Supporting Information

Fineran *et al*. 10.1073/pnas.0808832106

Table S1. Annotation of plasmid pECA1039

Name	Position (bp)*	% G + C	Size (aa)	PI	E value	% Id†	Significant BLAST hit‡ [organism]	Accession no.	Proposed function	Reference
Exc1	1–366*	51.09	121	9.88	3.00E-14	41	Plasmid ColE1 entry exclusion protein 1	NP_040377	Exclusion of ColE1 plasmids	Inoue, Uchida (1991) J Bacteriol 173:1208–1214
Rom	446–643	47.47	65	4.71	4.00E-10	52	Plasmid ColE1 RNA I inhibition modulator protein	NP_040367	RNA one modulator protein; inhibitor of plasmid replication	Inoue, Uchida (1991) J Bacteriol 173:1208–1214
OriV	789–791*	33.3	NA	NA				L39794	Origin of replication	Keenleyside, Whitfield (1995) J Bacteriol 177:5247–5253
RNAII	791–1383*	48.22	NA	NA	5.00E-43	68	Plasmid pWQ799 RNAII and RNAI genes [*Salmonella enterica* serovar Borreze]	L39794	Primer for plasmid replication	Keenleyside, Whitfield (1995) J Bacteriol 177:5247–5253
RNAI	1201–1343	44.05	NA	NA	3.00E-09	73	Plasmid pEC3 oriV-type replication unit [*Pectobacterium carotovorum* subsp. *carotovorum*]	D21263	Suppressor of plasmid replication; antisense RNA to RNAII	Nomura, Murooka (1994) J Ferment Bioeng 78:250–254
ToxI	1706–1957	36.5	NA	NA			No significant homologue			
ToxN	1958–2473	37.98	171	5.98	1.00E-10	31	Plasmid pSRQ900 AbiQ [*Lactococcus lactis* W-37]	NP_862552	Abortive infection	Emond, et al. (1998) Appl Environ Microbiol 64:4748–4756
ORF4	2585–2686	41.09	34	10.48			No significant homologue			
ORF5	2676–2837*	38.88	53	10.51			No significant homologue			
ORF6	2841–3170*	34.84	109	9.33	0.16	26	Resolvase domain [*Bacillus cereus* 03BB108]	ZP_02605729	Plasmid replication resolution	Direct submission of sequence (Dodson, et al. 2008)
ORF7	3331–3417	52.87	28	5.38	2.00E-05	68	Hypothetical protein ESA_03859 [*Enterobacter sakazakii* ATCC BAA-894]	YP_001439881		Direct submission of sequence (McClelland, et al. 2007)
ORF8	3794–3949	47.43	51	9.1	5.00E-13	67	Hypothetical protein CBO1652 [*Clostridium botulinum* A str. ATCC 3502]	YP_001254152		Sebaihia, et al. (2007) Genome Res 17:1082–1092
ORF9	3831–4118	44.09	95	9.87	1.00E-12	71	Hypothetical protein CKO_pCKO2p07159 [*Citrobacter koseri* ATCC BAA-895]	YP_001456637		Direct submission of sequence (McClelland, et al. 2007)
ORF10	3644–4183*	42.96	179	10.1	5.00E-06	59	Hypothetical protein SCB49_13150 [unidentified eubacterium SCB49]	ZP_01890101		Direct submission of sequence (Hagstrom, et al. 2007)
VapI	5261–5557*	53.19	98	4.98	7.00E-07	33	Plasmid pEU30p03 hypothetical protein similar to plasmid maintenance antidote protein VapI [*Erwinia amylovora*]	NP_943270	Virulence-associated protein; antidote protein to toxin-antitoxin system	Foster, et al. (2004) Appl Environ Microbiol 70:7539–7544

NA, not applicable.

*Denotes reading from a complementary strand.

†% identity determined using the GCG gap pairwise comparison.

‡BLASTp or BLASTn in cases of no predicted aa sequence. Each entry is the most relevant to this work and/or a significant hit.

Table S2. ToxIN functions as an abortive infection system

Phage	Plasmid	Adsorption (%)	Cell survival (%)	ECOI (%)	Burst size (%)
ϕA2	*toxI*, *toxN*	99.44 ± 0.04	0.23 ± 0.15	0.25 ± 0.13	0.002 ± 0.004
	toxI, *toxN*-FS	99.45 ± 0.04	0.55 ± 0.25	79.08 ± 0.13	95.32 ± 43.76
	Vector	99.59 ± 0.16	0.47 ± 0.08	100.00 ± 18.55	100.00 ± 18.40
ϕM1	*toxI*, *toxN*	88.44 ± 4.11	1.93 ± 0.51	10.79 ± 8.55	0.43 ± 0.03
	toxI, *toxN*-FS	86.75 ± 9.68	2.54 ± 1.51	101.15 ± 62.20	93.33 ± 39.46
	Vector	84.95 ± 9.28	3.12 ± 2.04	100.00 ± 45.56	100.00 ± 66.14

Table S3. ToxIN-type loci are present in diverse bacteria

Bacterium	ToxN, % aa id	Antitoxin promoter	Terminator	No. repeats	Consensus length, bp	Repeat consensus (5′-3′)	Accession no.
Alkaliphilus oremlandii OhILAs	32.7	Yes	Yes	2.2	40	TAATTCAAGGTATTTCGCAAC-CTTCTAAAAATGCGACTTG	YP_001513384
Bacillus thuringiensis serovar *konkukian* 97–27, pBT9727	22.9	Yes	Yes	1.9	51	GATAGGAGCAGTTTTTCGGAG-TACGCTCGGCTAAATA-AAAAACCGATACTT	CP000047
Bacillus thuringiensis serovar *kurstaki* HD73, pAW63	35.8	Yes	Yes	3.1	34	GTAACCTTACCGTAGTAGGT-GCTAAAAAATTGGT	AAZ06636
Bacillus weihenstephanensis KBAB4, pBWB402	35.7	Yes	Yes	2.9	34	GTAGCCTTACCGTAGTAGGCA-CAAAAAAATTGGT	YP_001642731
Clostridium botulinum C Eklund	36.0	Yes	Yes	ND	ND	ND	ZP_02621778
Clostridium perfringens C JGS1495	30.6	Yes	Yes	ND	ND	ND	ZP_02629773
Enterococcus faecium BM4147, pHTβ	25.9	No	No	ND	ND	ND	YP_398644
Erwinia carotovora **subsp *atroseptica* 1039 pECA1039**	100.0	Yes	Yes	5.6	36	AGGTGATTTGCTACCTTTAAGT-GCAGCTAGAAATTC	FJ176937
Fusobacterium nucleatum subsp *polymorphum* ATCC10953	39.1	Yes	Yes	3.2	35	AAAAAGCGGAAAACCCTGC-CGTTAAGTGGGAGTTA	ZP_02272649
Haemophilus influenzae biotype *aegyptius*, pF3031	30.5	Yes	Yes	2.7	31	GGCGCAAGCTGCCATAAGTGC-TACTTAATCA	NP_858107
Haemophilus influenzae biotype *aegyptius*, pF1947	31.8	Yes	Yes	2.7	31	GGCGCAAGCTGCCATAAGTGC-TACTTAATCA	YP_271719
Haemophilus influenzae biotype *aegyptius*, pF3028	30.5	Yes	Yes	2.7	31	GGCGCAAGCTGCCATAAGTGC-TACTTAATCA	NP_660242
Histophilus somni 2336	33.8	Yes	Yes	2.7	31	GGCGCAAGCTGCCATAAGTGC-TACTTAATCA	ZP_00133190
Lactococcus lactis subsp *lactis* W-37, pSRQ900	31.5	Yes	Yes	2.8	35	TATCCAAGCCATAGTTGGAAT-TGATAAAATTGGAG	NP_862552
Mouse cecal gut metagenome	34.9	5′ sequence unavailable	-	-	-	AATA01004437	
Ruminococcus torques ATCC 27756	26.4	Yes	Yes	ND	ND	ND	ZP_01968178
Ruminococcus torques ATCC 27756	37.3	Yes	Yes	3.2	38	AGGTGAGTCTCCACCATAAA-GAGAGAACCTAAAAAAGC	ZP_01968975
Thermosinus carboxydivorans Nor1	32.7	Yes	Yes	3.2	37	AAAAAAGCGGGGAAAC-CTCGCCCCTAAACTGAG-GATG	ZP_01667702
Vibrio harveyi BAA-1116	34.5	No	No	ND	ND	ND	YP_001444871

ND, no significant repeats detected by tandem repeats finder.

Table S4. Primer sequences

Primer	Sequence (5′-3′)	Restriction site
KD02	TTTTAAGCTTTTACGATCTTCAGTATGGG	*Hin*dIII
MJ7	TTTTGAATTCGTTTTATCGACATTGTGAACC	*Eco*RI
MJ12	TTTTAAGCTTCATATTTTTCTCGTAAAAAAGGCGACTATG	*Hin*dIII
MJ13	TTTTAAGCTTTTATTACTTGTCATCGTCGTCCTTGTAGTCCTCGCCTTCTTCCGTATCCTC	*Hin*dIII
PF134	CACGCATCTTCCCGACAAC	-
PF137	GATGAATTCGAGAAAAATATGAAATTC	*Eco*RI
PF146	CAATGTATTTGCTTGATATAGTGTAG	-
PF147	GTATCTAGAGTAGTCGCCTCTTTTACTTTATTAC	*Xba*I
PF164	TTTGAATTCAAGGTGATTTGCTACCTTTAAG	*Eco*RI
PF172	TTTCCATGGCCGTAGAAGAACAGC	*Nco*I
PF173	TTTCCATGGTTAGACATTATTTGCCGAC	*Nco*I
PF183	TTTGAATTCAAGGTGATTTACTACCTTTAAG	*Eco*RI
PF184	TTTAAGCTT*GAATTTCTAGCTGCACTTAAAGGTAGCAAATCACCT*TTGAATCTATTATAATTGTTATCCG (*toxI* repeat in italics)	*Hin*dIII
PF185	AAACAAATAGGGGTTCCG	-
PF186	TTTGAATTCAGAGTGGCAAGCATGC	*Eco*RI
PF187	TTTAAGCTTAACTTAAGTATACCTGATCTTCGG	*Hin*dIII
PF188	TTTGAATTCTTAAGTTAAGGTGATTTGCTACC	*Eco*RI
PF189	TTTAAGCTTATTTTTCTCGTAAAAAAGGCGAC	*Hin*dIII
PF190	TTTAAGCTT*GAGTTCCTTGCAGCAGAAAGTGGAAGCAAATCACCT*TTGAATCTATTATAATTGTTATCCG (mutant *toxI* repeat in italics)	*Hin*dIII
PF194	TTTGAATTCACAAAATTAGAGTAACCTTACTTTAG	*Eco*RI
PF195	TTTAAGCTTATCTCTCACGCCCC	*Hin*dIII
PF196	TTTAAGCTTCAACTTTCTTCTCCCCC	*Hin*dIII
PF197	TTTGAATTCGGAGAAGAAAGTTGACTAATAAAG	*Eco*RI
PF202	TTTAAGCTTGTAGTCGCCTCTTTTACTTTATTAC	*Hin*dIII
PF260	TTTAAGC*TTGAATTTCTAGCTGCACTTCAAGGTAGCAAATCACCT*TTGAATCTATTATAATTGTTATCCG (mutant *toxI* repeat in italics)	*Hin*dIII