Phase variation of a Type IIIG restriction-modification enzyme alters site-specific methylation patterns and gene expression in *Campylobacter jejuni* strain NCTC11168

Awais Anjum¹, Kelly J. Brathwaite², Jack Aidley¹, Phillippa L. Connerton², Nicola J. Cummings², Julian Parkhill³, Ian Connerton² and Christopher D. Bayliss¹,*

¹Department of Genetics, University of Leicester, Leicester LE1 7RH, UK, ²Division of Food Sciences, School of Biosciences, University of Nottingham, Sutton Bonington LE12 5RD, UK and ³The Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK

Received September 2, 2015; Revised December 22, 2015; Accepted December 26, 2015

ABSTRACT

Phase-variable restriction-modification systems are a feature of a diverse range of bacterial species. Stochastic, reversible switches in expression of the methyltransferase produces variation in methylation of specific sequences. Phase-variable methylation by both Type I and Type III methyltransferases is associated with altered gene expression and phenotypic variation. One phase-variable gene of *Campylobacter jejuni* encodes a homologue of an unusual Type IIIG restriction-modification system in which the endonuclease and methyltransferase are encoded by a single gene. Using both inhibition of restriction and PacBio-derived methylome analyses of mutants and phase-variants, the cj0031c allele in *C. jejuni* strain NCTC11168 was demonstrated to specifically methylate adenine in 5′CCCGA and 5′CCTGA sequences. Alterations in the levels of specific transcripts were detected using RNA-Seq in phase-variants and mutants of cj0031c but these changes did not correlate with observed differences in phenotypic behaviour. Alterations in restriction of phage growth were also associated with phase variation (PV) of *cj0031c* and correlated with presence of sites in the genomes of these phages. We conclude that PV of a Type IIIG restriction-modification system causes changes in site-specific methylation patterns and gene expression patterns that may indirectly change adaptive traits.

INTRODUCTION

*Campylobacter jejuni* is a major cause of food-borne gastroenteritis in humans with more than 400 million cases of Campylobacteriosis reported annually worldwide (1). Currently, the disease burden exerted by *Campylobacters* in the UK is higher than for any other foodborne pathogen (2) with chickens being the main source of infection in over 80% of disease cases (3). *Campylobacters* are present in a very high fraction of farmed chickens with significant numbers of bacteria in each colonized bird (4,5). Rapid microevolution may be a major mechanism that facilitates persistent colonization of and rapid transmission between individual birds. One of the characteristic mechanisms for enabling rapid adaptation is phase variation (PV) mediated by hypermutable simple sequence repeats (6). Most of phase-variable genes of *C. jejuni* encode surface proteins or enzymes that modify host surface structures (7). However one of these phase-variable genes encodes a restriction modification (RM) system that has the potential to act as a stochastically-regulated modifier of expression of other genes or mediate alternation between phage susceptible and resistant states (8).

PV is an adaptive process characterized by high frequency, reversible ON/OFF switching of gene expression that is stochastic in nature and gives rise to heterogeneous population (9). The main mechanism of PV in *C. jejuni* involves insertion or deletion of repeat units during genome replication in polyG or polyC repeat tracts located in the reading frames of genes. In a recent survey, each *C. jejuni* genome was found to contain between 12 and 29 polyG/polyC tracts capable of mediating PV (8). Most of these phase-variable loci encode enzymes that modify the glycan components of the flagella, capsule and lipooligosaccharide. One of the tracts was present in a putative Type IIIG restriction-modification system. This is a general paradigm

*To whom correspondence should be addressed. Tel: +44 0 116 252 3465; Fax: +44 0 116 252 3378; Email: cdb12@le.ac.uk

© The Author(s) 2016. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
with phase-variable Type I and III RM systems being found in other host adapted bacterial pathogens including *Helicobacter pylori*, *Neisseria meningitidis*, *Neisseria gonorrhoeae* and *Haemophilus influenzae* (6).

The three major classes of RM systems (i.e. I, II and III) are distinguished by the composition and nature of the recognition sequences, position of cleavage, requirements for co-factors and subunit composition (10). A Type IIG RM system is a subclass of the Type II RM systems in the majority of which the endonuclease (ENase) and methyltransferase (MTase) activities are fused into a single polypeptide chain rather than being present as separate enzymes. These Type IIG enzymes methylate only one strand within either a symmetric or asymmetric recognition site while the endonuclease activity occurs 14–21 bp away from the recognition site in a 3’ direction (11). This biochemical property of Type IIG RM systems is similar to that of Type IIS and Type III RM systems. Two of the most highly characterized prototypical Type IIG enzymes are BpuSI and Eco57I (12,13). However, REBASE (rebase.neb.com/rebase/rebase.html; 14) predicts the existence of 6400+ putative Type IIG RM enzymes with 100+ recognition sites and a presence in a vast range of bacterial species.

RM systems are regarded as defensive weapons of pathogenic bacteria preventing infection by phages or invasion by foreign DNA (15). A putative role in epigenetic gene regulation through methylation-dependent control of the interactions of DNA binding regulatory proteins with their cognate recognition sites was suggested by analogy with the Dam methylase, a known epigenetic regulator of bacterial virulence gene expression (16–18). The role of RM-mediated epigenetic regulation was demonstrated by showing that ON/OFF PV of the *H. influenzae* modA gene was correlated with altered expression of several genes (19). This gene encodes the methyltransferase component of a Type III RM system whose PV-dependent co-regulated genes were termed the phasevarion. Subsequently, Type III RM systems in three other human pathogens, *H. pylori*, *N. meningitidis* and *N. gonorrhoeae* (20,21) and a Type I system in *Streptococcus pneumoniae* were found to have phasevarion activities (22). The alterations in expression of a ‘regulon’ are predicted to be mediated through differential methylation of promoter elements or other regulatory sequences. Critically, phasevarions include genes mediating methylation of promoter elements or other regulatory sequences. The regulon are predicted to be mediated through differential methylation of promoter elements or other regulatory sequences. The role of RM-mediated epigenetic regulation was demonstrated by showing that ON/OFF PV of the *H. influenzae* modA gene was correlated with altered expression of several genes (19).

The *jejuni* gene sequence was downloaded from the NCBI database, and then used in a BLAST search to identify homologues of this gene and of the allelic variant, *A911_00150*, present in *C. jejuni* PT14 (25). These sequences were submitted to the Phylogeny.fr platform in order to construct a phylogenetic tree. This platform uses MUSCLE for the alignment of the submitted sequences, PhyML for tree building and TreeDyn for tree rendering (26). For building phylogenetic tree of *A911_00150*, the default parameters set for MUSCLE, PhyML and TreeDyn were applied.

**Isolation of ON and OFF phase-variants**

Serial dilutions of a single colony of *C. jejuni* NCTC11168 strain were plated onto MHA plates and incubated for 3 days under micro-aerobic conditions at 42°C. A total of 110 colonies were randomly selected and sub-cultured. Boiled lysates were prepared in 100 μl of distilled water for DNA amplification. The repeat tract of *cj0031* was amplified using two primers, 0031F-Fam and 0031R and then analysed by GeneScan as described elsewhere (8). A total of 75 isolates had a tract length of G10 (an OFF-phenotype) and 13 had G9 (an ON phenotype). The tract lengths of three isolates of each type were confirmed by dideoxy sequencing.

**Identification of the Cj0031 methylation site by inhibition of restriction**

Changes in the restriction patterns of chromosomal DNA extracted from phase variants and mutants was identified as follows. Chromosomal DNA was prepared using either a phenol extraction/caesium chloride method or a DNasey
extraction kit (Qiagen). Two micrograms of DNA was digested for 2 h at 37°C with 20–40 units of a range of methylation sensitive restriction enzymes. The resulting fragments were resolved on 1% agarose gels at 80V for 2–3 h, and visualized under UV illumination. Southern blot and hybridization (27) were performed using DIG-labelled (Roche) polymerase chain reaction (PCR) products as probes. These probes were generated using primers listed in Supplementary Table SI and were designed to contain one or more of the relevant restriction sites within chromosomal fragments of 0.5–10 kb. Products were visualized using a DIG luminescent detection kit (Sigma-Aldrich, UK) and autoradiography.

Construction of deletion and complementation mutants of cj0031 in C. jejuni variants 11168 and CH11168

For the deletion mutant, the M0031-up fragment was amplified from upstream of cj0031 using primers 0031–2108-F-pst1 and 0031–2832-R-2832-Bam (Supplementary Table SI). A second fragment (S0031-down) having a size of 621 bp was amplified from the downstream region of cj0032 using primers 0032-Bam-F and 0032-R (Supplementary Table SI). These PCR products were cloned separately into pGEM-TEasy to create pGEM-M0031-up and pGEM-S0031-down and then these recombinant plasmids were digested with PstI and BamHI to release the M0031-up fragment or BamHI and EcoRI to release the S0031-down fragment, respectively. These fragments were subcloned into pUC18, digested with PstI and EcoRI, in a three-way ligation reaction to obtain pUC18-M0031-up-S0031-down. This plasmid was linearized with BamHI and ligated to a 1.4 kb fragment containing a kanamycin resistance cassette and recovered from pJKM30 by BamHI digestion. The resulting plasmid, pUC18-cj0031::kan, was recombined into competent cells from 11168 and CH11168, and transformants were verified by PCR amplification and sequencing.

The pCfdxA complementation vector (28) enables recombination into cj0046, a pseudogene of C. jejuni NCTC11168. The cj0031 gene was amplified with primers N-cj0031-F and N-cj0031-R (Supplementary Table SI). The pCfdxA vector was cleaved with BsmBI (Esp3I), which generates Esp3I and NcoI overhangs and this was ligated to PCR products cut with Esp3I and NcoI. The resulting recombinant vector pCfdxA0031 was transformed into the C. jejuni mutant strains to generate the complementation constructs NCTC11168Δ0031::kan::0031c and CH11168Δ0031::kan::0031c. PCR and sequencing reactions were performed to verify the integrity of cj0031 and the orientation relative to the CAT promoter in pCfdxA0031 using three pairs of primers: CAT-INV-F, 0031–2108-R; 0046-INVF, 0032-R; and CAT-INV-F, 0031–2832-R.

Adhesion and invasion assays

Campylobacter strains CH11168-Δ0031::kan, CH11168-cj0031-ON (G9) and CH11168-cj0031-OFF (G10) were cultured for 3 days on MHA plates with antibiotics and then a sweep of cells was inoculated into MHB (without antibiotics). After overnight incubation under shaking conditions, the OD600 of the culture was adjusted to 0.5 (~2 x 10⁶ cells) prior to adding 10 µl (1 x 10⁷ cells) into selected wells of a 24-well tissue culture plate containing a confluent monolayer of Caco-2 cells. After 4 h incubation at 37°C with 5% CO2, media was removed and wells were gently washed twice with 2 ml of phosphate buffered saline (PBS). After washing, 1 ml of PBS/1% Triton X100 was added to each well and incubated for 10 min at room temperature to allow for lysis of the Caco-2 cells. Lysates were mixed and then subject to serial dilution in PBS in triplicate. A 50 µl aliquot of each dilution was plated onto an MHA plate and, after 2–3 d, CFUs were counted and the number of adherent bacterial cells was calculated.

For the invasion assays, the Caco-2 monolayer was challenged with 2 x 10⁷ bacterial cells and incubated for 4 h at 37°C. After washing the monolayers, fresh media containing 200 µl gentamicin was added to each well and incubation continued for a further 4 h. Each well was then processed as described above in and the numbers of invasive bacterial cells were determined from duplicate serial dilutions.

Measurement of biofilm formation

Biofilm formation (29) was performed using C. jejuni strains, grown overnight in MHB under shaking conditions, which were then diluted in MHB broth to give a final OD600 of 0.025 (~ 2.5 x 10⁷ CFU). One ml of this suspension was added to each well of a 24-well polystyrene plates (Corning). Plates were incubated for either 48 or 72 h. Following incubation, the media was discarded and wells were washed twice with distilled water. Plates were dried at 55°C for 30 min and then 1 ml of 0.1% crystal violet (CV) was pipetted into each well followed by 10 min incubation at room temperature. The CV solution was removed and wells were washed twice with water followed by drying at 55°C for 15 min. Bound CV was solubilized in 1 ml of 80% ethanol-20% acetone. The absorbance at 570 nm of a 100 µl aliquot was determined using a micro-plate reader (BMG-Labtech) as a measure of the biofilm mass.

Motility assay

Bacterial cells harvested from MHA plates were suspended in MHB and the optical density at 600 nm was adjusted to 0.45 (~1 x 10⁹ CFU/ml). A 2 µl aliquot of this bacterial suspension was stabbed into a 0.4% MHA plate and incubated for 48~72 h.

Testing the efficacy of plaque formation of phages on C. jejuni host strains

Propagation of bacteriophages (Supplementary Table SII) was carried out using the following host C. jejuni strains: PT14, HPC5 and 11168 (30) as previously described with slight modification (31). Briefly, the C. jejuni host strains were sub-cultured on MHA plates for 24 h under microaerobic conditions, bacterial cells were scrapped off and suspended in 10 ml of 10 mM MgSO4 and adjusted to a cell density of ~1 x 10⁹ CFU/ml. A 500 µl aliquot of cell suspension was added into molten NZCYM top agar (0.6% agar w/v), and poured on to the basal NZCYM plates.
Once the droplets had been absorbed, the plates were in triplicate, on to the surface of the NZCYM overlay plates. Bacterial Isolation kit with Max™ Bacterial Enhancement prepared in SM buffer and applied as 10^2 (1.2% agar w/v). Serial 10-fold, dilutions of the phages were prepared in SM buffer and applied as 10 μl droplets, in triplicate, on to the surface of the NZCYM overlay plates. Once the droplets had been absorbed, the plates were incubated for 24 h, under micro-aerobic conditions, at 42°C. The plaques were counted and used to calculate the plaque forming units per ml (PFU/ml).

SMRT sequencing and detection of the methylome

Genomic DNA was prepared using a DNeasy extraction kit (Qiagen) and sequenced on the Pacific Biosciences RSII machine according to standard manufacturer’s conditions. Template preparation used kit version 3.0, polymerase binding used version P4 and sequencing reagents were version C2. Data was captured using SMRT cells, to give genome coverage of >200-fold per sample. Data has been deposited in the European Nucleotide Archive, and accession numbers are given in Supplementary Table SIII. The genomic positions of each methylation site motif for Cj0031 were determined by a word search.

The relative positions of each 5’CCYGA sequence in the genome sequence of C. jejuni strain NCTC11168 were detected using a script to determine position relative to the nearest coding sequence.

Analysis of gene expression by RNASeq

Total RNAs were extracted from three independent log phase cultures of CH11168 cj0031-ON and the deletion mutant CH11168-Δcj0031::kan using the TRIZol® Max™ Bacterial Isolation kit with Max™ Bacterial Enhancement Reagent (Invitrogen, Paisley, UK) according to the manufacturer’s instructions, before ethanol-precipitation and purification using the RNeasy® Mini kit of Qiagen (Crawley, UK) according to the manufacturer’s instructions. Oncolumn DNase treatment was included during the purification using the RNase-free DNase set (Qiagen). The purified total RNAs were assessed for quality and rRNA decontamination using the RNase-free DNase set (Qiagen). Total RNAs were converted to cDNA using the Omniscript cDNA synthesis kit (Qiagen) according to the manufacturer’s protocol. Specific primers for qRT PCR were designed with lengths of 18–24 nt with melting temperatures at 57°C. Aliquots of cDNA were used as the template for qRT PCR. An optical 46 well microtiter plate (Applied Biosystems) was used with a 20 μl reaction volume consisting of Power SYBR® Green PCR master mix (Life Technologies), 50 nM gene specific primers and the template. An ABI Prism 7000 sequence detector (Applied Biosystems) was programmed for an initial set up of 30 s at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 57–58°C. SYBR® Green detects double stranded DNA. A melt curve was obtained from a first step starting from 50–95°C to control specificities of quantitative PCR reaction for each primer pair. Cycle threshold (CT) values were determined using Step one software version 2.0 (Applied Biosystems). The comparative threshold cycle method was used to calculate change (n-fold) where samples were normalized to the internal control product of the sodB gene (36), which showed no change in expression levels between CH11168 cj0031-ON and CH11168-Δcj0031::kan. Reactions were performed in triplicate. The fold changes were calculated using the 2ΔΔCt method. The primer pairs utilized in the qRT PCR experiments to determine the change in expression for selected genes are given in Supplementary Table SIV.

RESULTS

Distribution and conservation of Cj0031 homologues

The cj0031 gene of C. jejuni strain NCTC11168 is 3732 bp in length and contains a polyG tract of nine residues at position 2573 bp. Deletions or insertions in this tract can separate the gene into two reading frames (called cj0031 and cj0032 in the original annotation of the genome sequence). The cj0031 amino acid sequence (1243 amino acids) has homology to known Type IIG RM enzymes with conserved domains present on either side of the polyG tract that includes conservation of amino acids involved in specific and non-specific DNA interactions in the M.TaqI crystal structure (Figure 1). An unusual feature of these systems is that the methyltransferase and endonuclease activities are fused into a single reading frame. PV mediated by alterations in
**Figure 1.** Alignments and sequence conservation of Cj0031 homologues. (A) The positions and sequences of the conserved domains of the amino acid sequence of Cj0031 (top sequence) are shown in comparison to M.TaqI (2IBS_A; GI:155126; bottom sequence) and the consensus sequences of the Eco57I (pfam07669; middle sequence) and TaqI (pfam12950; middle sequence). The amino acids marked with an asterisk are involved in DNA interactions, specific and non-specific, as shown by Goedecke et al. for M.TaqI (51). The position of the phase-variable G-tract in Cj0031 is marked as G9.

(B) Representation of the conserved regions in Cj0031 and homologues from Campylobacter jejuni strains whose methylomes have been characterized. The dark grey boxes indicate >90% homology between sequences. The light grey regions indicate >90% homology between the two sequences but divergence from the other two sequences.

Orthologues of Cj0031 were found in several other species including a distant association with a putative RM system of the cyanobacteria, *Microcystis aeruginosa* (data not shown). Homologues from *Campylobacter* and *H. pylori*, the closest neighbour of *C. jejuni*, fell into two groups, separated by homologues in more distant organisms, suggestive of the possibility of acquisition by horizontal transfer (Figure 2). Homologues were present in other *Campylobacters* including all other *C. jejuni* strains. A notable feature was the lack of a repeat tract in many of the other strains indicating that PV only occurs in a limited number of alleles and is not a universal feature of this gene. For example, strain 81–176 had an interrupted repeat tract (4G-A-4G) that will effectively prevent PV in this allele. Alignments were performed between the *cj0031* nucleotide and amino acid sequences and those of homologues from the four *C. jejuni* strains whose methylomes have been determined (37–39). Strain F38011 exhibited a ~300 bp deletion in the central portion of the gene, expected to abrogate methylation activity and explaining the lack of methylation at sites recognized by this protein in the genome. The other three strains showed high levels of conservation in the N-terminal (regions I and II) and C-terminal regions (region V) with the exception of significant divergence in amino terminal region 1 of strain 81–176. The sequences diverged into two classes in region III and were all highly divergent in region IV suggesting the presence of a variable domain responsible for observed differences in site-specific methylation by each of these Cj0031 homologues.

**Characterization of the recognition site of Cj0031 by inhibition of restriction**

A previous attempt to identify the methylation site of Cj0031 was compromised by use of a strain with an OFF
number of repeats in \textit{cj0031} (37). To overcome this difficulty, both ON and OFF phase-variants were isolated from a population of cells for \textit{C. jejuni} strain NCTC11168 (see Methods). A fixed OFF strain (11168/\textit{Delta}1\textit{cj0031}::kan) was generated by constructing a deletion in \textit{cj0031}. Complementation of this mutation (strain 11168/\textit{Delta}1\textit{cj0031-cj0031c}) was achieved by insertion of the wild-type, phase-variable gene into a pseudogene (with expression controlled by a heterologous promoter) and isolation of a variant with an ON number of repeats. Genomic DNA from these strains was digested with a range of methylation sensitive restriction enzymes and tested for differences in the restriction patterns using probes from three different regions of the genome (Figure 3). An obvious difference in the restriction pattern of \textit{cj0031}-ON versus 11168/\textit{cj0031} and 11168/\textit{Delta}1\textit{cj0031::kan} was detected using probe III with inhibition of HindIII digestion in ON but not OFF phase variants or a deletion mutant (Figure 3). Protection of this site was partially restored in the complementation strain, 11168/\textit{Delta}1\textit{cj0031c}, confirming that protection was due to \textit{Cj0031}, with the incomplete level of protection being due to a sub-optimal expression from the \textit{fdxA} promoter in the complementation vector (36). Additional phase-variants were tested with probe III and exhibited the same restriction pattern (data not shown). Critically, another HindIII site present in Probe I did not show any difference in restriction pattern (Figure 3) suggesting that the putative \textit{Cj0031} M\text{\textit{T}a}se recognition site did not correspond to but overlapped with one of the two end sequences of the HindIII 5′-AAGCTT-3′ recognition sequence in Probe III with inhibition by methylation on the first adenine.

As \textit{Cj0031} shares 87\% identity with \textit{CjeFV} of \textit{C. jejuni} strain 81–176 whose recognition sequence was characterized as 5′-GGGCA or 5′-GGACA (37), we hypothesized that the \textit{Cj0031} recognition sequence was GC-rich and had one of three sequences: 5′-CCGAA-3′; 5′-CGAAG-3′; or 5′-CCCCGA-3′ (Figure 2). The alternative was that the site was AT-rich and had a recognition sequence of 5′AAGAA or 5′AGAAG. The PSR-HindIII-CCG probe was designed to investigate if the recognition site was 5′CCGAA or 5′CGAAG. No inhibition of restriction was observed (Figure 3D). This result indicated that all three cytosine residues were required and that a \textit{Cj0031} recognition sites could be 5′-CCCGA-3′.

\textbf{Detection of the methylation motif of \textit{Cj0031} by SMRT sequencing}

SMRT is a robust and reliable method for detecting methylation patterns in DNA sequences. The PacBio RS sequencing platform was utilized to generate methylomes for an ON phase variant and deletion mutant of \textit{cj0031} in \textit{C. jejuni} strain NCTC11168 strain. Both genomes contained two degenerate motifs, a palindromic motif (RA\textit{ATTY}) and a non-palindromic motif (GKA\textit{AYG}; Table 1) that had been previously identified (37) as the recognition sites of \textit{CjeNII} (Cj1051c), \textit{M.CjeNIV} (Cj1553c), \textit{CjeNIII} (Cj0690c) and \textit{M.CjeNI} (Cj0208). In contrast, a novel methylation motif—5′CCYGA—was present in the genome sequences for the ON phase variant of \textit{Cj0031} but absent from the deletion mutant. This 5′CCYGA motif, as also observed for 5′GKAAYG, only exhibited methylation on the adenine of one of the DNA strands. A high level of methylation was observed with 1759 of 1777 sites being detected as having methylation with a mean modification QV of 75\%. These results confirmed the modification of 5′CCGA in \textit{Cj0031} ON phase variants, as detected by Southern blotting, and additionally indicated modification of 5′CCTGA sites by this enzyme.

\textbf{Investigation of the impact of phase variation or mutation of \textit{cj0031} on adherence, invasion and biofilm formation by \textit{C. jejuni}}

Motility is a critical phenotype required for colonization and invasion by \textit{C. jejuni}. As the laboratory strain utilized for the methylome studies is non-motile, phase variants and mutants of \textit{cj0031} were obtained for a hypermotile chicken-adapted variant (i.e. CH11168) of \textit{C. je-
Figure 3. Identification of the methylation target site for Cj0031. Chromosomal DNA was extracted from ON and OFF phase variants (i.e. 11168-cj0031-ON and 11168-cj0031-OFF, respectively), a deletion mutant (11168-Δcj0031-kan) and a complementation mutant (11168Δcj0031-cj0031) of cj0031 in Campylobacter jejuni strain NCTC11168. DNA was subjected to restriction with HindIII followed by probing of Southern blots with region specific probes. (A) Schematic diagram of the HindIII recognition sites (A, B and C) and overlapping putative Cj0031 recognition site present in Probe III. The figure indicates the sizes of fragments generated by complete and partial HindIII digestion with the latter resulting from protection of site B by Cj0031. (B) Analysis of HindIII restriction fragments using probe III. Inhibition of restriction is observed in the ON phase variant and complementation strain but not the OFF variant or deletion mutant. (C) Analysis of HindIII restriction fragments using probe I. (D) Analysis of HindIII restriction fragments using probe PSR-HindIII-CCG. This probe covers a region in which the 5′-CCGAA and 5′-CGAAG sequences (present in Probe III) overlap with the HindIII sequence but are not protected against HindIII digestion. Black rectangles, position of probes; Cj0031?, putative Cj0031 recognition site; HindIII, recognition site for HindIII.

Table 1. Methylation motifs identified within the genomes of a C. jejuni NCTC11168 cj0031 ON phase-variant and a Δcj0031 mutant

<table>
<thead>
<tr>
<th>Motifs</th>
<th>Modified position</th>
<th>Type</th>
<th>% Motifs detected</th>
<th># Of motifs detected</th>
<th># Of motifs in genome</th>
<th>Partner motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. jejuni NCTC11168 cj0031 ON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACNNNNNNGT</td>
<td>1</td>
<td>m6A</td>
<td>99.91%</td>
<td>1065</td>
<td>1066</td>
<td>GAGNNNNNGT</td>
</tr>
<tr>
<td>GAGNNNNNGT</td>
<td>2</td>
<td>m6A</td>
<td>99.25%</td>
<td>1058</td>
<td>1066</td>
<td>ACNNNNNCTC</td>
</tr>
<tr>
<td>TAAYNNNNNTGA</td>
<td>3</td>
<td>m6A</td>
<td>99.79%</td>
<td>468</td>
<td>469</td>
<td>GCANNNNNTTA</td>
</tr>
<tr>
<td>GCANNNNNRTTA</td>
<td>3</td>
<td>m6A</td>
<td>97.87%</td>
<td>459</td>
<td>469</td>
<td>TAAYNNNNNTGC</td>
</tr>
<tr>
<td>GKAAYG</td>
<td>4</td>
<td>m6A</td>
<td>99.68%</td>
<td>1255</td>
<td>1259</td>
<td></td>
</tr>
<tr>
<td>RAATTY</td>
<td>3</td>
<td>m6A</td>
<td>99.01%</td>
<td>27 011</td>
<td>27 280</td>
<td>RAATTY</td>
</tr>
<tr>
<td>C. jejuni NCTC11168 Δcj0031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACNNNNNNGT</td>
<td>1</td>
<td>m6A</td>
<td>99.91%</td>
<td>1065</td>
<td>1066</td>
<td>GAGNNNNNGT</td>
</tr>
<tr>
<td>GAGNNNNNGT</td>
<td>2</td>
<td>m6A</td>
<td>99.25%</td>
<td>1058</td>
<td>1066</td>
<td>ACNNNNNCTC</td>
</tr>
<tr>
<td>TAAYNNNNNTGA</td>
<td>3</td>
<td>m6A</td>
<td>99.79%</td>
<td>468</td>
<td>469</td>
<td>GCANNNNNTTA</td>
</tr>
<tr>
<td>GCANNNNNRTTA</td>
<td>3</td>
<td>m6A</td>
<td>97.87%</td>
<td>459</td>
<td>469</td>
<td>TAAYNNNNNTGC</td>
</tr>
<tr>
<td>GKAAYG</td>
<td>4</td>
<td>m6A</td>
<td>99.68%</td>
<td>1255</td>
<td>1259</td>
<td></td>
</tr>
<tr>
<td>RAATTY</td>
<td>3</td>
<td>m6A</td>
<td>99.01%</td>
<td>27 011</td>
<td>27 280</td>
<td>RAATTY</td>
</tr>
</tbody>
</table>
The functional consequences of PV or inactivation of *cj0031* on motility were examined using Mueller-Hinton-soft-agar plates. No obvious differences in motilities of these strains were observed, indicating that *cj0031* methyltransferase activity does not affect the motility of *C. jejuni* (Supplementary Figure S1). Having shown these strains to be motile, the adhesion and invasion phenotypes were investigated. In two separate experiments, both an OFF variant and the deletion mutant (i.e. CH11168Δcj0031::kan) exhibited significant 2-fold lower levels of adhesion to Caco-2 cells than an ON variant (Figure 4). Complementation restored the adhesive phenotype to the same level as the ON variant. Invasion was assayed using Caco-2 cells and gentamycin-treatment. Both the OFF variant and deletion mutant of *cj0031* exhibited a significant reduction of 2–3-fold as compared to the wild-type *cj0031* ON variant. Complementation of the *cj0031* deletion mutant with an ON variant of the native *cj0031* gene partially restored the invasion phenotype (Figure 4).

Additionally, the contributions of *Cj0031* to biofilm formation were examined using a CV staining method (29,40). Both OFF phase-variants and a *cj0031* deletion mutant demonstrated 7- and 4-fold decreases, respectively, in biofilm formation as compared to an ON phase-variant after both 48 and 72 h of growth (Figure 5). The complementation strain (i.e. 11168-Δcj0031-cj0031) partially restored the phenotype to a level that was significantly (*P < 0.001) different from the deletion mutant, indicating that absence of *cj0031* is deleterious to biofilm formation by *C. jejuni* strain NCTC11168. A further biofilm experiment was performed to test whether PV of *cj0031* or other genes occurred during the course of the biofilm assay. Differences
Transcriptional profiling of Cj0031 phase variants and mutants

Alterations in methylation of Cj0031 recognition sites could impact on transcription of specific genes and produce differences in phenotypic behaviour. RNASeq profiles were generated for an ON variant in strain NCTC11168 and the deletion mutant of cj0031. Examination of sequences generated from the RNA-Seq data confirmed that cj0031 contained an ON number of repeats and that the mutant lacked the gene. Major differences in gene expression were observed with 89 and 130 genes exhibiting a statistically-significant reduction or increase, respectively, of >1.5-fold in the mutant (Figure 6). A subset of 20 genes exhibited changes of >5-fold (Figure 6; with 14 having reduced and 6 increased expression). An qRT-PCR of nine genes confirmed the observed changes in relative gene expression levels (Supplementary Figure S2). Major changes in gene expression were observed in two tRNA genes and genes whose products contribute to metabolism (e.g. Uxa, altrolate hydrolase; Cj0073-Cj0075, iron-sulphur cluster proteins) or transport (e.g. PutP, a sodium-proline symporter; HydA2, Ni-Fe hydrogenase; LctP, L-lactate permease). Only one of nine genes (i.e. capA, cadF, plldA, pblA, flpA, cadA, cipA, cj0268c and cj1349) with evidence of roles in adhesion and invasion, was altered in expression with a 3-fold increase in expression of pblA.

Altered expression of genes in Δcj0031 may be due to differential methylation of 5’CCYGA sequences. The NCTC11168 genome contains 1776 of these sequences with an average of 1.1 sites per kbp and ~1 site per gene. Only 69 sites are located in intergenic regions and these sites are associated with 38 genes. Significant >1.5-fold alterations in expression were only observed for three of these genes in Δcj0031 (two upregulated and one downregulated). None of these genes had known roles in adhesion, invasion or biofilm formation. The loci that contained genes exhibiting large differences (>5-fold) in expression were examined for the presence of Cj0031 methylation sites. In some of these loci, 2-4 5’CCYGA sequences were present within the reading frames of multiple genes but in other loci most of the genes contained 0–1 sequences (Supplementary Figure S3). As an alternative, loci with 4–6 genic 5’CCYGA sequences were examined for alterations in gene expression and multiple examples were observed with <1.5-fold changes in expression. There was however an association between the presence of multiple recognition sites and an increase in gene expression with 23% of genes containing four or more repeats (n = 83) exhibiting a significant increase in expression whereas only 7% and 10% of genes with 1–3 (n = 1038) or 0 (706) sites, respectively, exhibited an increase in expression. However the average increase in expression for genes containing four or more sites was 2-fold with no genes showing a >3.5-fold increase in expression. Overall these results indicated that there is only a weak association between alterations in gene expression in Δcj0031 and the presence of 5’CCYGA sequences.

Assessment of the phage restriction activity of Cj0031

The classic activity of RM enzymes is to prevent phage replication. Stocks of Campylobacter bacteriophages propagated on C. jejuni strain PT14 were tested for plaque formation at the routine test dilution (Log10 6 PFU) on bacterial lawns of cj0031 phase-variants and the deletion mutant. Some of these bacteriophages exhibited plaque formation on the cj0031 OFF variant or deletion mutant but not on cj0031 ON. These phages were selected for further study in order to determine whether they were subject to restriction by the putative endonuclease activity of Cj0031. Phages were propagated on Δcj0031 to allow for methylation on sites for all of the NCTC11168 RM systems except cj0031 and then tested for growth on cj0031 phase-variants and mutants. The efficiency of plating (EOP) values of the deletion mutant and cj0031 OFF variant were similar, whereas the EOPs recorded for the cj0031 ON variant exhibited between 5- and 1000-fold reductions in PFU (Table 2). These data are consistent with the hypothesis that cj0031 encodes a functional endonuclease.

The potential of an RM system to restrict phage replication depends on the presence and frequency of recognition sites in phage genomes. The methylation site frequencies were determined for all of the C. jejuni RM systems of ten phages whose genome sequences were available and ranged in size from 132 to 178 kb (Table 3). The Cj0031 site was present in all phage genomes at similar frequencies including phages with sensitivity (i.e. CP8 and CP30) to restriction on Cj0031 ON variants. A comparison between enzymes indicated that the highest site prevalence was for the AT-rich recognition site of Cj0208 (3-4 -sites/kb for RAATTY) whilst both Cj0301 and Cj0690 exhibited frequencies of ~1/kb whereas the more complex sites of Cj1051 and Cj1553 were present at <0.5 sites/kb.

DISCUSSION

Campylobacter jejuni contains several phase variable genes with known or potential roles in survival and adaptation of this species to various niches in hosts and the environment. While phase-variable genes usually encode surface structures, a provocative association with non-surface coding genes is widely reported and in particular a distinctive PV signal is found for RM systems of several pathogenic bacteria (19–22,41). PV of surface molecules is an adaptive strategy enabling escape of not only innate and adaptive host immune responses but also a dynamic and evolving phage population (8,9). The phase-variable RM systems
provide another strategy to escape selection by phages. A more surprising finding was of a phase variation whereby PV of the methyltransferase of a phase-variable RM system coordinates the expression of multiple genes through changes in genome methylation (19). We define herein the recognition sites of a novel, phase-variable Type IIG RM system and investigated the contributions of this enzyme to a Campylobacter phase variation and phage restriction.

Previous SMRT sequencing of C. jejuni strain NCTC11168 detected methylation sites for four of the putative RM systems but not for Cj0031 as this gene was in an OFF state in the analysed variant of this strain (37). We isolated an ON variant and constructed both a deletion
Table 2. Relative efficiency of plating of phages on phase variants and mutants of cj0031 in C. jejuni strain NCTC11168

<table>
<thead>
<tr>
<th>Phage</th>
<th>Relative EOP&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cj0031 ON</td>
</tr>
<tr>
<td>CP8</td>
<td>0.05</td>
</tr>
<tr>
<td>CP25</td>
<td>0.04</td>
</tr>
<tr>
<td>CP30</td>
<td>0.01</td>
</tr>
<tr>
<td>A1b</td>
<td>0.2</td>
</tr>
<tr>
<td>X3</td>
<td>0.1</td>
</tr>
<tr>
<td>G3</td>
<td>0.01</td>
</tr>
<tr>
<td>G4</td>
<td>0.06</td>
</tr>
<tr>
<td>MC1a</td>
<td>0.008</td>
</tr>
<tr>
<td>11</td>
<td>0.009</td>
</tr>
<tr>
<td>19b</td>
<td>0.005</td>
</tr>
<tr>
<td>NCTC12671</td>
<td>0.001</td>
</tr>
<tr>
<td>NCTC12679</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<sup>a</sup>Phages were propagated on 11168Δcj0031;
<sup>b</sup>EOP of the sensitive phages on the test host relative to a non-restricting host.

Table 3. Prevalence of recognition sites for the RM systems of C. jejuni strain NCTC11168 in phage genomes

<table>
<thead>
<tr>
<th>Phage&lt;sup&gt;a&lt;/sup&gt;/Group&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Cj0208</th>
<th>Cj1553c</th>
<th>Cj0031</th>
<th>Cj1051c</th>
<th>Cj0690c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tot&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Fre&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Tot</td>
<td>Fre</td>
<td>Tot</td>
</tr>
<tr>
<td>CP220/II</td>
<td>665</td>
<td>3.7</td>
<td>59</td>
<td>0.3</td>
<td>175</td>
</tr>
<tr>
<td>CP10/II</td>
<td>668</td>
<td>3.8</td>
<td>59</td>
<td>0.3</td>
<td>177</td>
</tr>
<tr>
<td>CP21/II</td>
<td>713</td>
<td>3.9</td>
<td>66</td>
<td>0.4</td>
<td>183</td>
</tr>
<tr>
<td>CP81/III</td>
<td>403</td>
<td>3.0</td>
<td>57</td>
<td>0.4</td>
<td>149</td>
</tr>
<tr>
<td>NCTC12673/III</td>
<td>394</td>
<td>2.9</td>
<td>56</td>
<td>0.4</td>
<td>152</td>
</tr>
<tr>
<td>CPX/III</td>
<td>408</td>
<td>3.1</td>
<td>57</td>
<td>0.4</td>
<td>151</td>
</tr>
<tr>
<td>CP30A/III</td>
<td>388</td>
<td>2.9</td>
<td>53</td>
<td>0.4</td>
<td>142</td>
</tr>
<tr>
<td>CP8/III</td>
<td>408</td>
<td>3.1</td>
<td>57</td>
<td>0.4</td>
<td>151</td>
</tr>
</tbody>
</table>

<sup>a</sup>Accession numbers: FN667788, FN667789, HE815464, FR823450, GU296433, JN132397, JX569801, JK148616;
<sup>b</sup>Group, phages belonging to group II have genomes of 180–190 kb whilst those in group III have genome sizes of 130–140 kb (52,53);
<sup>c</sup>Tot, total number of sites per genome;
<sup>d</sup>Fre, frequency of sites per kb.

The phylogenetic analysis of cj0031 showed that homologues of cj0031 were present in H. pylori, a close relative to C. jejuni, as well as several other more diverse bacterial species. Homologues of cj0031 were present in other strains of C. jejuni and have a 90–99% sequence similarity with conservation of the motifs characteristic of restriction endonucleases. Despite this high overall level of conservation, one region (IV; Figure 1) of the protein exhibited high levels of variation between all alleles of Cj0031 for which methylomes have been determined. This region is predicted to encode the TRD as has been predicted for the hypervariable regions of other RM-associated methyltransferases. The methylation site of CfeFV, the homologue of cj0031 in strain 81–176, has been characterized by SMRT sequencing as 5′GGRCA (37) whereas that of Cj0031 was 5′CCYGA. The putative TRD regions of these homologues differ by 87–91 amino acids in the 136–140 residues of region IV suggesting that the amino acids responsible for site-specific DNA recognition are located therein. The IA3902 homologue, CjeIAORF32P, exhibits significant amino acid variation from Cj0031 in both regions III and IV (Figure 1) and methylates at a 5′GAAGAA sequence (38). A notable feature of these alignments is the absence of the polyG tract in the other Cj0031 homologues suggesting that PV may be restricted to a sub-set of alleles.

Populations of C. jejuni strain NCTC11168 will contain phase-variants of Cj0031 with either methylated (ON) or non-methylated (OFF) 5′CCYGA sequences, raising the potential for association of a phasevarion with PV of Cj0031. A phasevarion is a set of genes whose expression is regulated by ON-OFF switching of site-specific methylation by a phase-variable methyltransferase. The observation of reductions in adhesion, invasion and biofilm formation in cj0031 deletion mutants was indicative of association of Cj0031 PV with a phasevarion. These alterations in host cell attachment were not due to differences in motility (Supplementary Figure S1) but were in general minor differences in phenotype. The degree of these effects could have been reduced by PV of the ON variants so that starting popula-
tions contained variable proportions of OFF variants. An alternative explanation of differences in phenotype is the impact of PV of other genes. The potential importance of this caveat was reduced by the observation of similar phenotypes for the ON variant of the wild-type and complementation mutant and similarly for the OFF variant and deletion mutant. Furthermore, no influence on biofilm formation of PV of cj0031 during the experiment or of differences in expression of other phase-variable genes between strains was observed in additional repetitions of the biofilm assay. Inactivation of another putative methyltransferase, Cj1461, reduced both the motility and invasion by C. jejuni strain 81–176 (42). Similarly, Fields and Thompson (43) reported that inactivation of CsrA, a global post-transcriptional regulator, reduced motility, biofilm formation and adherence of C. jejuni strain 81–176 but enhanced invasion leading to the proposal that CsrA directly or indirectly represses invasion specific genes. The RNASeq analysis of Δcj0031 detected an altered level of expression for Peb1A (cj0921c, −2.8-fold) but not for any other known adhesins or invasins. Peb1A is a periplasmic protein that mediates uptake of dicarboxylic amino acids (44). A variable association of this gene with reductions in adhesion has been reported (45). Thus, the Cj0031 methyltransferase might act as a phasevarion with alterations in expression of Peb1A resulting in reductions in host cell adhesion and invasion. Alternatively the alterations in metabolic pathways, also observed in the Δcj0031 RNASeq data, could be responsible for the observed phenotypic variation. A final possibility is that observed differences in gene expression may have impacted on other types of phenotypic variation.

The proposed mechanism of regulation of a phasevarion is through differential methylation of promoter elements leading to alterations in binding of regulatory proteins and/or RNA polymerase to promoter elements in an analogous manner to Dam methylase-mediated regulation of bacterial gene expression (17 and 18). There was only weak evidence of a genome-wide association between presence of 5′CCYGA sequences and alterations in transcription as detected in the RNASeq data. An alternate possibility is that variable methylation at only some of the 5′CCYGA sites has any effect on transcription and that the differential regulation is through perturbation of one or more global regulators of gene expression. However, there were no significant alterations in expression of known two-component regulators (i.e. RacRS, DccRS, cbrRS and FlgRS), sigma factors (FliA and RpoN), CsrA or of Fur (a significant 2-fold change in fliR was observed but this was disregarded as there was no overall change in expression of the fliA genes). Finally, the Peb1A gene contained a single 5′CCYGA sequence within the reading frame, which is the number expected by chance for any gene of this genome. Thus, neither the overall nor specific distribution of CJ0031 methylation sites provides a clear insight into the mechanism for altering gene expression as implied by differences in phenotypic variation or as detected in the RNASeq data.

An alternate putative function of all RM systems is restriction of phage replication. Testing of a phage panel revealed that inactivation of cj0031 resulted in a higher EOP for 12 phages. This loss of inhibition of phage replication provides evidence that Cj0031 is an active restriction endonuclease with the ability to restrict phage DNA. The efficiency of restriction by RM systems is, for some systems, directly proportional to the number of recognition sites located in the genome of the invading element (46–48) with low numbers of restriction sites for a specific RM system correlating with a greater probability to escape restriction. The high prevalence of sites for Cj0031 in two of the restricted phages, CP8 and CP30, supports the view that the restriction activity is correlated with the number of restriction sites. Intriguingly, the cj0031 site has a higher prevalence (1.5- to 6-fold) in all phage genomes than that of the homologue present in strain PT14 (i.e. A911_00150; GAACAA). If the main function of Cj0031 is restriction of phage growth and the number of recognition sites is correlated with higher restriction of phage growth, then the Cj0031 and CfeFV alleles may be more widely distributed in Campylobacter genomes than the alleles with 6 bp recognition sites. Evolution of PV of Cj0031 could then be driven by selection for phage resistance in the ON state and selection for expression of one state of the phasevarion in the alternate state. Alternatively selection for the OFF state may be driven by receptiveness to adaptive mutations associated with natural transformation (assuming in-coming DNA can be restricted by Cj0031 endonuclease activity) or fluctuating levels of Cj0031-resistant methylated phage genomes as speculated elsewhere (49).

Our study provides evidence of a dual role for PV of the Type IIG enzyme, Cj0031, of C. jejuni strain NCTC11168 as a limiter of phage spread and a regulator of a gene expression. The underlying mechanism of action for the alteration in gene expression is not known and does not appear to involve differential methylation of specific recognition sites by Cj0031. Nevertheless PV of Cj0031, as detected during infections in chickens and mice (8,50), produces changes in expression of multiple genes that may lead to phenotypic variation and generation of highly differentiated variants with differing capabilities of adaptation to diverse niches.

**SUPPLEMENTARY DATA**
Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENT**
The authors are grateful to Dr Richard Haigh for provision of the CH11168 strain and for his assistance with the adhesion and invasion assay.

**FUNDING**
Wellcome Trust Grant [098051]. Funding for open access charge: University of Leicester [M38DF47].

**Conflict of interest statement.** None declared.

**REFERENCES**


