

## Supplementary Data

Supplementary Table I. Location of repeat tracts in four *Campylobacter jejuni* genome sequences

Location <sup>1</sup>	Homolog in strain 1221 <sup>2</sup>	PV <sup>3</sup>	Homolog in strain 81-176 <sup>2</sup>	PV <sup>3</sup>	Homolog in strain 81116 <sup>2</sup>	PV <sup>3</sup>	Length/Tract/Position In gene <sup>4</sup>
Phase Variable Tracts in CjNCTC11168							
<i>cj0031</i> 1243 aa (48,995-49,006)	<i>CJE0031</i> (90%)	-	<i>CJJ81176_0068</i> (81%)	-	<i>C8J_0034</i> (66%)	-	10/G/M
<i>cj0045c*</i> 240 aa (65,747-65,758) <sup>5</sup>	<i>CJE0044</i> (80%)	+	<i>CJJ81176-0083</i> (pseudo) (check region)	+	<i>C8J_0048</i> (78%)	+	9-10/G/E
<i>cj0046</i> , pseudo 1520 nts (67,707-67,719)	<i>CJE0046</i> (pseudo) (92%)	-	<i>CJJ81176_0085 to 0089</i> (pseudo) (98%)	+	<i>C8J_0049</i> (pseudo?) (96%)	-	9-11/G/NR
<i>cj0170</i> 247 aa (167,050-167,794)	none	-	<i>CJJ81176_0206/0207</i> (92%)	+	none	-	9/G/M
<i>cj0275</i> , <i>clpX</i> 417 aa (252,477-252,486)	<i>CJE0324</i> (99%)	+	<i>CJJ81176_0302</i> (99%)	-	<i>C8J_0252</i> (98% nts)	-	7-8/G/
in front of <i>cj0565*</i> 614 nts (527,377-527,390)	In front of <i>CJE0670</i> (363/614 nts)	+	In front of <i>CJJ81176_0590</i> (327/614 nts)	+	in front of <i>C8J_0526</i> (323/614 nts)	+	10-12/G/NR
<i>cj0617</i> 418 aa 1200nts (577,585-577,595)	<i>CJE0720</i> (98% nts)	+	<i>CJJ81176_0646</i> (98% nts)	+	none	-	9/G/M
<i>cj0628</i> 1144 aa (588,368-588,379)	none	-	none	-	none	-	10/G/M
<i>cj0676</i> , <i>kdpA</i> , pseudo 1657 nts (628,169-628,179)	<i>CJE0774</i> (pseudo) (99%)	(+)	<i>CJJ81176_0699</i> (98%)	-	<i>C8J_0628</i> (98%)	-	9/G/NR
<i>cj0685c</i> , <i>cipA*</i> 450 aa (639,006-639,016)	<i>CJE0783</i> (94%)	+	<i>CJJ81176_0708</i> (95%)	+	<i>C8J_0652</i> (91%)	+	9-11/C/M
in front of rRNA* 600 nts (695,942-695,953)	Conserved 774,689 (99%)	+	Conserved 692,448 (96%)	+	Conserved 706,131 (97%)	+	9-11/C/NR
<i>cj1139c</i> , <i>wlaN</i> 303 aa (1,074,192-1,074,201)	None	-	none	-	None	-	8/G/M
<i>cj1144c</i> 281 aa (1,079,739-1,079,748)	None	-	none	-	None	-	8/G/M
<i>cj1295*</i> 435 aa (1,227,120-1,227,130)	<i>CJE1487</i> (95%)	+	<i>CJJ81176_1312/1313</i> (92%)	+	<i>C8J_1238</i> (93%)	+	9-15/G/S

<i>cj1296/1297</i> 263 aa (1,228,590-1,228,600)	none	-	none	-	none	-	9/G/M
<i>cj1305c</i> 405 aa 1241 nts (1,234,921-1,234,931)	none	-	CJJ81176_1321 (90% nts)	+	none	-	9/G/M
<i>cj1306c</i> 408 aa 1194 nts (1,236,160-1,236,170)	CJE1495/1494 (99% nts)	+	none	-	none	-	8-9/G/M
<i>cj1310c</i> 404 aa (1,239,673-1,240,887)	CJE1504 (80%)	+	none	-	none	-	8-9/G/M
<i>cj1318 maf1</i> 649 aa (1,246,845-1,246,857)	none	-	none	-	none	-	11/G/S
In front of <i>cj1321</i> (1,250,805-1,250,816)	none	-	none	-	none	-	10/G/NR
<i>cj1325</i> 224 aa (1,253,417-1,254,092)	none	-	none	-	none	-	10/G/M
<i>cj1335, maf4</i> 649 aa (1,263,507-1,265,454)	CJE1524 (82%)	-	none	-	none	-	9/G/S
<i>cj1342c, maf7</i> 413 aa (1,274,694-1,275,935)	CJE1531 (96%)	+	CJJ81176_1341 (75%)	-	C8J_1258 (91%)	+	9/G/M
<i>cj1420c</i> 257 aa (1,353,375-1,354,418)	none	-	CJJ81176_1419 (99%)	+	C8J_0641 (98%)	+	9/G/M
<i>cj1421c</i> 612 aa (1,354,215-1,356,053)	none	-	none	-	none	-	9/G/
<i>cj1422c</i> 625 aa (1,356,108-1,357,985)	none	-	none	-	none	-	9/G/
<i>cj1426c</i> 283 aa (1,360,321-1,361,172)	none	-	none	-	none	-	10/G/M
<i>cj1429c</i> 308 aa (1,363,196-1,364,122)	none	-	CJJ81176_1428/1429 (89%)	+	none	-	10/G/M
<i>cj1437c</i> (1,373,917-1,375,020)	none	-	none	-	none	-	9/G/M
Novel Phase Variable Tracts in Cj1221							
Location	Homolog in NCTC11168		Homolog in strain 81- 176		Homolog in strain 81116		
In front of <i>rrsA</i> 900 nts (36,916-36,926)	conserved region (38,285) (97%)	-	conserved region (38,191) (94%)	-	conserved region (36,600) (94%)	-	10/C/NR
CJE0775 (pseudo) 2060 nts	Cj0677 (kdpB) (97% nts)	-	Cjj81176_0700 (97% nts)	-	C8j_0629 (97% nts)	-	10/G/NR

(707855-709915)							
CJE0835 205 aa (767,754-768,371)	Cj0735 (96%)	-	Cjj81176_0758 (98%)	+	C8j_0703/0702 (98% nts)	+	9- 13/G/E
CJE1105 184 aa (1,025,669-1,026,223)	none	-	none	-	none	-	10/G/M
CJE1280 323 aa (1,198,131-1,199,102)	none	-	none	-	none	-	9/G/M
CJE1281 295 aa (1,199,095-1,199,982)	none	-	none	-	none	-	9/G/M
Intergenic 600 nts (1,307,026-1,307037)	conserved region (1,185,847)	-	conserved region (1,187,505)	-	conserved region (1,206,775)	-	11/C/NR
CJE1498 239 aa 719 nts (1,393,364-1,394,083)	none	-	Cjj81176_1325 (74% nts, 120 nts missing at 5')	+	C8j_1241 (83%)	-	9/G/M
CJE1500 (1,394,318-1,395,232)	none	-	none	-	none	-	10/G/M
CJE1502/CJE1503 (1,396,224-1,397,013)	none	-	none	-	none	-	11/G/M
CJE1515 (1,406,353-1,406,724)	none	-	none	-	none	-	9/G/E
Intergenic 525 nts (1,448,401-1,448,408)	none	-	none	-	in front of C8J_1278 (1,292,915) (99%)	+	8/C/NR
CJE1602 (1,502,884-1,504,734)	none	-	none	-	none	-	9/G/E
CJE1603 (1,504,703-1,506,655)	none	-	none	-	none	-	9/G/E
Novel Phase Variable Tracts in Cj81-176							
Location	Homolog in NCTC11168		Homolog in strain 1221		Homolog in strain 81116		
CJJ81176_1160 (1,078,500-1,079,447)	none	-	none	-	none	-	10/G/M
CJJ81176-1327 403 aa 1211 nts (1,241,740-1,240,529)	<i>Cj1305c</i> (80% nts)	+	none	-	C8J_1243 (85% aas) (86% nts)	+	9/G/M
CJJ81176_1420/1421 (1,341,847-1,343,626)	none	-	none	-	none	-	10/G/S
CJJ81176_1432 (1,353,939-1,352,230)	none	-	none	-	none	-	9/G/E
CJJ81176_1435 (1,358,459-1,356,648)	none	-	none	-	none	-	9/G/S
Novel Phase Variable Tracts in Cj81116							
Location	Homolog in NCTC11168		Homolog in strain 1221		Homolog in strain 81-176		
in front of C8J_0464	in front of	-	in front of CJE1048	-	in front of	-	10/C/NR

1000 nts (467,660-467,669)	Cj0967 (95%)		(84%)		CJJ81176_0528 (96%)		
in front of C8J_0909 (919,404-919,413)	duplication absent	-	duplication absent	-	duplication absent		10/C/NR
C8J_1078 (1,087,684-1,088,109)	none	-	none	-	none	-	10/G/E
C8J_1080/1081 (1,088,528-1,089,809)	none	-	none	-	none	-	9/G/M
C8J_1246 172 aas (1,255,231-1,255,749)	Cj1313 (77%)	-	CJE1508 (80%)	-	CJJ81176_1330 (80%)	-	10/G/E
C8J_1334 (1,348,646-1,349,179)	none	-	none	-	none	-	11/G/E
C8J_1350 276 aa (1,368,135-1,369,082)	Cj1443c (98%)	-	CJE1617 (97%)	-	CJJ81176_1437 (97%)	-	7/G/M

<sup>1</sup>Genes/loci marked with an asterisk (\*) are phase variable in all four genomes.

<sup>2</sup>Homologous gene = >75% amino acid identity over 90% or more of gene (if matching gene has a higher homology to another phase variable gene then this match is ignored; this occurs because some genes are duplicated in some genomes); a homologous pseudogene = >90% identical nucleotides; a homologous region = >90% identical nucleotides in sequences flanking repeat tract.

<sup>3</sup>Phase variation (PV) is present (+) or absent (-); an indication of PV was dependent on whether a tract of 7 or more Cs or Gs in tandem was located in the locus/gene.

<sup>4</sup>Position in gene. S, start, 5' end, first 200 nts; E, end, 3' end last 50 nts; M, middle, rest of gene.

<sup>5</sup>The tracts in *cj0045* of strain NCTC11168 mediate switching between addition of 0 amino acids (10G), 14 amino acids (11G), and 2 amino acids (DR, 12G). The 10G tract results in termination of translation immediately adjacent to the initiation codon of *cj0044* whilst the other lengths will result in overlap of the reading frames of these genes.

Statistical Analysis of PV Rates

Analyses were performed using a Mann-Whitney non-parametric rank sum test (InStat 1.0).

All 1139 constructs

G8-cat v G11 cat,  $P < 0.0001$

G8-cat v G9-cat,  $P = 0.0001$

G8-cat v G10-cat,  $P = 0.75$

G8-cat v G8-kan,  $P < 0.0001$

G8-cat v G7-kan,  $P = 0.0002$

G9-cat v G10-cat,  $P = 0.0005$

G9-cat v G7-kan,  $P = 0.0132$

G8-kan v G11-cat,  $P < 0.0001$

1139-cat v 0031

1139-G8 v 0031-G9,  $P < 0.0001$

1139-G10 v 0031-G10,  $P < 0.0001$

0031-G9 v 0031-G10,  $P = 0.4303$

1139-cat v capA

1139-G8 v CapA G11,  $P < 0.0001$

1139-G11 v CapA G11,  $P$  is  $0.0005$

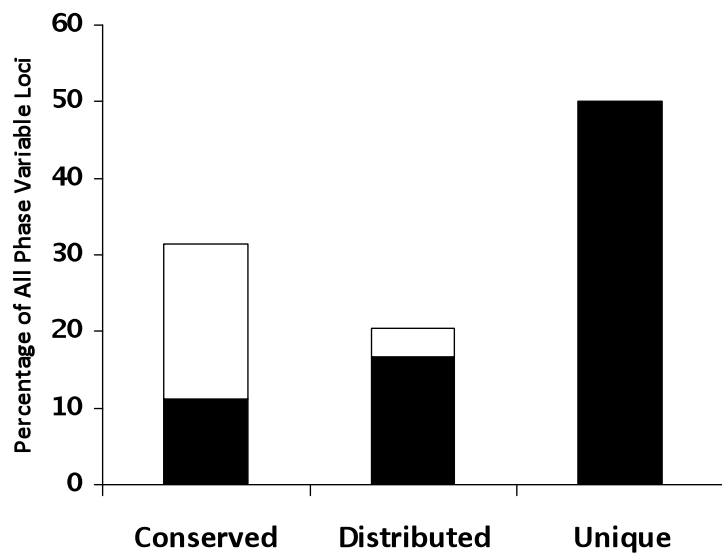
1139-G10 v CapA-G10,  $P = 0.6382$

1139-G11 v Cap G12,  $P = 0.0025$

CapA v capA

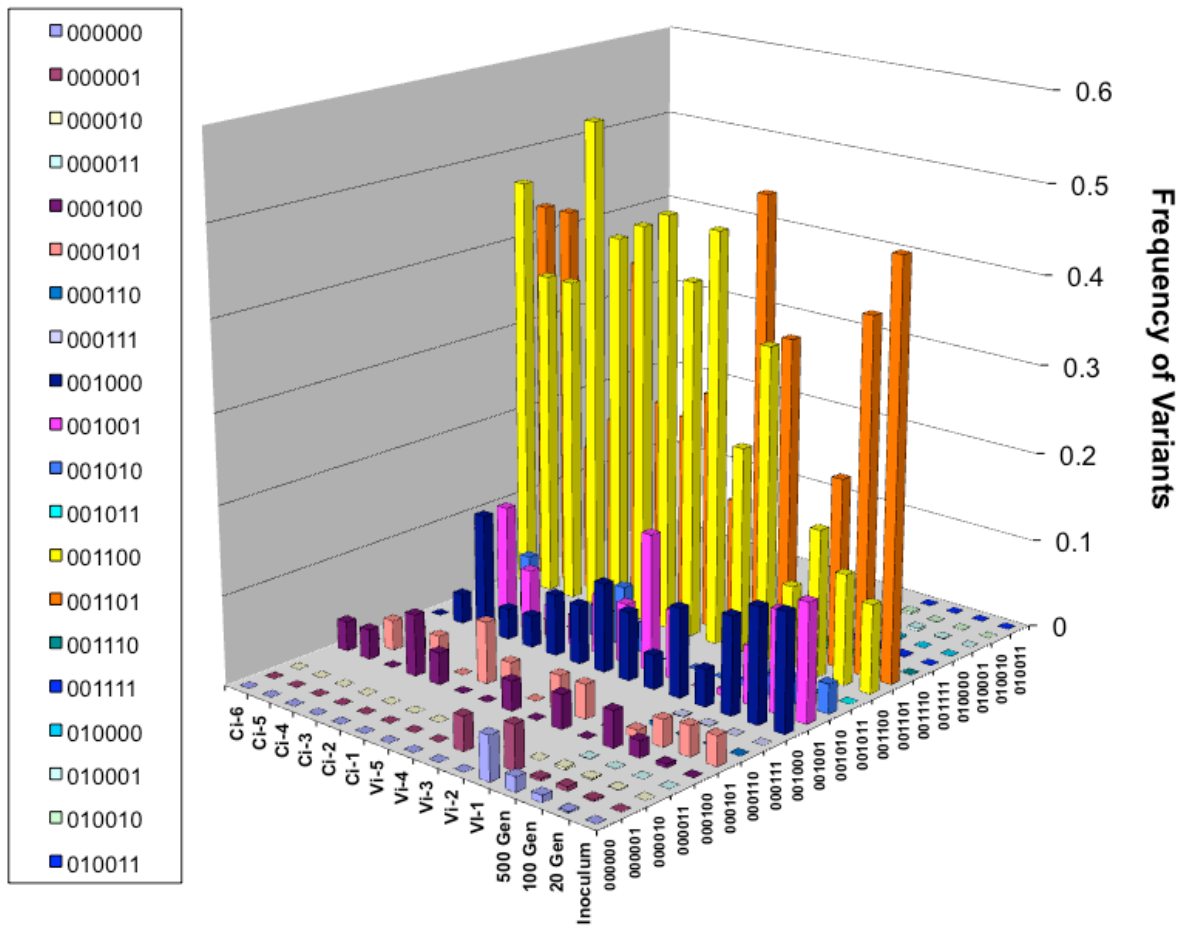
CapA G11 v Cap-G12,  $P = 0.0273$

CapA G11 v Cap G10,  $P < 0.0001$

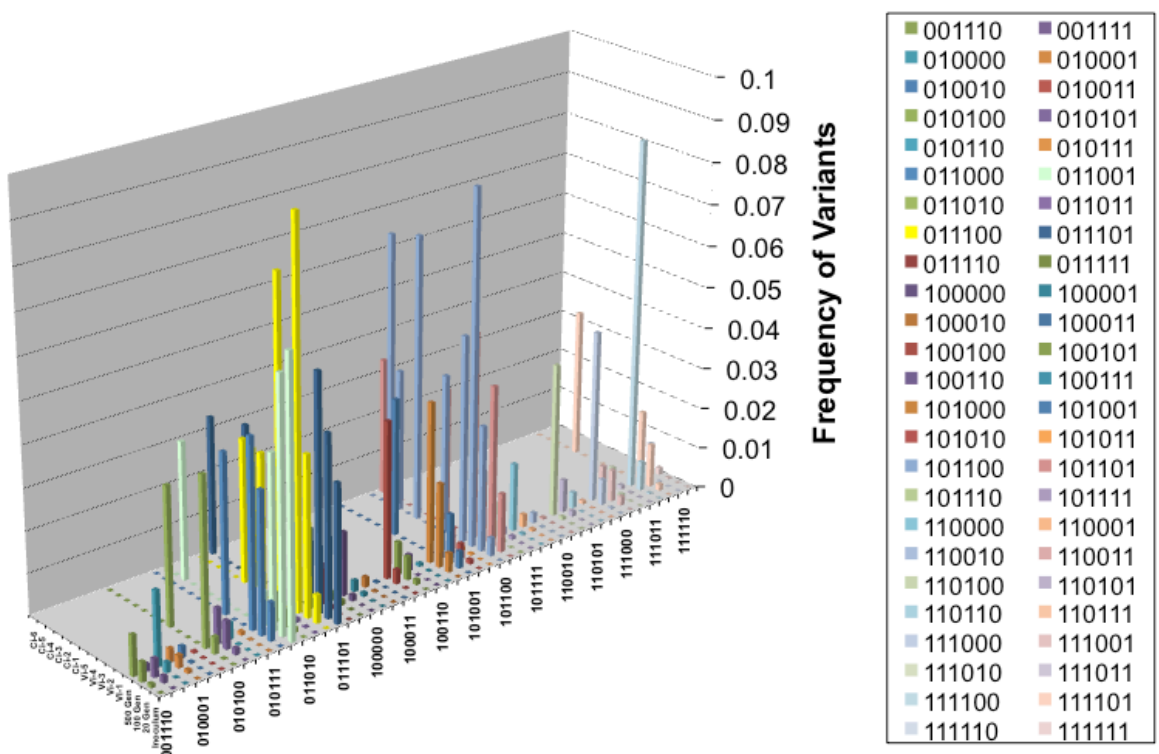


Supplementary Fig. 1. Distribution of phase variable loci in *Campylobacter jejuni* genomes. The conservation of phase variable loci (n = 55) containing polyC or polyG repeat tracts was compared for four published *C. jejuni* genome sequences (from strain NCTC11168, 81-176, 1221 and 81116). Conserved loci exhibited an amino acid identity of >75% or, where a gene was absent or disrupted (e.g. pseudogenes), a nucleotide sequence conservation of >90%. Conserved loci were present in all four genome sequences, distributed loci in 2-3 genomes and unique loci in a single genome. Black bars shows loci in which the repeat tract is conserved in all homologous regions.

(A)

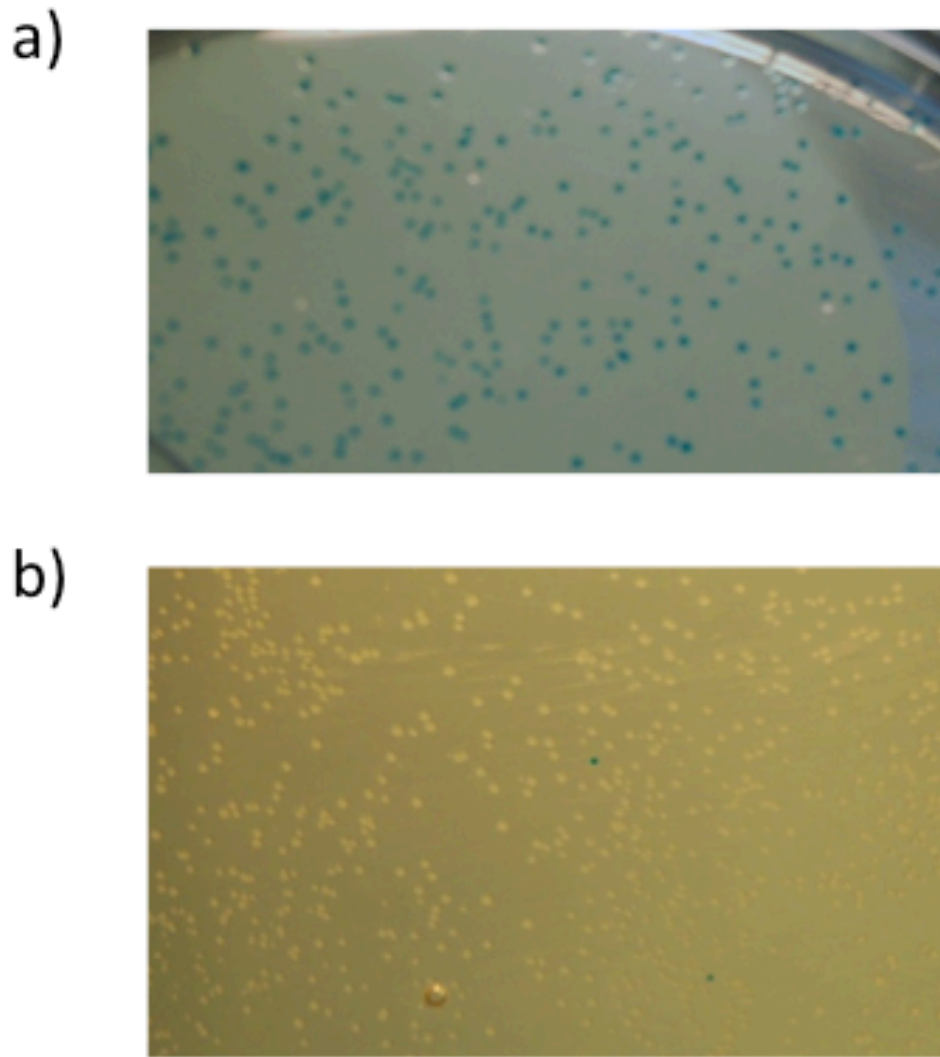


(B)



Supplementary Fig. 2. Modelling Changes in the Proportions of Genotypes for In Vitro Passage of *C. jejuni*. The proportions of genotypes for the inoculum were used as input to the theoretical model of phase variation. Genotypes were for the six genes with 0 representing an OFF phase variant and 1 an ON variant. The order of the genes and the ON-to-OFF and OFF-to-ON switching rates ( $\times 10^{-4}$ ) were as follows:- cj1326, 10.3, 17.9; cj0031, 10.3, 17.9; cj1139, 6.9, 2.1; cj0685, 2.1, 6.9; cj0045, 6.9, 2.1; capA, 38.8, 3.7. Panel A shows genotypes from 0-0-0-0-0-0 to 0-1-0-0-1-1 and panel B from 0-0-1-1-1-0 to 1-1-1-1-1-1. Note the differing scales on panels A and B. Inoculum, inoculum; 20 Gen, 100 Gen, and 500 Gen are output data from the model for runs of 20, 100 and 500 generations; Vi-1 to Vi-5 are output experimental data for variable size inoculums; Ci-1 to Ci-6, are output experimental data for a constant inoculum.





Supplementary Fig. 3. Demonstration of expression and phase variation of the *cj1139c-lacZ* reporter gene in *Campylobacter jejuni* strain NCTC11168. Panel a) shows ON-to-OFF switching of colonies from the reporter construct grown on Muller-Hinton agar plates containing X-gal. Panel b) shows OFF-to-ON switching of this construct.

Supplementary Table II. Phase variation of six genes in multiple single colonies derived from different starter colonies

Gene	<u>Start Colonies</u>								
	<u>A1</u>			<u>A2</u>			<u>A3</u>		
	Start	Non-Variants	Variants	Start	Non-Variants	Variants	Start	Non-Variants	Variants
<i>cj0031</i>	G9	29	0	G10	28	0	G10	30	0
<i>cj0045</i>	G9	25	0	G9	27	1 (G10)	G9	30	0
<i>cj0685</i>	G8	25	0	G8	28	0	G8	30	0
<i>cj1139</i>	G8	29	0	G8	28	0	G8	30	0
<i>cj1326</i>	G9	25	0	G10	28	0	G10	29	1 (G9)
<i>capA</i>	G12	28	1 (G11)	G11	26	0	G11	29	1 (G10)

## In Vitro Passage Experiment

The inoculum was prepared from strain NCTC11168 grown overnight on MHA plates. A suspension was prepared in MHB and the OD<sub>600</sub> was 0.8. A sample of 100 ul was inoculated into 5 ml of MHB for 6 separate cultures (Ci1-Ci6; i.e. constant inoculum). A further 6 tubes (Vi1-Vi6; i.e. variable inoculum) were prepared with the first tube being inoculated with 100 ul of the suspension and the 5 subsequent tubes with 10-fold dilutions of this suspension in 100 ul of MHB. Serial dilutions of the inoculum were also plated onto MHA and the numbers of colonies were counted after four days growth. Analysis of these plates indicated that the inoculum contained  $3.5 \times 10^9$  cfu/ml. After overnight growth 300 (Ci1-Ci6, Vi1-Vi3), 600 (Vi4) or 1 (Vi5) ml of the culture was transferred into 5 ml of MHB and again incubated overnight (Note that Vi6 did not grow). A second passage was performed using an inoculum of 250 ul into 5 ml. After overnight growth, the bacterial cells were pelleted and resuspended in 1 ml of MHB prior to plating of serial dilutions on MHA plates. Colonies from the inoculum and output plates were subject to immunoblotting with anti-CapA antisera. The frequency of OFF variants in the inoculum was 0.29.

Supplementary Table III. Inoculum and output population sizes during in vitro growth of *C. jejuni* strain NCTC11168 in Mueller-Hinton broth

Sample	Inoc	P1 OD	P2 OD	P3 OD	P3 (Total cfu)	CapA Fre OFF variants
Ci-1	$3.5 \times 10^8$	0.77	0.97	1.01	$1.1 \times 10^9$	0.24
Ci-2	$3.5 \times 10^8$	0.76	0.96	1.01	$9.4 \times 10^8$	0.28
Ci-3	$3.5 \times 10^8$	0.75	0.94	0.97	$4.4 \times 10^8$	0.28
Ci-4	$3.5 \times 10^8$	0.75	1.03	1.02	$6.9 \times 10^8$	0.29
Ci-5	$3.5 \times 10^8$	0.76	0.99	1.01	$9.2 \times 10^8$	0.29
Ci-6	$3.5 \times 10^8$	0.76	0.92	1.04	$1.2 \times 10^9$	0.36
Vi-1	$3.5 \times 10^8$	0.77	0.98	0.97	$8.2 \times 10^8$	0.36
Vi-2	$3.5 \times 10^7$	0.61	0.94	1.00	$4.9 \times 10^8$	0.27
Vi-3	$3.5 \times 10^6$	0.47	0.90	0.98	$5.6 \times 10^8$	0.27
Vi-4	$3.5 \times 10^5$	0.25	0.85	0.95	$3.7 \times 10^8$	0.33
Vi-5	$3.5 \times 10^4$	0.05	0.77	0.98	$5.1 \times 10^8$	0.27
Vi-6	$3.5 \times 10^3$	0.02	NR	NR	NR	NR

In Vitro Passage – Tract lengths

Supplementary Table IV. Repeat Tract Lengths of Colonies Derived Following In Vitro Passage of *C. jejuni* strain NCTC11168

Sample	<i>cj1326</i>				<i>cj0031</i>			<i>cj1139</i>			<i>cj0685</i>			<i>cj0045</i>			<i>capA</i>		
	9	10	11	12	9	10	11	7	8	9	8	9	10	8	9	10	10	11	12
Inoc	30				2	28		29 1			10	19 1		1	29		8 22		
Vi-1	2	25	1		2	26		2	23	1	4	21	3	25 1			8	9	3
Vi-2	27 3				2	28		28 2			4	25 1		30			8	17 3	
Vi-3	2	24	3		1	26	2	25 2			11	17		26	1	12 10 4			
Vi-4	1	29			2	28		28 2			6	24		1	27	2	15 11 4		
Vi-5	1	29			1	28	1	29 1			7	23		2	28		14 11 5		
Ci-1	3	26	1		1	29		28 2			4	25	1	1	29		16 10 4		
Ci-2	1	27	1		2	28		28 1			2	27		26	3	10 14 4			
Ci-3	1	28			29			26 3			3	26		29			15 9 5		
Ci-4	2	28			1	28	1	29 1			11	19		2	27	1	15 11 4		
Ci-5	30				29 1			28 2			4	26		29	1	12 17 1			
Ci-6	1	28	1		1	29		29 1			30			29	1	14 15 1			

Note:- 'on' repeat numbers for *cj1326*, *cj0031*, *cj1139*, *cj0685*, *cj0045* and *capA* are 9, 9, 8, 9, 8 and 11, respectively.

Supplementary Table V. Relative frequencies and confidence intervals for 'on/off' phenotypes of populations derived by in vitro passage

Sample	<i>cj1326</i>			<i>cj0031</i>			<i>cj1139</i>			<i>cj0685</i>			<i>cj0045</i>			<i>capA</i>		
	on	off	CI <sup>1</sup>	on	off	CI <sup>1</sup>	on	off	CI <sup>1</sup>	on	off	CI <sup>1</sup>	on	off	CI <sup>1</sup>	on	off	CI <sup>1</sup>
Inoc	0	1	0	0.07	0.93	0.07	0.97	0.03	0.05	0.63	0.37	0.14	0.03	0.97	0.05	0.73	0.27	0.13
Vi-1	0.07	0.93	0.08	0.07	0.93	0.08	0.88	0.12	0.1	0.75	0.25	0.13	0	1	0	0.45	0.55	0.18
Vi-2	0	1	0	0.07	0.93	0.07	0.93	0.07	0.07	0.83	0.17	0.11	0	1	0	0.61	0.39	0.15
Vi-3	0.07	0.93	0.08	0.03	0.97	0.06	0.93	0.07	0.08	0.61	0.39	0.15	0	1	0	0.38	0.62	0.16
Vi-4	0.03	0.97	0.05	0.07	0.93	0.07	0.93	0.07	0.07	0.8	0.2	0.12	0.03	0.97	0.05	0.37	0.63	0.14
Vi-5	0.03	0.97	0.05	0.03	0.97	0.05	0.03	0.97	0.05	0.77	0.23	0.13	0.07	0.93	0.07	0.37	0.63	0.14
Ci-1	0.1	0.9	0.09	0.03	0.97	0.05	0.93	0.07	0.07	0.83	0.17	0.11	0.03	0.97	0.05	0.33	0.67	0.14
Ci-2	0.07	0.93	0.08	0.07	0.93	0.07	0.97	0.03	0.06	0.93	0.07	0.08	0	1	0	0.5	0.5	0.16
Ci-3	0.03	0.97	0.06	0	1	0	0.9	0.1	0.09	0.9	0.1	0.09	0	1	0	0.31	0.69	0.14
Ci-4	0.07	0.93	0.07	0.03	0.97	0.05	0.97	0.03	0.05	0.63	0.37	0.14	0.07	0.93	0.07	0.37	0.63	0.14
Ci-5	0	1	0	0	1	0	0.93	0.07	0.07	0.87	0.13	0.1	0	1	0	0.57	0.43	0.15
Ci-6	0.03	0.97	0.05	0.03	0.97	0.05	0.03	0.97	0.05	1	0	0	0	1	0	0.5	0.5	0.16

<sup>1</sup>CI, 90% confidence intervals

Evaluation of the independence of selection of genes for the in vitro passaged populations  
Proportions of each genotype were calculated from the individual proportions of on/off for each gene and compared to the observed proportions. In the majority of cases the observed proportions for the major genotypes (30 in total with an average of 2.5 per culture) were within the confidence intervals of the expected values indicating that independence of mutation/selection held for each case. The proportion of minor genotypes, observed due to 1 or 2 colonies (i.e. <0.07), were often higher and outside the confidence intervals for the expected values. There were a few exceptions:- Vi-3, 0-0-1-1-0-0, expected 31%, observed 46%; Vi-3, 0-0-1-0-0-0, expected 20%, observed 8%; Vi-1, 0-0-1-0-0-1, expected 9%, observed 0.

In vivo passage of *C. jejuni* strain NCTC11168 – Tract Lengths

Supplementary Table VI. GeneScan Analysis of Repeat Tract Lengths on DNA Extracts from Caecal Material of Chickens Colonised with *C. jejuni* strain NCTC11168H

Sample	<i>cj1326</i>		<i>cj0031</i>		<i>Cj1139</i>		<i>cj0685</i>		<i>cj0045</i>		<i>capA</i>	
	Major Peak	Minor Peak <sup>1</sup> (ratio)	Major Peak	Minor Peak <sup>1</sup> (ratio)	Major Peak	Minor Peak <sup>1</sup> (ratio)	Major Peak	Minor Peak <sup>1</sup> (ratio)	Major Peak	Minor Peak <sup>1</sup> (ratio)	Major Peak	Minor Peak <sup>1</sup> (ratio)
<b>Inoc</b>	G10	G9 (2.3)	G10	G9 (2)	G8	G7 (1.4)	C8	C9 (4.3)	G9	G10 (2.2)	G11	G12 (2.7)
<b>B6</b>	G10	G9 (3.3)	G9	G10 (4)	G9	G8 (3.0)	C9	C8 (3.9)	G10	G9 (4.7)	G11	G12 (2.6)
<b>B7</b>	G10	G9 (3.2)	G9	G10 (6.2)	G9	G10 (6.9)	C9	C10 (5.3)	G10	G9 (5.2)	G11	G10 (2.5)
<b>B8</b>	G10	G9 (3.3)	G9	G10 (5.7)	G9	G10 (7.8)	C9	C8 (8.4)	G10	G11 (1.3)	G11	G10 (3.0)
<b>B9</b>	G10	G9 (3.1)	G9	G10 (3.9)	G9	G10 (6.9)	C9	C8 (8.4)	G10	G11 (1.1)	G11	No data
<b>B10</b>	G10	G9 (3.2)	G9	G10 (3.3)	G9	G8 (3.3)	C9	C8 (3.4)	G10	G11 (1.6)	G11	G10 (2.8)
<b>B11</b>	G10	G9 (3.3)	G9	G10 (5.9)	G9	G10 (5.1)	C9	C10 (4.0)	G10	G9 (5.0)	G11	G12 (1.4)

<sup>1</sup>Multiple peaks are observed in GeneScan analyses of PCR products spanning repeat tracts. The size of the largest peak is defined as the major peak and represents the major repeat type present in the sample. The second largest peak was defined as the minor peak and this may represent and be proportional to a minor repeat type present in the sample but note that these peaks can also be generated as a result of replication slippage during the PCR. The number in brackets is the ratio between the major and minor peak.

Supplementary Table VII. GeneScan Analysis of Repeat Tract Lengths of Colonies Derived Following In Vivo Passage of *C. jejuni* strain NCTC11168H

Sample	<i>cj1326</i>				<i>cj0031</i>			<i>cj1139</i>				<i>cj0685</i>			<i>cj0045</i>				<i>capA</i>				
	9	10	11	12	9	10	11	7	8	9	10	8	9	10	9	10	11	12	9	10	11	12	13
<b>Inoc</b>	8	21	0	1	5	23	2	3	25	2	0	29	1	0	13	17	0	0	0	3	11	10	5
<b>B6</b>	1	27	0	0	21	9	0	0	7	22	0	4	24	0	0	25	3	0	0	6	12	10	0
<b>B7</b>	1	21	2	0	28	2	0	0	1	29	0	0	24	0	0	16	8	0	0	8	20	2	0
<b>B8</b>	1	24	1	0	25	1	0	0	0	26	0	1	25	1	1	11	15	0	0	3	19	3	0
<b>B9</b>	0	30	0	0	24	6	0	0	0	29	1	0	30	0	0	12	16	2	0	9	11	9	0
<b>B10</b>	0	25	2	0	18	1	0	0	0	0	0	0	25	1	2	19	5	0	0	1	12	2	0
<b>B11</b>	1	28	0	0	26	2	0	0	3	25	0	0	29	0	0	27	2	0	1	1	8	17	1

Note:- 'on' repeat numbers for *cj1326*, *cj0031*, *cj1139*, *cj0685*, *cj0045* and *capA* are 9, 9, 8, 9, 11 and 11, respectively.

Supplementary Table VIII. Relative frequencies and confidence intervals for 'on/off' phenotypes of populations derived from in vivo passage of strain NCTC11168H

Sample	<i>cj1326</i> <sup>1</sup>			<i>cj0031</i> <sup>1</sup>			<i>cj1139</i> <sup>1</sup>			<i>cj0685</i> <sup>1</sup>			<i>cj0045</i> <sup>1</sup>			<i>capA</i> <sup>1</sup>		
	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>
<b>Inoc</b>	0.29	0.71	0.14	0.11	0.89	0.1	0.82	0.18	0.12	0.04	0.96	0.06	0	1	0	0.39	0.61	0.15
<b>B6</b>	0.04	0.96	0.06	0.69	0.31	0.15	0.23	0.77	0.14	0.85	0.15	0.12	0.08	0.92	0.09	0.42	0.58	0.15
<b>B7</b>	0.04	0.96	0.07	0.92	0.08	0.09	0.04	0.96	0.07	1	0	0	0.33	0.67	0.16	0.71	0.29	0.15
<b>B8</b>	0.04	0.96	0.07	1	0	0	0	1	0	1	0	0	0.57	0.43	0.17	0.78	0.22	0.14
<b>B9</b>	0	1	0	0.79	0.21	0.12	0	1	0	1	0	0	0.52	0.48	0.15	0.38	0.62	0.15

B11	0.04	0.96	0.06	0.92	0.08	0.09	0.08	0.92	0.09	1	0	0	0.08	0.92	0.09	0.31	0.69	0.15
-----	------	------	------	------	------	------	------	------	------	---	---	---	------	------	------	------	------	------

<sup>1</sup>Phenotypes were derived assuming the 'on' repeat numbers for cj1326, cj0031, cj1139, cj0685, cj0045 and capA were 9, 9, 8, 9, 8 and 11, respectively. Phenotypes were only derived for colonies in which unambiguous data was generated for every gene of a colony.

<sup>2</sup>CI, 90% confidence intervals

Evaluation of the independence of selection of genes for the in vivo passaged populations  
Proportions of each genotype were calculated from the individual proportions of on/off for each gene and compared to the observed proportions. In the majority of cases the observed proportions for the major genotypes were within the confidence intervals of the expected values indicating that independence of mutation/selection held for each case. The proportions for minor genotypes, observed due to 1 or 2 colonies (i.e. a frequency of <0.07), were often higher and outside the confidence intervals of the expected values. Samples from bird 8 (B8) exhibited a number of significant variations from this general trend with two major genotypes over-represented and one under-represented (see Supplementary Table IX). This may be due to selection against a genotype in which *cj0045* is expressed (i.e. contains 11 repeats) and *capA* is switched off or may be due to sampling error. Intriguingly a similar genotype in which both these genes are switched 'off' is over-represented in two samples further suggesting that when *capA* is switched 'off' then there is selection against a *cj0045* 'on' phenotype. However, a test for dependent selection (using a Chi-squared test) of these two genes in the B8 sample did not detect any statistically-significant divergence between observed and expected proportions of the combined genotypes of these two genes.

Supp. Table IX. Genotypes exhibiting a significant divergence from expected frequencies assuming independent selection of genes.

Sample	Over-represented in observed <sup>1</sup>	Under-represented in observed <sup>1</sup>
Inoculum	000001 (obs 0.11; exp 0.04 +/-0.04)	001001 (obs 0.07, exp 0.2 +/-0.09)
B6	010100 (obs 0.5; exp 0.23 +/-0.1) 001101 (obs 0.12, exp 0.02 +/-0.02)	000100 (obs 0.04, exp 0.1 +/-0.06)
B7	none	None
B8	010111 (obs 0.39, 0.16 +/-0.08) 010101 (obs 0.26, 0.15 +/-0.08)	010110 (obs 0.09, exp 0.25 +/-0.11)
B9	010100 (obs 0.34, exp 0.24 +/-0.1)	
B11	none	none

<sup>1</sup>Only genotypes observed or expected to exceed a frequency of 0.1 were included in this table. Numbers in brackets are 90% confidence intervals.

In vivo passage of *C. jejuni* strain 81-176 – Tract Lengths and Genotypes

Supp Table X. Repeat Tract Lengths of Colonies Derived Following In Vivo Passage of *C. jejuni* strain NCTC11168H

Sample	<b>81176-0083<sup>1</sup></b>				<b>81176-0646<sup>1</sup></b>			<b>81176-0708<sup>1</sup></b>			<b>81176-1160<sup>1</sup></b>			<b>81176-1312<sup>1</sup></b>			<b>81176-1325<sup>1</sup></b>		
	9	10	11	12	9	10	11	8	9	10	8	9	10	8	9	10	8	9	10
<b>Inoc-S</b>	1	8	21	0	1	27	2	0	30	0	0	27	2	0	28	2	1	28	1
<b>Inoc-L</b>	0	10	20	0	4	25	1	0	29	1	0	29	1	1	26	3	0	29	0
<b>B03</b>	1	5	21	3	6	21	3	1	21	8	2	24	0	0	27	3	0	24	3
<b>B06</b>	0	3	22	4	3	26	0	0	30	0	1	23	4	0	29	1	0	28	2
<b>B09</b>	1	26	3	0	4	25	1	0	29	0	0	29	1	0	23	7	0	23	7
<b>B014</b>	0	28	2	0	3	27	0	0	30	0	0	29	1	0	27	3	0	29	1

<sup>1</sup>The ‘on’ repeat numbers for *81176-0646*, *81176-0708*, *81176-1160*, *81176-1312* and *81176-1325* are 10, 9, 10, 9 and 9, respectively. Gene *81176-0083* is a pseudogene that is homologous to *cj0045* of strain NCTC11168. Genes *81176-0646*, *81176-0708* and *81176-1312* are homologs of the strain NCTC11168 genes *cj0617*, *cj0685* and *cj1295*, respectively.

Supp. Table XI. Relative frequencies and confidence intervals for ‘on/off’ phenotypes of populations derived from in vivo passage of strain NCTC11168H

Sample	<b>81176-0083<sup>1</sup></b>			<b>81176-0646<sup>1</sup></b>			<b>81176-0708<sup>1</sup></b>			<b>81176-1160<sup>1</sup></b>			<b>81176-1312<sup>1</sup></b>			<b>81176-1325<sup>1</sup></b>		
	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>	on	off	CI <sup>2</sup>
<b>Inoc-S</b>	0.69	0.31	0.14	0.90	0.10	0.09	1	0	0	0.07	0.93	0.08	0.93	0.07	0.08	0.93	0.07	0.08
<b>Inoc-L</b>	0.69	0.31	0.14	0.83	0.17	0.12	0.96	0.03	0.06	0.03	0.97	0.06	0.86	0.14	0.11	1	0	0
<b>B03</b>	0.81	0.19	0.13	0.69	0.31	0.15	0.73	0.27	0.14	0	1	0	0.88	0.12	0.12	0.88	0.12	0.10
<b>B06</b>	0.79	0.21	0.13	0.93	0.07	0.08	1	0	0	0.14	0.86	0.11	0.96	0.04	0.06	0.93	0.07	0.08
<b>B09</b>	0.10	0.90	0.09	0.83	0.17	0.12	1	0	0	0.03	0.97	0.06	0.79	0.21	0.12	0.76	0.24	0.13
<b>B014</b>	0.07	0.93	0.07	0.90	0.10	0.09	1	0	0	0.03	0.97	0.05	0.90	0.10	0.09	0.97	0.03	0.05

<sup>1</sup>Phenotypes were derived assuming the ‘on’ repeat numbers for *81176-0083*, *81176-0708*, *81176-1160*, *81176-1312* and *81176-1325* are 11, 10, 9, 10, 9 and 9, respectively. Note, that *81176-0083* is a pseudogene and that a 11 repeats was arbitrarily chosen as an ON number of repeats to match data for strain NCTC11168. Phenotypes were only derived for colonies in which unambiguous data was generated for every gene of a colony.

<sup>2</sup>CI, 90% confidence intervals.

Test for independence of in vivo selected 81-176 genotypes

Expected proportions of each genotype were calculated from observed proportions of ON/OFF variants for each gene and compared to observed proportions of each genotype. Non-overlapping confidence intervals were detected for a number of the minor genotypes observed in a single colony but no differences for any of the major genotypes (i.e. with a frequency >0.01).