Please cite this article in press as: Sberro et al., Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning, Molecular Cell (2013), http://dx.doi.org/10.1016/j.molcel.2013.02.002

Molecular Cell
Article

Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning

Hila Sberro,[1,4] Azita Leavitt,[1,4] Ruth Kiro,[3,4] Eugene Koh,[1] Yoav Peleg,[2] Udi Qimron,[3] and Rotem Sorek[1,*]
[1]Department of Molecular Genetics
[2]Israel Structural Proteomics Center, Faculty of Biochemistry
Weizmann Institute of Science, Rehovot 76100, Israel
[3]Department of Clinical Microbiology and Immunology, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel
[4]These authors contributed equally to this work
*Correspondence: rotem.sorek@weizmann.ac.il
http://dx.doi.org/10.1016/j.molcel.2013.02.002

SUMMARY

Toxin-antitoxin (TA) modules, composed of a toxic protein and a counteracting antitoxin, play important roles in bacterial physiology. We examined the experimental insertion of 1.5 million genes from 388 microbial genomes into an *Escherichia coli* host using more than 8.5 million random clones. This revealed hundreds of genes (toxins) that could only be cloned when the neighboring gene (antitoxin) was present on the same clone. Clustering of these genes revealed TA families widespread in bacterial genomes, some of which deviate from the classical characteristics previously described for such modules. Introduction of these genes into *E. coli* validated that the toxin toxicity is mitigated by the antitoxin. Infection experiments with T7 phage showed that two of the new modules can provide resistance against phage. Moreover, our experiments revealed an "antidefense" protein in phage T7 that neutralizes phage resistance. Our results expose active fronts in the arms race between bacteria and phage.

INTRODUCTION

In many ecosystems, phages are 10-fold more abundant than bacterial cells, posing significant predation pressure on bacterial populations (Bergh et al., 1989; Chibani-Chennoufi et al., 2004). To survive in the face of perpetual phage attacks, bacteria have developed a variety of antiphage defense systems (Labrie et al., 2010; Stern and Sorek, 2011). These systems include restriction enzymes that recognize and cleave foreign DNA (King and Murray, 1994), abortive infection (Abi) mechanisms that lead the bacterial cell, upon phage invasion, to commit "suicide," thus protecting the colony against phage spread (Chopin et al., 2005); and the recently identified adaptive defense system called CRISPR/Cas, which uses small RNAs to target invading phage DNA (Deveau et al., 2010; Horvath and Barrangou, 2010; Sorek et al., 2008; van der Oost et al., 2009). Due to the rapid evolution and elaborated biological novelty associated with the bacteria-phage arms race, it is estimated that many additional, yet-uncharacterized antiphage defense systems are encoded by bacteria and archaea (Makarova et al., 2011; Stern and Sorek, 2011). As part of this continuous arms race, successful phages had also developed numerous counterresistance mechanisms to overcome bacterial defense (Labrie et al., 2010; Stern and Sorek, 2011).

The growing availability of genomic sequences has elucidated the vast dispersion of toxin-antitoxin (TA) systems in prokaryotic genomes (Shao et al., 2011). These modules, composed of a toxic gene and a neutralizing gene, were first suggested to function as plasmid "addiction molecules" (Van Melderen and Saavedra De Bast, 2009; Wozniak and Waldor, 2009), but their prevalent existence on chromosomes (Aizenman et al., 1996; Makarova et al., 2009; Shao et al., 2011) has led to the understanding that it is unlikely that this is their major role. Accumulating evidence suggests that TA modules play pivotal roles in prokaryotic cellular biology, including programmed cell death (Hazan et al., 2004), stress response (Christensen et al., 2001), generation of persister cells (Schumacher et al., 2009), biofilm formation (Kim et al., 2009), and phage defense via Abi (Fineran et al., 2009; Hazan and Engelberg-Kulka, 2004; Koga et al., 2011; Pecota and Wood, 1996).

The most prevalent kind of TA systems is type II systems, where both toxin and antitoxin are proteins (as opposed to types I and III, where the antitoxin is a noncoding RNA [Fineran et al., 2009; Fozo et al., 2010]). The two genes, which reside on the same operon, code for small proteins, and inhibition of the toxin is carried out through protein-protein interaction. As a rule, the toxin is a stable protein, and the antitoxin is unstable and is degraded rapidly by one of the housekeeping bacterial proteases, usually Lon or ClpP (Aizenman et al., 1996; Cherny and Gazit, 2004; Christensen et al., 2001, 2003, 2004; Lehnherr and Yarmolinsky, 1995; Roberts et al., 1994; Van Melderen et al., 1996). As a result, continuous production of the antitoxin is required to prevent the toxin's deleterious effects (Van Melderen and Saavedra De Bast, 2009). Most toxins target the translation process by cleaving cellular mRNA (Amitai et al., 2009; Daines et al., 2007; Hurley and Woychik, 2009; Jørgensen et al., 2009; Koga et al., 2011; Neubauer et al., 2009). Other toxin types exert their toxicity by phosphorylating the elongation factor EF-Tu (Schumacher et al., 2009), associating with the ribosome (Liu et al., 2008), inhibiting DNA replication (Bernard and Couturier, 1992; Jiang et al., 2002), or targeting cytoskeletal proteins

Please cite this article in press as: Sberro et al., Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning, Molecular Cell (2013), http://dx.doi.org/10.1016/j.molcel.2013.02.002

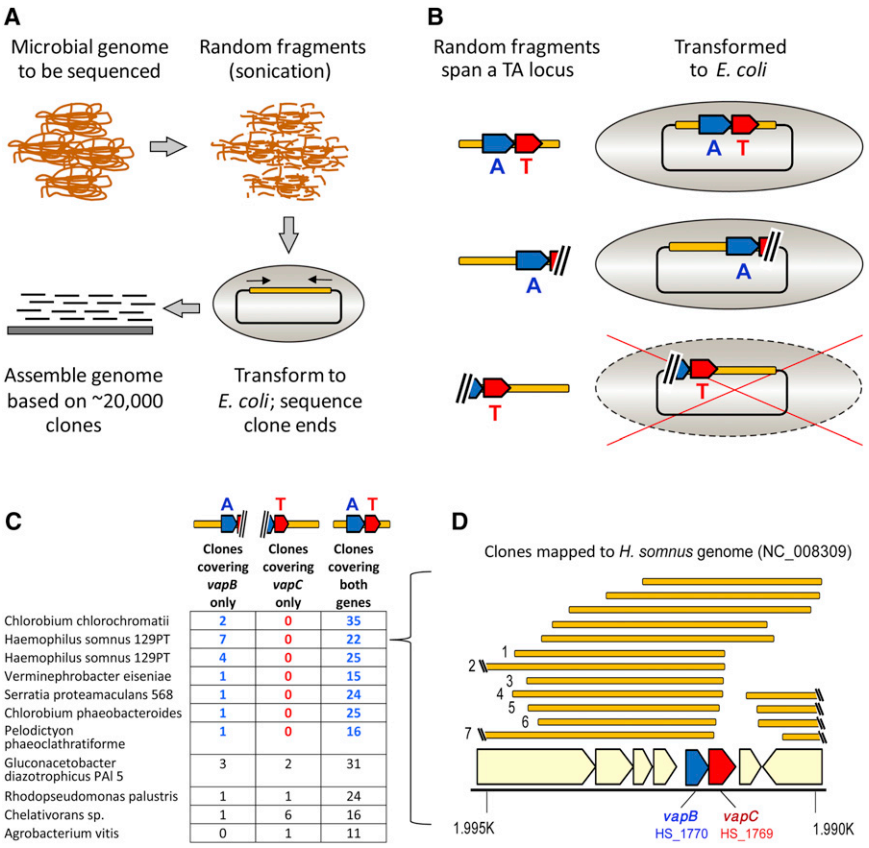

Figure 1. Data Derived from Whole-Genome Shotgun Sequencing Expose Toxin-Antitoxin Pairs

(A) The "Sanger"-based process of DNA sequencing involves random genome fragmentation and transformation of DNA fragments into *E. coli*.

(B) In a DNA locus spanning a toxin-antitoxin (TA) gene pair, random fragmentation leaving the toxin detached from its cognate antitoxin leads to *E. coli* growth arrest, whereas a fragment containing both genes, or the antitoxin alone, will be propagated and sequenced.

(C) A known family of TA gene pairs, of the VapBC type, was found in 11 of the analyzed genomes. In 7 of 11 cases the gene pair follows the "TA cloning pattern," significantly higher than the number expected by chance ($p = 4 \times 10^{-3}$).

(D) The *vapBC* locus in *Haemophilus somnus* 129T (accession number NC_008309, locus tags HS_1769-HS_1770). Shown are clones (brown) mapped to the reference genome at that locus. Clones covering the antitoxin but not the toxin are numbered. Only five of the 22 clones covering both toxin and antitoxin are shown.

(Masuda et al., 2012; Tan et al., 2011). Although several major families of type II TA systems have been described to date (Leplae et al., 2011; Makarova et al., 2009; Masuda et al., 2012; Shao et al., 2011), the extent and roles of such systems in bacterial genomes are far from fully understood.

In numerous cases where TA systems were studied experimentally, cloning of the toxin was nearly impossible in the absence of the cognate antitoxin (Fico and Mahillon, 2006; Goulard et al., 2010; Zhang et al., 2006). Based on this concept, we reasoned that data derived from Sanger-based whole-genome shotgun sequencing can experimentally and systematically reveal active TA pairs. In such a genome sequencing process, randomly fragmented DNA pieces of the genome are serially cloned and propagated within *E. coli* prior to sequencing. The ends of the cloned fragments are then sequenced, and overlapping sequences are used for genome assembly (Figure 1A). We have previously shown that analysis of clone distribution patterns can reveal genes toxic to bacteria, which are uncloneable and cause gaps (Kimelman et al., 2012; Sorek et al., 2007). However, the toxin in a TA pair is not expected to cause a gap, since the adjacent antitoxin will sometimes be found on the same clone, neutralizing the toxic effect. Nevertheless, random fragments that contain the toxin but not the antitoxin will cause cell death and will be absent from the set of clones covering the genome (Figure 1B).

Here we took advantage of this typical biased cloning pattern to systematically detect TA gene pairs within hundreds of microbial genomes. Our analyses, while retrieving many known TA systems, have also exposed several other families of TA modules widespread in numerous bacterial species. These systems were subsequently experimentally validated as TA systems. Infection experiments with wild-type (WT) and multiple T7 phage deletion mutants further showed that two of these new TA pairs provide resistance against T7, and also revealed a general anti-TA mechanism encoded by the phage.

RESULTS

Systematic Discovery of TA Families Based on Large-Scale Cloning Experiments

We analyzed 360 bacterial and 28 archaeal genomes that were sequenced using the clone-based Sanger approach, and for which the raw sequencing data were accessible and mapped to the assembled genome (Kimelman et al., 2012). For the sequencing of each genome, an average of 22,313 different randomly fragmented clones (typically sized between 3 and 8 kb, thus typically spanning approximately three to eight genes) were inserted into *E. coli*. Cumulatively, the analyzed genomes span more than 1.5 million genes that were sequenced using more than 8.5 million clones. We recorded for each gene, and for each pair of consecutive genes, the number of cloned DNA fragments that fully contain it.

To detect families of gene pairs in which one of the genes (putative toxin) is absent from clones unless the adjacent gene (putative antitoxin) is also present, we first searched for homologous gene pairs that repeatedly appear adjacent to each other in multiple genomes and clustered them into families of pairs (Experimental Procedures). To avoid the analysis of

Please cite this article in press as: Sberro et al., Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning, Molecular Cell (2013), http://dx.doi.org/10.1016/j.molcel.2013.02.002

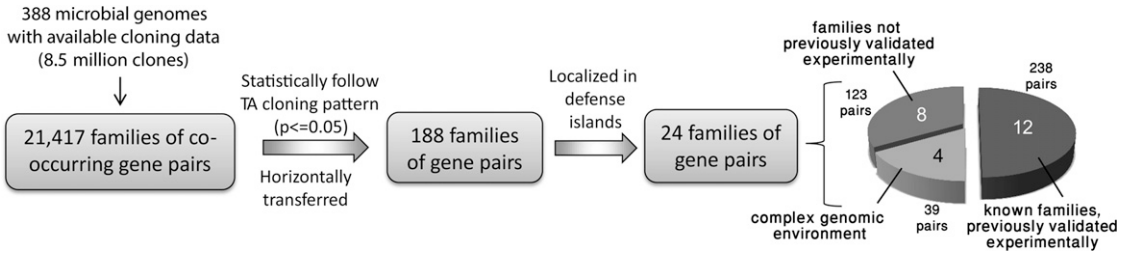

Figure 2. Workflow for Systematic Discovery of Families of Toxin-Antitoxin Associated with Antiphage Defense

Families are divided to "known," representing families that were previously experimentally established as TA systems; "not previously validated," representing families for which there was so far no experimental support of being TA systems; and "complex genomic environment," representing families of genes pairs embedded in a larger operon. Cumulative number of gene pairs belonging to each subset of families is indicated.

housekeeping genes that appear in conserved operons (e.g., ribosomal protein genes), we focused on families showing high tendency to undergo horizontal gene transfer (Experimental Procedures). Each pair of genes (X and Y) in each family was considered as following the "TA cloning pattern" if the number of clones covering gene X (toxin) but not Y (antitoxin) was 0, the number of clones covering gene Y but not X was >0, and the number of clones covering both X and Y was >0 (Figures 1B and 1C).

Not all toxins are expected to manifest their toxicity when cloned in *E. coli*, because their expression depends on the ability of the *E. coli* host to recognize their native promoters and translate them using available tRNA pool (Sorek et al., 2007). Therefore, for a given TA family of gene pairs, we expect a significant fraction of pairs, but not necessarily all pairs, to follow the TA cloning pattern (Figure 1C). To assign a statistical significance for a given family as possibly coding for a bona fide TA family, we performed, for each pair in each family 1,000 random simulations, where the clones used for sequencing of the relevant genome were randomly shuffled on the genome (Experimental Procedures; see Figure S1 online). The results were then used to assign an empirical p value per family, revealing families in which the fraction of pairs that follow the TA cloning pattern is significantly above the fraction expected by chance ($p \leq 0.05$). This yielded 188 candidate families (Table S1, Figure 2).

The identified families may include genes that follow the TA cloning pattern as a byproduct of their functions, and not for reasons associated with classical TA systems. For example, a metabolic enzyme whose expression results in accumulation of toxic intermediates might be neutralized by an accompanying transcriptional repressor. This may be the case for the argininosuccinate synthase and the ArgR repressor of the arginine regulon that obey the TA cloning patterns in 4 out of 12 homologous pairs (Table S1). Additional aspects of the clones distribution may also lead to false-positive predictions (see the Discussion). To identify TA systems more likely to play phage defense-related roles, we focused on those families that had high tendency to appear within bacterial "defense islands." It was recently shown that bacterial antiphage immune systems such as restriction enzymes, CRISPR, and Abi genes aggregate in such defense island loci in bacterial genomes (Makarova et al., 2011). We therefore selected those families in which the genomic neighborhood was enriched for defense genes (Experimental Procedures). This analysis resulted in a set of 24 putative families of TA gene pairs, overall containing 400 pairs from 176 genomes (Figure 2, Table 1, Table 2, and Table S2).

Of the 24 identified families, 12 families (50%) were already previously experimentally described as TA systems, providing strong validation to our cloning-based approach for TA discovery (Table 1). Although a diverse set of known TA families is represented in the set retrieved by our algorithm, due to limitations of our approach not all known families were represented. For example, in the family consisting of *HipAB* gene pairs, 6 out of 19 pairs were found to conform with the TA cloning pattern, but since the toxin is relatively large and the antitoxin is a short gene, such a pattern has high probability to occur by chance in the random simulations, and hence this family did not pass our statistical threshold ($p = 0.43$; see the Discussion).

Experimental Validation of TA Families Identified in This Study

Our analysis retrieved eight putative, previously unidentified families of binary TA systems for which no study has so far showed experimental support (although some of these systems have previously been predicted bioinformatically to function as TA systems, see below) (Table 2). Six of these families were selected for further experimental characterization (Table 3). A representative pair was selected from each family and cotransformed into *E. coli* BL21(DE3) on a compatible two-vector system, so that the putative toxin was under the control of an IPTG-induced promoter and the antitoxin was under the control of an arabinose-induced promoter (Figure 3A). *E. coli* bacteria carrying these two plasmids were plated on agar plates containing IPTG, arabinose, or both IPTG and arabinose. In all six tested pairs, induction of toxin expression inhibited bacterial growth, while coinduction of the toxin and antitoxin resulted in bacterial survival (Figure 3A). Experiments in batch cultures, where bacteria were grown with or without toxin/antitoxin induction, confirmed the plate-based experiments (Figure 3B). These results validate our genome-wide approach for discovery of TA systems.

We named the six validated families based on the species from which the validated system was taken: *pmenTA* (*P. mendocina*), *sanaTA* (*S. sp.* ana-3), *rlegTA* (*R. leguminosarum*), *psyrTA* (*P. syringae*), *sdenTA* (*S. denitrificans*), and *hhalTA* (*H. halophila*) (Table 2). We further attempted to test whether the toxins in these new systems have a bacteriocidal (cell-killing) or bacteriostatic (growth-inhibiting) effect. For this, we used

Please cite this article in press as: Sberro et al., Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning, Molecular Cell (2013), http://dx.doi.org/10.1016/j.molcel.2013.02.002

Table 1. Families of Previously Known Toxin-Antitoxin Systems Retrieved by the TA Discovery Algorithm

	Number of Pairs in Family	Antitoxin Superfamily	Toxin Superfamily	COGs Associated with Antitoxin	COGs Associated with Toxin	P Value for TA Cloning Pattern
1	16	RelB	RelE	COG3905	COG3668	5.0×10^{-3}
2	25	RelB	RelE	COG3077	COG3041	3.0×10^{-3}
3	13	RelB	RelE	n/a	n/a	1.1×10^{-2}
4	33	Xre	RelE	COG3620	COG4679	2.5×10^{-2}
5	28	Xre	RelE	n/a	COG2944	2.5×10^{-2}
6	32	Xre	RelE	COG5499	COG4680	7.0×10^{-3}
7	27	Phd	RelE	COG2161	COG2026	3.4×10^{-2}
8	9	HicB	HicA	COG1598	COG1724	9.0×10^{-3}
9	11	HicB	HicA	n/a	COG1724	3.4×10^{-2}
10	11	VapB	VapC	COG4456	COG1487	4.0×10^{-3}
11	10	HigA	HigB	COG3093	COG3549	1.0×10^{-3}
12	22	MosA	MosT	n/a	COG2253	0

IPTG to induce toxin expression for different time intervals (ranging from 30 to 300 min), and then plated the cells on agar plates containing arabinose to activate antitoxin expression (Figure 3C). For the *rlegTA* system, colony-forming units dropped by five orders of magnitude following 120 min of toxin induction, implying that the rleg toxin has a bactericidal effect on the cells. These results were also supported by the kinetic assays, where induction of antitoxin expression 2.5 hr after toxin induction did not result in cell regrowth for *rlegTA* (Figure 3B). A milder effect was observed for the remaining TA systems, with *sdenTA* and *hhalTA* showing almost no reduction in colony-forming units following toxin induction, suggestive of a bacteriostatic effect for these systems (Figure 3C). We note, however, that there is a general debate whether TA systems are bacteriostatic or bacteriocidal, and factors such as inducer concentrations of toxin and antitoxin and duration of toxin induction prior to antitoxin induction (''point of no return'') may affect our interpretations.

Characteristics of Detected TA Families

Most type II TA modules described to date share several typical characteristics: the antitoxin appears upstream of the toxin, both the toxin and antitoxin are small proteins (typically ~100 aa), and the antitoxin contains a DNA binding domain (Makarova et al., 2009). Since our approach does not rely on such attributes for TA modules discovery, it has the potential to expand the premises of TA modules properties. Indeed, some of the new families we experimentally validated deviate significantly from the previously described characteristics. For example, the sizes of many new toxins and antitoxins are significantly larger than 100 aa, with a maximum of 698 aa in the toxin of the *psyrTA* system; in two families the toxin is located upstream of the antitoxin (Figure 4A); and several antitoxins do not contain a known DNA binding domain (although we cannot rule out the possibility that some antitoxins code for such yet-uncharacterized domains) (Figure 4A).

Although in most of the known TA systems the toxin is a ribonuclease, diverse domains within the new toxins we detected, including DNA helicase, phosphoribosyl-transferase, and nucleotidyl-transferase, suggest novel mechanisms of toxicity (Figure 4A). Similarly, the presence of ADP-ribose-binding and nucletotide-binding domains in some of the antitoxins suggests that these antitoxins perform a more complex function than simply masking the activity of the toxin by protein-protein interactions (Figure 4A and Table 2).

Finally, in four additional families that we detected, it seems that the putative TA system is part of a larger operon that is cohorizontally transferred between genomes in the context of defense islands, suggesting their involvement in more complex defense mechanisms (Table 2 and Figure S2). As such, these are probably not bona fide TA systems but might represent false-positive predictions of our approach (see the Discussion).

Previous analyses have shown that TA systems can be modular, such that members of one toxin family may be associated with several different types of antitoxin, and vice versa (Leplae et al., 2011; Makarova et al., 2009). Indeed, the toxins of three of our new families (*rlegTA*, *sdenTA*, and *sanaTA*) carry the same domain, DUF1814, with a different antitoxin associated with this domain in each of the three families (Table 2 and Figure 4A). One of these families, *rlegTA*, where the DUF1814 toxin is accompanied by COG5340 as an antitoxin, was previously shown to be enriched in defense islands in bacterial genomes and was suggested as a new TA system based on its two-gene nature (Makarova et al., 2011). Although the COG5340 domain was not identified in the *sden* and *sana* antitoxins, HHpred comparisons suggest that the antitoxins of the three families share distant homology, and hence may belong to a single superfamily. The DUF1814 domain was recently classified as a nucleotidyl-transferase domain based on structural information, but its specific substrates are yet unknown (Kuchta et al., 2009). Interestingly, the DUF1814 domain was also documented in AbiG, a two-gene system involved in Abi in *Lactococcus lactis* via an unknown mechanism (Makarova et al., 2011; O'Connor et al., 1996), although there is no direct homology between the AbiG system and any of the genes in the new TA systems we detected. Therefore, our results point to DUF1814 domain-containing proteins as a widespread superfamily of toxins that might be involved in antiphage defense (see below).

Overall, members of the systems we detected in this study appeared in 21% of the genomes analyzed, and in the vast majority of cases (93%) they appeared on chromosomal DNA rather than

Please cite this article in press as: Sberro et al., Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning, Molecular Cell (2013), http://dx.doi.org/10.1016/j.molcel.2013.02.002

Table 2. Families of Toxin-Antitoxin Systems Retrieved by the TA Discovery Algorithm that Were Not Validated Previously

Number of Pairs in Family	Antitoxin	Toxin	Antitoxin Annotation	Toxin Annotation	Domains Associated with Antitoxin	Domains Associated with Toxin	p Value for TA Cloning Pattern	Exp[a]
11	psyrA	psyrT	Nucleotide-binding protein	RecQ family DNA helicase	COG0758	COG0514	0	Yes
10	sanaA	sanaT	Hypothetical protein	Hypothetical protein	n/a	DUF1814	0	Yes
10	pmenA	pmenT	ADP-ribose binding domain protein	Hypothetical protein	COG2110	n/a	0	Yes
20	rlegA	rlegT	Predicted transcriptional regulator	Hypothetical protein	COG5340	DUF1814	0	Yes
7	sdenA	sdenT	Hypothetical protein	Hypothetical protein	n/a	DUF1814	3×10^{-3}	Yes
44	hhalA	hhalT	Nucleotidyl- transferase family protein (MNT domain)	Hypothetical protein (HEPN domain)	COG1669	COG2361	2×10^{-3}	Yes
7			DNA polymerase, beta-like region (MNT domain)	HEPN domain-containing protein	n/a	COG1895	4×10^{-3}	
14			Hypothetical protein (RHH domain)	Hypothetical protein	COG5304	COG2929	2×10^{-3}	
8			Hypothetical protein (associated with cas/cmr genes)	RAMP domain protein (Cmr6-like)	n/a	DUF324	2.9×10^{-2}	*
15			Uncharacterized membrane protein	Membrane protein, TraG-like N-terminal domain	n/a	PF07916	2×10^{-3}	*
9			Hypothetical protein	Adenine specific DNA methylase	n/a	COG2189	0	*
7			T/G mismatch-specific endonuclease	Type II restriction endonuclease	COG3727	PF09019	0	*

Asterisk indicates TA system that is part of a larger operon putatively involved in bacterial defense (Figure S2).
[a]Validated experimentally in this study.

on plasmids. These results suggest that such systems have important roles in bacterial physiology/defense rather than functioning as plasmid addiction molecules. The distribution of most families is spread across multiple bacterial phyla including several important human pathogens (Table S3 and Figures 4B and 4C). For example, the *psyrTA*, *sanaTA*, and *pmenTA* systems are abundant in enteropathogenic and uropathogenic *E. coli* strains; the *pmenTA* system also exists in many *Mycobacterium tuberculosis* isolates; *psyrTA* exists in *Shigella* and *Pseudomonas aeruginosa* strains; and the *rlegTA* and *sanaTA* systems exist in several pathogenic *Legionella* species. In addition, many resident bacteria of the human gut carry one or more of these TA modules, including bacteria belonging to the *Bifidobacterium* and *Prevotella* genera (Table S3 and Figure 4C). This underscores the toxin-antitoxin families identified in this study as potentially contributing to persistence, phage defense and stress responses of clinically important bacteria.

Phage Resistance

We next set out to explore whether any of the TA systems we detected can provide defense against phage. For this, efficiency-of-plating (EOP) assays of T7 phage on *E. coli* hosts were performed (Figure 5A). Since these new TA systems are widespread in *E. coli* strains (but are not found in the lab strains *E. coli* K-12 and *E. coli* BL21), we hypothesized that a successful coliphage, such as T7, might hold antidefense mechanisms that mitigate the defense conferred by the TA systems. We therefore tested, in addition to the WT T7 phage, 12 additional T7 mutants lacking genes that are nonessential for infection of *E. coli* K-12 (Table S4). Each of these T7 mutants was used to infect *E. coli* K-12 expressing the verified new TA systems, as well as control clones expressing only the antitoxin of each system (Figure 5A). One of the tested systems, *sanaTA*, was found to provide *E. coli* with resistance against T7Δ4.5 and T7Δ4.3Δ4.5Δ4.7, reducing sensitivity to these phage strains by about three orders of magnitude (Figure 5B). A second system, *rlegTA*, resulted in opaque plaques with plaque diameters reduced almost 4-fold for the WT T7 strain (diameters of 0.47 ± 0.06 mm for *E. coli* expressing both the toxin and the antitoxin, as compared to 1.77 ± 0.03 mm for bacteria expressing the antitoxin only).

Since the *sanaTA* system provides resistance to the T7Δ4.5 mutant phage but not to the WT T7 phage, we hypothesized that the 4.5 gene codes for an antidefense mechanism that overcomes the Abi imposed by the TA system. Indeed, complementation assay in *E. coli* K-12 expressing the 4.5 gene from a plasmid verified this gene as encoding the anti-Abi mechanism (Figure 5C). The 4.5 gene codes for a short protein (89 aa) with no functional annotation.

We next asked whether the defense provided by the *sanaTA* system against the phage is Lon dependent. The Lon protease is one of the major proteolytic machineries in the bacterial cell (Gottesman, 2003) and was implicated in degradation (“destabilization”) of many types of antitoxins in *E. coli*, thus enabling toxin activity (Christensen et al., 2001; Van Melderen et al., 1996; Wang et al., 2011). Indeed, T7 mutant growth on *E. coli* containing the *sanaTA* system was restored by two orders of magnitude

Please cite this article in press as: Sberro et al., Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning, Molecular Cell (2013), http://dx.doi.org/10.1016/j.molcel.2013.02.002

Table 3. Gene Pairs Selected for Experimental Verification in a Dual-Plasmid Arabinose/IPTG Expression Induction System

Family	Antitoxin Locus Tag	A Size (aa)	Toxin Locus Tag	T Size (aa)	Organism	Number of Clones that Cover Antitoxin	Number of Clones that Cover Toxin	Number of Clones that Cover Both	Reversible Toxicity?
sdenTA	Sden_0299	174	Sden_0300	295	Shewanella denitrificans OS217	8	0	31	Yes
psyrTA	Psyr_3805	455	Psyr_3804	698	Pseudomonas syringae pv. syringae B728a	13	0	1	Partial
sanaTA	Shewana3_4160	136	Shewana3_4161	313	Shewanella sp. ANA-3	5	0	29	Yes
pmenTA	Pmen_0566	360	Pmen_0565	217	Pseudomonas mendocina ymp	3	0	10	Partial
rlegTA	Rleg_6340	205	Rleg_6339	289	Rhizobium leguminosarum bv. trifolii WSM1325	11	0	31	No
hhalTA	Hhal_0686	96	Hhal_0685	119	Halorhodospira halophila SL1	2	0	23	Yes

when the *E. coli* also lacked *lon* (Figure 5D), suggesting that the *sanaTA* protection from T7 phage depends on Lon activity.

We hypothesized that the phage gene product (Gp) 4.5 interacts with Lon to prevent antitoxin degradation and thus hinders the *sanaTA* Abi activity. To test this hypothesis, we coexpressed Lon (Flag-tagged) and Gp4.5 within *E. coli*. Indeed, we found that Lon and Gp4.5 coimmunoprecipitate, indicating that 4.5 tightly binds Lon (Figure 5E). A reciprocal coimmunoprecipitation assay, in which pull-down was performed on Flag-tagged Gp4.5 protein, produced similar coimmunoprecipitation patterns (Figure 5F). Overall, these results imply that the T7 Gp4.5 neutralizes TA-system-mediated Abi by inhibiting the Lon protease activity, thus preventing antitoxin degradation and toxin activation.

DISCUSSION

Due to the rapid evolution and the functional novelties frequently associated with TA modules, discovery of novel TA families is challenging and often depends on serendipity (Hayes and Van Melderen, 2011; Stern and Sorek, 2011). Even when a member of a known family is identified by sequence similarity, there is usually no experimental evidence to support functionality of that specific member. We used gene clonability in *E. coli* as an experimental readout for TA module functionality, and performed a massive-scale analysis of millions of cloned genes to detect hundreds of functional TA modules within microbial genomes. Many functional members belonging to 12 known families were detected, as well as several additional families associated with defense islands, expanding the set of TA systems verified to date.

Three of the systems we identified (those containing MNT and RHH domains in their antitoxin) were previously predicted to be TA systems by a computational analysis based on comparative genomics and genomic context analyses (Makarova et al., 2009; Table 2). However, the cloning patterns we recorded for two of these systems that belong to the *HEPN-MNT* superfamily suggested that the HEPN is the toxin and the MNT is the antitoxin, and not the opposite, as previously predicted (Makarova et al., 2009). Our experiments with one such system, *hhalTA*, verified that the gene with the HEPN domain (*hhalT*) is indeed the toxin and *hhalA* is the antitoxin, pointing to the strength of our cloning-based approach over purely computational approaches.

Some of the systems we detected in this study contain domains that were not described in TA systems before. The antitoxin in the *pmenTA* system, for example, shares homology with the macrodomain (COG2110), which was shown to bind the metabolite ADP-ribose (Han et al., 2011). It is therefore possible that this cofactor is important for the pmenA antitoxin activity. The psyrT shares homology with domains of the RecQ helicase, a family of proteins implicated in DNA repair (Bernstein et al., 2010); and the antitoxin of the same system, psyrA, has a nucleotide binding domain (COG0758) that was previously described in proteins involved in DNA uptake (Mortier-Barrière et al., 2007) and recently also in the *shosT/shosA* toxin/antitoxin system (Kimelman et al., 2012). The presence of these domains within the *psyrTA* system may suggest that the toxicity of this system is conferred by DNA manipulation of some sort.

Not all known TA families were retrieved by our approach. First, families represented in the studied genomes by fewer than seven members were not analyzed due to lack of statistical power. Second, in families where the antitoxin is encoded by a very small gene, the occurrence of a randomly fragmented clone that contains the toxin but not the antitoxin has low probability, and this again reduces the statistical power for detection of such families. This has resulted in high p values for several bona fide TA families, such as the *HipAB* family, and these did not pass our statistical thresholds. Our discovery method is therefore biased toward detection of systems characterized by larger genes than those found in most known families. It is therefore conceivable that additional families of TA systems, not detected by our algorithms, are yet to be discovered.

Our study had focused on type II TA systems, where both the toxin and antitoxin are protein entities. However, our general cloning-based approach can in principle detect other types of TA systems, where the antitoxin is a noncoding RNA molecule (types I and III). To detect such cases, one may search for genes (toxin) that cannot be cloned unless the nearby intergenic region (encoding the noncoding antitoxin) is also found on the same clone. However, such an application is out of the scope of the current study.

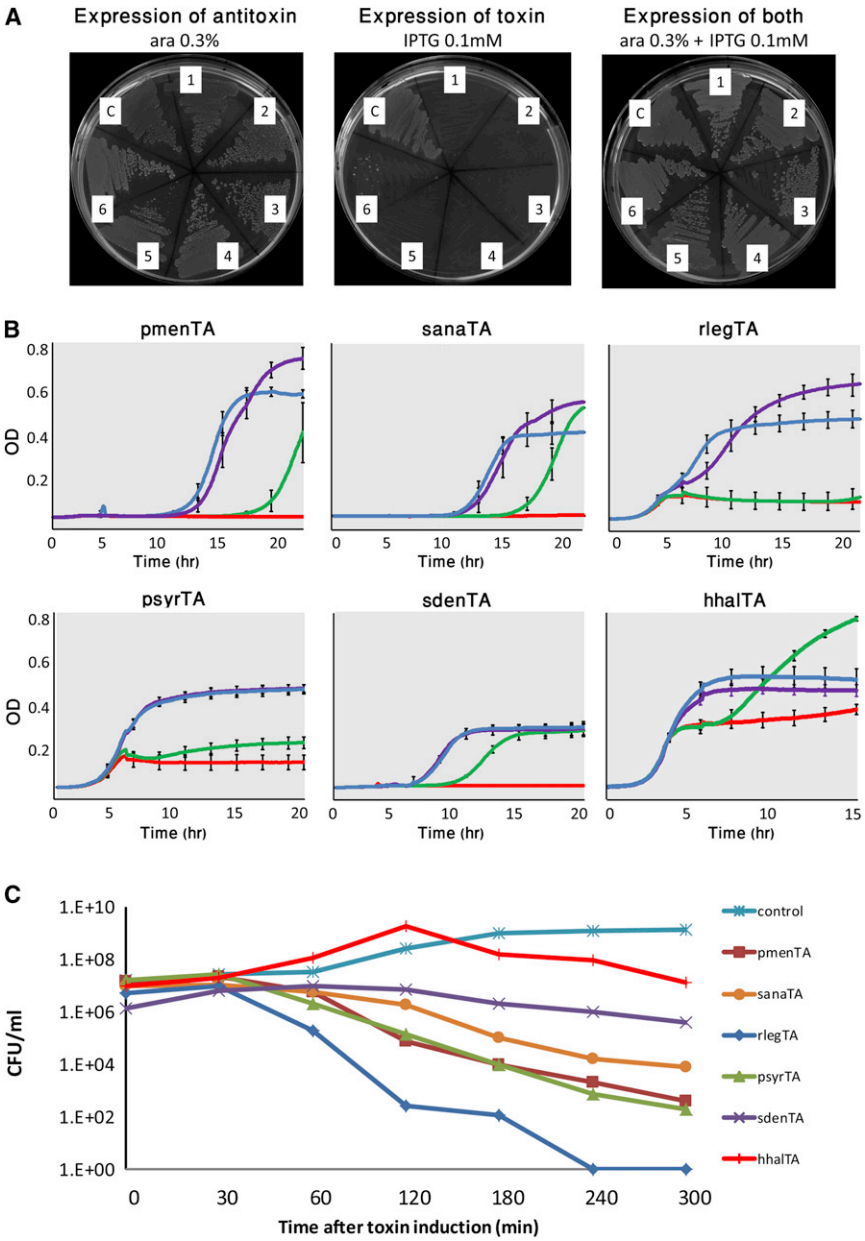

Figure 3. Properties of Experimentally Verified TA Systems

(A) Growth of bacteria when only antitoxin is induced (left), only toxin is induced (middle), and both are induced together (right). Toxin and antitoxin were cloned on pRSF (IPTG inducible promoter) and pBAD (arabinose inducible), respectively, in *E. coli* BL21(DE3)pLysS. C, control bacteria with empty plasmids; 1-pmen system; 2-sana system; 3-psyr system; 4-rleg system; 5-sden system; 6-hhal system.

(B) Kinetics of *E. coli* BL21(DE3)pLysS growth when toxin and antitoxin are coexpressed simultaneously (purple), alone (red and blue for toxin and antitoxin, respectively), or when antitoxin is induced 2.5 hr following toxin induction (green). Kinetic measurements were performed on biological triplicates in technical duplicates. Error bars represent standard deviation.

(C) Viability assays for cells following exposure to toxin. Transcription of toxin was induced by 100 μM IPTG. At increasing time points following toxin induction (30, 60, 120, 180, 240, and 300 min), cells were plated on LB plates containing 0.3% arabinose and no IPTG, to activate antitoxin expression. Colony-forming units (CFUs) were determined by colony counting.

Although we focused our validation and analyses on TA families that preferentially appear in defense islands, other toxin-immunity modules may also be revealed by our algorithm. For example, secreted toxins that are targeted to other bacteria via the type VI secretion system in *P. aeruginosa* have an adjacent immunity gene to prevent the toxin from affecting the producing cell (Hood et al., 2010). Another example is of the contact-dependent inhibition protein A (CdiA), which kills neighboring cells upon cell-cell contact; the toxicity of this protein is neutralized by the immunity protein CdiI, encoded by a small gene found immediately downstream to *cdiA* (Aoki et al., 2010; Poole et al., 2011). Since such antimicrobial toxin-immunity systems are expected to yield a cloning bias similar to that observed for TA systems, but are not expected to localize in phage defense islands, it is possible that the broader set of families retrieved by our algorithm (Table S1) may contain such systems related to antimicrobial activities rather than to phage defense. Nevertheless, some of these predicted families possibly represent false-positive predictions. Such false positives may stem from long-range effects of neighboring genes that change the clone-pattern distribution at the locus, leading to a distribution that may falsely appear as representing a TA system. Indeed, in four cases out of the 24 systems that our algorithm had retrieved, the genomic organization of the locus indicates on operons that are larger than two genes, which is not typical of TA systems. Therefore, single predictions need to be treated with caution until experimentally validated.

Our results show that the *sanaTA* system provides resistance against T7 phage lacking the nonessential 4.5 gene. We also found that expression of the *rlegTA* system results in significant reduction in T7 plaques diameter, although this result might be attributed to the partial toxicity of the *rlegTA* system when expressed in *E. coli*. To date, only two other type II TA systems were shown to provide phage resistance (Hazan et al., 2004; Koga et al., 2011). However, the distribution of numerous TA systems in bacterial defense islands (Makarova et al., 2011) implies that the involvement of TA systems in bacterial battle against phages might be underestimated. It is possible that the systems we tested which did not show antiphage activity do

Please cite this article in press as: Sberro et al., Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning, Molecular Cell (2013), http://dx.doi.org/10.1016/j.molcel.2013.02.002

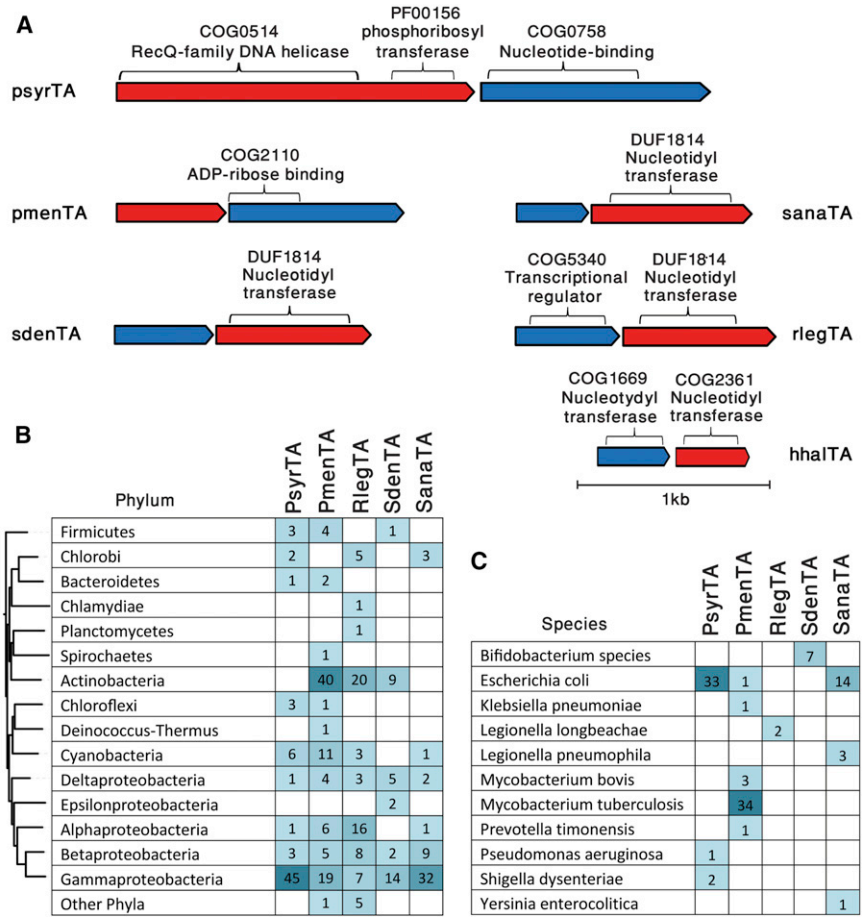

Figure 4. Properties of Families Experimentally Validated in This Study

(A) Operon and domain organization of the validated families. Representative pair of each family is shown. For each pair, red and blue genes denote toxin and antitoxin, respectively.

(B and C) (B) Distribution of TA families among different bacterial phyla and (C) human-associated bacteria. Number of instances of each system within a phylum/bacterial species is indicated, with darker colors indicating higher number of instances (the phylogenetic distribution of the *hhalTA* family does not appear here, as it was described previously by Makarova et al. [2009]).

indeed function in phage resistance in their genome of origin. Since the TA systems we tested were not originally derived from *E. coli*, the exogenous antitoxin might lack the sequence signals to be recognized and degraded by the *E. coli* proteases. In the absence of antitoxin destabilization, no antiphage activity is possible. Another possibility is that, similar to the 4.5 gene reported here, T7 carries additional genes encoding anti-TA activities that were not tested in our study; these genes might have mitigated the defensive activity of the tested systems. Infection with additional mutant phages might expose antiphage activity for additional systems, and might shed more light on the evolution of the arms race between bacteria and phage.

Two phage strategies to counteract bacterial activity of TA systems have been described to date. One strategy is to bring an external antitoxin to substitute for the degraded antitoxin. This was shown for phage T4 which encodes a protein, Dmd, a suppressor of the RnlA toxin (Koga et al., 2011; Otsuka and Yonesaki, 2012). Another strategy is to inhibit the bacterial protease, thus overcoming the rapid destabilization of the antitoxin. This mode of action was demonstrated for T4 PinA protein, which blocks Lon protease (Skorupski et al., 1988), and for the λ RexB protein, which inhibits degradation of the Phd and MazE antitoxins by hindering activity of the ClpP protease (Engelberg-Kulka et al., 1998). Our results imply that the T7 Gp4.5 also resists TA-mediated Abi by inhibiting the Lon protease. This suggests that protease inhibition is a general anti-Abi mechanism used by phages. The complete lack of homology between the anti-protease phage proteins described to date might suggest that the protease inhibition strategy has been evolutionarily invented multiple times in parallel. Indeed, the bacterial proteases seem to be the weak link in the TA-mediated defense strategy, as a single protease usually dominates the destabilization of antitoxins belonging to multiple different TA systems in the same organism (Aizenman et al., 1996; Cherny and Gazit, 2004; Christensen et al., 2001; Lehnherr and Yarmolinsky, 1995; Roberts et al., 1994; Van Melderen et al., 1996).

Several studies suggest the involvement of TA systems and the Lon protease in generating bacterial persister cells (Maisonneuve et al., 2011; Rotem et al., 2010; Schumacher et al., 2009; Vázquez-Laslop et al., 2006). Persistence, in which bacteria enter a dormant, metabolically silent state, is a major obstacle in effective antibiotic treatment against pathogens such as *Mycobacterium tuberculosis*, *Staphylococcus aureus*, and *Pseudomonas aeruginosa* (Gomez and McKinney, 2004; Tuchscherr et al., 2011). Deletion of Lon in *E. coli* was shown to mitigate TA-system-mediated persistence and resulted in higher sensitivity to antibiotics (Maisonneuve et al., 2011). It appears that phages may be a rich source for molecules that inhibit bacterial proteases, and it is possible that with some molecular adaptations, such molecules could be used to treat bacterial persistence. One can envision that in the future such phage-derived compounds may enter clinical use as synergistic supplements to treatment with antibiotics. Thus, the arms race between bacteria and phage could be harnessed for the benefit of mankind.

EXPERIMENTAL PROCEDURES

Coverage Analysis of Pairs of Genes

Mapping of sequencing clones on 360 bacteria and 28 archaea for which raw clone sequencing data were available (http://www.weizmann.ac.il/pandatox) was performed as described in Kimelman et al. (2012). Gene positions and

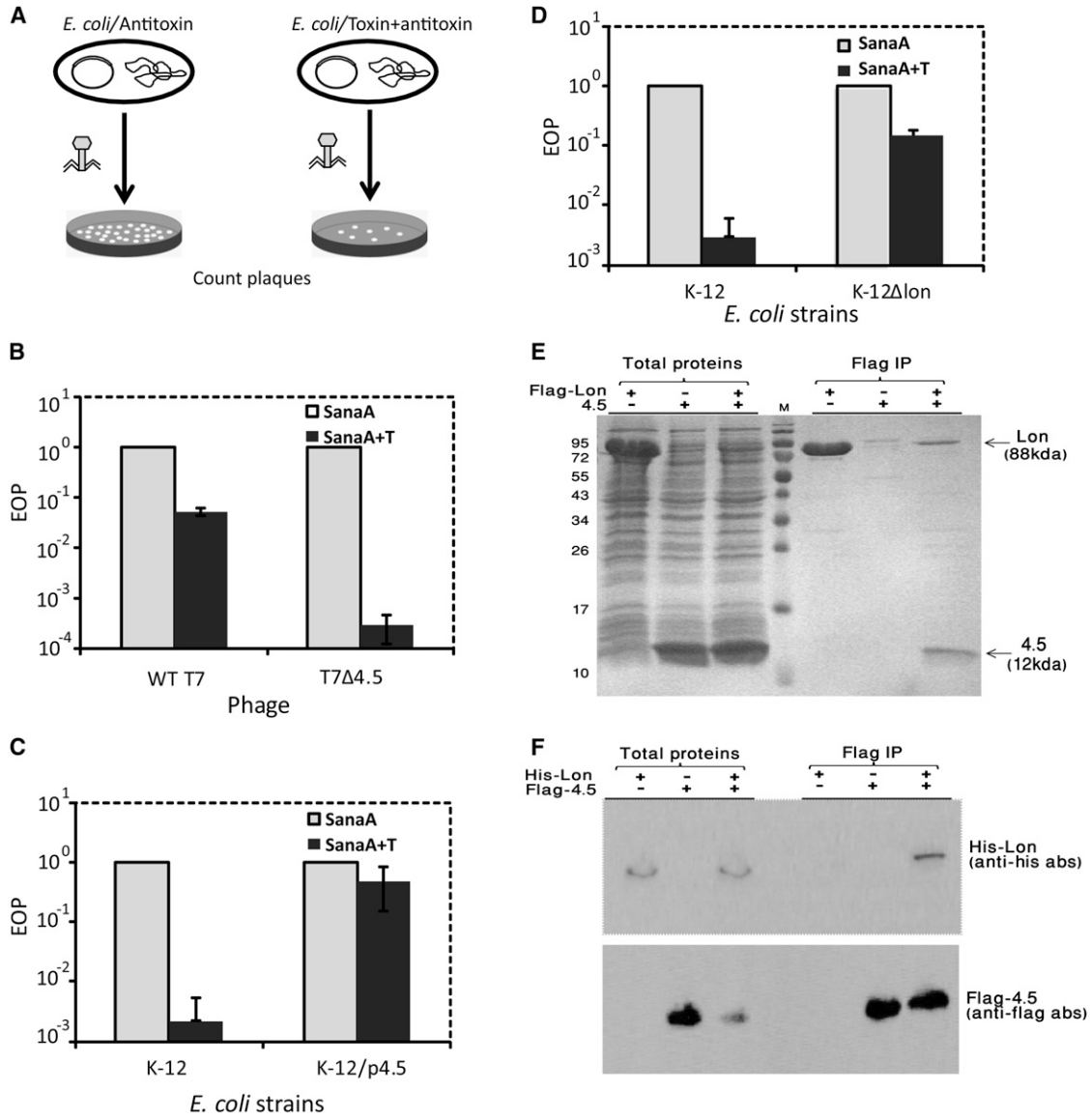

Figure 5. T7 Gp4.5 Antagonizes Abortive Infection and Interacts with Lon Protease

(A) Illustration of the plaque-forming unit (PFU) assays on *E. coli* harboring toxin/antitoxin systems. Efficiency of plating (EOP) was calculated by dividing the number of PFUs obtained for a bacterial lawn expressing the toxin + antitoxin by the number of PFUs obtained on a lawn expressing the antitoxin alone. Bars in (B), (C), and (E) represent average ± SD of three independent EOP experiments.

(B) EOP experiments with *E. coli* harboring the *sanaTA* system when infected by WT T7 (left) or by T7Δ4.5 (right). Error bars represent standard deviation between replicates.

(C) EOP experiments with WT *E. coli* (left) and *E. coli* expressing Gp4.5 (right), when infected by T7Δ4.3Δ4.5Δ4.7. Error bars represent standard deviation between replicates.

(D) EOP experiments with WT *E. coli* (left) and *E. coli* lacking *lon* (right), when infected by T7Δ4.3Δ4.5Δ4.7. Error bars represent standard deviation between replicates.

(E) Coimmunoprecipitation (coIP) of Lon and Gp4.5. The *E. coli* Lon protease was Flag tagged at the N terminus and expressed within *E. coli* BL21(DE3) with or without coexpression of gene 4.5. Samples were analyzed by 15% SDS-PAGE. Three left lanes, total soluble proteins; three lanes on the right, following immunoprecipitation with anti-Flag antibody. CoIP of Lon and Gp4.5 was validated by mass spectrometry analysis. Numbers on the left denote protein marker sizes in kDa.

(F) Western blot analysis on reciprocal coIP with Flag-tagged Gp4.5 and his-tagged Lon. Immunoprecipitation was done with anti-Flag antibody.

annotations were downloaded as described (Kimelman et al., 2012). For each consecutive pair of genes in each genome, three numbers were recorded: x = [number of clones fully spanning the first gene but not the second gene], y = [number of clones fully spanning the second gene but not the first gene], and z = [number of clones fully spanning both genes].

A pair of genes conforming with {x = 0; y > 0; z > 0} or {x > 0; y = 0; z > 0} was declared as "following the TA cloning pattern." Pairs in which the putative toxin was identified as a "hitchhiker" (Kimelman et al., 2012) were eliminated from further counts in order to avoid cases in which this clonability pattern was a byproduct of a nearby single, standalone toxic gene. Pairs

residing in replicons lacking sufficient clone coverage were also ignored in further counts.

Clustering of Pairs into Families

Clustering of individual genes based on sequence similarity was retrieved from IMG (http://img.jgi.doe.gov/cgi-bin/w/main.cgi) on November 2010 (based on the "IMG cluster" field). Cluster IDs were recorded for every consecutive pair of genes analyzed. Pairs containing a gene that lacked a cluster ID were discarded. All pairs having the same two cluster IDs (regardless of the order and the strand) were aggregated into a single "family of pairs." Families containing fewer than seven pairs were ignored, to ensure statistical power in next steps of the analysis. This resulted in 21,417 families, containing at least seven pairs of consecutive genes sharing the same two cluster IDs, which were further analyzed. The following "TA cloning fraction" (F) parameters were calculated for each family: F1 = the fraction of family members that follow the TA cloning pattern {x = 0; y > 0; z > 0}, and F2 = the fraction of family members that follow the TA cloning pattern {x > 0; y = 0; z > 0}.

Directionality of putative family (i.e., determining whether gene "x" is the putative toxin or the putative antitoxin) was determined using F_{max} = max {F1,F2}, such that F_{max} = F1 → x is the putative toxin, and F_{max} = F2 → y is the putative toxin.

For each family, the genus names of all containing organisms were extracted. A "family diversity" (FD) measure was defined as the number of different genus names divided by the total number of family members (pairs). This measure was used to roughly assess the tendency of the family to undergo horizontal gene transfer (HGT) within a wide array of organisms, with higher FD corresponding to higher HGT tendency. For example, a ten-member family in which all members are found in strains of *Escherichia coli* will receive a low FD of 0.1. Only families having FD > = 0.5 were further analyzed, based on the empirical FD distribution among known families of TA systems (Figure S3A).

A Statistical Framework to Detect TA Families

Since a given pair of genes has the potential to follow the TA cloning pattern merely by chance (i.e., due to the random clone fragmentation) rather than reflecting a real toxin/antitoxin activity, clone distribution simulations were performed to assess statistical significance per family (Figure S1). For every pair of genes in each family, all clones covering its genome of origin were randomly distributed on the genome, shuffling clone positions but maintaining their number and sizes. Based on these random clone distributions, the x, y, and z values were measured for the gene pair, and a simulated "TA cloning fraction" (F_{sim}) was then calculated for the family. This procedure was repeated 1,000 times for each family, generating a distribution of $F_{sim(i)}$ (i = 1..1000). The real F_{max} of the family was then compared to the distribution of F_{sim} values obtained from the simulations, generating an empirical p value for a family (Figure S1).

Since this procedure is computationally demanding to perform for >21,000 families, only families where F_{max} > = 0.3 (i.e., at least 30% of family members followed the TA cloning pattern) were thus analyzed. Families presenting p value $\leq$ 0.05 were considered as following the TA cloning pattern in a statistically significantly manner (Table S1).

To identify families significantly localized to defense islands (DIs) (Makarova et al., 2011), a value of "mean number of defense island genes in proximity" (DI_{val}) was calculated for each family based on a list of 132 COGs (DI genes) that were shown to be enriched in defense island regions (Makarova et al., 2011). For this, the average number of DI genes within a range of ±5 genes from each family member was calculated. Families having DI_{val} > 0.5 were defined as DI associated (Figure 2), based on the empirical DI_{val} distribution among known families of TA systems (Figure S3B and Figure S4). Families were further reviewed manually to remove transposon-containing families possibly associated with defense islands due to their transposition-related properties rather than being genuine TA systems.

Analysis of Known TA Systems within the Identified Set

The locus tags of all genes from the 24 final families (Figure 2) were checked against a previously compiled list of known TA systems (Makarova et al., 2009). For a given family, if at least one pair was found in the list of known TA systems, the family was declared as "known." Families that were not declared as "known" were further similarly checked against a list of predicted TA genes downloaded from TADB (http://bioinfo-mml.sjtu.edu.cn/TADB/) (Shao et al., 2011). Families in which at least one pair was found in this TADB list were declared as "predicted."

Domain analysis of genes in the identified families was performed by searching their sequences against the Conserved Domain Database (CDD) (Marchler-Bauer et al., 2011) using rpsblast (ftp://ftp.ncbi.nih.gov/blast/documents/rpsblast.html) with e value threshold of 0.05.

For the phylogenetic distribution analysis of families (Figures 4B and 4C), the IMG cluster IDs (http://img.jgi.doe.gov/cgi-bin/w/main.cgi) of the toxin and antitoxin for each family were used to retrieve all adjacent pairs of genes in IMG that have the same two cluster IDs. For each family, number and identity of organisms carrying members of the family were extracted (Table S3 and Figures 4B and 4C).

Experimental Evaluation of "TA Cloning Pattern"

Toxins and antitoxins were amplified from their genomes of origin (in the case of *sanaTA*, *psyrTA*, *sdenTA*, and *hhalTA*) or synthesized (GenScript) for the families *pmenTA* and *rlegTA*. The toxin was then directionally ligated into the pRSFDuet-1 vector (EMD Chemicals Inc.) and the antitoxin ligated into the pBAD/HisA plasmid (Invitrogen). Since transformation of the toxin gene alone resulted in mutations in the toxin due to toxicity, the two plasmids (carrying the toxin and antitoxin) were cotransformed into *E. coli* BL21(DE3) pLysS (Stratagene) in the presence of 0.3% arabinose to induce the antitoxin. The clones were verified by direct sequencing with primers on the pRSFduet-1 and pBAD/HisA vectors.

For the toxicity assay on plates, clones were cultured in LB medium with 100 μg/ml ampicillin, 50 μg/ml kanamycin, 34 μg/ml chloramphenicol, and 0.3% arabinose overnight. The next day, a portion of each overnight culture was inoculated into fresh medium (10-fold dilution), and 10 μl was spotted on LB plates supplemented with 100 μg/ml ampicillin, 50 μg/ml kanamycin, and 34 μg/ml chloramphenicol. Toxin, antitoxin, or both were induced by 100 μM IPTG and 0.3% arabinose, respectively.

For the growth kinetics experiments, three different colonies of each system were cultured in LB medium with 100 μg/ml ampicillin, 50 μg/ml kanamycin, 34 μg/ml chloramphenicol, and 0.3% arabinose overnight. The next day cells were diluted 1:20 and measured for OD using 1 cm path cuvettes. Samples were equilibrated to the same OD, and 5 μl of these samples were added to 175 μl LB medium supplemented with 100 μg/ml ampicillin, 50 μg/ml kanamycin, and 34 μg/ml chloramphenicol in a 96-well microplate. Cells were placed in a microplate reader (Infinite M200) in a script employing OD measurement approximately every 5 min. Measurements were done using 486 ex/516 em bandwidth (Gain 95, 105, and 110) and OD at 600 nm overnight. For each of the colonies, the following treatments were applied: no induction, induction of 100 μM IPTG and/or 0.3% arabinose after ~4 hr, and induction of 100 μM IPTG after ~4 hr and of 0.3% arabinose after another ~2.5 hr. The overnight growth replicate values were averaged, and the measured OD values were plotted against time.

The viability assay (Figure 3C) was performed as described by Pedersen et al. (2002). Briefly, each strain was grown overnight at 37°C in LB medium containing 100 μg/ml ampicillin, 50 μg/ml kanamycin, and 34 μg/ml chloramphenicol. The next morning, cells were then diluted 1:1,000 in the same medium as above and grown for 3 hr. At time zero, cells were then washed once with LB, and then transcription of toxin was induced by 100 μM IPTG. At increasing time points after toxin induction (30, 60, 120, 180, 240, and 300 min), cells were plated in several dilutions on LB plates containing 100 μg/ml ampicillin, 50 μg/ml kanamycin, 34 μg/ml chloramphenicol, and 0.3% arabinose. CFUs were determined in the next morning by colony counting.

Plaque Assays

E. coli strains harboring the antitoxin only or the TA of *psyrTA*, *sanaTA*, *pmenTA*, *rlegTA*, and *sdenTA* binary TA systems were grown overnight in LB liquid medium supplemented with 35 μg/ml chloramphenicol, 100 μg/ml ampicillin, and 0.3% L-arabinose with or without 50 μg/ml kanamycin, respectively. Overnight cultures were diluted 1:100 in fresh LB medium supplemented with

inducers and antibiotics as above and aerated with shaking at 37°C until reaching OD 600 ≈ 0.5. Cultures were then centrifuged for 10 min and resuspended in LB to an OD 600 of exactly 0.5. Volumes of 200 μl of culture and 10 μl of the indicated T7 phages were mixed in 3 ml of warm 0.8% agar LB supplemented with 35 μg/ml chloramphenicol, 100 μg/ml ampicillin and 0.3% L-arabinose or 35 μg/ml chloramphenicol, 100 μg/ml ampicillin, 50 μg/ml kanamycin, 0.3% L-arabinose, and 50–100 μM IPTG for the antitoxin-only or TA-harboring cultures, respectively. The mixtures were immediately overlaid on LB plates supplemented with the above indicated inducers and antibiotics. Overlaid plates were incubated at 37°C for 3 hr, and plaques were then counted. For a given TA pair, EOP was calculated by dividing the number of plaque-forming units (PFUs) obtained for bacterial lawn expressing the toxin + antitoxin by the number of PFU obtained on the bacterial lawn expressing the antitoxin alone.

Expression of Lon and Gp4.5 Proteins and Coimmunoprecipitation Assay

Expression of Lon and Gp4.5 was performed using the expression vector pRSFDuet-1 (EMD Chemicals, Inc.). For coexpression experiments, the 4.5 gene (3′ His-tagged) and full-length *Lon* protease gene (5′ Flag-tagged) were cloned into the first and second expression cassettes, respectively. For reciprocal coIP experiments, the 4.5 gene (5′ Flag-tagged) and the full-length Lon protease (5′ His-tagged) were cloned into pRSFDuet-1 as described above. Induction was carried out at 37°C for 3 hr by addition of 200 μM of IPTG. Cell pellets were lysed by sonication in a buffer containing 50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1 mM EDTA, 1%(v/v) Triton X-100, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 1 μl/mL protease inhibitor cocktail (Set IV, EMD Chemicals, Inc.). Cell debris was removed by centrifugation at 4°C for 15 min at 18,000 *g*. Clear supernatants were transferred to 1.5 ml tubes and incubated on a rotator shaker at 4°C for 1 hr with 80 μl prewashed anti-Flag M2-agarose beads (Sigma, #A2220). The beads were washed three times with the buffer described above, and the Flag-tagged protein or protein complex was eluted using Flag-peptide (Sigma, #F3290) using the manufacturers' recommendations. Anti-Flag tag and anti-His tag Abs were purchased from Sigma.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and four tables and can be found with this article at http://dx.doi.org/10.1016/j.molcel.2013.02.002.

ACKNOWLEDGMENTS

We thank Shany Doron, Eyal Weinstock, Daniel Dar, Uri Gophna, Omri Wurtzel, Tal Dagan, Gil Amitai, and Debbie Lindell for comments and stimulating discussions. We also thank Ada Dantes for excellent technical support. R.S. was supported, in part, by NIH grant R01AI082376-01, ISF-FIRST program (grant 1615/09), ISF (grant 1303/12), ERC-StG program (grant 260432), and the Leona M. and Harry B. Helmsley Charitable Trust, and by a DIP grant from the Deutsche Forschungsgemeinschaft.

Received: June 20, 2012
Revised: November 21, 2012
Accepted: January 31, 2013
Published: March 7, 2013

REFERENCES

Aizenman, E., Engelberg-Kulka, H., and Glaser, G. (1996). An Escherichia coli chromosomal "addiction module" regulated by guanosine [corrected] 3′,5′-bispyrophosphate: a model for programmed bacterial cell death. Proc. Natl. Acad. Sci. USA *93*, 6059–6063.

Amitai, S., Kolodkin-Gal, I., Hananya-Meltabashi, M., Sacher, A., and Engelberg-Kulka, H. (2009). Escherichia coli MazF leads to the simultaneous selective synthesis of both "death proteins" and "survival proteins". PLoS Genet. *5*, e1000390. http://dx.doi.org/10.1371/journal.pgen.1000390.

Aoki, S.K., Diner, E.J., de Roodenbeke, C.T., Burgess, B.R., Poole, S.J., Braaten, B.A., Jones, A.M., Webb, J.S., Hayes, C.S., Cotter, P.A., and Low, D.A. (2010). A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. Nature *468*, 439–442.

Bergh, O., Børsheim, K.Y., Bratbak, G., and Heldal, M. (1989). High abundance of viruses found in aquatic environments. Nature *340*, 467–468.

Bernard, P., and Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. J. Mol. Biol. *226*, 735–745.

Bernstein, K.A., Gangloff, S., and Rothstein, R. (2010). The RecQ DNA helicases in DNA repair. Annu. Rev. Genet. *44*, 393–417.

Cherny, I., and Gazit, E. (2004). The YefM antitoxin defines a family of natively unfolded proteins: implications as a novel antibacterial target. J. Biol. Chem. *279*, 8252–8261.

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M.L., and Brüssow, H. (2004). Phage-host interaction: an ecological perspective. J. Bacteriol. *186*, 3677–3686.

Chopin, M.C., Chopin, A., and Bidnenko, E. (2005). Phage abortive infection in lactococci: variations on a theme. Curr. Opin. Microbiol. *8*, 473–479.

Christensen, S.K., Mikkelsen, M., Pedersen, K., and Gerdes, K. (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. Proc. Natl. Acad. Sci. USA *98*, 14328–14333.

Christensen, S.K., Pedersen, K., Hansen, F.G., and Gerdes, K. (2003). Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. J. Mol. Biol. *332*, 809–819.

Christensen, S.K., Maenhaut-Michel, G., Mine, N., Gottesman, S., Gerdes, K., and Van Melderen, L. (2004). Overproduction of the Lon protease triggers inhibition of translation in Escherichia coli: involvement of the yefM-yoeB toxin-antitoxin system. Mol. Microbiol. *51*, 1705–1717.

Daines, D.A., Wu, M.H., and Yuan, S.Y. (2007). VapC-1 of nontypeable Haemophilus influenzae is a ribonuclease. J. Bacteriol. *189*, 5041–5048.

Deveau, H., Garneau, J.E., and Moineau, S. (2010). CRISPR/Cas system and its role in phage-bacteria interactions. Annu. Rev. Microbiol. *64*, 475–493.

Engelberg-Kulka, H., Reches, M., Narasimhan, S., Schoulaker-Schwarz, R., Klemes, Y., Aizenman, E., and Glaser, G. (1998). rexB of bacteriophage lambda is an anti-cell death gene. Proc. Natl. Acad. Sci. USA *95*, 15481–15486.

Fico, S., and Mahillon, J. (2006). TasA-tasB, a new putative toxin-antitoxin (TA) system from Bacillus thuringiensis pGI1 plasmid is a widely distributed composite mazE-doc TA system. BMC Genomics *7*, 259.

Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S., and Salmond, G.P. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. Proc. Natl. Acad. Sci. USA *106*, 894–899.

Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V., and Storz, G. (2010). Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. Nucleic Acids Res. *38*, 3743–3759.

Gomez, J.E., and McKinney, J.D. (2004). M. tuberculosis persistence, latency, and drug tolerance. Tuberculosis (Edinb.) *84*, 29–44.

Gottesman, S. (2003). Proteolysis in bacterial regulatory circuits. Annu. Rev. Cell Dev. Biol. *19*, 565–587.

Goulard, C., Langrand, S., Carniel, E., and Chauvaux, S. (2010). The Yersinia pestis chromosome encodes active addiction toxins. J. Bacteriol. *192*, 3669–3677.

Han, W., Li, X., and Fu, X. (2011). The macro domain protein family: structure, functions, and their potential therapeutic implications. Mutat. Res. *727*, 86–103.

Hayes, F., and Van Melderen, L. (2011). Toxins-antitoxins: diversity, evolution and function. Crit. Rev. Biochem. Mol. Biol. *46*, 386–408.

Hazan, R., and Engelberg-Kulka, H. (2004). Escherichia coli mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. Mol. Genet. Genomics *272*, 227–234.

Hazan, R., Sat, B., and Engelberg-Kulka, H. (2004). Escherichia coli mazEF-mediated cell death is triggered by various stressful conditions. J. Bacteriol. *186*, 3663–3669.

Hood, R.D., Singh, P., Hsu, F., Güvener, T., Carl, M.A., Trinidad, R.R., Silverman, J.M., Ohlson, B.B., Hicks, K.G., Plemel, R.L., et al. (2010). A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. Cell Host Microbe *7*, 25–37.

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. Science *327*, 167–170.

Hurley, J.M., and Woychik, N.A. (2009). Bacterial toxin HigB associates with ribosomes and mediates translation-dependent mRNA cleavage at A-rich sites. J. Biol. Chem. *284*, 18605–18613.

Jiang, Y., Pogliano, J., Helinski, D.R., and Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of Escherichia coli gyrase. Mol. Microbiol. *44*, 971–979.

Jørgensen, M.G., Pandey, D.P., Jaskolska, M., and Gerdes, K. (2009). HicA of Escherichia coli defines a novel family of translation-independent mRNA interferases in bacteria and archaea. J. Bacteriol. *191*, 1191–1199.

Kim, Y., Wang, X., Ma, Q., Zhang, X.S., and Wood, T.K. (2009). Toxin-antitoxin systems in Escherichia coli influence biofilm formation through YjgK (TabA) and fimbriae. J. Bacteriol. *191*, 1258–1267.

Kimelman, A., Levy, A., Sberro, H., Kidron, S., Leavitt, A., Amitai, G., Yoder-Himes, D.R., Wurtzel, O., Zhu, Y., Rubin, E.M., and Sorek, R. (2012). A vast collection of microbial genes that are toxic to bacteria. Genome Res. *22*, 802–809.

King, G., and Murray, N.E. (1994). Restriction enzymes in cells, not eppendorfs. Trends Microbiol. *2*, 465–469.

Koga, M., Otsuka, Y., Lemire, S., and Yonesaki, T. (2011). Escherichia coli rnlA and rnlB compose a novel toxin-antitoxin system. Genetics *187*, 123–130.

Kuchta, K., Knizewski, L., Wyrwicz, L.S., Rychlewski, L., and Ginalski, K. (2009). Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. Nucleic Acids Res. *37*, 7701–7714.

Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nat. Rev. Microbiol. *8*, 317–327.

Lehnherr, H., and Yarmolinsky, M.B. (1995). Addiction protein Phd of plasmid prophage P1 is a substrate of the ClpXP serine protease of Escherichia coli. Proc. Natl. Acad. Sci. USA *92*, 3274–3277.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Drèze, P., and Van Melderen, L. (2011). Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. Nucleic Acids Res. *39*, 5513–5525.

Liu, M., Zhang, Y., Inouye, M., and Woychik, N.A. (2008). Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. Proc. Natl. Acad. Sci. USA *105*, 5885–5890.

Maisonneuve, E., Shakespeare, L.J., Jørgensen, M.G., and Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. Proc. Natl. Acad. Sci. USA *108*, 13206–13211.

Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. Biol. Direct *4*, 19. http://dx.doi.org/10.1186/1745-6150-4-19.

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.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. *39*(Database issue), D225–D229.

Masuda, H., Tan, Q., Awano, N., Yamaguchi, Y., and Inouye, M. (2012). A novel membrane-bound toxin for cell division, CptA (YgfX), inhibits polymerization of cytoskeleton proteins, FtsZ and MreB, in Escherichia coli. FEMS Microbiol. Lett. *328*, 174–181.

Mortier-Barrière, I., Velten, M., Dupaigne, P., Mirouze, N., Piétrement, O., McGovern, S., Fichant, G., Martin, B., Noirot, P., Le Cam, E., et al. (2007). A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. Cell *130*, 824–836.

Neubauer, C., Gao, Y.G., Andersen, K.R., Dunham, C.M., Kelley, A.C., Hentschel, J., Gerdes, K., Ramakrishnan, V., and Brodersen, D.E. (2009). The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. Cell *139*, 1084–1095.

O'Connor, L., Coffey, A., Daly, C., and Fitzgerald, G.F. (1996). AbiG, a genotypically novel abortive infection mechanism encoded by plasmid pCI750 of Lactococcus lactis subsp. cremoris UC653. Appl. Environ. Microbiol. *62*, 3075–3082.

Otsuka, Y., and Yonesaki, T. (2012). Dmd of bacteriophage T4 functions as an antitoxin against Escherichia coli LsoA and RnlA toxins. Mol. Microbiol. *83*, 669–681.

Pecota, D.C., and Wood, T.K. (1996). Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. J. Bacteriol. *178*, 2044–2050.

Pedersen, K., Christensen, S.K., and Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. Mol. Microbiol. *45*, 501–510.

Poole, S.J., Diner, E.J., Aoki, S.K., Braaten, B.A., t'Kint de Roodenbeke, C., Low, D.A., and Hayes, C.S. (2011). Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. PLoS Genet. *7*, e1002217. http://dx.doi.org/10.1371/journal.pgen.1002217.

Roberts, R.C., Ström, A.R., and Helinski, D.R. (1994). The parDE operon of the broad-host-range plasmid RK2 specifies growth inhibition associated with plasmid loss. J. Mol. Biol. *237*, 35–51.

Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shoresh, N., Biham, O., and Balaban, N.Q. (2010). Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. Proc. Natl. Acad. Sci. USA *107*, 12541–12546.

Schumacher, M.A., Piro, K.M., Xu, W., Hansen, S., Lewis, K., and Brennan, R.G. (2009). Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. Science *323*, 396–401.

Shao, Y., Harrison, E.M., Bi, D., Tai, C., He, X., Ou, H.Y., Rajakumar, K., and Deng, Z. (2011). TADB: a web-based resource for Type 2 toxin-antitoxin loci in bacteria and archaea. Nucleic Acids Res. *39*(Database issue), D606–D611.

Skorupski, K., Tomaschewski, J., Rüger, W., and Simon, L.D. (1988). A bacteriophage T4 gene which functions to inhibit Escherichia coli Lon protease. J. Bacteriol. *170*, 3016–3024.

Sorek, R., Zhu, Y., Creevey, C.J., Francino, M.P., Bork, P., and Rubin, E.M. (2007). Genome-wide experimental determination of barriers to horizontal gene transfer. Science *318*, 1449–1452.

Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat. Rev. Microbiol. *6*, 181–186.

Stern, A., and Sorek, R. (2011). The phage-host arms race: shaping the evolution of microbes. Bioessays *33*, 43–51.

Tan, Q., Awano, N., and Inouye, M. (2011). YeeV is an Escherichia coli toxin that inhibits cell division by targeting the cytoskeleton proteins, FtsZ and MreB. Mol. Microbiol. *79*, 109–118.

Tuchscherr, L., Medina, E., Hussain, M., Völker, W., Heitmann, V., Niemann, S., Holzinger, D., Roth, J., Proctor, R.A., Becker, K., et al. (2011). Staphylococcus aureus phenotype switching: an effective bacterial strategy to escape host immune response and establish a chronic infection. EMBO Mol. Med. *3*, 129–141.

van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M., and Brouns, S.J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem. Sci. *34*, 401–407.

Van Melderen, L., and Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: more than selfish entities? PLoS Genet. *5*, e1000437. http://dx.doi.org/10.1371/journal.pgen.1000437.

Van Melderen, L., Thi, M.H., Lecchi, P., Gottesman, S., Couturier, M., and Maurizi, M.R. (1996). ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions. J. Biol. Chem. *271*, 27730–27738.

Vázquez-Laslop, N., Lee, H., and Neyfakh, A.A. (2006). Increased persistence in Escherichia coli caused by controlled expression of toxins or other unrelated proteins. J. Bacteriol. *188*, 3494–3497.

Wang, X., Kim, Y., Hong, S.H., Ma, Q., Brown, B.L., Pu, M., Tarone, A.M., Benedik, M.J., Peti, W., Page, R., and Wood, T.K. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. Nat. Chem. Biol. *7*, 359–366.

Wozniak, R.A., and Waldor, M.K. (2009). A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. PLoS Genet. *5*, e1000439. http://dx.doi.org/10.1371/journal.pgen.1000439.

Zhang, X.Z., Yan, X., Cui, Z.L., Hong, Q., and Li, S.P. (2006). mazF, a novel counter-selectable marker for unmarked chromosomal manipulation in Bacillus subtilis. Nucleic Acids Res. *34*, e71. http://dx.doi.org/10.1093/nar/gkl1358.