacking Ec48 (Figures 5B and 5C). Expression of the mutated versions of the phage RecBCD inhibitor genes, which enabled them to escape from Ec48 defense, did not affect bacterial growth in Ec48-containing cells (Figures 5D and 5E). To verify that the observed reduction in the bacterial cell density corresponds to cell death, we induced the expression of the T7 RecBCD inhibitor gp5.9 in bacteria that contain the Ec48 defense system and plated bacteria to count colony-forming units (CFUs). Counts of viable bacteria dropped dramatically following induction of T7 gp5.9 in bacteria that contained the retron system, consistent with cell death (Figure S5A). Altogether, these results indicate that once the Ec48 retron defense system senses the presence of phage-encoded RecBCD inhibitors, its toxicity becomes activated and inflicts abortive infection.

Notably gp1.7, the second gene found to be mutated in the Ec48-escaping T7 phages, did not cause growth arrest when expressed in Ec48-containing cells (Figure S5B). This gene is a dNMP kinase that participates in the synthesis of dNTP nucleotides used by the phage for DNA replication (Tran et al., 2010). Perhaps the mutation in gp1.7 allows the phage to escape a different step in the Ec48 defense process, or, alternatively, this mutation may be a passenger mutation not involved in escape from retron defense.

Although the Gam protein of phage λ and the gp5.9 protein of phage T7 both activate the Ec48 retron defense system to induce growth arrest, these proteins share no detectable sequence similarity. We therefore hypothesized that rather than directly sensing the phage proteins, the Ec48-containing system monitors the integrity of RecBCD itself and becomes activated when RecBCD is tampered with. If this hypothesis is correct, then perturbations of RecBCD that are not linked to phage infection would also trigger the Ec48 defense system and activate its toxicity. To test this hypothesis, we used strains of *E. coli* in which components of RecBCD are deleted. Such cells are impaired in DNA repair but are viable in non-stress conditions (Biek and Cohen, 1986; Chaudhury and Smith, 1984). We attempted to transform a plasmid that encodes the Ec48 defense system or a control plasmid (lacking the system) into *E. coli* strains deleted for one of the three components of RecBCD (either Δ*recB*, Δ*recC*, or Δ*recD*) or a control strain. While the control plasmid was transformed with high efficiency into all strains, transformation of the Ec48-containing plasmid into Δ*recB* consistently and repeatedly failed, yielding zero colonies or a single colony in three independent transformation attempts (Figure 5F). The same Ec48-containing plasmid was transformed with high efficiency into the Δ*recC*, Δ*recD*, and the control strains (Figure 5F). These results suggest that the Ec48 system becomes activated specifically when the RecB protein is missing or impaired. Notably, the Gam protein of phage λ has been shown to inhibit RecBCD by directly binding to RecB (Wilkinson et al., 2016), supporting this model.

DISCUSSION

Our results show that retrons broadly function in anti-phage defense, solving a three-decade-old mystery on the function of these peculiar genetic elements. The functional retron anti-phage system includes three components: the RT, the ncRNA, and an effector protein. Accessory open reading frames have recently been predicted as genetically linked to some retrons and were suggested to assist the retron function (Simon et al., 2019). Our data indicate that these accessory genes function as effector proteins essential for the anti-phage activity.

Our experiments with retron Ec48 lead to a mechanistic model in which the Ec48 retron defense system "guards" the normal activity of the bacterial immunity hub RecBCD. Inactivation of the RecB protein by phage inhibitors triggers the retron system and leads to growth arrest before the phage is able to complete its replication cycle (Figure 5G). As the growth arrest function is likely carried out by the effector transmembrane-spanning protein, affecting membrane permeability, we further hypothesize that the role of the reverse-transcribed retron RNA, perhaps together with the RT, is to sense or monitor the integrity of RecB. Notably, the RecB protein has a domain that binds ssDNA, to which the Gam protein has been shown to bind. Perhaps the reverse transcribed ssDNA in the retron binds RecB, and during infection, it is displaced by the phage Gam protein, releasing the retron to somehow activate the effector protein.

It is noteworthy that a plasmid encoding the new retron defense system Retron-Eco8 was readily transformed into all three RecBCD deletion strains, including Δ*recB*, suggesting that the Retron-Eco8 defense system does not guard RecBCD and likely senses phage infection in a different manner (Figure 5F). Therefore, retron-containing defense systems other than Ec48 may sense alternative signals or guard other central components of the cell to mitigate phage infection.

Interestingly, retron systems that have a similar genetic composition defend against different phages. For example, the Ec73 and Ec86 systems, both of which contain a ribosyltransferase effector, protect against a completely different set of phages (Figures 3 and S3). The fact that these systems both encode genes with a similar effector domain, but encode retron ncRNA sequences that are significantly different and cannot be aligned, suggests that the nucleic acid component of the retron defense system and/or the RT participate in the recognition of phage infection. The exact mechanism by which the retron RNA-DNA hybrid takes part in this recognition, and the mechanism by which the signal is transferred to the effector protein to induce

abortive infection, remains to be elucidated. It is possible that the retron ssDNA component attaches to DNA-binding cellular complexes (such as RecBCD in Ec48) or phage proteins (in other retrons) to monitor their activities. It is also possible that the msDNA or some fragments of it physically bind the effector protein and control its activity. Indeed, a recent study has shown that in the St85 retron, the msDNA together with the RT directly interact with the effector protein and curb its toxicity (Bobonis et al., 2020a).

The discovery that retrons broadly function as anti-phage defense systems can shed light on previous studies performed on retrons. For example, msDNA was shown to be important for the survival of pathogenic strains of *Salmonella* Typhimurium and *E. coli* in the host intestine (Elfenbein et al., 2015). In light of our results, a possible explanation for this observation may be that these retrons play a role in defending the bacteria against phages it encounters in the host intestine, enabling the bacteria to better colonize the intestine when the retron is present. Furthermore, two recent studies have found that the retron-Sen2 (St85) functions as a toxin/antitoxin system, with the effector toxin being activated by proteins of phage origin (Bobonis et al., 2020a, 2020b). These studies provide further support to our observations and suggest that St85, too, is an anti-phage system that functions through an abortive infection mechanism.

Retron RTs form a diverse and distinct group within the larger collection of RTs found in microbial genomes. Other classes of RTs (Simon and Zimmerly, 2008), distantly related to retron RTs, such as AbiK and CRISPR-associated RTs, have been previously shown to function in anti-phage defense. Our findings reveal that retron RTs, which are the second largest class of bacterial RTs, also play a role in protection against phage. It remains to be discovered whether other types of RTs participate in anti-phage defense as well. Although most of the retron systems we tested provided defense against phages (Figure 3) and many retrons tend to appear in defense islands, two clades of RT homologs do not follow this pattern and are only rarely found next to known defense systems (Figure 2A). This suggests that some retrons may have adapted to perform functions other than anti-phage defense, perhaps in a manner analogous to toxin-antitoxin systems, some of which are involved in anti-phage defense while others play a role in the bacterial stress response (Harms et al., 2018).

Our results suggest that Ec48 functions as a ''guardian'' of RecBCD, a central cellular hub of anti-phage immunity in *E. coli*. Such a defensive strategy makes immunological sense, because different phages can inhibit RecBCD by multiple different mechanisms (Bobay et al., 2013; Dillingham and Kowalczykowski, 2008). Instead of recognizing the phage inhibitors themselves, which can be highly divergent and share no similarity, Ec48 monitors the integrity of the bacterial immune system. If the integrity of RecBCD is impaired the molecular ''conclusion'' is that the cell has been infected by phage, activating Ec48 as a second line of defense. An analogous strategy is exerted by the PrrC protein in *E. coli*, which similarly guards the normal function of type I restriction-modification systems and inflicts cell-suicide when phage proteins inhibit the restriction enzyme (Kaufmann, 2000). Remarkably, such a defensive strategy has been documented as a central aspect of the immune system of plants, where it was termed ''the guard hypothesis'' (Dangl and Jones, 2001). In plants, pathogen resistance immune proteins are known to guard central cellular processes and detect their disruption by pathogen effectors as a signature for infection (Lopes Fischer et al., 2020). Our discovery that retron Ec48 similarly guards the RecBCD complex in bacteria therefore shows that similar immunological principles govern the design of immune systems across bacteria and plants.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacterial strains and phages
- METHOD DETAILS
  - Detection of RT genes in defense islands
  - Prediction of ncRNA structure
  - Identification and analysis of retron RT homologs
  - Plasmids and strain construction
  - Plaque assays
  - Detection of ncRNA expression using RNA-seq
  - Isolation of msDNA
  - Phage-infection dynamics in liquid medium
  - Microscopy of infected cells
  - Isolation of mutant phages
  - Amplification of mutant phages
  - Sequencing and genome analysis of phage mutants
  - Bacterial growth upon induction of phage genes
  - CFU count experiments
  - Transformation efficiency assay
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.09.065.

ACKNOWLEDGMENTS

We thank the Sorek laboratory members for comments on earlier versions of this manuscript. We also thank Hyeim Jung for providing us with a protocol for msDNA isolation. A.M. was supported by a fellowship from the Ariane de Rothschild Women Doctoral Program and, in part, by the Israeli Council for Higher Education via the Weizmann Data Science Research Center. A.B. is the recipient of a European Molecular Biology Organization (EMBO) Long Term Fellowship (EMBO ALTF 186-2018). R.S. was supported, in part, by the Israel Science Foundation (personal grant 1360/16), the European Research Council (ERC-CoG 681203), the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine, the Minerva Foundation with funding from the Federal German Ministry for Education and Research, and the Knell Family Center for Microbiology.

AUTHOR CONTRIBUTIONS

A.M., A.B., and A.S.-A. led the study and performed all analyses and experiments unless otherwise indicated. A.M. performed the computational analyses. A.B. and A.S.-A. performed the experimental assays and their analysis. Y.O.-S. performed the microscopy analysis that appears in Figure 4B. T.F. assisted with isolation of mutant phages that appear in Figure 5A and Table S2. M.V. performed RNA-seq analysis for *Paenibacillus polymyxa* that appears in Figure 1D. A.L. performed plaque assays that appear in Figure 1C. R.S. supervised the study and wrote the paper together with the team.

DECLARATION OF INTERESTS

R.S. is a scientific cofounder and consultant of BiomX Ltd., Pantheon Ltd., and EcoPhage.

Received: April 16, 2020
Revised: August 4, 2020
Accepted: September 28, 2020
Published: November 5, 2020

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 2006.0008.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. *19*, 455–477.

Baym, M., Kryazhimskiy, S., Lieberman, T.D., Chung, H., Desai, M.M., and Kishony, R. (2015). Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS ONE *10*, e0128036.

Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. Nat. Rev. Microbiol. *18*, 113–119.

Biek, D.P., and Cohen, S.N. (1986). Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in Escherichia coli. J. Bacteriol. *167*, 594–603.

Bobay, L.-M., Touchon, M., and Rocha, E.P.C. (2013). Manipulating or superseding host recombination functions: a dilemma that shapes phage evolvability. PLoS Genet. *9*, e1003825.

Bobonis, J., Mateus, A., Pfalz, B., Garcia-Santamarina, S., Galardini, M., Kobayashi, C., Stein, F., Savitski, M.M., Elfenbein, J.R., Andrews-Poymenis, H., et al. (2020a). Bacterial retrons encode tripartite toxin/antitoxin systems. bioRxiv. https://doi.org/10.1101/2020.06.22.160168.

Bobonis, J., Mitosch, K., Mateus, A., Kritikos, G., Elfenbein, J.R., Savitski, M.M., Andrews-Polymenis, H., and Typas, A. (2020b). Phage proteins block and trigger retron toxin/antitoxin systems. bioRxiv. https://doi.org/10.1101/2020.06.22.160242.

Chaudhury, A.M., and Smith, G.R. (1984). Escherichia coli recBC deletion mutants. J. Bacteriol. *160*, 788–791.

Chen, I.A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., Huntemann, M., Varghese, N., White, J.R., Seshadri, R., et al. (2019). IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic Acids Res. *47* (D1), D666–D677.

Cheng, K., Wilkinson, M., Chaban, Y., and Wigley, D.B. (2020). A conformational switch in response to Chi converts RecBCD from phage destruction to DNA repair. Nat. Struct. Mol. Biol. *27*, 71–77.

Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., and Sorek, R. (2019). Cyclic GMP-AMP signalling protects bacteria against viral infection. Nature *574*, 691–695.

Dangl, J.L., and Jones, J.D. (2001). Plant pathogens and integrated defence responses to infection. Nature *411*, 826–833.

Dar, D., Shamir, M., Mellin, J.R., Koutero, M., Stern-Ginossar, N., Cossart, P., and Sorek, R. (2016). Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. Science *352*, aad9822.

Das, R., Shimamoto, T., Hosen, S.M.Z., and Arifuzzaman, M. (2011). Comparative Study of different msDNA (multicopy single-stranded DNA) structures and phylogenetic comparison of reverse transcriptases (RTs): evidence for vertical inheritance. Bioinformation *7*, 176–179.

Deatherage, D.E., and Barrick, J.E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol. Biol. *1151*, 165–188.

Dhundale, A., Lampson, B., Furuichi, T., Inouye, M., and Inouye, S. (1987). Structure of msDNA from Myxococcus xanthus: evidence for a long, self-annealing RNA precursor for the covalently linked, branched RNA. Cell *51*, 1105–1112.

Dillingham, M.S., and Kowalczykowski, S.C. (2008). RecBCD enzyme and the repair of double-stranded DNA breaks. Microbiol. Mol. Biol. Rev. *72*, 642–671.

Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., and Sorek, R. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. Science *359*, eaar4120.

Doulatov, S., Hodes, A., Dai, L., Mandhana, N., Liu, M., Deora, R., Simons, R.W., Zimmerly, S., and Miller, J.F. (2004). Tropism switching in Bordetella bacteriophage defines a family of diversity-generating retroelements. Nature *431*, 476–481.

Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics *5*, 113.

Elfenbein, J.R., Knodler, L.A., Nakayasu, E.S., Ansong, C., Brewer, H.M., Bogomolnaya, L., Adams, L.G., McClelland, M., Adkins, J.N., and Andrews-Polymenis, H.L. (2015). Multicopy Single-Stranded DNA Directs Intestinal Colonization of Enteric Pathogens. PLoS Genet. *11*, e1005472.

Farzadfard, F., and Lu, T.K. (2014). Synthetic biology. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. Science *346*, 1256272.

Fortier, L.-C., Bouchard, J.D., and Moineau, S. (2005). Expression and site-directed mutagenesis of the lactococcal abortive phage infection protein AbiK. J. Bacteriol. *187*, 3721–3730.

Harms, A., Brodersen, D.E., Mitarai, N., and Gerdes, K. (2018). Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. Mol. Cell *70*, 768–784.

Herzer, P.J. (1996). Starvation-induced expression of retron-Ec107 and the role of ppGpp in multicopy single-stranded DNA production. J. Bacteriol. *178*, 4438–4444.

Hsu, M.Y., Inouye, S., and Inouye, M. (1989). Structural requirements of the RNA precursor for the biosynthesis of the branched RNA-linked multicopy single-stranded DNA of Myxococcus xanthus. J. Biol. Chem. *264*, 6214–6219.

Hsu, M.Y., Eagle, S.G., Inouye, M., and Inouye, S. (1992). Cell-free synthesis of the branched RNA-linked msDNA from retron-Ec67 of Escherichia coli. J. Biol. Chem. *267*, 13823–13829.

Jung, H., Liang, J., Jung, Y., and Lim, D. (2015). Characterization of cell death in Escherichia coli mediated by XseA, a large subunit of exonuclease VII. J. Microbiol. *53*, 820–828.

Kaufmann, G. (2000). Anticodon nucleases. Trends Biochem. Sci. *25*, 70–74.

Kropinski, A.M., Mazzocco, A., Waddell, T.E., Lingohr, E., and Johnson, R.P. (2009). Enumeration of bacteriophages by double agar overlay plaque assay. Methods Mol. Biol. *501*, 69–76.

Kuzmenko, A., Oguienko, A., Esyunina, D., Yudin, D., Petrova, M., Kudinova, A., Maslova, O., Ninova, M., Ryazansky, S., Leach, D., et al. (2020). pAgo-induced DNA interference protects bacteria from invader DNA. bioRxiv. https://doi.org/10.1101/2020.03.01.971358.

Lampson, B.C., Inouye, M., and Inouye, S. (1989). Reverse transcriptase with concomitant ribonuclease H activity in the cell-free synthesis of branched RNA-linked msDNA of Myxococcus xanthus. Cell *56*, 701–707.

Lampson, B., Inouye, M., and Inouye, S. (2001). The msDNAs of bacteria. Prog. Nucleic Acid Res. Mol. Biol. *67*, 65–91.

Lampson, B.C., Inouye, M., and Inouye, S. (2005). Retrons, msDNA, and the bacterial genome. Cytogenet. Genome Res. *110*, 491–499.

Lee, T.S., Krupa, R.A., Zhang, F., Hajimorad, M., Holtz, W.J., Prasad, N., Lee, S.K., and Keasling, J.D. (2011). BglBrick vectors and datasheets: A synthetic biology platform for gene expression. J. Biol. Eng. *5*, 12.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. *44* (W1), W242-5.

Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., and Sorek, R. (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature *520*, 505–510.

Lin, L. (1992). Study of bacteriophage T7 gene 5.9 and gene 5.5. PhD thesis (State University of New York).

Lopatina, A., Tal, N., and Sorek, R. (2020). Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. Annu. Rev. Virol. *7*, 371–384.

Lopes Fischer, N., Naseer, N., Shin, S., and Brodsky, I.E. (2020). Effector-triggered immunity and pathogen sensing in metazoans. Nat. Microbiol. *5*, 14–26.

Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol. Biol. *6*, 26.

Makarova, K.S., Wolf, Y.I., Snir, S., and Koonin, E.V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. J. Bacteriol. *193*, 6039–6056.

Mazzocco, A., Waddell, T.E., Lingohr, E., and Johnson, R.P. (2009). Enumeration of bacteriophages using the small drop plaque assay system. Methods Mol. Biol. *501*, 81–85.

Pacumbaba, R., and Center, M.S. (1975). Partial purification and properties of a bacteriophage T7 inhibitor of the host exonuclease V activity. J. Virol. *16*, 1200–1207.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. *26*, 1641–1650.

Rice, S.A., and Lampson, B.C. (1995). Bacterial reverse transcriptase and msDNA. Virus Genes *11*, 95–104.

Schiltz, C.J., Adams, M.C., and Chappie, J.S. (2020). The full-length structure of Thermus scotoductus OLD defines the ATP hydrolysis properties and catalytic mechanism of Class 1 OLD family nucleases. Nucleic Acids Res. *48*, 2762–2776.

Sharon, E., Chen, S.-A.A., Khosla, N.M., Smith, J.D., Pritchard, J.K., and Fraser, H.B. (2018). Functional Genetic Variants Revealed by Massively Parallel Precise Genome Editing. Cell *175*, 544–557.

Sikowitz, M.D., Cooper, L.E., Begley, T.P., Kaminski, P.A., and Ealick, S.E. (2013). Reversal of the substrate specificity of CMP N-glycosidase to dCMP. Biochemistry *52*, 4037–4047.

Silas, S., Mohr, G., Sidote, D.J., Markham, L.M., Sanchez-Amat, A., Bhaya, D., Lambowitz, A.M., and Fire, A.Z. (2016). Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. Science *351*, aad4234.

Simon, D.M., and Zimmerly, S. (2008). A diversity of uncharacterized reverse transcriptases in bacteria. Nucleic Acids Res. *36*, 7219–7229.

Simon, A.J., Morrow, B.R., and Ellington, A.D. (2018). Retroelement-Based Genome Editing and Evolution. ACS Synth. Biol. *7*, 2600–2611.

Simon, A.J., Ellington, A.D., and Finkelstein, I.J. (2019). Retrons and their applications in genome engineering. Nucleic Acids Res. *47*, 11007–11019.

Söding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. *33*, W244-8.

Solovyev, V., and Salamov, A. (2011). Automatic Annotation of Microbial Genomes and Metagenomic Sequences. Metagenomics and Its Applications in Agriculture, Biomedicine, and Environmental Studies (Nova Science Publishers), pp. 61–78.

Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. *35*, 1026–1028.

Torarinsson, E., and Lindgreen, S. (2008). WAR: Webserver for aligning structural RNAs. Nucleic Acids Res. *36*, W79–W84.

Tran, N.Q., Lee, S.J., Richardson, C.C., and Tabor, S. (2010). A novel nucleotide kinase encoded by gene 1.7 of bacteriophage T7. Mol. Microbiol. *77*, 492–504.

Unger, R.C., Echols, H., and Clark, A.J. (1972). Interaction of the recombination pathways of bacteriophage lambda and host Escherichia coli: effects on lambda recombination. J. Mol. Biol. *70*, 531–537.

Voichek, M., Maaß, S., Kroniger, T., Becher, D., and Sorek, R. (2020). Peptide-based quorum sensing systems in *Paenibacillus polymyxa*. Life Sci. Alliance *3*, e202000847.

Wang, C., Villion, M., Semper, C., Coros, C., Moineau, S., and Zimmerly, S. (2011). A reverse transcriptase-related protein mediates phage resistance and polymerizes untemplated DNA in vitro. Nucleic Acids Res. *39*, 7620–7629.

Wilkinson, M., Troman, L., Wan Nur Ismah, W.A., Chaban, Y., Avison, M.B., Dillingham, M.S., and Wigley, D.B. (2016). Structural basis for the inhibition of RecBCD by Gam and its synergistic antibacterial effect with quinolones. eLife *5*, e22963.

Yee, T., Furuichi, T., Inouye, S., and Inouye, M. (1984). Multicopy single-stranded DNA isolated from a gram-negative bacterium, Myxococcus xanthus. Cell *38*, 203–209.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
E. coli K-12 MG1655	American Type Culture Collection (ATCC)	ATCC 47076
NEB 5-alpha Competent *E. coli*	NEW ENGLAND BioLabs	Cat#C2987H
E. coli BW25141 Δ*iscR*	Keio collection; Baba et al., 2006	N/A
E. coli BW25141 Δ*recB*	Keio collection; Baba et al., 2006	N/A
E. coli BW25141 Δ*recC*	Keio collection; Baba et al., 2006	N/A
E. coli BW25141 Δ*recD*	Keio collection; Baba et al., 2006	N/A
Phage Lambda-vir	U. Qimron	N/A
Phage SECPhi17	Doron et al., 2018	N/A
Phage SECphi18	Doron et al., 2018	N/A
Phage SECphi27	Doron et al., 2018	N/A
Phage SECphi6	This paper	N/A
Phage SECphi4	This paper	N/A
Phage P1	U. Qimron	N/A
Phage T2	German Collection of Microorganisms and Cell Cultures GmbH (DSMZ)	DSM 16352
Phage T4	U. Qimron	N/A
Phage T5	U. Qimron	N/A
Phage T6	German Collection of Microorganisms and Cell Cultures GmbH (DSMZ)	DSM 4622
Phage T7	U. Qimron	N/A
Chemicals, Peptides, and Recombinant Proteins		
membrane stain FM1-43	Thermo Fisher Scientific	Cat#T35356
propidium iodide	Sigma-Aldrich	Cat#P4170
Critical Commercial Assays		
DNeasy blood and tissue kit	QIAGEN	Cat#69504
Nextera DNA Library Prep Kit	Illumina	Cat#15027865, Cat#15027866
RNase A (within QIAprep Spin Miniprep Kit)	QIAGEN	Cat#27106
DNase-I	Merck	Cat#11284932001
Deposited Data		
Phage SECphi4 genome sequence	This paper	GenBank: MT331608
Phage SECphi6 genome sequence	This paper	ENA: CADCZA010000001.1
Oligonucleotides		
Primers, see Table S3	This paper	N/A
Recombinant DNA		
pBbS8k-RFP	Lee et al., 2011	Addgene Cat#35276
pSG1-RFP	Doron et al., 2018	N/A
pBad/His A	Thermo Fisher Scientific	Cat#43001
pAA1-pAA22, pAA48-pAA53, see Table S4	This paper, Genscript Corp.	N/A
pAA23, pAA56-pAA74, see Table S5	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
The Integrated Microbial Genomes (IMG)	Chen et al., 2019	https://img.jgi.doe.gov/m/
MMseqs2	Steinegger and Söding, 2017	https://github.com/soedinglab/mmseqs2
RNAfold WebServer	Lorenz et al., 2011	http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi
BLASTclust from the BLAST NCBI package	Altschul et al., 1990	https://ftp.ncbi.nlm.nih.gov/blast/executables/legacy.NOTSUPPORTED/2.2.26/
WAR web server	Torarinsson and Lindgreen, 2008	https://rth.dk/resources/war
HHpred	Söding et al., 2005	https://toolkit.tuebingen.mpg.de/tools/hhpred
MUSCLE (v3.8.1551)	Edgar, 2004	http://www.drive5.com/muscle
FastTree	Price et al., 2009	http://microbesonline.org/fasttree
iTOL	Letunic and Bork, 2016	https://itol.embl.de/
BPROM	Solovyev and Salamov, 2011	http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb
Zen software version 2.0	Zeiss	https://www.zeiss.com/microscopy/int/products/microscope-software/zen.html
Breseq (version 0.29.0)	Deatherage and Barrick, 2014	http://barricklab.org/twiki/bin/view/Lab/ToolsBacterialGenomeResequencing

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to the Lead Contact, Rotem Sorek (rotem.sorek@weizmann.ac.il).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

Phage SECphi4 assembled genome sequence was deposited in GenBank under the accession number GenBank: MT331608. Phage SECphi6 assembled genome sequence was deposited in the European Nucleotide Archive (ENA) under accession number ENA: CADCZA010000001.1.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial strains and phages

Escherichia coli strains (MG1655, NEB 5-alpha, and Δ*iscR*, Δ*recB*, Δ*recC* and Δ*recD* from the Keio collection (Baba et al., 2006) were grown in LB or LB agar at 37°C shaking at 200 rpm unless mentioned otherwise. Whenever applicable, media were supplemented with ampicillin (100 $\mu g ml^{-1}$) or kanamycin (50 $\mu g ml^{-1}$) to ensure the maintenance of plasmids.

The phages used in this study are listed in the Key Resources Table. Infection was performed in LB medium supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ at RT or 37°C as detailed in Method Details. Phage SECphi4 and SECphi6 were isolated from sewage samples as previously described (Doron et al., 2018) on *E. coli* MG1655. Phage DNA was extracted using the QIAGEN DNeasy blood and tissue kit (cat #69504) and DNA libraries were prepared using a modified Nextera protocol (Baym et al., 2015) for Illumina sequencing. Phage DNA was assembled from sequenced reads using SPAdes v. 3.10.1 (with the –careful and –cov-cutoff auto modifiers) (Bankevich et al., 2012). The assembled genomes were deposited for SECphi4 in GenBank under the accession number MT331608, and for SECphi6 in the European Nucleotide Archive (ENA) under accession CADCZA010000001.1. Classification of the phage family was performed according to the most closely related known phage based on sequence similarity, as described previously (Doron et al., 2018). The sequence accession numbers on NCBI for the rest of the phages used in this study are: λ -vir NC_001416.1, SECPhi17 LT960607.1, SECphi18 LT960609.1, SECphi27 LT961732.1, P1 AF234172.1, T2 LC348380.1, T4 AF158101.6, T5 AY543070.1, T6 MH550421.1, T7 NC_001604.1.

METHOD DETAILS

Detection of RT genes in defense islands

Protein sequences of all genes in 38,167 bacterial and archaeal genomes were downloaded from the Integrated Microbial Genomes (IMG) database (Chen et al., 2019) in October 2017. These proteins were filtered for redundancy using the 'clusthash' option of MMseqs2 (release 2-1c7a89) (Steinegger and Söding, 2017) using the '–min-seq-id 0.9' parameter and then clustered using the 'cluster' option, with default parameters. Each cluster with > 10 genes was annotated with the most common pfam, COG, and product annotations in the cluster. For each cluster annotated as reverse transcriptase a defense score was calculated as previously described (Doron et al., 2018), recording the fraction of genes in each cluster that have known defense genes in their genomic environment spanning 10 genes upstream and downstream the inspected gene. Members of clusters with high tendency to be associated with known defense genes were scanned for conserved gene cassettes as previously described (Doron et al., 2018).

Prediction of ncRNA structure

The upstream boundaries of Retron-Eco8 ncRNA were estimated based on a promoter prediction upstream to the RT of *E. coli* 200499 using BPROM (Solovyev and Salamov, 2011). The start position was chosen 10nt downstream to the predicted −10 sequence. The last position in the intergenic region before the RT gene was taken as the end point. The resulting predicted ncRNA corresponds to coordinates 369202-369366 (positive strand) in the *E. coli* 200499 assembly, scaffold accession number Ga0119705_103 in the IMG database (Chen et al., 2019).

To predict the structure of the Retron-Eco8 ncRNA the extracted sequence was folded using the RNAfold web server (Lorenz et al., 2011). For the alignment and structure prediction appearing in Figure 1E, the 250nt upstream of the RTs of homologs of the system were extracted. The sequences were filtered for redundancy using blastclust (with options -S 100 -L 1) (Altschul et al., 1990). The remaining sequences were analyzed using WAR web server (Torarinsson and Lindgreen, 2008) and the MAFFT+RNAalifold representative results are shown.

Identification and analysis of retron RT homologs

A list of RTs of experimentally verified, known retrons (Mx162, Sa163, Mx65, Ec67, Ec48, Ec86, Vc137, Ec73, Vc81, St85, Ec78, Ec83, Yf79, Vc95, Vp96, Ec107, Se72) together with the RT of Retron-Eco8 were searched against the downloaded database of protein sequences from 38,167 bacterial and archaeal genomes using the 'search' option of MMseqs2 (release 6-f5a1c). Hits were mapped to the clusters mentioned above, and protein clusters with at least 10 hits were taken as homologs. To identify the various types of effector genes associated with the RT, the genomic environments spanning 5 genes upstream and downstream of each of the homologs were searched to identify conserved gene cassettes. The identity of the associated gene was determined by analysis of the protein sequence using HHpred (Söding et al., 2005), the Pfam and COG annotations in the IMG database (Chen et al., 2019) and the cluster the gene belongs to (see Methods section ''Detection of RT genes in defense islands'').

To generate the phylogenetic tree in Figure 2, sequences shorter than 200 aa were discarded, as these are likely cases in which the protein has been truncated. The 'clusthash' option of MMseqs2 (release 6-f5a1c) was then used to remove protein redundancies (using the '–min-seq-id 0.9' parameter). Six sequences of group II intron reverse transcriptase (NCBI accession numbers: WP_138709943, WP_137561357, WP_131190414, WP_130644660, WP_130706420, WP_130793790) were added and were used as outgroup. Sequences were aligned using MUSCLE (v3.8.1551) (Edgar, 2004) with default parameters. The FastTree (Price et al., 2009) software was used to generate a tree from the multiple sequence alignment using default parameters. The iTOL (Letunic and Bork, 2016) software was used for tree visualization. A defense score was calculated for each clade as described above, while removing the effector genes from the positive set to avoid artificial inflation of the scores.

Plasmids and strain construction

Primers used in this study are shown in Table S3. Plasmids built for this study and the process of their construction are presented in Tables S4 and S5. Retron systems and their mutants (pAA1-pAA22, pAA48-pAA53) were synthetized and cloned in plasmid pSG1-RFP (between the AscI and NotI sites of the multiple cloning sites) by Genscript Corp. as previously described (Doron et al., 2018) (Table S4). Plasmids used for msDNA isolation, pAA23 and pAA56, were constructed through Gibson assemblies that were first transformed into *E. coli* NEB 5-alpha and then into *E. coli* MG1655 (Table S5).

Plasmids pAA63, pAA64, pAA69, pAA73 and pAA74 used for the expression of phage genes in *E. coli*, were constructed through Gibson assemblies presented in Table S5. Plasmid pBbS8k-RFP (obtained from Addgene, Plasmid #35276 (Lee et al., 2011)) was used as vector and amplified by primers AS_328, AS_329. All phage genes (Table S5, the ''Insert'' rubrics) were cloned in the same position, under the control of the pBad inducible promoter. Phage genes were amplified by PCR on phage DNA from WT or mutant phage (Table S5) using the primers listed in Table S3. After Gibson assembly, these plasmids expressing phage genes were first transformed into *E. coli* NEB 5-alpha. As some toxicity could be expected, 1% glucose was added to the recovery media and in the selecting plates to avoid leaky expression of potentially toxic genes. Plasmids were then transformed into the relevant strains of *E. coli* MG1655 (with or without Ec48). 1% glucose was added also here to the recovery media and in the selecting plates to avoid leaky expression of potentially toxic genes.

Plaque assays

Phages were propagated by picking a single phage plaque into a liquid culture of *E. coli* MG1655 grown at 37°C to OD_{600} 0.3 in LB medium supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ until culture collapse. The culture was then centrifuged for 10 minutes at 4000 r.p.m and the supernatant was filtered through a 0.2 μM filter to get rid of remaining bacteria and bacterial debris. Lysate titer was determined using the small drop plaque assay method as described in Mazzocco et al. (2009).

Plaque assays were performed as previously described (Doron et al., 2018; Mazzocco et al., 2009). Bacteria (*E. coli* MG1655 with plasmids pAA1-pAA53) or negative control (*E. coli* MG1655 with a plasmid pSG1-RFP) were grown overnight at 37°C. Then 300 μl of the bacterial culture was mixed with 30 mL melted MMB agar (LB + 0.1 mM $MnCl_2$ + 5 mM $MgCl_2$ + 0.5% agar) and let to dry for 1 hour at room temperature. 10-fold serial dilutions in MMB were performed for each of the 12 tested phages and 10 μl drops were put on the bacterial layer. Plates were incubated overnight at 25°C (for phages SECphi4, SECphi6, SECphi18, SECphi27, SECphi17, and T7) or 37°C (for phages λ-vir, P1, T2, T4, T5, and T6). Efficiency of plating (EOP) was measured and compared between the defense and control strain.

Detection of ncRNA expression using RNA-seq

E. coli MG1655 cells containing plasmid pAA1 were diluted 1:100 in 5 mL LB medium supplemented with antibiotics (ampicillin). These cells were grown at 37°C with shaking at 250 rpm to an OD_{600} of 0.6. Samples were centrifuged for 10 minutes at 4000 r.p.m at 4°C. The supernatant was discarded, and pellets were used for RNA extraction.

Bacterial pellets were treated with 100 μL of 2 mg/ml lysozyme using Tris 10 mM EDTA 1 mM pH 8.0 as a buffer. Samples were incubated at 37°C for 5 minutes. 1 mL of TRI-reagent (Sigma-Aldrich, 93289) was added to each sample. Samples were then vortexed for 10 s to promote lysis before addition of 200 μL chloroform. Following another vortexing step, the samples were left at room temperature for 5 minutes to allow phase separation. Samples were then centrifuged at 12,000 g at 4°C for 15 minutes. The upper phase was collected and 500 μL of isopropanol was added. Samples were then incubated overnight at −20°C. The next day, samples were washed. Following 30 minutes of centrifugation at 12,000 g at 4°C, the supernatant was discarded leaving a small pellet. 750 μL of ice cold 70% ethanol was added without disturbing the pellet and the samples were then centrifuged for 10 minutes at 12,000 g at 4°C. This wash was repeated once after which the remaining ethanol was discarded and the tubes were left open for 5 minutes to remove residual ethanol. The pellet was then resuspended in 50 μL water and incubated for 10 minutes at 56°C to promote elution. RNA concentrations were measured using Nanodrop.

All RNA samples were treated with TURBO DNase (Life technologies, AM2238). Ribosomal RNA depletion and RNA-seq libraries were prepared as described in Dar et al. (2016), except that all reaction volumes were reduced by a factor of 4. RNA-seq libraries were sequenced using Illumina NextSeq platform. Reads were mapped as described in Dar et al. (2016) to the reference genome of *E. coli* MG1655 (GenBank accession number U00096.3) as well as to the sequence of plasmid pAA1 (Table S4). For *Paenibacillus polymyxa* ATCC 842, reads were found in the European Nucleotide Database (ENA), study accession number PRJEB34369, and mapped to the reference genome of *Paenibacillus polymyxa* ATCC 842 (genome ID 2547132099 in the IMG database) (Chen et al., 2019).

Isolation of msDNA

msDNA isolation was performed as described in Jung et al. (2015). Despite high expression of ncRNA, detection of msDNA from *E. coli* requires the overexpression of the RT and cognate ncRNA. For this, plasmids pAA23 and pAA56 were designed to encode the RT and ncRNA of Retron-Eco8 or retron Ec78, respectively, under the control of an inducible pBad promoter (pBad/His A, Thermofisher, Catalog number 43001) (Table S5). Retron Ec78 was chosen as a positive control because of the relatively high expression of its msDNA as compared to other *E. coli* retrons (Jung et al., 2015). The same plasmid encoding a GFP instead of the retron was used as negative control.

E. coli MG1655 cells containing plasmids pAA23, pAA56 or the negative control were diluted 1:100 in 25 mL LB medium supplemented with antibiotics (100 μg/ml ampicillin). After 30 min of growth at 37°C with shaking of 250 r.p.m, expression of the retron was induced with 0.2% arabinose (final concentration) and grown to an OD_{600} of 0.7. Samples were centrifuged for 10 minutes at 4000 r.p.m at 4°C, after which the supernatant was discarded.

To isolate the msDNA, each pellet was treated with three solutions: Solution I (50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl, pH 8.0), Solution II (0.2 N NaOH, 1% SDS) and Solution III (3 M potassium acetate, 2 M acetic acid). 200 μL of Solution I was added to each pellet followed by vortex. Then, 200 μL of Solution II was added and the tubes were mixed by inverting each tube 5 times. 400 μL of Solution III was then added and samples were mixed again by inverting the tubes 5 times. Samples were centrifuged for 10 minutes at 4000 r.p.m at 4°C, after which the supernatant was collected.

Nucleic acids were then precipitated using ethanol. 2.1 mL of 100% ethanol was added to 700 μL of the collected supernatant to allow precipitation. Samples were centrifuged at 4°C, 13,000 r.p.m for 30 minutes and the supernatant was discarded. To wash the pellets, 2.1 mL of 70% ethanol was added, without disturbing the pellet, and the samples were centrifuged at 4°C, 13,000 r.p.m for 10 minutes. Supernatant was discarded and samples were dried for 20 minutes before resuspension in 20 μL water.

An RNase A treatment was then applied. 2ul of RNase A 10 μg/μl (from QIAGEN Miniprep kit Cat No./ID: 27106) was added to each sample. Samples were incubated at 37°C for 20 minutes. To facilitate visualization on a gel, the ssDNA ladder was diluted 1:100 and the positive control (pAA56, Ec78) was diluted 1:10. 10ul of each of these samples was then mixed with 2X TBE urea loading dye and incubated at 70°C for 3 minutes. After 5 minutes of cooling down, samples were loaded on a 10% denaturing polyacrylamide gel and

run for 1h30 minutes at 180V. For visualization, the gel was washed once with TBE, then incubated for 10 minutes in 10% acetic acid, followed by a 20 minutes incubation in an ethidium bromide bath for staining.

Phage-infection dynamics in liquid medium

Overnight cultures of bacteria (*E. coli* MG1655 with plasmids pAA1-pAA20) or negative control (*E. coli* MG1655 with a an empty pSG1 plasmid) were diluted 1:100 in MMB medium supplemented with ampicillin and incubated at 37 °C while shaking at 200 rpm until early log phase (OD_{600} = 0.3). 180 μl of the diluted culture were transferred into wells in a 96-well plate containing 20 μL of phage lysate for a final MOI of 2 or 0.02, or 20 μL of phage buffer (50 mM Tris pH 7.4, 100 mM MgCl2, 10 mM NaCl) for uninfected control. Infections were performed in duplicate from overnight cultures prepared from two separate colonies. Plates were incubated at 37 °C (for phages λ-vir and T5) or 25°C (for phages T4 and SECPhi4) with shaking in a TECAN Infinite200 plate reader and OD_{600} was followed with measurement every 10 min. For plotting OD_{600}, a value of 0.09 was subtracted from all y axis values to remove the optical density obtained for blank growth media.

Microscopy of infected cells

E. coli MG1655 cells that contain the Ec48 retron system (plasmid pAA10) or the same system with a deletion of the two predicted transmembrane helices of the effector gene (plasmid pAA50) were grown in MMB medium supplemented with ampicillin and 0.2% maltose at 37 °C 200 rpm When growth reached an OD_{600} of 0.3, bacteria were infected with λ-vir phage (MOI of 3). 500 μl of the infected samples were centrifuged at 10,000 *g* for 2 min at 25 °C and resuspended in 5 μL of 1 × Dulbecco's phosphate-buffered saline (DPBS) (Thermo Fisher Scientific 14200075), supplemented with 1 μg/ml membrane stain FM1-43 (Thermo Fisher Scientific T35356) and 0.2 μg/ml propidium iodide (PI) (Sigma-Aldrich P4170). Cells were visualized and photographed using an Axioplan2 microscope (ZEISS) equipped with ORCA Flash 4.0 camera (HAMAMATSU). System control and image processing were carried out using Zen software version 2.0 (Zeiss).

Isolation of mutant phages

To isolate mutant phages that escape Ec48 defense, phages were plated on bacteria expressing the Ec48 defense system (*E. coli* MG1655 with plasmid pAA10, Table S4) using the double-layer plaque assay (Kropinski et al., 2009). MMB 1.1% agar was used as the bottom layer, and MMB 0.3% or 0.5% agar was used for the top layer (for λ-Vir and T7, respectively). For the λ-vir phage, the double-layer plates were incubated overnight at 37°C, and 10 single plaques were picked into 90 μl phage buffer. For phage T7, the double-layer plate was incubated overnight at room temperature and the entire top layer was scraped into 2 mL of phage buffer to enrich for phages that escape Ec48 defense. The phages were left for 1 h at room temperature during which the phages were mixed several times by vortex to release them from the agar into the phage buffer, after which the phages were centrifuged at 3200 g for 10 min to get rid of agar and bacterial cells, and the supernatant was transferred to a new tube.

To test the phages for the ability to escape from Ec48 defense, the small drop plaque assay was used (Mazzocco et al., 2009). 300 μl bacteria - Ec48 (*E. coli* MG1655 with plasmid pAA10) or negative control (*E. coli* MG1655 with a plasmid pSG1, lacking the Ec48 defense system) were mixed with 30 mL melted MMB 0.3% or 0.5% agar (for λ-Vir and T7, respectively) and let to dry for 1 hour at room temperature. 10-fold serial dilutions in phage buffer was performed for the ancestor phages (WT phage used for the original double layer plaque assay) and the phages formed on Ec48 and 10 μl drops were put on the bacterial layer. The plates were incubated overnight at 37°C (for λ-vir phage) or at room temperature (for T7 phage). Efficiency of plating (EOP) was measured and compared between the defense and NC strain.

Amplification of mutant phages

Isolated phages for which there was decreased defense compared to the ancestor phage were further propagated by picking a single plaque formed on Ec48 in the small drop plaque assay into a liquid culture of Ec48 cells grown in 1 mL MMB to an OD_{600} of 0.3. The phages were incubated with the bacteria at 37°C 200 r.p.m for ~3 hr, and then an additional 9 mL of bacterial culture grown to OD_{600} 0.3 in MMB was added, and incubated for an additional ~3 h at 37°C with shaking at 200 rpm The lysate was then centrifuged at 3200 g for 10 min and the supernatant was filtered through a 0.2 μM filter to get rid of remaining bacteria.

Phage titer was then checked using the small drop plaque assay on the negative control strain and in cases where the titer was less than 10^7 pfu/ml, the phage titer was raised using either propagation in liquid culture (see plaque assays section above for detailed protocol) or plate lysate (Kropinski et al., 2009) on Ec48 cells. For plate lysate, 100 μl of phage lysate containing 10^3-10^5 plaque forming units (pfu) was mixed with 100 μl Ec48 cells and left at room temperature for 10 minutes, after which 5 mL of pre-melted 0.3% MMB was added and then poured onto a bottom layer of MMB. The phages were incubated with the bacteria overnight at 37°C (for λ-vir phage) or at room temperature (for T7 phage), letting the phage to lyse the cells. Then the entire top layer was scraped into 5 mL of phage buffer, and for 1 h at room temperature during which the phages were mixed several times by vortex to release them from the agar into the phage buffer. The phages were centrifuged at 3200 g for 10 min to get rid of agar and bacterial cells, and the supernatant was filtered through a 0.2 μM filter.

Sequencing and genome analysis of phage mutants

High titer phage lysates (> 10^7 pfu/ml) of the ancestor and isolated phage mutants were used for DNA extraction. 500 μl of the phage lysate was treated with DNase-I (Merck cat #11284932001) added to a final concentration of 20 μg/ml and incubated at 37°C for 1

hour to remove bacterial DNA. DNA was extracted using the QIAGEN DNeasy blood and tissue kit (cat #69504) starting from the Proteinase-K treatment step to lyse the phages. Libraries were prepared for Illumina sequencing using a modified Nextera protocol as previously described (Baym et al., 2015). Reads were aligned to the phage reference genomes (GenBank accession numbers: NC_001416.1 (λ phage), NC_001604.1 (T7 phage)) and mutations compared to the reference genome were identified using Breseq (version 0.29.0) with default parameters (Deatherage and Barrick, 2014). Only mutations that occurred in the isolated mutants, but not in the ancestor phage, were considered. Silent mutations within protein coding regions were disregarded as well.

Bacterial growth upon induction of phage genes

A single fresh colony was picked into 5 mL of LB supplemented with ampicillin (100 $\mu g ml^{-1}$), kanamycin (50 $\mu g ml^{-1}$) and 1% glucose to avoid leaky expression from the pBad promoter, and grown at 37°C overnight. Bacteria were diluted 1:50 into 3 mL fresh media (LB with ampicillin, kanamycin and 1% glucose) and grown at 37°C until the bacterial cultures reached an OD_{600} of 0.1.

For the induction of T7 gp5.9 and λ-vir Gam, after reaching an OD_{600} of 0.1, 180 μl of the bacteria were then dispensed into a 96-well plate and incubated at 37°C with shaking in a TECAN Infinite200 plate reader with OD_{600} measurement every 6 minutes. When bacteria reached an OD_{600} of 0.3, 20 μl of ultra-pure water (for uninduced samples) or 3% arabinose (for induced samples, for a final concentration of 0.3%), was added to the bacterial cultures and the bacteria were then incubated further at 37°C with shaking in a TECAN Infinite200 plate reader with OD_{600} measurement every 6 minutes.

For the induction of T7 gp1.7, after reaching an OD600 of 0.1, the bacteria were centrifuged at 3200 g 4°C for 10 min, the supernatant was removed and the bacteria were resuspended in 1 mL fresh LB to wash away remaining glucose (for optimal induction of gp1.7 under the pBad promoter). The bacteria were then centrifuged again, after which the supernatant was discarded and the bacteria were resuspended in 10% of the initial volume (300 μl) of LB. 20 μl of the concentrated bacteria were then transferred to a 96-well plate containing 180 μl LB supplemented with ultra-pure water (for uninduced samples) or 0.3% arabinose (for induced samples). The plate was then incubated at 37°C with shaking in a TECAN Infinite200 plate reader with OD_{600} measurement every 6 minutes.

CFU count experiments

For CFU count experiments, bacteria were grown overnight at 37°C in 5 mL of LB supplemented with ampicillin (100 $\mu g ml^{-1}$), kanamycin (50 $\mu g ml^{-1}$) and 1% glucose. Bacteria were diluted 1:50 into fresh media (as above), incubated at 37°C with shaking at 200 r.p.m and OD was monitored every ~30 minutes. When bacteria reached an OD_{600} of 0.3, 3 mL of the 6 mL culture was transferred to a new tube, and 333 μl of either ultra-pure water or 3% arabinose was added to each of the tubes (for uninduced and induced respectively). At several time points throughout the experiment, samples were taken for CFU counts. For this, 10 μl of the bacterial culture was taken into 90 μl LB, and 10 μl of serial dilutions of the bacteria in LB were dropped on LB agar plates supplemented with ampicillin (100 $\mu g ml^{-1}$), kanamycin (50 $\mu g ml^{-1}$) and 1% glucose. The plates were then tilted to allow the drops to spread into a line on the plate, and incubated overnight at 37°C after which single colonies were counted.

Transformation efficiency assay

Strains from the *E. coli* Keio collection (Baba et al., 2006) (Δ*recB*, Δ*recC*, Δ*recD* and Δ*iscR* as a negative control) were made electrocompetent as follows: cells were grown until OD_{600} of 0.4. Cells were then washed twice with ice-cold water, then once with 10% cold glycerol, and resuspended in 1/100 of their volume in 10% cold glycerol. 100 ng of plasmid pAA1 (containing Retron-Eco8), pAA10 (Ec48) or pSG1 (control) were electroporated in 30 μL of competent cells. Cells were then incubated in 500 μL LB for 1 h at 37°C and plated on LB plates supplemented with kanamycin (50 μgml−1) and ampicillin (100 μgml−1). Transformation efficiency was assessed by counting single colonies formed after overnight incubation at 37°C.

QUANTIFICATION AND STATISTICAL ANALYSIS

The average of three triplicates is shown throughout with individual points overlaid, unless stated otherwise.

Supplemental Figures

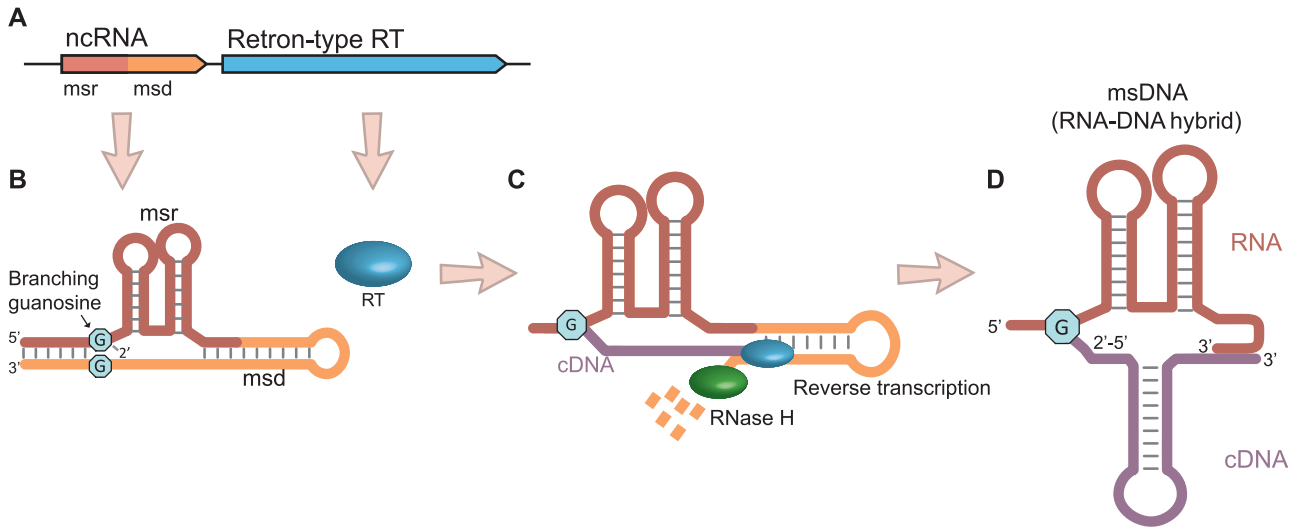


Figure S1. A Schematic Depiction of Retron msDNA Synthesis, Related to Figure 1

(A) Retrons encode a non-coding RNA (ncRNA) and a reverse transcriptase (RT). The ncRNA is composed of two parts termed msr and msd. (B) The ncRNA folds such that the edges of the msr and msd base-pair to form a stem that ends with two conserved unpaired guanosine residues on both strands. (C) The RT recognizes the structured ncRNA, uses the 2′OH of the conserved guanosine as primer, and reverse transcribes the msd section which serves as a template, starting from the other conserved guanosine. While the RT forms the cDNA, the template is degraded by RNase H. (D) This process results in a unique covalently linked RNA-cDNA hybrid called msDNA, which is branched from the guanosine nucleotide.

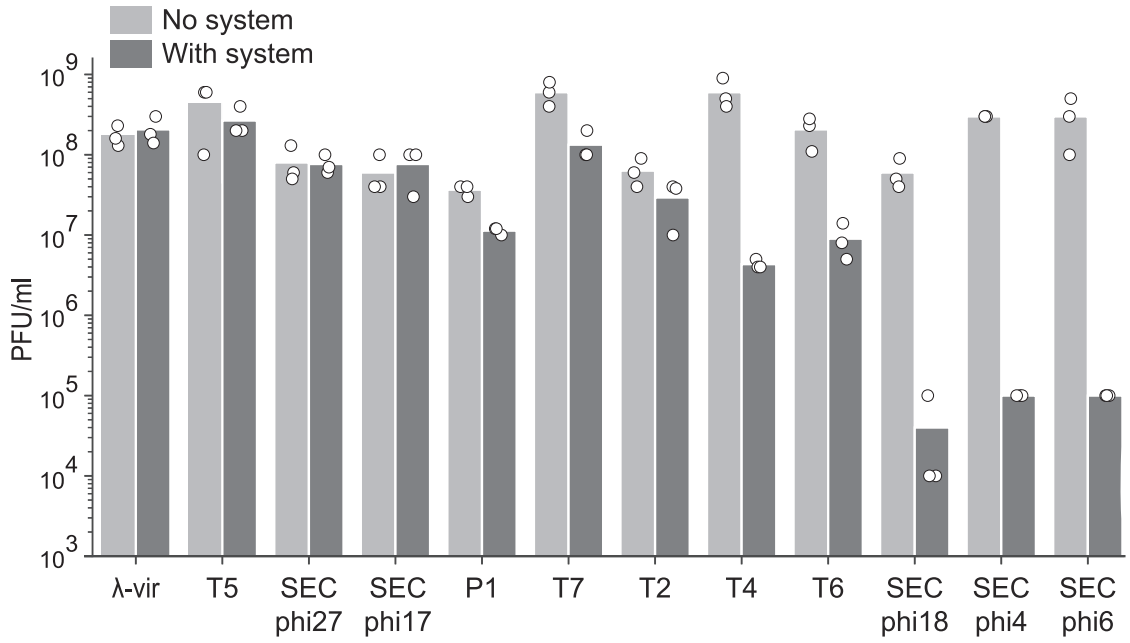
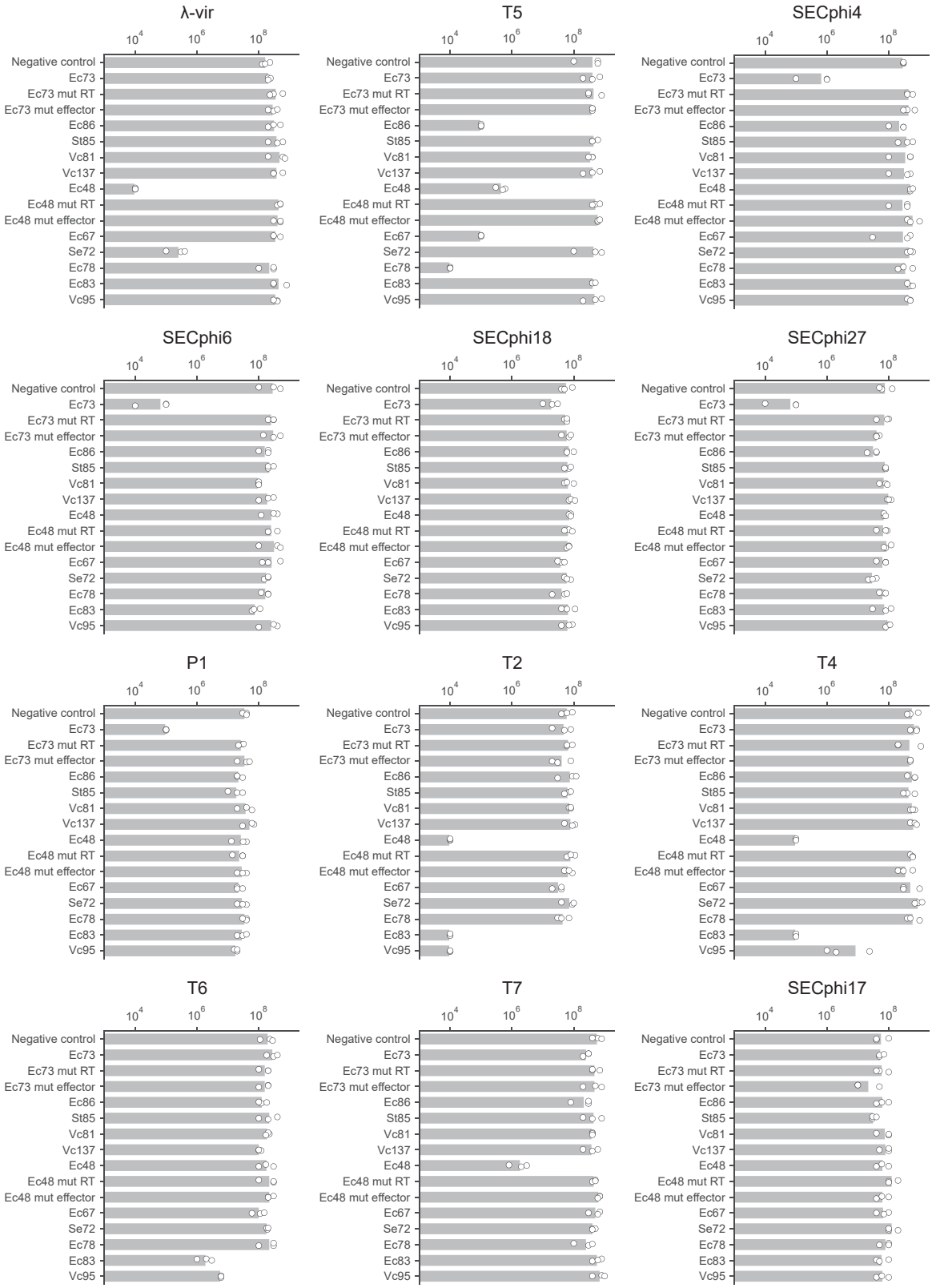


Figure S2. Efficiency of Plating of Phages Infecting *E. coli* with and without the Endonuclease-RT Defense System, Related to Figure 1.
The system was cloned from *E. coli* strain 200499, together with its flanking intergenic regions, into the laboratory strain *E. coli* MG1655 (Methods). The efficiency of plating is shown for 12 phages infecting the control *E. coli* MG1655 strain (with a plasmid encoding RFP as negative control) (no system, light gray) and the two-gene cassette cloned from *E. coli* 20099 (with system, dark gray). Data represent plaque-forming units (PFU) per ml; bar graph represents an average of three replicates, with individual data points overlaid.



(legend on next page)

Figure S3. Efficiency of Plating of Phages Infecting Retron Containing Strains, Related to Figure 3

The efficiency of plating is shown for 12 phages infecting *E. coli* MG1655 strains cloned with retron defense systems, or with a plasmid encoding RFP as negative control. Data represent plaque-forming units (PFU) per ml; bar graph represents an average of three replicates, with individual data points overlaid.

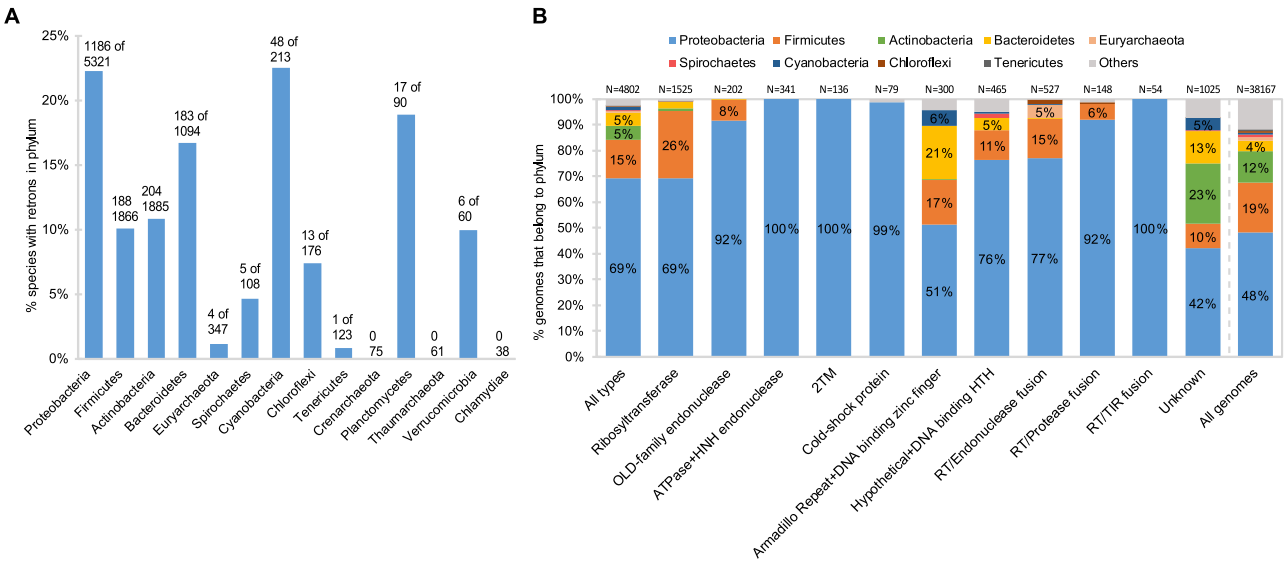


Figure S4. Phylogenetic Distribution of Retron RT Homologs and Their Associated Effectors, Related to Figure 2

(A) Presence of retron RT homologs in species, divided by phyla. Data for phyla with > 100 genomes in the database are shown. A retron RT homolog was counted as present in a species if at least one genome of that species contained such a gene. The numbers above each bar represent the number of species from the specific phylum that contain a retron RT homolog out of the total number of species from the specific phylum that are present in the analyzed database. (B) Phyletic distribution of genomes per effector type. Data are shown for phyla with > 200 genomes in the database. Rightmost bar depicts the phyletic distribution of all 38,167 genomes in the database for comparison purposes. The system type "RT/Endonuclease fusion" contains cases where the RT is fused to either a predicted OLD-family endonuclease (TOPRIM domain) or a FokI endonuclease, as detailed in Table S1.

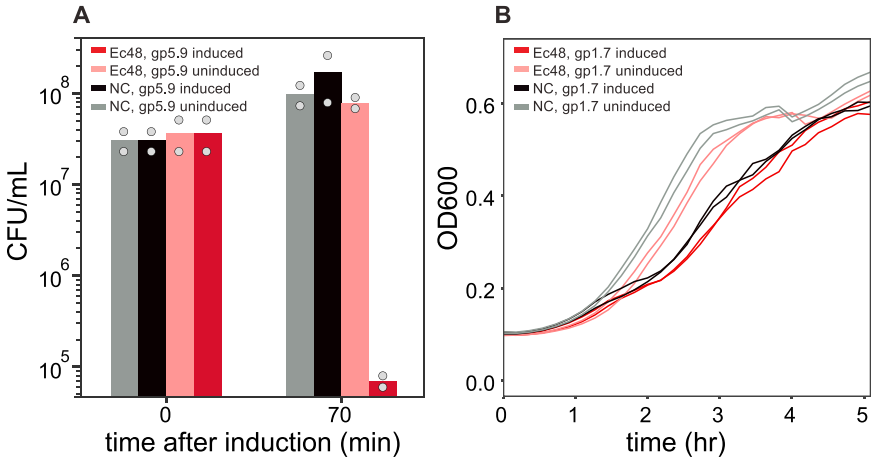


Figure S5. Related to Figure 5

(A) Overexpression of T7 gp5.9 in cells with the Ec48-containing defense system activates cell death. CFU/ml measurements taken immediately upon (0 min) or 70 minutes after the induction of gp5.9 expression. Induction of gp5.9 was performed by the addition of 0.3% arabinose to exponentially growing cells at OD_{600} 0.1. Ec48, cells containing the Ec48 defense system. NC, negative control cells containing an empty vector instead. Bar graph represents an average of two replicates, with individual data points overlaid. (B) Growth curves of *E. coli* expressing the Ec48 system or a negative control vector that lacks the system (NC), upon induction of phage dNMP kinase (T7 gp1.7 protein) by the addition of 0.3% arabinose at OD_{600} 0.1. Two biological replicates are presented as individual curves.