

Figure S1. Phage-inducible chromosomal island genomes. Genomes are aligned according to the prophage convention, with the integrase gene (*int*) at the left end. Genes are coloured according to their sequence and function: *int* is yellow; transcription regulator (*alpA* or *merR*) is dark blue; replication genes are purple; encapsidation genes are green, with the terminase small subunit gene (*terS*) in light green; superantigen and other virulence genes are pink; genes encoding putative phage resistance proteins are black; other accessory genes are red; genes encoding hypothetical proteins are white.



Figure S2. Locations of the *att* sites in *Escherichia coli*. Phage *att* sites are indicated in orange, PICI *att* sites are in blue and P4 *att* sites are in black. The gene next to the *int* encoded by the different elements is indicated; in the outer are the genes flanking the phage *att* sites while in the inner are those genes flanking either the EcCI or the P4 *att* sites.



Figure S3. Location of the *tet*A and *cat* markers in the EcCI genomes.



Figure S4. Testing EcCICFT073 *int* **function. (a)** Detection of excision and circularization. DNA from *E. coli* CFT073 was extracted and PCR-amplified using specific primers recognizing the external and internal sequence of the element (integration: I), primers recognizing the chromosomal flanking sequences (excision: E) or PCR amplified using a pair of primers that hybridize divergently at both termini of the island (circularization: C). M: molecular weight marker. **(b)** pMAK700 derivative plasmids *int-att*_{PI}; *int-att*_L and $\Delta int-att_{PI}$. Strains carrying the different plasmids were grown and tested for integration in DH5 α (*rec*A-mutant) on selective agar at the restrictive (44°C) temperature.

cosQ = <u>TTTACGGGTCCTTTCC</u> cosN = <u>GGGCGGCGACCT</u> cosB = R3 = <u>AAGGCGTTTCCGTTCT</u> R2 = <u>AGAAAGGAAACGACAG</u> R1 = <u>CTGTCGTTTCCTTTCT</u>

cos phage

Lambda

Lambda

phi80

CFT073 cos.1

TTTATGGGTCCTTTCCGGCATATGGACCCGTTACG<mark>GGGCGGCGACCT</mark>CGCGGGTTTTCGCTATTTATGACGTTTTTCCGTGAAGGTGACACCACCACCACCACCACTATTATAATA TTTAACCATGCAGTTAAGGTAACATTATGATTGATAAAGCTTGTTTTGTAAGTCAGCAGGAAATAGCTGAACATTTCAAGGTTAACAGAACCACTATTCGCGCATGGACC AAACAGGGGATGCCGTATCTTAATGCGGATCGCGGAAAGTCTGGCGGTTATCACATCGGGCATACATTGCTTTGGTCTTCA

CFT073 cos.2

phi4 CFT073

TTTTCGGGTCCTTTCCGTCGATCCAACAGGTTACG<mark>GGGCGGCGACCT</mark>CGCGGGTTTTTCACTATTTATGAAAATTTTTCAGGGAAAATCGTGTCGGTACTTCTCGAATATA ACTTTTTGTTTTTTTTAATATTTGCATCCGTAAAGGTCCGACATGAAAGTGTCCGAAAATGCCTTTTTCTGGCGTTTTCATGTCGGGCCCTTGTATTTGATAATGGGTTGTT TTCATGAAGGTTAATAAAAAGAGGCCTGCCGAAATTTTCAACGTGGACCCGCGGACGATTGAACGCTGGCAGTCTCAGGGAC



Figure S5. Functionality of the EcCICFT073 *cos* **sites. (a)** Sequences of the different putative *cos* sites. While all the elements carry similar *cosQ* and *cosN* sequences, the PICI *cosB* sequences are completely divergent. **(b)** pET28a derivative plasmids, containing the different *cos* sequences diagrammed in panel (a), including the λ , ϕ 80, phi4 and EcCICFT073 *cos1* or *cos2* sequences were introduced in the lysogenic strains for phages λ or ϕ 80. The generated strains were MC induced (2 µg/ml) and the transfer (transduction) of the plasmid analysed.







Figure S6. Growth curve of the *E. coli* C600 derivative carrying EcCICFT073 and plasmid pBAD18 or pBAD-*alp*A. The growth curve was done over the course of 24 hours at 30 °C using OD_{600} measurements (a) or CFU/mL counting (b). 0.02% arabinose was added at time 0 to express *alp*A. Each point represents the means of results from three independent experiments.



Figure S7. Map of the *P. multocida* **phages present in strains Pm86 (MH238467) and Pm172 (MH238466).** Genomes are aligned according to the prophage convention, with the integrase gene (*int*) at the left end. Genes are coloured according to their sequence and function: *int* is yellow; transcription regulator is dark blue; replication genes are purple; encapsidation genes are green, with the terminase small subunit gene (*terS*) in light green; superantigen and other virulence genes are pink; other accessory genes are red; genes encoding hypothetical proteins are white.



Figure S8. Comparison of PICI and P4 genomes. Genomes are aligned according to the prophage convention, with the integrase gene (*int*) at the left end. Genes are coloured according to their sequence and function: *int* is yellow; transcription regulator (*alpA* or *vis*) is dark blue; replication genes are purple; encapsidation genes are green, with the terminase small subunit gene (*terS*) in light green; superantigen and other virulence genes are pink; genes encoding putative phage resistance proteins are black; other accessory genes are red; genes encoding hypothetical proteins are white.

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

	10	0	20	. 30	40	50
O42_AlpA (CBG34595.1)	<mark>M L S</mark>	T D R F <mark>T</mark> R <mark>E</mark> K <mark>E</mark>	C E <mark>KL</mark> TGLS	- <mark>RT CRY</mark> RLI	EKA <mark>G</mark> Q F <mark>P</mark> SRR	K L <mark>G</mark> G R
ATCC25922_AlpA (EOR52814.1)	<mark>MEA</mark>	IRKI <mark>I</mark> FRQE	V K <mark>KI</mark> IHIK	A <mark>ds t</mark> lqsm:	INA <mark>g</mark> e F <mark>P</mark> QgF	R <mark>V G L</mark> R
IHE3034_AlpA (ADE91716.1)	<mark>MR</mark> EIN	E D R V <mark>I</mark> R E D E	C R <mark>KL</mark> TGVC	- <mark>rt t</mark> ryeli	EKK <mark>g</mark> s F <mark>p</mark> srr	N L <mark>G</mark> G R
RM13514_AlpA (AHG11287.1)	M T A N H Q E <mark>A I</mark> P	T T G Y <mark>I R R C N</mark>	L A <mark>KL</mark> LDTS	- <mark>VS T</mark> IDRW	VRI <mark>G</mark> K LPKPT	K L <mark>G</mark> E K
Consistency	<mark>0000000</mark> 443	4 3 4 4 <mark>*</mark> 5 <mark>6</mark> 2 7	4 4 <mark>*8</mark> 5354	2 <mark>7 6</mark> 43554	4 5 4 <mark>*</mark> 5 7 * 5 3 3	<mark>5 8 * 2</mark> 8
	60	0	70			
O42_AlpA (CBG34595.1)	S V G W S L S <mark>E V</mark> L	A <mark>W</mark> KD <mark>SC</mark> KAV	Н			
ATCC25922_AlpA (EOR52814.1)	R R G W Y E D <mark>D V</mark> L	A <mark>W</mark> QKEREQE	A RG <mark>T</mark> AA			
IHE3034_AlpA (ADE91716.1)	S V G W L L S <mark>E V</mark> M	E <mark>W</mark> VK <mark>SR</mark> DRI	N SG <mark>K</mark> AA			
RM13514_AlpA (AHG11287.1)	N T V F N A V E I N	H <mark>W</mark> LNERRNN	т ко			
Consistency	<mark>5 4 5 7 2 3 4 8 9</mark> 4	4 <mark>*</mark> 34 <mark>65</mark> 433	3 22011			

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

				. 10)			:	20.					30.					40)				. 50
O42_AlpA (CBG34595.1)									- 1	M L S	S T I	D <mark>r f</mark>	IR	ΕI	K E C	EK		T G <mark>I</mark>	S	- R	ТС	RY	R L	E K
ATCC25922_AlpA (EOR52814.1)									- 1	٩ <mark>Ε</mark>	AI	R K J	IF	R	2 <mark>E</mark> V	K K	I	с н <mark>з</mark>	Γĸ	A D	S 1	ĽÇ	SM	IN
IHE3034_AlpA (ADE91716.1)								- M	R I	EII	NE	D <mark>R \</mark>	/ I R	ΕI	D <mark>E</mark> C	R <mark>K</mark>	L	۲G <mark>۱</mark>	/C	- R	T٦	RY	ΕL	E K
Vis_P4 (P12552.1)	MQA	VFS	SP	S P	AP	VTI	PLM	PL	P	D <mark>I</mark> .	тę	e r f	LR	νı	P <mark>E</mark> V	MH	I L	C G <mark>I</mark>	S	- R	S 1	ΙΥ	ΕL	I R
RM13514_AlpA (AHG11287.1)						- M -	T A N	ΗQ	E /	A I I	P T	T <mark>G \</mark>	I R	R	C <mark>n</mark> l	AK	L	L D <mark>1</mark>	r s	- V	S 1	ID	RW	<mark>V</mark> R
Consistency	000	000	000	00	000	000	900	01	03	3 <mark>5</mark> 4	4 3 ·	4 <mark>5 5</mark>	5 <mark>9</mark> 6	4	2 <mark>7</mark> 5	37	' 9 <mark>4</mark>	44 <mark>6</mark>	55	<mark>0</mark> 4	77	44	55	<mark>4</mark> 6
				. 60)				70.					80.										
O42_AlpA (CBG34595.1)	A <mark>G</mark> Q	F <mark>P</mark> S	RR	. 60 K L	G <mark>G</mark> I	R S V	∕G <mark>k</mark>	S L	70. S <mark>I</mark>	E V I	LA	N K C	 <mark>SC</mark>	80. K/	A V H	 			-					
O42_AlpA (CBG34595.1) ATCC25922_AlpA (EOR52814.1)	A <mark>G</mark> Q A <mark>G</mark> E	FPS FPQ	R R I Q G F I	. 60 K L R V	G G I	R S V R R I	V G <mark>M</mark> R G M	S L Y E	70. SI	E V I D V I	L A L A	и к с и к с и Q к	S C (E R	80. K/	Q E A	 R 6	 	• • •	-					
O42_AlpA (CBG34595.1) ATCC25922_AlpA (EOR52814.1) IHE3034_AlpA (ADE91716.1)	A <mark>G</mark> Q AGE KGS	FPS FPS	R R G F R R	. 60 K L R V N L	G G I G L I G G I	R S V R R I R S V	VGW RGW VGW	S L Y E L L	70. SI DI	E V I D V I E V I	L A L A M E	M K E M K E	S C (E R (S R	80. K / E (D	AVH QEA RIN	 R (i T /	4 A - 4 A -	-					
O42_AlpA (CBG34595.1) ATCC25922_AlpA (EOR52814.1) IHE3034_AlpA (ADE91716.1) Vis_P4 (P12552.1)	A <mark>G</mark> Q A G E K G S K G E	FPS FPQ FPS FP	R R G F R R Q V	. 60 K L R V N L S L	G G I G C I G G I G G I	R S V R R I R S V K N V	/GW RGW /GW /AW	S L IY E I L L	70. SI DI SI	E V I D V I E V I E V I	L A L A M E T A	W K C W Q K W V K W V K	S C (E R (S R (G R	80. K / E (D I /	AVH QEA RIN AGR	R 0 S 0	i T / i K /	ΔΑ ΔΑ Υ D /	- -					
O42_AlpA (CBG34595.1) ATCC25922_AlpA (EOR52814.1) IHE3034_AlpA (ADE91716.1) Vis_P4 (P12552.1) RM13514_AlpA (AHG11287.1)	AGQ AGE KGS KGE IGK	FPS FPQ FPS FPP LPK	RR QGF RR QV CPT	. 60 K L R V N L S L K L	G G I G L I G G I G G I G G I	R S V R R I R S V K N V K N V	/GW RGW /GW /AW FVF	SL YE LL LH	70. SI DI SI SI		L A L A M E T A N H	W K C W Q K W V K W M A W M A	S C (E R (S R (G R I E R	80. K/ E (D I / R	AVH QEA RIN AGR	R 0 S 0 K R	i T / i K / i G ` 0	4 A 4 A 7 D /	-					

Figure S9. Comparison of different EcCl AlpA and Vis P4 proteins. (a) Alignment generated by PRALINE (matrix BLOSUM62) of AlpA protein sequences from EcClO42 (CBG34595), EcClATCC25922 (EOR52814), EcClIHE3034 (ADE91716), and EcClRM13514 (AHG11287). **(b)** Alignment generated by PRALINE (matrix BLOSUM62) of AlpA protein sequences from EcClO42, EcClATCC25922, EcClIHE3034, EcClRM13514 and Vis protein from P4 (P12552). **(c)** Radial phylogenetic tree resulting from alignment of AlpA and Vis sequences using CLC Genomics Workbench. The accession number for the protein sequences are as follow: EcClCFT073 (Identical to EIL17504.1), EcClRM12579 (AEZ39958), EcClO42 (CBG34595), EcClK1516 (EZB64731), EcClATCC25922 (EOR52814), EcClED1a (CAR08823.2), EcClIHE3034 (ADE91716), EcCl11128 (BAI36743), EcCl1368 (BAI26316), EcClD14 (identical to EGB43710.1), EcCl224 (Identical to ACB19277), EcClEC2733.1 (Identical to ERB13914.1), EcCl18098 (Identical to EFJ60051), EcClRM13514 (AHG11287) and Vis protein from P4 (P12552). The tree was generated using the following parameters; algorithm = UPGMA, distance measure = Jukes-Cantor, bootstrap = 1000 Replicates.

b



Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

			. 80 90	110
CFT073_pri (AAN79964.1)MKLAP_NVKQQSRGIK	<mark>HKE</mark> TEVIIFA GSDAW <mark>SHAKQ W</mark> QE <mark>H</mark> DARMAG	DN <mark>EPPVW</mark> LGE QQL <mark>S</mark> EL <mark>DK</mark> LQ IVPEGRKS	V <mark>R IFRAGYLA</mark> PV MIKA <mark>IGQKLA AAGV</mark> QDANF	Y P <mark>DGMH</mark> GQKVE NWREYL <mark>ARER QNL</mark> SDGLVIE LPVKQ
RM13514_pri (AHG11293.1)MKSAP NLKKQPRGKK	<mark>had</mark> teviifa gsdaw <mark>ahakq w<mark>leq</mark>dgklag</mark>	DN <mark>IPPVV</mark> LAD EQL <mark>K</mark> DIGNLQ IVPDGRKS	AR IYKAGHL <mark>D</mark> QV MVKGIGQKLA AAGVQDADY	Y PEGMH <mark>hner</mark> q nwrnyl <mark>eter kn</mark> isdglvie lpvkkkergk
ED1a_pri (CAR08829.2) MERAEMKKAP NLKHQPR C	<mark>dkm</mark> teviifa gsdawahakq w <mark>qeqd</mark> grlag	DN <mark>VPPVWL</mark> GE QQL <mark>AELDK</mark> LQ IVPEGRKS	VR IFRAGHL <mark>E</mark> PV MIKAIGQKLA AAGVQDANF	Y PDGMH <mark>GQEV</mark> E NWREYLARER <mark>QNLSDGLVIE FPVKKKDTGS</mark>
	<mark>DKM</mark> TEVIIFA GSDAWAHAKQ W <mark>QEQDGRL</mark> AG	DNVPPVVLAD DQLDELADLR IIDEGRYC	VR LYKAGHI <mark>r</mark> ps ninaiahkla aagvtdany	Y PEGMH <mark>SHMR</mark> E NWREYLERVR <mark>GKEPAEEKKH QRKT</mark>
Consistency 00000**44*****8*5*66101	<mark>563</mark> ****** ****8**** *6*6*788**	** <mark>5***</mark> 3*67 6** <mark>4</mark> 88 <mark>44</mark> *7 *968**66	7 <mark>* 877**78</mark> 366 6977*76*** ****6**77	* *7 <mark>***</mark> 45448 ***7**566* 4746665656 4666 <mark>311010</mark>
160 170	180 190 20	0 210 220	230 240 250	260 270 280 200 200
CET073 $pri(\Delta\Delta N79964.1) = = = = KAOLS OMADSERAOL$		GVWCPVS TMELSREMVA TVSEHRAT		
RM13514 pri (AHG11293.1) GDDATSLALN OMGASORGEV			YS ONATKSAVD	V MONTARNITO ESNOVE DIRTODEREH DKNDWLITAS
ED1a pri (CAR08829.2) HSDDELKPRV ESRADGVFWV 1	TPKVDKOSGE IIRPETWLCS PLELLGTGTI	GKEHYRVMRW KKLANHEVIT MAIPCGGI	GD RDGWRLLKDH GLNVTTNGKY RAILADWMO	L SGSHEEWOLS TTTGWHEDAY IMPDGSIIGD SEKPILETGK
O42_pri (CBG34588.1)TLPMS VGSTGYDTQL -	D YVVKGIIP AVSLCSIYG	ASGSYKSFLA G	<mark>SWACHV</mark> A	R R <mark>VAHGAVLYV VGE<mark>g</mark>gigv</mark> pr rvkawevvhd eqvkn <mark>l</mark> ylvn
Consistency 0010034454 5344522247 2	210000000 <mark>5</mark> 24114 <mark>5</mark> 2412 2 <mark>544<mark>7</mark>23123</mark>	2 <mark>000</mark> 212221 34243 <mark>5555</mark> 4 31220111	13 242322 <mark>5</mark> 430 000000000 00 <mark>545</mark> 2233	2 4 <mark>543431</mark> 652 445 <mark>*0000</mark> 43 2 <mark>554565</mark> 32 <mark>5 44332</mark> 83525
		0	. 380	410
CFT073_pri (AAN79964.1) GIEYTPPAPG ENIRDNAPNF H	HKWLEHAAGK DPRK-MMR <mark>I</mark> C A <mark>AL</mark> YMIMANR	YDWQ-MFIEA TGDGG <mark>SGKS</mark> T FTH <mark>I</mark> ASLL	A <mark>g kontvsaemt slddagg-ra ovvg</mark> sr <mark>li</mark> v	L <mark>ADQPKYTGEGTGIK</mark> KI TGGDPVE <mark>I</mark> NP KYEKRFTAVI
RM13514_pri (AHG11293.1) ELPFTPPAEG ETLATHAPNE M	NKWLRRSVAE NDRK-ADRVL AALFMVLANR	YDWQ-LFIEV TGPGGSGKSV MAEICTML	A <mark>G KANTVSASMK ALEDARE-RA LVVG</mark> FS <mark>LI</mark> I	L PDMTRYAGDG <mark>AGIKAI TGGDKVAI</mark> DP KHKAPYSTRI
ED1a_pr/ (CAR08829.2) SAAINGYSV AGTAEGW R	RDSVARLAGG NPSM-MLGIA TSLAAPLIGL	VGADGFGVHL FEQSSAGKTT TQNIASSL	W <mark>G EPDSQRLTWY GTALGIANEA ESHN</mark> DGLLP	L DEIGQAGNAREVSTSAYT LFNGSGKLQG AKDGGNREIK
Consistency 222 E442 E21 1012127444	DEALARQVERE IGREVENTIL DILARCEGON	DENDSRD MGAFIRGC	DELK RRIGATVLV	V HHSGKDEIKG ARGSSAFRAS LDA-EYRIRR EDAGSEAL
Consistency 3323443321 10121374444 1	1 555454444 5554044254 40 4450404	5215022422 22154 <mark>567</mark> 54 45574550	2 <mark>5 5155222221 22215010</mark> 4 <mark>6 5557</mark> 55 <mark>88</mark> 5	6 2555 6500 46 00 456 445 54 554245 48 45 5244224512
		0	530	560
CFT073_pri (AAN79964.1) RAVVLATNNN PMIFTERAGG	<mark>VARR RVIFRFDN</mark> <mark>IVSEAE</mark> KD	RELPEKIAAE IPVIIRRLLA NFADPEKA	RA LLIEQRDGDE ALAIKQOTDPVIEF	C QF <mark>l</mark> nfleea r _{glmmggggds vkytt<mark>r</mark>nsly rv<mark>ylafm</mark>aya}
RM13514_pri (AHG11293.1) PAVVLAVNNS AMSFSDRSGG -	<mark>ISRR RVIFNF</mark> SE <mark>VVPENE</mark> RD	PMLPEKIEGE LA <mark>v</mark> virh <mark>l</mark> lt rfa <mark>d</mark> qdea	RR LLYEQQKSEE AL <mark>AIKR</mark> EGDSLVDF	C GY <mark>L</mark> MSSVMC D G <mark>LLVG-NAEI IPFSPR</mark> RYLY HA <mark>YLSY</mark> MRAH
ED1a_ <i>pri</i> (CAR08829.2) <mark>hwrtvaistg emdvet</mark> flkt <mark>e</mark>	<mark>GIKVK</mark> AGQL VR <mark>l</mark> lnvpm <mark>ek ath</mark> fh <mark>e</mark> ystg	KA <mark>hadalkda wt<mark>e</mark>nhga<mark>a</mark>gr ewvkwlag</mark>	HQ QEAKDTVREC RE <mark>rwrnlipe syge</mark> q <mark>v</mark> hrv	G ER <mark>F</mark> AILEAAL V <mark>L</mark> SGH-VTGW AAQEC <mark>R</mark> DAIQ HN <mark>FNAW</mark> VKEF
О42_ <i>pri</i> (СВ G34588.1) <mark></mark>VISCTKMКDAEELKE -	<mark>AAYD LR</mark> VVEL <mark>FT</mark> <mark>DA</mark> DGELI	TS <mark>LVVVDDPR PP<mark>v</mark>eler<mark>i</mark>ee Agnktenh</mark>	T <mark>Alwgc Ir<mark>srtqngdk ctip</mark>l<mark>l</mark>rd</mark> d	M <mark>KK</mark> lgy <mark>emknf rrwlyk</mark> lekd
Consistency 1156565542 1523453443 0	000000 <mark>65</mark> 43 34 <mark>8</mark> 55523 <mark>00 00133</mark> 82 <mark>7</mark> 33	33 <mark>5</mark> 4444424 24 <mark>6</mark> 3444 <mark>6</mark> 23 334 <mark>5</mark> 2434	42 2213342243 43 <mark>5255</mark> 32 <mark>00 0005</mark> 2 <mark>7</mark> 253	3 43 <mark>7</mark> 2322120 2 <mark>5</mark> 21201110 21121 <mark>7</mark> 2345 52 <mark>6543</mark> 7432
CET072 pri (A ANZ0064 4) P o v P v V V P v V P V V P V V P V V V V V)	. 680 690 700	/10/20/30
RM13514 pri (AHG11293.1) prs v pyti tp. r c apmp c ana r				
ED1a pri (CAR08829.2) PTCN-PEEKO MVEDAEAELS S		TENEDDEEDE VIEDUVEEGE TAGENDS		
O42 pri (CBG34588.1) GVTR-TDGDD VAPL	TOTOTICET FROMERE ROLAGTRADO	TRUEDELINI IIIIINILOL IAQUINPS	ALUKIUK	
Consistency 235144234 5432211222 1	L3 <mark>5</mark> 1322000 000000011 3121210000	000000122 1212211310 4000000	00000000 000000000 0000000 00000000	0 00000000 000000000

Figure S10. Comparison of different EcCl Pri-Rep and α P4 proteins. (a) Alignment generated by PRALINE (matrix BLOSUM62) of Pri-Rep protein sequences from EcClCFT073 (AAN79964), EcClRM13514 (AHG11293), EcClO42 (CBG34588) and EcClED1a (CAR08829). (b) Alignment generated by PRALINE (matrix BLOSUM62) of Pri-Rep protein sequences from EcClCFT073, EcClRM13514, EcClO42, EcClED1a and α protein from P4 (P10277). (c) Radial phylogenetic tree resulting from alignment of Pri-Rep and α sequences using CLC Genomics Workbench. The accession number for the protein sequences are as follow: EcClCFT073 (AAN79964), EcClRM12579 (AEZ39963), EcClO42 (CBG34588), EcClK1516 (EZB64739), EcClATCC25922 (EOR52820), EcClED1a (CAR08829), EcClIHE3034 (ADE92165), EcCl11128 (BAI36753), EcCl11368 (BAI26324), EcClD14 (AER87766), EcCl224 (Identical to KRR57888), EcClEC2733.1 (Identical to WP_053285629.1), EcCl18098 (Identical to EQQ34345), EcClRM13514 (AHG11293) and α protein from P4 (P10277). The tree was generated using the following parameters; algorithm = UPGMA, distance measure = Jukes-Cantor, bootstrap = 1000 Replicates.



CFT073_pri (AAN79964.1) RM13514_pri (AHG11293.1) α_P4 (P10277.1) ED1a_pri (CAR08829.2) O42_pri (CBG34588.1) Consistency or		0 30 PALGIQVLKN RHQPCPV 0000000000 00000000	. 40 50		.80	100110120	MAKTRTGTGN AYLTRKGFPG RECRML	150 TGTH 00000
 CFT073_ <i>pri</i> (AAN79964.1) - RM13514_ <i>pri</i> (AHG11293.1) - α_P4 (P10277.1) R ED1a_ <i>pri</i> (CAR08829.2) N O42_ <i>pri</i> (CBG34588.1) - Consistency B		70	. 190 200 HA KQWQEHDARM AGDNEP HA KQWLEQDGKL AGDNIP KG GQVRGTCHTL EQQQAG HA KQWQEQDGRL AGDNVP HA KQWQEQDGRL AGDNVP 77 6 65656568 7 7 4 7	. 210	.230240 GR KSVRIFRAGY LAPVMIKAI GR KSARIYKAGH LDQVMVKGI T- VMVALSSVNL LSLASLARQ GR KSVRIFRAGH LEPVMIKAI GR YCVRLYKAGH IRPSNINAI G6 4587856875 944558556	25026027(GQKLA GQKLA K HPACQIVLAA DRDLSGDGQK GQKLA GQKLA A		300 A R E R E T E R T Q Y G A R E R E R V R 5 6 4 6
CFT073 _pri (AAN79964.1) RM13514_pri (AHG11293.1) α_P4 (P10277.1) ED1a_pri (CAR08829.2) O42_pri (CBG34588.1) Consistency 44	310	20	. 340	360370. F0 – GVCVHPESEI VHVWRGGV YGG ELAINADSDT VHVWRSV YGE ALAVDANQQL LSRYENGV LCS PLELLGTGTI GKEHYRVM TGY DTQLDYVVKG IIPAVSLC 441 3637333443 43232435	.380	400		
CFT073_pri (AAN79964.1) p RM13514_pri (AHG11233.1) p a_P4 (P10237.1) p ED1a_pri (CAR08829.2) I O42_pri (CBG34588.1) R Consistency 4	4604 KTGEFSPH TPENWITTHN RTGDFREH DKNDWLLIAS FQNGTFHPH SPSHWRTLC APDGSIIGD SEKPILFTGK /KAWEVVHD EQVKNLYLVN 554656235 5433373523	70	. 490	. 510	.530540 EA TGDGGSGKST FTHIASLLA EV TGPGGSGKSV MAELCTMLA EV TGPGGSGKSI MAELATLLA HL FEQSSAGKTT TQNIASSLW NDSRD MGAFIRGCD 43 3324467764 544755373	550		600 KKIT KAIT KAIT YTLF FRAS 44 <mark>6</mark> 5
CFT073_ <i>pri</i> (AAN79964.1) g RM13514_ <i>pri</i> (AHG11293.1) g a_P4 (P10277.1) g ED1a_ <i>pri</i> (CAR08829.2) w O42_ <i>pri</i> (CBG34588.1) - Consistency at	6106 DPVEINPK YEKRFTAVIR DBVAIDPK HKAPYSTRIP DAVSVDPK YRDAYSTHIP SGKLQGAK DGGNREINHW LDAEYR IRREDAGSEA 541534448 3433353342	20	. 640	660670 IVSEAEKDRE LPEKIAAE VVPENERDPM LPEKIEGE QIAPQERDPQ LKDKITRE KAT HFHEYSTGKA HADALKDA ADGELITS LVVVDDPR 000 1346274434 63555325	.680	700710720 L IEQRDGDEAL AIKQQT L YEQQKSEEAL AIKREG L QSQQNSDEAL NIKRDA L AKDTVRECKE NWRNLIPESY ALWGCIR SRTQNGDKCT 3 1344235454 4365330000		750 DSVK EIIP SIIP GWAA
CFT073_ <i>pri</i> (AAN79964.1) RM13514_ <i>pri</i> (AHG11293.1) a_P4 (P10277.1) Ri ED1a_ <i>pri</i> (CAR08829.2) <u>pr</u> O42_ <i>pri</i> (CBG3458.1) Consistency 12		70	. 790. . 800. VY GHEY EY GREY EY GREY SF GFSRYLPYPN SDERDLP 25 6133000000		.830	850	LGKIGGKQHV FYVLMFQPEA EDYPPC	 2 0 0





Figure S11. Genome comparison of the EcCl and P4 elements integrated at the same *attB* site. Genomes are aligned according to the prophage convention, with the integrase gene (*int*) at the left end. Genes are coloured according to their sequence and function: *int* is yellow; transcription regulator (*alpA*) is dark blue; replication genes are purple; encapsidation genes are green, with the terminase small subunit gene (*terS*) in light green; superantigen and other virulence genes are pink; genes encoding putative phage resistance proteins are black; other accessory genes are red; genes encoding hypothetical proteins are white.