A novel paralogous gene family involved in phase-variable flagella-mediated motility in Campylobacter jejuni

Andrey V. Karlyshev,1 Dennis Linton,1 Norman A. Gregson2 and Brendan W. Wren1

Author for correspondence: Brendan W. Wren. Tel: +44 207 927 2288. Fax: +44 207 637 4314. e-mail: brendan.wren@lshtm.ac.uk

1 Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, University of London, Keppel Street, London WC1E 7HT, UK
2 Department of Neuroimmunology, Division of Clinical Neurosciences, KCL, Guy’s Campus, London SE1 1UL, UK

Flagella-mediated motility is recognized to be one of the major factors contributing to virulence in Campylobacter jejuni. Motility of this bacterium is known to be phase variable, although the mechanism of such variation remains unknown. C. jejuni genome sequencing revealed a number of genes prone to phase variation via a slipped-strand mispairing mechanism. Many of these genes are hypothetical and are clustered in the regions involved in formation of three major cell surface structures: capsular polysaccharide, lipooligosaccharide and flagella. Among the genes of unknown function, the flagellar biosynthesis and modification region contains seven hypothetical paralogous genes designated as the motility accessory factor (maf) family. Remarkably, two of these genes (maf1 and maf4) were found to be identical and both contain homopolymeric G tracts. Using insertional mutagenesis it was demonstrated that one of the genes, maf5, is involved in formation of flagella. Phase variation of the maf1 gene via slipped-strand mispairing partially restored motility of the maf5 mutant. The maf family represents a new class of bacterial genes related to flagellar biosynthesis and phase variation. Reversible expression of flagella may be advantageous for the adaptation of C. jejuni to the varied in vivo and ex vivo environments encountered during its life cycle, as well in evasion of the host immune response.

Keywords: C. jejuni, flagellar biosynthesis, phase variation, slipped-strand mispairing, genetic instability

INTRODUCTION

The flagellum is an essential virulence factor of Campylobacter jejuni and is involved in colonization of the gastro-intestinal tract (Nuijten et al., 1992). Flagella-associated motility of C. jejuni has been found to be variable (Nuijten et al., 1995), and reversible non-motile variants have been reported after extended incubation (Caldwell et al., 1985). Different mechanisms for high-frequency phase variation in other bacteria, including those mediated by the slipped-strand mispairing mechanism, are well documented (Henderson et al., 1999). In Neisseria gonorrhoeae, phase variation of expression of pili is attributed to homopolymeric tracts present in a gene involved in pilin biosynthesis (Jonsson et al., 1991). Among other examples are the control of Neisseria meningitidis pilin modification by a phase-variable gene (Jennings et al., 1998) and phase variation of a gene encoding a galactosyltransferase of C. jejuni (Linton et al., 2000a). Recently, in the closely related pathogen, Campylobacter coli, phase-variable flagellar expression due to reversible frameshift mutation in a short homopolymeric tract of T residues in the flhA gene has been reported (Park et al., 2000).

The flagellar biogenesis apparatus in bacteria is complex and, despite extensive studies, functions of some flagella-related genes remain unknown (Macnab, 1996). Analysis of the C. jejuni genome sequence (Parkhill et al., 2000) revealed a number of putative flagella-related genes similar to genes of Escherichia coli and other bacteria. In addition, genes involved in Campylobacter flagellin...
glycosylation were found (Guerry et al., 1996; Linton et al., 2000b). The genes involved in flagellin glycosylation are part of a large cluster also containing genes thought to be involved in sugar biosynthesis and transport, as well as seven closely related genes of unknown function (Parkhill et al., 2000). We have termed these genes (Fig. 1) the motility accessory factor (maf) family of flagellin-associated proteins. Two maf genes (maf1 and maf4) appear to be identical, both containing homopolymeric G tracts. In this study we demonstrate that the genes of the maf family are involved in the variation of motility via the slipped-strand mispairing mechanism.

**METHODS**

**Bacterial strains, plasmids and growth conditions.** C. jejuni NCTC 11168 was received from J. Ketley (University of Leicester, UK). C. jejuni strains were grown at 37 °C on 7% blood agar or Mueller–Hinton agar plates in a microaerobic atmosphere (CampyGen, Oxoid) for 2 d. *E. coli* XL2 Blue MRF’ (Stratagene), used in cloning experiments, was grown overnight at 37 °C on LB agar plates. Unless stated otherwise, 0.3% Mueller–Hinton agar swarm plates were used in motility experiments and the plates were incubated at 37 °C in microaerobic conditions for either 2 d (motility studies) or 5 d (for isolation of spontaneous derivatives with increased
motility). Where necessary the agar plates were supplemented with kanamycin at 50 μg ml⁻¹, chloramphenicol at 10 μg ml⁻¹ and ampicillin at 100 μg ml⁻¹.

**Construction of mutants.** In order to mutate the maf5 gene, a 2 kb fragment of this gene was amplified using primers AK100 (GCT TTA AGC GGT TTT GAG TAT AAC AAC TTA CGC) and AK101 (CGT TCG TGT GCA TAA ACC CAA GC) (Fig. 2) and cloned into the pGEM-T vector. A unique BsaBI site in the cloned fragment was used for insertion of blunt-ended BamHI kan' cassette (Trieu-Cuot et al., 1985). The size and origin of the inserts were confirmed using PCR and restriction mapping techniques (data not shown). The construct was used for transformation of C. jejuni via electroporation (Wassenaar et al., 1993). Mutants with direct orientation of the cassette were selected using PCR with primers DL3 (kan' cassette specific, ACC CAG CGA ACC CAA GC) and DL38 (flaB specific, GTC AGG CTA ATG CAG TGC AG) (Fig. 2).

Double mutants were constructed using a chloramphenicol-resistance (cm') cassette (Wang & Taylor, 1990) isolated from the pAV35 delivery plasmid (van Vliet et al., 1998) and used as a selection marker. A recombinant plasmid cam11f9 was constructed using the SmaI fragment containing the cm' cassette was inserted into the BsaBI site of plasmid cam11f9 and chloramphenicol-resistant colonies were selected after transformation of the revertant strain Cj1337::kan' rev. The size and origin of the inserts were confirmed using PCR and restriction mapping (data not shown).

**Sequencing of the homopolymeric tracts.** The poly(G)-tract-containing regions of each gene were amplified from chromosomal DNA using primer DL8 (CAT AAT AAT GAC TCA TCA TTA CGC) and AK101 (GCT TCG TGT GCA TAA ACC CAA GC) (Fig. 2) and cloned into the pGEM-T vector. The construct was used for transformation of C. jejuni via electroporation (Wassenaar et al., 1993). Mutants with direct orientation of the cassette were selected using PCR with primers DL3 (kan' cassette specific, ACC CAG CGA ACC CAA GC) and DL38 (flaB specific, GTC AGG CTA ATG CAG TGC AG) (Fig. 2).

Double mutants were constructed using a chloramphenicol-resistance (cm') cassette (Wang & Taylor, 1990) isolated from the pAV35 delivery plasmid (van Vliet et al., 1998) and used as a selection marker. A recombinant plasmid cam11f9 was constructed using the SmaI fragment containing the cm' cassette was inserted into the BsaBI site of plasmid cam11f9 and chloramphenicol-resistant colonies were selected after transformation of the revertant strain Cj1337::kan' rev. The size and origin of the inserts were confirmed using PCR and restriction mapping (data not shown).

**PCR conditions.** Crude cell lysates were amplified using Gibco-BRL Taq polymerase in 20 μl volumes containing 0.1 μg primers and DNA; 94 °C for 1 min, 25 cycles of 94 °C 45 s, 50 °C 45 s, 72 °C 2 min, followed by 7 min extension at 72 °C. The PCR products were purified using S300 microspin columns (Bio-Rad) and/or analysed on agarose gels.

**Western blotting.** Polyacrylamide gels were Western blotted onto PVDF membrane (Millipore) using a semi-dry electroblotting apparatus (Hoefer Scientific Instruments). Blots were blocked overnight at 4 °C in Tris-buffered saline containing...
0.01% Tween 20 (TBST) and 3% skimmed milk. After blocking, blots were incubated for 1 h with anti-flagellin serum at a dilution of 1:100 in TBST containing 1% bovine serum albumin (TBST-BSA) and washed three times in TBST, followed by incubation with peroxidase-labelled anti-rabbit IgG (Sigma) diluted 1:1000 in TBST-BSA for a further 1 h. Following washing as above, blots were developed using the DAB staining kit with nickel enhancement according to the manufacturer’s instructions (Vector Laboratories).

**Electron microscopy.** *C. jejuni* colonies from a 2 d culture grown on Columbia agar containing 7% horse blood were gently resuspended in phosphate-buffered saline. Copper grids were placed Formvar-coated side down onto a drop of 3% ammonium molybdate for a further 1 min. Grids were left to dry and viewed by transmission electron microscopy.

### RESULTS

**Characterization of maf genes and their putative products**

The *maf* genes are located in a large (37 kb) gene cluster, consisting of two adjacent regions, *Cj1317* (*neuB3*)–*maf5* and *maf6*–*Cj1347*, and belong to a family of seven paralogous genes, *maf1* (*Cj1318*), *maf2* (*Cj1333*), *maf3* (*Cj1334*), *maf4* (*Cj1335/6*), *maf5* (*Cj1337*), *maf6* (*Cj1341*) and *maf7* (*Cj1342*) (Fig. 1A). A similarity matrix between the genes and their products is presented in Table 1; two of the genes (*maf1* and *maf4*) are identical. Although the presence of homopolymeric tracts in both *maf1* and *maf4* implied that the genes might be involved in a slipped-strand mispairing mechanism of phase variation, no clues to their function could be found via database similarity searches. As such they were designated ‘function unknown genes’ in the annotation of the genome sequence (http://www.sanger.ac.uk/Projects/C._jejuni/).

Using ProSite (http://www.expasy.ch/tools/scnpsit1.html), most of the *maf* gene products are predicted to be cytoplasmic, with *maf3* having a potential transmembrane region. Proteins from this region with similarities in the database include flagellar modification proteins (*Cj1331* and *Cj1332*), sugar modification/biosynthesis proteins (*Cj1319* and *Cj1328*), sugar-, amino- or nucleotidyl-transferases (*Cj1321*, *Cj1329* and *Cj1331*) and an aminotransferase (*Cj1320*). Some genes in this region, such as *neuB3* (*Cj1317*) and *ptmA* (*Cj1332*), are also known to be involved in flagellar biosynthesis and modification (Guerry et al., 1996; Linton et al., 2000b). Since the stop codons of these genes partially overlap with N-terminal regions of some of the *maf* genes (Parkhill et al., 2000), we hypothesized that *maf* genes might also be involved in flagella biosynthesis/modification.

Closer inspection of the *maf* genes revealed orthologues in three other species: *Helicobacter pylori*, *Clostridium acetobutylicum* and *Bacillus stearothermophilus* (Fig. 1B–D). In all cases the *maf*-like genes are linked to either flagellar biosynthesis genes or/and genes involved in sugar biosynthesis and transport. For example, in *H. pylori* an orthologue HP0114 is present in the same operon and immediately downstream from the *flaB* gene encoding flagellin (Fig. 1D). Although the genes present in the operon containing paralogous gene HP0465 are hypothetical, the product of one of them (HP0466) is predicted to interact with FlgB and may be a part of the flagellar biosynthesis apparatus (Fig. 1D). Four paralogous genes similar to the *maf* family (CAC2168, CAC2196, CAC2200 and CAC2202) are present in *Cl. acetobutylicum* and are associated with the genes related

### Table 1. Percentage identity between *maf* genes and (in parentheses) their respective products

<table>
<thead>
<tr>
<th><em>maf1</em></th>
<th><em>maf2</em></th>
<th><em>maf3</em></th>
<th><em>maf4</em></th>
<th><em>maf5</em></th>
<th><em>maf6</em></th>
<th><em>maf7</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>maf1</em></td>
<td>82 (74)</td>
<td>60 (42)</td>
<td>100 (100)</td>
<td>62 (49)</td>
<td>52 (32)</td>
<td>53 (32)</td>
</tr>
<tr>
<td><em>maf2</em></td>
<td>59 (43)</td>
<td>82 (74)</td>
<td>62 (49)</td>
<td>52 (33)</td>
<td>52 (32)</td>
<td></td>
</tr>
<tr>
<td><em>maf3</em></td>
<td>60 (42)</td>
<td>61 (45)</td>
<td>53 (35)</td>
<td>53 (35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>maf4</em></td>
<td>62 (49)</td>
<td>52 (32)</td>
<td>53 (32)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>maf5</em></td>
<td>54 (34)</td>
<td>54 (32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>maf6</em></td>
<td>87 (83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
to both sugar and flagellar biosynthesis. Similarly, both flagellar and sugar/biosynthesis-related genes are associated with a single maf orthologue present in B. stearothermophilus (Fig. 1C). All Maf-like proteins share extensive local regions of high similarity, particularly in the N-terminal and central regions. One such region is presented in Fig. 1(E). Despite the different organization of the regions containing maf-like genes, the striking similarity is that they are always associated with genes involved in biosynthesis/transport of sugars and flagella.

Subclonal variation in motility of C. jejuni NCTC 11168

We found that motility of the original strain NCTC 11168 was significantly lower than that of fresh clinical isolates (data not shown). However, a small fraction of the clonal population of this strain revealed variable motility. Variants with motility ranging from almost non-motile to hypermotile, occurring with a frequency of approximately $10^{-3}$, can be isolated from the wild-type strain (Fig. 3). A hypermotile variant (11168H) was selected for mutagenesis along with the wild-type strain NCTC 11168 (see below).

In some cases such variation could be a result of reversible changes in a number of genes involved in chemotaxis, motility and flagellar biosynthesis. Due to the presence of homopolymeric G tracts, genes maf1 and maf4 could potentially be involved in such variation. However, sequencing of the regions containing homopolymeric tracts in these genes revealed that both genes were out of frame in all variants, and therefore not involved in this type of motility variation.

Gene maf5 is involved in motility

During analysis of naturally occurring non-motile mutants we also investigated the integrity of other maf genes and found that some of these mutants contained deletions, resulting from recombination between homologous regions of the adjacent paralogous genes. In one such case the region was PCR amplified and sequenced, confirming recombination between adjoining maf4 and maf5 genes (Fig. 4). These preliminary studies suggested that some genes of the maf family might be involved in motility. However, as these changes may be coincidental with independent mutations in other genes, defined mutants were required. We selected one of the genes, maf5, for further studies.

In order to establish that maf5 is involved in motility, non-polar site-directed insertional mutants were constructed. Insertion of the kanr cassette in the maf5 gene of strain 11168H (see Methods) resulted in the inability

---

**Fig. 4.** Determination of recombination site in a spontaneous non-motile variant. (A) The black areas in the genes (shown as arrows) represent regions of highest similarity. (B) Alignment between the maf4 and maf5 sequences corresponding to dark areas in (A). The hybrid gene contains a region including the 5'-terminal part of the maf4 gene (top sequence, underlined) and a region including the 3'-terminal part of the maf5 gene (lower sequence, underlined). These areas are separated by a region identical in both genes (in bold) where the recombination has occurred.
of the bacteria to swarm, confirming the hypothesis that this gene is indispensable for full motility (Figs 2B and 5).

Several independent clones resulting from the mutagenesis of the maf5 gene in strain 11168H were non-motile, confirming that the phenotype was not coincidental with possible simultaneous genetic variation of another locus. To further prove that maf5 is essential for motility we extracted chromosomal DNA from the 11168H maf5::kan\textsuperscript{r} mutant and transformed the wild-type strain NCTC 11168. All transformants in this strain were also non-motile, giving additional confirmation that gene maf5 is associated with motility.

The maf5::kan\textsuperscript{r} mutation affects flagellar biosynthesis

Electron microscopy revealed the complete lack of the flagella in the maf5::kan\textsuperscript{r} mutant, whereas the hook organelle could be seen (Fig. 6B). Analysis of cell lysates demonstrated reduced intensity of a band corresponding to flagellin in the maf5::kan\textsuperscript{r} mutant both on the

![Effect of inactivation of the maf5 gene on motility of strain 11168H. The plate (0.45% agar) was incubated overnight.](image)

![Electron micrographs of the wild-type strain of C. jejuni 11168H (A) and its maf5::kan\textsuperscript{r} derivative (B). The arrows indicate full-sized flagella in the wild-type strain and a hook-like structure in the mutant. Bars, 0.5 \( \mu \)m.](image)

![Effect of flagellin production in the 11168H maf5::kan\textsuperscript{r} mutant. (A) Coomassie-stained gel; (B) blotting with anti-flagellin antibodies. Lane 1, molecular mass standards; lane 2, 11168H; lane 3, 11168H maf5::kan\textsuperscript{r} mutant. The samples contained similar total amount of proteins as judged by the intensity of common bands in different samples in (A). The position of flagellin is indicated by an arrow.](image)
polyacrylamide gel (Fig. 7A) and on a Western blot with anti-flagellin antibodies (Fig. 7B).

**Motility in the maf5::kan' mutant can be partially restored by the maf1 gene product**

Prolonged incubation of the 11168H maf5::kan' mutant resulted in the formation of outgrowths of motile cells from the non-motile inoculum in motility plates of reduced agar concentration (Fig. 8, 3). Analysis of two such pseudorevertants obtained in two independent experiments revealed that in both cases the reading frame of the maf1 gene was restored due to the presence of an extra G (G10→G11) in the homopolymeric tract. The maf4 frame was unchanged (inactive, 10 Gs). Motility of the pseudorevertants was significantly lower than that of the parent 11168H strain.

In order to exclude the possibility of other coincidental mutations that could contribute to the partial restoration of motility, we constructed double mutants in which the restored copy (maf1) of a pseudorevertant was knocked out by insertion of a cm' cassette (see Methods). Due to the identical sequences of genes maf1 and maf4, the cm' cassette was integrated in either the maf1 or the maf4 gene, resulting in two classes of double mutants. As expected, motility of those containing inserts in the inactive maf4 gene was unaffected (Fig. 8, 4). However, the mutants containing inserts into the restored copy (maf1) were non-motile (Fig. 8, 1). These results confirm that partial restoration of motility in the 11168H maf5::kan' was the result of activation of the maf1 gene via a presumed slipped-strand mispairing mechanism.

**DISCUSSION**

In this report we describe a new class of genes involved in *C. jejuni* flagellar biosynthesis and phase variation. The genes appear to represent a unique class of flagella-related genes. A recent study has also suggested that a member of the maf gene family is involved in motility of *C. jejuni* 81116 (Bleumink-Pluym et al., 1999). The closest orthologue to the maf family of proteins is a hypothetical protein of *H. pylori*, HP0114. As in *C. jejuni*, the *H. pylori* orthologue is positioned close to the flaB gene, which encodes the major structural subunit of the *H. pylori* flagellar filament. Similarly to the *C. jejuni* maf genes, no significant similarity to any known protein with experimentally determined function could be found for the putative products encoded by orthologous genes present in other bacteria. However, comparative analysis of the genetic organization of the respective loci demonstrates that the maf-like genes in other species are also associated with the genes involved in flagellar biosynthesis and might be related in function.

Over 40 genes are involved in flagellar biogenesis of *E. coli* and *Salmonella typhimurium* (Macnab, 1996). Since Maf proteins do not have similarity in the non-redundant amino acid sequence database, their exact role in flagellar biogenesis remains an enigma. However, based on their position in the *C. jejuni* genome, and, more importantly, partial overlap with sugar biosynthesis (neuB for maf1) or flagellar modification (ptmA for maf2) genes, one can hypothesize that the maf genes may also be involved in flagellin modification. In *C. coli* VC167, motility is modulated by flagellin modification (Guerry et al., 1992). We found that, similarly to NCTC 11168, in a number of other *C. jejuni* strains the representatives of the maf family are linked to either the neuB3 gene or the ptmA gene (data not shown). Both neuB3 and ptmA are known to be involved in flagellin modification (Doig et al., 1996; Linton et al., 2000b). Flagellin modification is thought to occur in the cytoplasm (Doig et al., 1996). ProSite analysis of the putative maf gene products suggests that they are likely to be cytoplasmic or inner membrane bound. Our Western blot results demonstrate decreased intensity of the flagellin band in the maf5::kan’ mutant. It seems likely that the maf genes are involved in post-translational processing of the flagellin and/or assembly of the flagella. The lack of modification or transport may result in proteolytic degradation of the flagellin.

Reversible expression of flagella in *C. coli* resulting from a slipped-strand mispairing mechanism of phase variation in *flhA* gene has recently been demonstrated (Park et al., 2000). In contrast to *C. coli*, *C. jejuni* does not have a homopolymeric tract in this gene. The results presented here indicate the presence of alternative mechanisms of variable flagellar expression in *C. jejuni*.
One mechanism involves slipped-strand mispairing involving the contingency genes mafl and mafl, whereas the other involves recombination between homologous genes. As demonstrated in this report, homologous recombination between adjacent genes can also be involved in variation. Further evidence of homologous recombination between mafl genes is suggested by the absence of genes between mafl and mafl in C. jejuni strains 81116 and 81-76 (unpublished observation). The presence in strain NCTC 11168 of four mafl genes, organized in a tandem fashion, indicates a potential for gene deletion/duplication.

The formation of non-motile deletion derivatives in the natural C. jejuni population may confer flexibility upon bacteria in adapting to changing environmental conditions when flagellar expression and motility are undesirable. The presence of the remaining homologous gene copies might potentially allow restoration of the missing copy through gene duplication. The need for flexibility in the expression of the flagellum could be explained by its possible dual function. As an adhesin, the glycosylated flagellum might be required for initial attachment to the host cells (Nuijten et al., 1992). Once colonization is established, flagella may not be required and their formation may be switched off. When the nutrients at the infection site become limiting, a fraction of bacteria expressing flagella may acquire an advantage, as they can move towards a new colonization site. In addition, reversible expression of flagella may be beneficial in evading the host immune response (Nuijten et al., 1995). For an organism with a relatively small genome size (1-6 Mb for C. jejuni NCTC 11168), the availability of such a large number of genes dedicated to flagellar biosynthesis, modification and phase variability seems extraordinary. The presence of a complex mechanism for flagellar expression may provide C. jejuni with a selective advantage in the ecological niches it occupies.

ACKNOWLEDGEMENTS

We thank Neil Oldfield and Arnold van Viet for plasmid pAV35, and Charles Penn for flagellin monoclonal antibodies. The Wellcome Trust and the BBSRC supported this work.

REFERENCES


Received 18 June 2001; revised 1 October 2001; accepted 8 October 2001.