Homonucleotide stretches in chromosomal DNA of *Campylobacter jejuni* display high frequency polymorphism as detected by direct PCR analysis

Trudy M. Wassenaar a,*, Jaap A. Wagenaar b, Alan Rigter b, Cathrine Fearnley c, Diane G. Newell c, Birgitta Duim b,1

a Molecular Microbiology and Genomics Consultants, Tannenstrasse 7, D-55576, Zotesheim, Germany
b Division of Infectious Diseases and Food Chain Quality, Institute for Animal Science and Health (ID-Lelystad), P. O. Box 65, 8200 AB Lelystad, The Netherlands
c Veterinary Laboratories Agency (Weybridge), Addlestone, Surrey KT15 3NB, UK

Received 26 March 2002; accepted 24 April 2002

First published online 28 May 2002

Abstract

Homopolymeric nucleotide tracts have been previously identified in the genome sequence of *Campylobacter jejuni* 11168 [Parkhill et al., Nature 403 (2000) 665–668]. These tracts are believed to regulate contingency genes but as yet no phenotypic variation has been identified associated with many of these genes. To investigate homopolymeric tracts for genes for which there is no observable phenotype, a method was designed to visualise profiles of the various tract lengths directly at the genomic level by means of PCR and denatured polyacrylamide gel electrophoresis. Six of the seven contingency genes investigated displayed variation in the length of the respective homonucleotide tracts. Surprisingly, each contingency gene gave a typical peak profile that represented a conserved size distribution of polymorphic forms. For each gene studied, peak profiles were conserved between strains of *C. jejuni*. Duplicated genes, containing homonucleotide stretches, displayed locus-specific peak distributions for each gene copy. Contingency genes were polymorphic within single colonies, and the observed complex peak profiles suggested a frequency of slippage several orders of magnitude higher than reported for other organisms. No G7 (or C7) stretch was ever observed, and their absence from the complete genome suggests strong selection against their presence. In view of the predictable outcome of the process leading to these polymorphisms, it is hypothesised that the formation and/or selection of these tracts is not a random process, but is driven by as yet unknown mechanism(s). High-frequency polymorphism of these genes may be a mechanism by which *C. jejuni* survives selection bottlenecks between opportunities for growth within a host. © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Microbiological Societies.

Keywords: Contingency gene; Homopolymorphic repeat; Phase-variable gene

1. Introduction

*Campylobacter jejuni* is a major cause of human acute bacterial enteritis worldwide [1]. This pathogen also colonises the gastrointestinal tracts of a wide variety of domestic and wild animals and birds. Contaminated meat products are considered a major source of human infection. To survive a wide range of habitats, campylobacters must have evolved mechanisms to respond to various environmental stresses.

Several bacterial species use binary switches (phase variation) of gene expression mediated by changes in polymeric DNA repeat sequences [2,3]. Their single units may consist of up to seven nucleotides and are repeated in variable numbers, or, in the case of homonucleotide stretches, comprise runs of a single nucleotide. Variation in the number of repeat units, present within or upstream of a gene, results in variable expression of so-called contingency genes. Such variation is introduced by slippage during replication, and phenotypic selection can result in adaptation for best-fit survival as an alternative to transcriptional regulation. In most recognised cases, contingency genes result in antigenic or phase variation to escape
host defences or change virulence properties such as adhesion or invasiveness [3,4].

In the genome sequence of _C. jejuni_ 11168, multiple polymeric G or C (and in one case T) homonucleotide stretches have been identified [5]. The biological significance of these tracts, in terms of the functions of the genes in which they occur, is currently largely unknown. To date, the only gene containing a polymorphic homonucleotide stretch in _C. jejuni_ 11168 to which a phase-variable phenotype was assigned is Cj1139c, encoding a glycosyl transferase involved in LOS biosynthesis [6].

Since phenotypic variation of _C. jejuni_ 11168, resulting from homopolymeric tract variation, has not yet been identified (other than Cj1139c), an alternative approach to study such variation is required. A novel method, based on the acrylamide gel electrophoresis of the PCR products of these tracts was therefore developed. This technique enabled the lengths and distribution profiles of the homopolymeric tracts to be determined directly from chromosomal DNA. The method has been used to determine the extent of conservation of some of these tracts between strains and variants, and to determine if variation after selection could be demonstrated.

2. Materials and methods

2.1. Bacterial strains and growth conditions

_C. jejuni_ strain 11168S was obtained, as the genome sequence strain, directly from C.W. Penn (Birmingham University, UK). A low colonising variant (11168L) was obtained from B. Wren (London School of Tropical Hygiene, London, UK). A higher colonising variant (11168H) was derived from 11168L following passage in vivo through the chick gut. _C. jejuni_ strain 11168T was obtained from the NCTC culture collection. Other strains included were Penner sero-reference strains _C. jejuni_ O:2, O:3, O:4 and O:7 (provided by A. Lastovica, University of Cape Town, South Africa), and _C. jejuni_ 2412 (P. de Boer, ID-Lelystad, The Netherlands). _C. jejuni_ strains were grown at 37°C or 43°C on 7% sheep blood agar, in heart infusion (HI) broth, or in HI/YT biphasic medium, under microaerobic conditions (Anoxomat System) for 48 h. _Escherichia coli_ was grown overnight at 37°C on LB agar plates. Where necessary, the agar plates were supplemented with 50 μg ml⁻¹ ampicillin.

2.2. PCR analysis of polymorphic regions

PCR primers (Table 1) were designed for seven polymorphisms of strain NCTC11168 based on the genome sequence [5]. The PCR conditions were optimised for strain 11168 and comprised 2.5 mM MgCl₂, 200 μM of each dNTP, 50 mM KCl, 10 mM Tris–HCl pH 9.0, 1 pmol μl⁻¹ forward primer (carboxyfluorescein-labeled) and reverse primer, and 0.25 μl _PfuTurbo_ DNA polymerase (Stratagene, La Jolla, CA, USA). Chromosomal DNA was isolated using the PUREGENE kit (Gentra Systems, Minneapolis, MN, USA) or by CTAB/NaCl lysis [7]. The optimal DNA template concentration ranged from 25 to 250 ng per reaction. The PCR was performed with 94°C for 90 s, 25 cycles at 94°C for 45 s, 55°C for 1 min and 72°C for 45 s, ending with 72°C for 4 min.

2.3. Analysis of PCR products

PCR products were checked on agarose gels for purity of bands and then analysed on a 6.7% polyacrylamide gel using an ABI 373A automated sequencer. One microlitre of PCR products, diluted 5 or 10 times, was mixed with an internal size standard (Genescan-500, labeled with red 6-carboxyl-x-rhodamine) for analysis. Electrophoresis conditions (in 1X TBE) were 2.0 kV at 45°C for 12 h.

2.4. Cloning and sequencing of polymorphic regions

Two PCR products of each gene obtained from strain 11168T were cloned in the pCR-II-TOPO vector using the TOPO TA cloning kit (Invitrogen, Leek, The Netherlands) and inserts were sequenced using an ABI 373A automated sequencer.

2.5. In vitro invasion, heat stress and cold stress

In vitro invasion experiments with 11168S were carried out as described previously [8]. Prior to invasion, bacteria were cultured in biphasic brain heart infusion (BHI) medium. Invaded bacteria were recovered from plates and inoculated in biphasic medium for the next invasion experiment. After five consecutive invasion experiments, DNA was extracted.

For temperature stress, bacteria were grown on biphasic medium for 21 h at 43°C and divided into three lots. For a control, one lot was plated out directly; the heat shock was performed in a water bath at 52°C for 5 min and bacteria were either plated out directly, or stored at 7°C for 24 h before plating out. After growth for 21 h bacteria were harvested, chromosomal DNA isolated, and analysed as before.

2.6. Detection of polymorphic regions during colonisation of chicken

One-day-old chicks were orally inoculated with 10⁸ CFU of strain 11168L, as previously described [9]. Six days post-challenge the birds were killed and colonising bacteria were recovered. Further colonisation experiments indicated an increase in colonisation potential of the recovered isolates, resulting in 11168H. The caecal contents of chicks colonised with 11168H were harvested as before, diluted five times in phosphate-buffered saline (PBS) and...
centrifuged at 400 × g for 15 min at 4°C to remove debris. The supernatant was then centrifuged at 3000 × g for 20 min at 4°C to recover bacteria. The resultant pellet was washed in PBS and DNA was extracted using the CTAB/NaCl method. As a positive control, Campylobacter-negative chick caecal contents were spiked with C. jejuni 11168L from a confluent plate, to give a final concentration of 1.95 × 10^6 CFU ml⁻¹ of caecal contents, after which samples were treated as described above.

3. Results

From the 23 genes containing polymorphic homonucleotide tracts as reported by Parkhill et al. [5], genes were selected for further analysis on the following basis: the polymorphism was present in the 5'-end of the gene; the gene product had a significant level of amino acid identity to a protein with a predicted function; and the gene product was expected to have a selective advantage in conditions or that could be tested. On this basis, the polymeric tracts in Cj0031/0032, Cj0685c, and Cj1139c were analysed (Table 1). In addition four genes, being part of two DNA duplications in which the homonucleotide stretches were embedded in the duplcon (Fig. 1), were included, to determine if the lengths of the polymers were conserved between the duplicons (Table 1). For each locus, PCR primers were designed to flank the polymorphic tract and produce a PCR product with an expected size between 250 and 450 bp. In the case of the duplicated genes Cj1318/Cj1336 and Cj1421c/Cj1422c, only one of the primers was selective for the gene; the other primer was part of the duplcon (Fig. 1). For each primer set, one of the primers was fluorescently labeled. Fluorescently labeled amplicons were derived by PCR from chromosomal DNA of C. jejuni 11168S.

The size and purity of the PCR product were confirmed by agarose gel electrophoresis, after which they were loaded on denaturing acrylamide gels with automated fluorescent analysis. A characteristic peak profile was obtained for each amplicon (Fig. 2A), comprising several peaks for each gene studied, with the exception of Cj1139c, which gave only one peak. Since all PCR products were designed to be of different sizes, the amplicons could be combined and analysed in a single gel lane.

Duplicate PCR samples obtained from the same batch of chromosomal DNA gave identical results (Fig. 2B), as did different batches of C. jejuni 11168 obtained from various laboratories 11168S, 11168L (Fig. 2C) and 11168T (not shown). In addition variations in template DNA concentration (25–200 ng) had no effect on polymorphic profiles (not shown).

To confirm the specificity and size of the PCR products obtained, they were cloned in to an E. coli vector and two clones per amplicon were sequenced. The sizes of the homoplymeric tracts obtained from these sequences (Table 1) in some cases differed from the reported sizes [5]. When a PCR reaction was performed with the cloned vector DNA as the template, only a single band was detected (Fig. 2D), indicating that the polymorphic profiles were not an artifact of the PCR reaction but accurately represent the size variation present in chromosomal DNA from C. jejuni.

The strong conservation of the peak patterns could result from either a stable heterogeneous population in which variation in size distribution of the polymorphic stretches did not occur, or from a population that formed a conserved size distribution by mutation. To test which of these explanations was correct, the technique was adapted to obtain the PCR profiles directly from single colonies. These PCR products (obtained from approximately 10^6 CFU), demonstrated polymorphic peak profiles (Fig.

### Table 1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Annotation</th>
<th>Forward primer/Reverse primer</th>
<th>Product size, stretch of sequenced clone(s)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cj0031/</td>
<td>Type II restriction/modification gene. Cj0031 is the restriction, Cj0032 is the modification gene. Open reading frames are separated by the polymeric G stretch.</td>
<td>GGTTTTGATTTCCCTGAAAA GGAAAATTCCTCCCATATCT</td>
<td>267, G10</td>
</tr>
<tr>
<td>Cj0685c</td>
<td>Possible sugar transferase with one domain of similarity to Cj1318 family. Reported ‘invasion phenotype protein’ [10].</td>
<td>TTTTTGATTTCCCATTGGAAGGCTC TCAGAATTACCTTCTTTTAATGCGG GATTTTGGATATGGTGGGGAG</td>
<td>248, C8</td>
</tr>
<tr>
<td>Cj1139c</td>
<td>Probable galactosyl transferase.</td>
<td>AATTTGGATATGGGAGGGAATGC</td>
<td>330, G8</td>
</tr>
<tr>
<td>Cj1318</td>
<td>Unknown function. One of five genes with over 40% shared identity</td>
<td>Cj1318</td>
<td>344, G9</td>
</tr>
<tr>
<td>Cj1336</td>
<td>Identical to Cj1318</td>
<td>Cj1336</td>
<td>428, G9; 429, G10</td>
</tr>
<tr>
<td>Cj1421c</td>
<td>Possible glycosyl transferase.</td>
<td>Cj1421c</td>
<td>466, G8; 467, G9</td>
</tr>
<tr>
<td>Cj1422c</td>
<td>Homologous to Cj1421c with two domains of identity (Fig. 1).</td>
<td>Cj1422c</td>
<td>300, C9</td>
</tr>
</tbody>
</table>

*The lengths of PCR products and the size of the homonucleotide stretch of clones obtained in this study.

*Duplicate clones were sequenced for each gene product. In indicated cases, two different clones were obtained.
These results indicate that the bacterial population undergoes variation in size distribution. Moreover, in several colonies, the observed peaks had resulted from more than one mutation (resulting in more than two peaks), indicating that more than two mutations accumulated during the approximately 20 generations required to generate a colony. This implies a high frequency of switching. In some cases, this number of generations was insufficient to produce the polymeric profile typically observed in a streak population. The predominant peak in the profiles of single clones probably represents the form present in the original cell that formed the colony.

In attempts to detect a relationship between polymorphism profiles and growth conditions or selective adaptation, the effects of environmental conditions were investigated. No significant differences were detected when bacteria were cultured on agar plates containing sheep blood at temperatures of 37°C and 42°C, in liquid broth without sheep blood, or on BHI/YE biphasic medium (Fig. 3B shows representative results for gene Cj0031/0032). Neither did variation occur in the polymorphism profiles following selective pressures, such as growth after heat- or cold stresses or multiple in vitro passages via invasion through tissue culture cells (data not shown).

The effect of growth under natural conditions within the chicken gut was also tested. Although C. jejuni 11168L poorly colonises the 1-day-old chick model, a spontaneous variant, C. jejuni 11168H, with an increased colonisation potential, was identified (D.G. Newell, unpublished). The polymorphic profiles of this variant, following recovery from chicken caecum, showed no significant difference from in vitro grown organisms. To eliminate the possibility that recovery of the bacteria, by in vitro culture from the caecal contents, obscured any variation, PCR products were also obtained directly from the caecal contents. No significant changes in profiles were observed (results not shown).

The conservation of the polymeric homonucleotide tracks between strains of C. jejuni was determined. Because the PCR primers were designed for C. jejuni 11168, a PCR product was not obtained for each primer set. In those strains giving a PCR product, the results (Fig. 4) indicated that the sizes of the PCR products could vary between strains. Nevertheless the gene-specific profiles were remarkably conserved between the strains tested (representative data shown in Fig. 4). Thus, the size differences were probably due to small insertions/deletions elsewhere in the DNA sequences amplified in the PCR fragments. In all strains tested, the PCR product of gene Cj1139c always comprised one peak (results not shown). When strain 2412 was analysed before and after 50 in vitro passages on agar plates, no significant changes were detected in the genes for which a PCR product could be produced (Fig. 4).

4. Discussion

The paucity of known mechanisms of gene regulation in

![Fig. 1](https://academic.oup.com/femsle/article-abstract/212/1/77/451349/17451349)
Fig. 2. Polymorphic homonucleotide tracts in selected *C. jejuni* genes. The x-axis represents the distance of the bands run on acrylamide gel, the y-axis represents intensity of the fluorescence signal. The exact size of the bands could not be determined. A: Polymorphic profiles obtained with 25 ng chromosomal DNA of *C. jejuni* 11168S. B: Products as shown in A, from duplicate PCR analysis. C: Polymorphic profiles obtained with 250 ng chromosomal DNA of *C. jejuni* 11168L. D: Monomorphic profiles of cloned amplicon DNA, derived from *C. jejuni* 11168S and cloned in *E. coli*. The PCR reaction was performed as for A. The DNA sequence of the cloned inserts were determined and the sizes of the homonucleotide stretches are given in Table 1.
Fig. 3. Polymorphic homonucleotide tracts in single colonies and cells grown under different growth conditions as indicated. A: Single colonies of *C. jejuni* 11168S. The results of four separate colonies are shown for genes Cj0031/32, Cj0685c, and Cj1318. B: Polymeric profiles of gene Cj0031/0032 obtained from *C. jejuni* 11168S grown under different growth conditions: HI sheep blood plates at 37°C (1); HI sheep blood plates at 42°C (2); HI broth at 37°C (3); BHI/YE biphasic medium (4).

A  Single colonies

B  Different growth conditions, Cj0030/31

<table>
<thead>
<tr>
<th>HI Sheep blood agar 37°C</th>
<th>HI Sheep blood agar 42°C</th>
<th>BHI/YE Sheep blood agar 43°C</th>
<th>HI broth 37°C</th>
</tr>
</thead>
</table>

Fig. 3. Polymorphic homonucleotide tracts in single colonies and cells grown under different growth conditions as indicated. A: Single colonies of *C. jejuni* 11168S. The results of four separate colonies are shown for genes Cj0031/32, Cj0685c, and Cj1318. B: Polymeric profiles of gene Cj0031/0032 obtained from *C. jejuni* 11168S grown under different growth conditions: HI sheep blood plates at 37°C (1); HI sheep blood plates at 42°C (2); HI broth at 37°C (3); BHI/YE biphasic medium (4).
Fig. 4. Polymorphic profiles obtained with various *C. jejuni* strains. Note that no PCR product was obtained from some genes with the primers used. Strain 2412 was analysed before and after 50 passages in vitro.
the genome sequence of *C. jejuni* 11168 was a surprising finding [5], however the presence of several polymorphic homonucleotide tracts indicates that alternative gene transcription can be mediated, in certain genes of *Campylobacter*, by mutation of the number of nucleotide repeats. Unfortunately, no phenotype has yet been identified for many of these genes. Thus, an alternative approach to phenotypic selection, based on PCR, was adopted which enabled determination of the size distribution of such polymorphisms directly at the chromosomal level. The results confirmed the presence of polymorphisms in six of the seven sequences selected for investigation in this study, with more complex polymorphic profiles than expected from the variation previously reported by Parkhill et al. [5].

There was striking interstrain conservation in the profiles obtained for each homopolymeric tract investigated. This conservation was gene and locus specific, even when the tracts were present in duplicated gene fragments. These results suggest that rather than a stochastic process being involved, the mutations are driven and selected. It cannot be determined at present what the driving force is for the predominantly observed distribution profiles. Selection based exclusively on the on/off switch of reading frame is unlikely. In that case, there would be no preferential difference between a +1 or a −1 frameshift since both result in a ‘gene off’ configuration. However, the observed peak profiles suggest that there is a conservation of size distribution even for out-of-frame nucleotide stretches. Analysis of the complete genome sequence of *C. jejuni* 11168 furthermore showed that the guanidyl/cytidyl septimer (G7/C7) is absent, and it is also absent in all polymorphic profiles generated. These observations suggest that the homonucleotide tracts are formed and/or maintained in a regulated manner.

The biological role of the gene-specific polymorphic profiles is speculative. Under natural conditions, *C. jejuni* replicates exclusively in vivo. Because the bacteria appear genetically poorly equipped to survive conditions of environmental stress, each transmission to a new host is a major bottleneck. The generation of diversity, at a high frequency, is one potential mechanism to survive during such severe bottlenecks. We therefore hypothesise that the presence of homopolymeric tracts allows *C. jejuni* populations to create diversity by the multiple switching of contingency genes during growth. Even if only a few of the variants generated would survive any given hostile environment to infect a subsequent host, during growth diversity is constantly generated. Thus, this hypothesis explains the lack of observable shifts in the polymorphic profiles under the various growth conditions tested. It seems likely that *C. jejuni* has evolved to refine this strategy to drive mutations of contingency genes towards preferred directions.

The results of this study indicate a high frequency of mutation in *C. jejuni* contingency genes. In other bacteria the switching frequency is sufficiently low to isolate colonies or strains in a polymorphic form [11]. In *C. jejuni* 11168, in some of the loci analysed, more than two variants were observed in one colony (10³ CFU), suggesting a mutation frequency of at least two per 20 generations. Recently colonies of different monomorphs of the G stretch present in Cj1139c have been reported [6]. If the frequency of switching in this specific homonucleotide tract is significantly lower than that observed with the other genes investigated here, then the absence of a detectable polymorphic profile (this study), and the formation of a seemingly monomorphic population [6], would be explainable.

Our method allows the direct chromosomal analysis of bacteria, including species other than *Campylobacter*, displaying high-frequency mutation of homopolymeric nucleotide tracts despite lack of prior knowledge of the phenotypic effects of the contingency genes. The method may be of particular value in investigation of contingency genes without available selective antisera, for example in *Helicobacter pylori* [12].

Acknowledgements

Part of this research was supported by the Dutch Ministry of Agriculture, and by the UK Ministry of Agriculture, Fisheries and Food. We thank Rick Meinersmann for many stimulating discussions.

References

by homologous recombination demonstrates that flaA but not flaB is required for invasion. EMBO J. 10, 2055–2061.


