Transposon mutagenesis in a hyper-invasive clinical isolate of *Campylobacter jejuni* reveals a number of genes with potential roles in invasion.

Muhammad Afzal Javed[1], Andrew J. Grant[2], Mary. C. Bagnall[3], Duncan J. Maskell[2], Diane G. Newell[3] and Georgina Manning[1].

[1] School of Science and Technology, Nottingham Trent University, Nottingham, NG11 8NS, UK
[2]Department of Veterinary Medicine, University of Cambridge, CB3 0ES, UK.
[3]Veterinary Laboratories Agency (Weybridge), New Haw, Addlestone, Surrey, KT15 3NB, UK.

Running title: *Campylobacter jejuni* mutants with reduced invasion.

Contents Category: Microbial Pathogenesis

Corresponding author:
Dr G Manning
School of Science and Technology
Nottingham Trent University
Clifton Lane
Clifton
Nottingham. NG11 8NS. UK.
Tel: 00 44 115 848 3373
Fax: 00 44 115 848 6636
georgina.manning@ntu.ac.uk

Summary

Transposon mutagenesis has been applied to a hyper-invasive clinical isolate of *C. jejuni*, 01/51. A random transposon mutant library was screened in an *in vitro* assay of invasion and 26 mutants with a significant reduction in invasion were identified. Given that the invasion potential of *C. jejuni* is relatively poor compared to other enteric pathogens the use of a hyper-invasive strain was advantageous as it greatly facilitated the identification of mutants with reduced invasion. The location of the transposon insertion in 23 of these mutants has been determined, all but three of which are in genes also present in the genome sequenced strain NCTC11168. Eight of the mutants contain transposon insertions in one region of the genome (ca. 14kb) which when compared with the genome of NCTC11168 overlaps with one of the previously reported plasticity regions and is likely to be involved in genomic variation between strains. Further characterisation of one of the mutants within this region has identified a gene that might be involved in adhesion to host cells.

Introduction

Campylobacter jejuni is the main cause of human acute bacterial enteritis in the developed world. In England and Wales in 2008 there were 49,880 reported cases of *C. jejuni* gastroenteritis (Health Protection Agency Centre for Infections), which is thought to be a significant underestimate (Tompkins *et al.*, 1999). The disease itself is usually self-limiting and the main symptoms in developed countries, such as the UK and US, are abdominal pain and diarrhoea, often with mucous and blood in the stool (Ketley, 1997; Tompkins *et al.*, 1999; Wassenaar & Blaser, 1999). Nevertheless this organism is still the most common cause of food-borne intestinal infectious disease and is a significant public health burden.

C. jejuni is an invasive organism and there is much evidence, from both *in vivo* and *in vitro* studies, that invasion is a virulence mechanism used by campylobacters (De Melo *et al.*, 1989; Fauchere *et al.*, 1986; Garrity *et al.*, 2005; Klipstein *et al.*, 1985; Konkel & Joens, 1989). In the absence of a suitable animal model that mimics human disease (Newell, 2001) invasion has been studied using *in vitro* cell culture. There have been many reports of *in vitro* cell culture models of invasion for *Campylobacter* and these have recently been reviewed (Friis *et al.*, 2005). As in other enteropathogens, variation in virulence also exists between strains of *C. jejuni*. This has been observed for invasion and adhesion as well as toxicity (Abuoun *et al.*, 2005; Everest *et al.*, 1992; Gilbert & Slavik, 2004; Konkel & Joens, 1989; Newell *et al.*, 1985; Wassenaar, 1997). A number of hyper-invasive clinical strains of *C. jejuni* have recently been identified (Fearnley *et al.*, 2008). These isolates were found to invade the human epithelial cell lines INT-407 and Caco-2 to significantly higher levels (>25 fold) compared with a low-invasive control strain of *C. jejuni* 81116 (NCTC11828).

These strains therefore provide a unique opportunity to investigate the molecular basis of invasion as any reduction in invasion would be much easier to identify compared to using low invasive strains like 81116. It should however be noted that these hyper-invasive strains may use different invasion strategies to other non-hyper-invasive strains and so the role of any genes in invasion should be confirmed using strains with varying levels of invasion.

Molecular techniques to investigate virulence mechanisms are now available for use in *C. jejuni*. Over the last few years the development of transposon mutagenesis for *C. jejuni* has advanced significantly and several methods are now available for the random mutagenesis of this organism (Colegio *et al.*, 2001; Golden *et al.*, 2000; Golden & Acheson, 2002; Hendrixson *et al.*, 2001; Hendrixson & DiRita, 2004). One method in particular utilises an *in vitro* method of transposition, using a *mariner*-based transposon, followed by natural transformation to introduce the mutated genomic DNA back into the host strain (Grant *et al.*, 2005). This is a particularly useful technique as many strains of *C. jejuni* will not take up heterologous DNA. Previously transposon mutagenesis has been limited to a handful of those strains for which genetic manipulation is relatively easy and many of these strains have a relatively low invasion potential. However using this *in vitro* system a transposon mutant library has been constructed in a hyper-invasive clinical isolate of *C. jejuni*, strain 01/51 (Fearnley *et al.*, 2008), providing a unique opportunity to investigate the molecular basis of invasion in a strain with a significantly high level of invasion.

The aim of this study was to investigate the molecular basis of host cell invasion in *C. jejuni* by applying transposon mutagenesis to a hyper-invasive strain. We report the

screening of this mutant library in an *in vitro* assay of invasion and the identification of a number of previously uncharacterised genes that have a role in invasion. Mutants were selected for confirmatory assays and further study on the basis of their reduced level of invasion and maintenance of motility compared to the wild-type strain 01/51. We also identify a region of the *C. jejuni* genome known to be variable between strains (Pearson *et al.*, 2003) in which several genes with a role in invasion are located.

Materials and Methods

Bacterial strains and growth media. Six clinical *Campylobacter jejuni* isolates were used in this study: 01/38, 01/41, 01/51, 01/35, 01/10, 01/04, all of which were human faecal isolates apart from 01/10 and 01/04, which were isolated as a blood culture from patients with bacteraemia. These clinical strains were selected on the basis that they were all hyper-invasive (>25 fold as invasive as the control strain 81116) in an *in vitro* assay of invasiveness, apart from 01/38 which was found to be highly invasive (>10 fold as invasive as the control strain 81116) (Fearnley *et al.*, 2008). Strain 01/38 was included just in case it proved difficult to generate a random mutant library in one of the hyper-invasive strains. The non-motile, non-invasive mutant *C. jejuni* 81116 *fla*A$^-$/*fla*B$^-$ (Wassenaar *et al.*, 1991) was used as a negative control in the invasion assays. NCTC11168 and RM1221 were both used as reference strains for investigation of the TTT trinucleotide sequence (see later). All *C. jejuni* strains were routinely grown on blood agar (BA) plates (Columbia agar supplemented with 5% (v/v) sheep blood) or mCCDA agar (Oxoid) at 37°C in a microaerobic atmosphere (10% (v/v) CO_2, 5% (v/v) O_2, 85% (v/v) N_2) for 24-48 hours. When required the medium was also supplemented with 10 μg chloramphenicol per ml, or 50 μg

kanamycin per ml. The transposon mutants (n=864) were grown in each well of a 96-well plate containing 100 µl BA supplemented with 10 µg chloramphenicol per ml. *E. coli* strains were grown in Luria-Bertani (LB) broth with shaking, or on LB agar plates, aerobically at 37°C. The following antibiotics were added when required: 100 µg ampicillin per ml, 50 µg kanamycin per ml.

Generation of a random transposon mutant library. The random transposon mutant library was generated and the randomness assessed as previously described (Grant *et al.*, 2005).

In vitro adhesion and invasion assay. The semi-quantitative invasion assay used was an adaptation of the gentamicin protection assay, using the human intestinal epithelial cell line, INT-407 (Elsinghorst, 1994; Fearnley *et al.*, 2008). Note that it is now generally recognised that the INT-407 cell line was contaminated with HeLa cells in the 1970s and therefore has cellular markers consistent with this contamination (Lacroix, 2008), however it is still widely used for studying invasion and in our study it was used as a preliminary screen with confirmation of any mutants of interest using the alternative cell line Caco-2. For preliminary screening of the transposon mutants a 96-well plate assay was adapted from that described by Golden and Acheson (Golden & Acheson, 2002). Briefly, the frozen transposon mutant library was thawed on ice and 10 µl of each mutant transferred to BA (100 µl) supplemented with chloramphenicol (10 µg/ml) in each well of a 96-well plate. The plates were incubated at 37°C microaerobically for 48 hours. Meanwhile 200 µl of INT-407 cells at a density of 1 x 10^5 cells per ml was seeded to each well of a fresh 96-well plate and incubated for 48 hours at 37°C in 5 % (v/v) CO_2 to allow the cells to grow to

160 confluency. On the day of the assay, the cell culture media covering the INT-407 cell
161 monolayers was replaced with 200 µl fresh pre-warmed complete cell culture media
162 (CCCM). Fresh CCCM (100 µl) was also added to each well of the plate containing
163 bacterial growth and left for 5-10 min for the colonies to soften. The bacterial cells
164 were resuspended by pipetting and 20 µl of the suspension was added to the 200 µl of
165 media in each well covering the INT-407 monolayers. Given the number of mutants
166 tested in this way it was very difficult to normalise the starting bacterial cell
167 concentration, however any mutants that did not grow well prior to the assay were
168 noted and removed from further study. Infected plates were incubated for 3 hours at
169 37°C in 5% (v/v) CO_2. After incubation, monolayers were washed and then 200 µl
170 CCCM supplemented with 250 µg gentamicin per ml was added. Following 2 hours
171 of incubation, monolayers were washed and the cells lysed with 100 µl 1 % (v/v)
172 Triton-X100. The total number of bacteria per well was determined by viable count
173 and plating onto BA plates. *C. jejuni* 01/51 and *C. jejuni* 81116 *fla*A$^-$/*fla*B$^-$
174 (Wassenaar *et al.*, 1991) were also used as reference parent strain and negative
175 control, respectively, in the assays. The preliminary invasion screening assay was
176 repeated three times independently. Mutants that consistently showed a reduced
177 recovery compared to the parent strain 01/51 following the assay were selected for
178 further characterisation including motility and confirmation of the reduced invasion
179 phenotype using a standardised assay as described below.
180 For a more quantitative analysis and confirmation of the invasion of selected mutants
181 standard gentamicin protection assays (confirmatory assay) using both INT-407 and
182 Caco-2 cells in 24-well plates were carried out, in which each mutant was assayed in
183 triplicate. The mutants were grown on BA plates for 48 hours and were added to the
184 monolayers at a multiplicity of infection (MOI) of 100. To determine the number of

185 associated bacteria, the INT-407 cells were lysed after the initial 3 hour incubation
186 period and the bacterial cells enumerated by plate count. This gives the total number
187 of bacteria that are associated and internalised. The number of associated bacteria was
188 then calculated by subtracting the number of internalised bacteria from the total
189 number counted. Invasion efficiency was expressed as the percent of the inoculum
190 that survived the gentamicin treatment.
191
192 *Motility assay*. Bacterial motility was tested as described previously (Fearnley *et al.*,
193 2008). The parent strain 01/51 was assumed to have 100% motility (with an average
194 diameter of the zone of growth of 5.5 cm) and only mutants with $\geq$ 75% motility as
195 compared to the parent strain were selected for further study.
196
197 *Location of transposon insertions*. To determine the location of the transposon
198 insertion point a plasmid rescue technique was carried out (Grant *et al.*, 2005).
199 Briefly the genomic DNA from each of the mutants was isolated and digested to
200 completion with *Bgl*II or *Ssp*I. The genomic DNA fragments were then self-ligated
201 and transformed by electroporation into *E. coli* S17-λ*pir* (Simon *et al.*, 1983).
202 Plasmids were prepared using the QIAprep spin miniprep kit (Qiagen, Crawley, UK)
203 from a 10 ml overnight culture. The protocol was carried out according to the
204 manufacturer's recommendations for large plasmids, which meant that the DNA was
205 eluted from the spin-column using water heated to 70°C to maximise recovery.
206 Recovered plasmid was then concentrated by ethanol precipitation and resuspended in
207 5 µl of dH_2O. The insertion site was identified by DNA sequencing using a
208 transposon specific primer (5'-CCCGGGAATCATTTGAAG- 3').
209

SNP Detection assay. A Single Nucleotide Polymorphism (SNP) detection assay to detect the observed polymorphisms in one of the mutated genes (Cj0490) was developed based on one previously reported (AbuOun *et al.*, 2005). Primers were designed to the region flanking the TTT trinucleotide sequence (Figure 3A) (FWD: 5'-AAAGAGCGATTGAAGC-3'; REV: 5'-CATTAAAACTTCGGTTAAGA-3') and the probe was generated from the 01/51 sequence *i.e.* containing the TTT trinucleotide sequence (underlined), but on the complimentary strand (5'-Cy5-GCATTTTTGCGTATTAAACTAGCT-biotin-3'). The amplification was performed as follows: an initial denaturation step for 10 minutes at 95°C; which was followed by 41 cycles, with 10 s at 94°C, 15 s at 55°C and 10 s at 74°C. Melting curve analysis was performed immediately after amplification by heating the product to 94°C (20°C/s), cooling to 45°C for 15 s, and then heating to 85°C (0.1°C/s). DNA from 01/51 and NCTC11168, and a no template DNA control were included in each run as controls. NCTC11168 was found to have a melting temperature of 61°C indicating the presence of the TT dinucleotide sequence and 01/51 had a higher melting temperature of 64 °C which indicated that the probe and sequence were identical, *i.e.* that 01/51 possesses the TTT trinucleotide sequence.

Mutagenesis of Cj0497. For mutagenesis of Cj0497 the gene was amplified from *C. jejuni* 01/51 and cloned into pBluescript (Stratagene) via pCR2.1TOPO™ (Invitrogen), a T-tailed cloning vector. The following primers were used for the amplification of Cj0497: Cj0497F: 5'- TTGATTTAAGGGTTATGAAGGC-3' and Cj0497R: 5'- AGCCTTAATCACATCTTTTGG-3'). There was a unique *Bgl*II site in Cj0497 into which the *C. coli* kanamycin cassette, from pJMK30 (van Vliet *et al.*, 1998) (kindly provided by Professor Julian Ketley, University of Leicester, UK), was

235 cloned. Constructs with the kanamycin cassette in the same orientation as the
236 disrupted gene (as determined by PCR analysis) to minimise the risk of polar effects
237 since the kanamycin cassette does not bear transcription termination sequences were
238 introduced into *C. jejuni* 01/51 by electrotransformation. Mutants were confirmed by
239 PCR and Southern blot analysis. The approach of inserting the kanamycin cassette,
240 lacking transcription termination sequences, in the same orientation as the gene to be
241 inactivated has been reported previously (Elvers *et al.*, 2004; Ge *et al.*, 2005; Linton
242 *et al.*, 2002) and has been shown to be non-polar on downstream genes (Hickey et al.,
243 2000).
244
245 *Phenotypic assays to characterise Cj0497 mutant.* In order to further characterise the
246 mutant in Cj0497 a number of phenotypic assays were carried out including
247 microaerobic growth, aeration survival and autoagglutination. For all of these
248 additional assays the bacteria were grown at 37°C for 48 hours on blood agar or
249 mCCDA with 50 μg of kanamycin added per ml for the mutant. Bacterial growth was
250 harvested from these plates and resuspended in 2.5 ml sterile PBS prior to use in the
251 assays. All assays were carried out in triplicate.
252
253 For the growth curve an appropriate volume of this suspension was added to Mueller-
254 Hinton (MH) broth in a 96-well plate to make a final volume of 100 μl in a well and
255 an absorbance of ~ 0.1 (OD 570nm). The plate was then incubated microaerobically
256 with gentle shaking at 37°C. Regular absorbance readings were taken up to 30 hours.
257
258 For the aeration stress assay 2 ml of the resuspended culture was used to inoculate 200
259 ml of pre-warmed Mueller Hinton broth (MHB). The culture was incubated

microaerobically with gentle shaking at 37°C for 24 hours. Following overnight growth (approximately 10^8 –10^9 CFUml^{-1}), the culture was exposed to atmospheric oxygen at 37°C incubation continued for a further 6 hours. Samples were removed at hourly time points and viable counts performed.

The autoagglutination assay was performed following the protocol described elsewhere (Golden & Acheson, 2002). The absorbance of harvested cultures was adjusted to ~1.0 at 600 nm in PBS and the actual absorbance was measured again. The bacterial suspension (2.0 ml) was transferred into sterile bijoux tubes and incubated undisturbed at 37°C microaerobically for 24 hours to allow the bacterial cells to autoagglutinate and settle to the bottom of the tube. One millilitre of the upper aqueous phase was then aspirated and the absorbance was measured at 600 nm. The level of autoagglutination was calculated by subtracting the absorbance of the aspirate collected after 24 hours from the initial absorbance measured at the start of incubation.

Statistical analysis. For this a paired Student's T test was performed using MS-Excel software. A probability value $P < 0.01$ indicated statistical significance.

Results

Generation of a random transposon mutant library in a hyper-invasive strain. Six clinical strains (01/38, 01/41, 01/51, 01/35, 01/10 and 01/04) were tested for their ability to undergo random transposition using the *mariner*-based *in vitro* transposon system, which had been optimised for transposition of *C. jejuni* NCTC11168, 81-176 and M1. The generation of a random transposon mutant library was only found to be

286 successful in one of the clinical hyper-invasive strains, 01/51. This was a faecal
287 isolate from a patient with *C. jejuni* gastroenteritis. No other information regarding
288 the disease severity or symptoms is available. In order to hit every gene in the genome
289 of 01/51 it would be necessary to screen around 4-5000 mutants however it was
290 decided to screen an initial batch of up to 1000 mutants in the first instance. The
291 initial batch of mutants was picked and stored for future use.
292
293 The randomness of the transposon insertions was determined by randomly selecting
294 ten mutants and carrying out Southern blot hybridisations, using a fragment of the
295 transposon as a probe (data not shown), as well as sequence analysis to determine the
296 location of the transposon insertion point. From this initial screen there did not appear
297 to be any bias towards a particular region of the genome or particular strand of the
298 chromosome for the insertion point (data not shown). This level of randomness was in
299 keeping with other strains such as the genome sequenced strain, NCTC11168 and
300 strain M1. In addition the risk of siblings within the library was minimised by
301 keeping the recovery time of the mutants, following transformation, to a minimum. It
302 was anticipated, following analysis of 01/51 and the other strains to which this
303 technique has been applied, that the maximum risk of siblings within the library is 2
304 %.
305
306 *Screening of the 01/51 transposon mutant library for defects in invasion and motility.*
307 The mutant library (n=864) was screened in an invasion assay using INT407 cells and
308 174 mutants that showed minimal invasion in the preliminary screening assays were
309 selected to assess motility. The mutants showed a wide range of motility phenotypes
310 with nine mutants being non-motile and the remainder retaining motility ranging from

311 20% - 174% of the parent strain 01/51. As motility is important for invasion, an
312 arbitrary cut-off for mutant selection was chosen and 40 mutants with ≥ 75% motility
313 as compared to wild-type strain, 01/51, were selected. The reduced invasion capacity
314 of these selected mutants was confirmed using INT407 cells in a confirmatory
315 standard invasion assay in which each mutant was tested in triplicate. A selection of
316 mutants that showed a ≥ 60% reduction in invasion in INT-407 cells were also
317 checked for their ability to invade Caco-2 cells. All the tested mutants showed a
318 similar reduction in invasion of Caco-2 cells compared to INT-407 cells (Table 1).
319 The localisation of the transposon insertions was investigated in 26 mutants.
320
321 *Location of the transposon insertions*. The transposon insertion point was identified
322 in 23 out of the 26 mutants by the plasmid rescue technique and sequencing using a
323 transposon derived primer (Table 1). The transposon insertion site could not be
324 determined in the remaining three mutants even after several attempts with alternative
325 restriction enzymes. Southern blot analysis was also performed on a random selection
326 of mutants to ensure that a single transposon had inserted into each mutant. Analysis
327 showed one band in every mutant lane confirming that the transposon had inserted at
328 one site in the genome of each mutant (data not shown).
329
330 All but three of the transposon insertions identified were found to be in genes which
331 were also present in NCTC11168 (Table 1). Of the remaining three insertions one was
332 in the gene, *dtp*T (di-/tripeptide transporter), which is present in other *C. jejuni* strains
333 for which the genome sequences are now available, including *C. jejuni* strain RM1221
334 (CJE0757) and *C. jejuni* strain 81116 (C8J-0613). The second was annotated as a
335 capsule polysaccharide biosynthesis protein in *C. jejuni* subspecies *doylei* strain

269.97. In the third mutant (10D12) the sequence obtained showed homology to a putative *rlo*E gene (CJJ26094_0063) in *C. jejuni* strain 260.94 whose sequence was incomplete at the time of searching (July 2009). The function of this gene is unknown and in other *C. jejuni* strains, namely 84-25 and Tgh133, it is annotated as a putative hypothetical protein.

In 8 out of the 23 mutants for which the transposon insertion point was determined, the transposon was located within a ca. 14 kb region of the genome between Cj0483 and Cj0499, according to the genome sequence of NCTC11168 (Figure 1). The gene order in 01/51 was found to be similar to that in NCTC11168 as determined by PCR analysis across this region (Figure 1). Two of the genes located within this 14 kb region of the genome, Cj0490 and Cj0497 were identified as the site of transposon insertion in 3 and 2 separate mutants respectively. The position of the transposon in these two genes seems to have little effect on invasion, but a greater effect on motility (Table 1). For example the three mutants in Cj0490 were between 16-18% as invasive and between 78-174% as motile as the parent strain.

Genomic comparison of the 14kb region amongst other strains of C. jejuni revealed some conservation in gene order. Comparison of the 14 kb region (Cj0483 - Cj0499) of NCTC11168 with strain RM1221, another available *C. jejuni* genome sequence (TIGR), indicated that this region is similar in the two strains. However in RM1221 between Cj0493 and Cj0494 there is an insertion of a *c.*40kb region of DNA (Figure 1), which appears to have very few similarities to known sequences. This 40kb region has a GC content of 27.42% compared with an average of 30.31% for the RM1221 genome (Fouts *et al.*, 2005). Interestingly on further analysis of the NCTC11168

genome there is an additional small ORF located between Cj0494 and *fus*A (Cj0493) annotated as tRNA-Arg, however it is not known whether this tRNA gene also exists in 01/51 at this position.

Cj0490 is annotated as two ORFs in NCTC11168, but is a single ORF in 01/51. Cj0490 was one of the genes for which there were 3 individual transposon mutants. In NCTC11168 Cj0490 encodes the aldehyde dehydrogenase C terminus whereas Cj0489 encodes the N- terminus of this same protein. Comparison of this sequence with that in other bacteria, such as *E. coli*, *Shigella flexneri* and *Neisseria meningitidis*, indicated that it is unusual to find the two parts of this protein encoded by separate genes. The sequence of this aldehyde dehydrogenase was therefore determined in the hyper-invasive strain, 01/51, and interestingly this gene was also found to be a single ORF with both the N-terminus and C-terminus together, as in other bacteria. At the DNA level in NCTC11168 a single T deletion appears to have resulted in the generation of a stop codon (Figure 2A), which is followed by an intergenic region of 36 nucleotides before the start of the next ORF. At the amino acid level (Figure 2B) the stop codon causes a frameshift and the appearance of the separate N and C termini.

The presence of the TTT trinucleotide sequence was investigated in seven *C. jejuni* strains by sequencing across the region between Cj0489 and Cj0490 and in a further 20 strains with known invasion potential using a SNP detection assay. There was no correlation between the presence of the TTT trinucleotide and the invasion potential therefore further analysis of this gene was not carried out.

386 *Targeted mutagenesis of Cj0497 in 01/51 reveals a potential adhesin.* As the mutants
387 with transposon insertions in Cj0497 (1B5 and 1D1) both showed reduced invasion,
388 yet retained motility, compared to the wild-type 01/51, this gene was selected for
389 further study. This gene was independently inactivated in *C. jejuni* strain 01/51 by
390 insertion of a kanamycin resistance cassette to confirm the observations with the
391 transposon mutant. The resulting targeted mutant was tested in assays of association
392 and invasion and was significantly reduced in overall association with INT-407 (4%
393 of inoculum associated for mutant vs. 9.5% for 01/51; $p< 0.01$) and Caco-2 cells
394 (0.7% of inoculum associated for mutant vs. 3% for 01/51, $p< 0.01$) and invasion into
395 INT-407 (0.11% of inoculum internalised vs. 1% for 01/51; $p< 0.01$) and Caco-2 cells
396 (0.004% of inoculum internalised for mutant vs. 0.39% for 01/51; $p< 0.01$) compared
397 to the wild-type strain 01/51 which suggests that the reduction in invasion in the
398 Cj0497 mutant may be due to a reduction in adhesion. The mutant was also found to
399 grow as well under microaerobic conditions as the wild-type strain and to survive as
400 well under atmospheric oxygen conditions. In addition the targeted mutant was found
401 to be 1.3 times more motile ($p<0.001$) and 1.3 times better able to autoagglutinate
402 ($p<0.001$) than the wild-type strain.
403
404 **Discussion.**
405 *C. jejuni* is the most common cause of bacterial food-borne diarrhoea worldwide, but
406 its pathogenic mechanisms are not clear. However, previous studies indicate that
407 invasion and motility are important for *Campylobacter* pathogenesis and not all
408 strains have the same virulence potential. *C. jejuni* strain dependent variability in
409 invasion into eukaryotic cells has been reported by many researchers (Everest *et al.*,
410 1992; Konkel & Joens, 1989; Malik-Kale *et al.*, 2007; Newell *et al.*, 1985). The

411 invasiveness of *C. jejuni* strains is generally low making investigation of this property
412 difficult, however, we recently reported the identification of hyper-invasive strains of
413 *C. jejuni* (Fearnley *et al.*, 2008). A transposon mutant library was constructed in one
414 of these hyper-invasive strains (01/51) in order to investigate the molecular basis of
415 invasion and 26 mutants with reduced invasion were selected for further study and
416 identification of the genes inactivated.
417
418 Only one of the genes identified was found to be previously associated with invasion.
419 Mutant 3A10 has a transposon insertion in *cip*A (Cj0685) a putative *Campylobacter*
420 invasion protein). A previous study reported that a mutation in the *cip*A gene of
421 *C. jejuni* TGH9011 resulted in the reduced invasion of HEp-2 cells but there was no
422 change in invasion of INT-407 and Caco-2 cells (Lynett, 1999). This gene was
423 identified in our study as being involved in invasion in both INT-407 and Caco-2
424 cells. In previous annotations of the *Campylobacter* genome this gene was annotated
425 as a sugar transferase with similarity to two genes involved in capsule biosynthesis
426 (Cj1421c and Cj1422c). It is possible therefore that a mutation in this gene has altered
427 some surface property of the organism resulting in reduced invasion and this gene is
428 currently being investigated further.
429
430 The fact that 8 of the 23 mutants in this study had transposon insertions within one
431 region of the genome is interesting. There was no apparent bias in the insertion point
432 when an initial ten mutants were screened to check whether the transposition was
433 random and none of those ten mutants possessed a transposon insertion in this 14 kb
434 region. It is possible that this 14 kb region has a role in the interaction of *C. jejuni*
435 with its host; however this region does appear to be associated with genomic

variability. The sequence between Cj0483 and Cj0499, overlaps with one of the seven hypervariable plasticity regions, PR, previously described in the genome of *C. jejuni* which are likely to reflect the high level of phenotypic variation seen amongst the *C. jejuni* population and account for the ability of this organism to exist in a wide range of ecological niches (Pearson *et al.*, 2003). Moreover further analysis of this region revealed that in *C. jejuni* RM1221 there appears to be a large DNA insertion at this point (CJIE2) which has integrated into the 3'- end of an arginyl tRNA gene (Fouts *et al.*, 2005) and may represent an intact prophage or a genomic island which highlights the genetic diversity within this region.

Many of the genes into which the transposons have inserted are genes associated with metabolism and survival *e.g. put*A (Cj1503c), a putative proline dehydrogenase/delta-1-pyrroline-5-carboxylate dehydrogenase, which catalyses the oxidation of proline into glutamate; *ald* (Cj0490), aldehyde dehydrogenase, involved in energy acquisition and amino acid transport; *uxa*A (Cj0483), altronate hydrolase, involved in carbohydrate metabolism; Cj0519, involved in molybdopterin biosynthesis; *sur*E, a putative stationary-phase survival protein; *dtp*T encoding a di-tripeptide transporter protein.

Interestingly the aldehyde dehydrogenase gene (Cj0490) is polymorphic in those strains in which it is present and it is only present in five of the eight *C. jejuni* genome sequences currently available suggesting that this gene is non-essential for metabolism and may even be a pseudogene (http://xbase.bham.ac.uk/campydb/).

460 The lack of well-defined virulence mechanisms in *C. jejuni* and the involvement of
461 metabolism associated genes with virulence phenotypes may reflect the possibility
462 that *C. jejuni* is an opportunistic pathogen and does not possess specific virulence
463 factors as in other bacteria, with disease resulting as a consequence of the need of the
464 organism to grow and survive within the human host. The fact that many of these
465 genes are part of different metabolic pathways highlights that invasion is a multi-
466 faceted phenotype, involving many different pathways.
467
468 The transposon in two mutants (1B5 and 1D1) was located within Cj0497, which is
469 annotated as a putative lipoprotein. Further analysis of this gene sequence indicates
470 that it contains a signal peptide, suggesting that it might be located in the periplasm
471 and is likely to be membrane-bound. It also contains a tetratricopeptide repeat (TPR)
472 region found in multi-protein complexes and transmembrane segments. TPR motifs
473 are thought to mediate inter- and intra-molecular protein interactions and occur widely
474 in nature (Ohara *et al.*, 1999). In bacteria TPR repeat containing proteins are thought
475 to have a role in gene regulation, flagellar motor function and virulence (Newton *et*
476 *al.*, 2007). Moreover in *Legionella pneumophila*, two TPR containing genes *lpn*E and
477 *enh*C have been shown to be associated with entry into human tissue culture cell lines
478 (Cirillo *et al.*, 2000; Newton *et al.*, 2006).
479
480 Targeted insertional inactivation of Cj0497 resulted in a reduction in bacterial
481 association and invasion which suggests that this gene may have a role in host cell
482 adhesion, which may lead to invasion. Further phenotypic studies indicated that the
483 mutant was unaffected in its ability to grow under microaerobic conditions or to
484 survive in air compared to the wild-type and so the reduction in invasion could not be

attributed to either of these factors. The mutant was however more motile and better able to autoagglutinate than the wild-type, suggesting that a loss of this gene does not result in reduced motility or autoagglutination. A similar finding was reported previously whereby a mutation in a known adhesin gene, *peb*1A, failed to reduce the level of autoagglutination compared to the wild-type strain (Misawa & Blaser, 2000). This could mean that the mechanisms used by the organism to adhere to cells are different to those used to adhere to each other and in the case of Cj0497, if it were to encode an adhesin, by removing it, autoagglutination was increased. The reason for this is unclear but may indicate enhanced exposure of other surface molecules that are involved in the autoagglutination process.

A number of adhesins have been identified in *C. jejuni* including PEB1 (Kervella *et al.*, 1993; Pei & Blaser, 1993), CadF (Konkel *et al.*, 1997) and JlpA (Jin *et al.*, 2001). The role of Cj0497 as a lipoprotein and adhesin in *C. jejuni* warrants further investigation as it does appear to be present in all C. *jejuni* strains tested to date and is not species-specific like *jlp*A, as it is present in the genome sequence of *C. coli*, *C. upsaliensis* and *C. lari* as well as other members of the epsilon subdivision of the proteobacteria (http://xbase.bham.ac.uk/campydb/).

We have identified a number of previously uncharacterised genes with a potential role in host cell invasion. The advantage of this study was that a hyper-invasive strain of *C. jejuni* was used for transposon mutagenesis which facilitated the detection of mutants with reduced invasion. Many of the genes are annotated as metabolism-associated rather than "virulence" *per se* and many appear to be conserved within the species. In addition this study has also highlighted a region of the genome that may be

involved in genomic variability between strains. Further investigation of some of the genes described through targeted mutation as well as complementation studies is currently underway which will help to provide greater insight into the molecular basis of the observed variation in virulence in *C. jejuni*.

Acknowledgements.

Thanks go to Dr David Ussery at the Danish Technical University for the genome comparisons and Dr Trudy Wassenaar for help with some of the bioinformatics. Thanks also go to Dr Manal AbuOun, Dr Anne Ridley and Dr Catherine Fearnley for assistance with the SNP detection assay, and Dr Alan McNally for critical review of the manuscript. This work was partly funded by the Department for Environment, Food and Rural Affairs, UK and partly by Nottingham Trent University.

548
549
550
551
552

553 **References.**
554
555 **Abuoun, M., Manning, G., Cawthraw, S. A., Ridley, A., Ahmed, I. H.,**
556 **Wassenaar, T. M. & Newell, D. G. (2005).** Cytolethal distending toxin (CDT)-
557 negative *Campylobacter jejuni* strains and anti-CDT neutralizing antibodies are
558 induced during human infection but not during colonization in chickens. *Infect Immun*
559 **73**, 3053-3062.
560
561 **Cirillo, S. L., Lum, J. & Cirillo, J. D. (2000).** Identification of novel loci involved in
562 entry by *Legionella pneumophila*. *Microbiology* **146 (Pt 6)**, 1345-1359.
563
564 **Colegio, O. R., Griffin, T. J. t., Grindley, N. D. & Galan, J. E. (2001).** *In vitro*
565 transposition system for efficient generation of random mutants of *Campylobacter*
566 *jejuni*. *J Bacteriol* **183**, 2384-2388.
567
568 **De Melo, M. A., Gabbiani, G. & Pechere, J. C. (1989).** Cellular events and
569 intracellular survival of *Campylobacter jejuni* during infection of HEp-2 cells. *Infect*
570 *Immun* **57**, 2214-2222.
571
572 **Elsinghorst, E. A. (1994).** Measurement of invasion by gentamicin resistance.
573 *Methods Enzymol* **236**, 405-420.
574
575 **Elvers, K. T., Wu, G., Gilberthorpe, N. J., Poole, R. K. & Park, S. F. (2004).** Role
576 of an inducible single-domain hemoglobin in mediating resistance to nitric oxide and
577 nitrosative stress in *Campylobacter jejuni* and *Campylobacter coli*. *J Bacteriol* **186**,
578 5332-5341.
579
580 **Everest, P. H., Goossens, H., Butzler, J. P., Lloyd, D., Knutton, S., Ketley, J. M.**
581 **& Williams, P. H. (1992).** Differentiated Caco-2 cells as a model for enteric invasion
582 by *Campylobacter jejuni* and *C. coli*. *J Med Microbiol* **37**, 319-325.
583
584 **Fauchere, J. L., Rosenau, A., Veron, M., Moyen, E. N., Richard, S. & Pfister, A.**
585 **(1986).** Association with HeLa cells of *Campylobacter jejuni* and *Campylobacter coli*
586 isolated from human feces. *Infect Immun* **54**, 283-287.
587
588 **Fearnley, C., Bagnall, M., Manning, G., Javed, M. A., Wassenaar, T. M. &**
589 **Newell, D. G. (2008).** Identification of hyperinvasive *Campylobacter jejuni* strains
590 isolated from poultry and human clinical sources. *J Med Microbiol* **57**, 570-580.
591
592 **Fouts, D., Mongodin, E., Mandrell, R. & other authors (2005).** Major structural
593 differences and novel potential virulence mechanisms from the genomes of multiple
594 *Campylobacter* species. *PLoS Biol* **3**, e15.
595
596 **Friis, L. M., Pin, C., Pearson, B. M. & Wells, J. M. (2005).** *In vitro* cell culture
597 methods for investigating *Campylobacter* invasion mechanisms. *J Microbiol Methods*
598 **61**, 145-160.
599
600 **Garrity, G. M., Bell, J. A. & Lilburn, T. (2005).** Class V. *Epsilonproteobacteria*. In
601 *Bergey's manual of systematic bacteriology*, pp. 1145. Edited by D. J. Brenner, N. R.
602 Krieg, J. T. Staley & G. M. Garrity. New York: Springer.

Ge, B., McDermott, P. F., White, D. G. & Meng, J. (2005). Role of efflux pumps and topoisomerase mutations in fluoroquinolone resistance in *Campylobacter jejuni* and *Campylobacter coli*. *Antimicrob Agents Chemother* **49**, 3347-3354.

Gilbert, C. & Slavik, M. (2004). Determination of toxicity of *Campylobacter jejuni* isolated from humans and from poultry carcasses acquired at various stages of production. *J Appl Microbiol* **97**, 347-353.

Golden, N. J., Camilli, A. & Acheson, D. W. (2000). Random transposon mutagenesis of *Campylobacter jejuni*. *Infect Immun* **68**, 5450-5453.

Golden, N. J. & Acheson, D. W. (2002). Identification of motility and autoagglutination *Campylobacter jejuni* mutants by random transposon mutagenesis. *Infect Immun* **70**, 1761-1771.

Grant, A. J., Coward, C., Jones, M. A., Woodall, C. A., Barrow, P. A. & Maskell, D. J. (2005). Signature-tagged transposon mutagenesis studies demonstrate the dynamic nature of cecal colonization of 2-week-old chickens by *Campylobacter jejuni*. *Appl Environ Microbiol* **71**, 8031-8041.

Hendrixson, D. R., Akerley, B. J. & DiRita, V. J. (2001). Transposon mutagenesis of *Campylobacter jejuni* identifies a bipartite energy taxis system required for motility. *Mol Microbiol* **40**, 214-224.

Hendrixson, D. R. & DiRita, V. J. (2004). Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Mol Microbiol* **52**, 471-484.

Jin, S., Joe, A., Lynett, J., Hani, E. K., Sherman, P. & Chan, V. L. (2001). JlpA, a novel surface-exposed lipoprotein specific to *Campylobacter jejuni*, mediates adherence to host epithelial cells. *Mol Microbiol* **39**, 1225-1236.

Kervella, M., Pages, J. M., Pei, Z., Grollier, G., Blaser, M. J. & Fauchere, J. L. (1993). Isolation and characterization of two *Campylobacter* glycine-extracted proteins that bind to HeLa cell membranes. *Infect Immun* **61**, 3440-3448.

Ketley, J. M. (1997). Pathogenesis of enteric infection by *Campylobacter*. *Microbiology* **143**, 5 - 12.

Klipstein, F. A., Engert, R. F., Short, H. & Schenk, E. A. (1985). Pathogenic properties of *Campylobacter jejuni:* assay and correlation with clinical manifestations. *Infect Immun* **50**, 43-49.

Konkel, M. E. & Joens, L. A. (1989). Adhesion to and invasion of HEp-2 cells by *Campylobacter* spp. *Infect Immun* **57**, 2984-2990.

Konkel, M. E., Garvis, S. G., Tipton, S. L., Anderson, D. E. & Cieplak, W. (1997). Identification and molecular cloning of a gene encoding a fibronectin-binding protein (CadF) from *Campylobacter jejuni*. *Mol Microbiol* **24**, 953-963.

Lacroix, M. (2008). Persistent use of "false" cell lines. *Int J Cancer* **122**, 1-4.

Linton, D., Allan, E., Karlyshev, A. V., Cronshaw, A. D. & Wren, B. W. (2002). Identification of *N*-acetylgalactosamine-containing glycoproteins PEB3 and CgpA in *Campylobacter jejuni*. *Mol Microbiol* **43**, 497-508.

Lynett, J. (1999). Defining the role of *cip*A in the pathogenesis of *Campylobacter jejuni* infection. In *MSc thesis submitted to department of laboratory medicine and pathobiology*, pp. 64-101. Toronto: University of Toronto.

Malik-Kale, P., Raphael, B. H., Parker, C. T., Joens, L. A., Klena, J. D., Quinones, B., Keech, A. M. & Konkel, M. E. (2007). Characterization of genetically matched isolates of *Campylobacter jejuni* reveals that mutations in genes involved in flagellar biosynthesis alter the organism's virulence potential. *Appl Environ Microbiol* **73**, 3123-3136.

Misawa, N. & Blaser, M. J. (2000). Detection and characterization of autoagglutination activity by *Campylobacter jejuni*. *Infect Immun* **68**, 6168-6175.

Newell, D. G., McBride, H., Saunders, F., Dehele, Y. & Pearson, A. D. (1985). The virulence of clinical and environmental isolates of *Campylobacter jejuni*. *J Hyg (Lond)* **94**, 45-54.

Newell, D. G. (2001). Animal models of *Campylobacter jejuni* colonization and disease and the lessons to be learned from similar *Helicobacter pylori* models. *Symposium series (Sfam)*, 57S-67S.

Newton, H. J., Sansom, F. M., Bennett-Wood, V. & Hartland, E. L. (2006). Identification of *Legionella pneumophila*-specific genes by genomic subtractive hybridization with *Legionella micdadei* and identification of *lpnE*, a gene required for efficient host cell entry. *Infect Immun* **74**, 1683-1691.

Newton, H. J., Sansom, F. M., Dao, J., McAlister, A. D., Sloan, J., Cianciotto, N. P. & Hartland, E. L. (2007). Sel1 repeat protein LpnE is a *Legionella pneumophila* virulence determinant that influences vacuolar trafficking. *Infect Immun* **75**, 5575-5585.

Ohara, M., Wu, H. C., Sankaran, K. & Rick, P. D. (1999). Identification and characterization of a new lipoprotein, *Nlp*I, in *Escherichia coli* K-12. *J Bacteriol* **181**, 4318-4325.

Parkhill, J., Wren, B. W., Mungall, K. & other authors (2000). The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665-668.

Pearson, B. M., Pin, C., Wright, J., I'Anson, K., Humphrey, T. & Wells, J. M. (2003). Comparative genome analysis of *Campylobacter jejuni* using whole genome DNA microarrays. *FEBS letters* **554**, 224-230.

Pei, Z. & Blaser, M. J. (1993). PEB1, the major cell-binding factor of *Campylobacter jejuni*, is a homolog of the binding component in gram-negative nutrient transport systems. *J Biol Chem* **268**, 18717-18725.

Simon, R., Periefer, U. & Puhler, A. (1983). A broad host range mobilization system for *in vitro* genetic engineering: transposon mutagenesis in gram negative bacteria. *Biotechnology* **1**, 784-791.

Tompkins, D. S., Hudson, M. J., Smith, H. R. & other authors (1999). A study of infectious intestinal disease in England: microbiological findings in cases and controls. *Commun Dis Pub Health / PHLS* **2**, 108-113.

van Vliet, A. H., Wooldridge, K. G. & Ketley, J. M. (1998). Iron-responsive gene regulation in a *Campylobacter jejuni fur* mutant. *J Bacteriol* **180**, 5291-5298.

Wassenaar, T. M., Bleumink-Pluym, N. M. & van der Zeijst, B. A. (1991). Inactivation of *Campylobacter jejuni* flagellin genes by homologous recombination demonstrates that *flaA* but not *flaB* is required for invasion. *EMBO J* **10**, 2055-2061.

Wassenaar, T. M. (1997). Toxin production by *Campylobacter* spp. *Clin Microbiol Rev* **10**, 466-476.

Wassenaar, T. M. & Blaser, M. J. (1999). Pathophysiology of *Campylobacter jejuni* infections of humans. *Microbes Infect / Institut Pasteur* **1**, 1023-1033.

731 **Figure Legends**
732
733 **Figure 1: Location of the transposon insertions in the 14 kb region between**
734 **Cj0483 and Cj0499.** ORFs with transposon insertions are shown as black arrows.
735 The shaded inverted triangle represents the position in *C. jejuni* strain RM1221 where
736 there is an insertion of *c.* 40 kb. The direction of the arrow denotes the predicted
737 direction of transcription. The dotted arrow at the bottom represents the region
738 confirmed by PCR analysis in *C. jejuni* 01/51 as having the same gene order as
739 NCTC11168. This region, flanked by altronate hydrolase C-terminus (Cj0483) and a
740 putative histidine triad (HIT) family protein (Cj499), contains genes encoding a
741 putative MFS (Major Facilitator Superfamily) transport protein (Cj0484), a putative
742 oxidoreductase (Cj0485), a putative sugar transporter (Cj0486), a putative
743 amidohydrolase (Cj0487), two conserved hypothetical proteins (Cj0488 and Cj0496),
744 a putative aldehyde dehydrogenase N and C terminus (Cj0489 and Cj0490), 30S
745 ribosomal proteins S12 and S7 (Cj0491 and Cj0492), *fus*A elongation factor G
746 (Cj0493), a putative exporting protein (Cj0494), a putative methyltransferase domain
747 protein (Cj0495), a putative lipoprotein (Cj0497) and *trp*C, an indole-3-glycerol
748 phosphate synthase (Cj0498).
749
750 **Figure 2: (A) DNA sequence of the Cj0489 - Cj0490 region in NCTC11168**
751 **compared with that in 01/51.** The dark grey boxes highlight single nucleotide
752 polymorphisms in the sequence between the two strains and the TTT trinucleotide
753 present in 01/51 compared with the TT dinucleotide in NCTC11168. The arrow
754 highlights the deletion of a single T nucleotide in NCTC11168 which results in a stop
755 codon (TAA). The light grey and clear boxes highlight the primers and probe used in
756 the melting curve assay respectively. **(B) The predicted amino acid sequence of the**
757 **Cj0489-Cj0490 region in NCTC11168 compared with 01/51.** The whole reading
758 frame is open for 01/51 from the beginning to end. The underlined sequence is the
759 NCTC11168 sequence for Cj0489, terminating in a stop codon (**.**), which then brings
760 about a frame shift before the start (arrow) of the Cj0490 sequence. The shaded box
761 highlights an intervening sequence in NCTC11168 which is not translated into the
762 final protein sequence.
763
764

765

Table 1: Location of the transposon insertion point in 26 selected mutants with reduced invasion compared with the parent strain 01/51. Invasion potential and motility are also given. Mutants are arranged according to functional classification (http://xbase.bham.ac.uk/campydb/).

Mutant	INT-407 Invasion[a]	Caco-2 Invasion[a]	Motility[b]	Location and annotation	
Adaptations and atypical conditions					
10D2	40	n/d	105	Cj0293	*sur*E, stationary phase survival
Central intermediary metabolism					
1C1	19	1	104	Cj0490	*ald*, aldehyde dehydrogenase C-terminus (571/1182)[c]
1C5	16	n/d	174	Cj0490	*ald*, aldehyde dehydrogenase C-terminus (139/1182)
1F4	18	n/d	78	Cj0490	*ald*, aldehyde dehydrogenase C-terminus (929/1182)
1F3	26	n/d	165	Cj0483	*uxa*X altronate hydrolase C-terminus
Conserved hypothetical proteins					
10D10	3	2	76	Cj1555c	hypothetical protein
Degradation- amino acids					
10H3	6	4	88	Cj1503c	putative proline dehydrogenase
DNA replication, restriction modification, recombination and repair					
6A9	37	40	82	Cj0690c	putative restriction/ modification enzyme
Energy metabolism- Electron transport					
3H7	2	<1	82	Cj1020c	putative cytochrome C
Membranes, lipoproteins and porins					
1B5	14	n/d[c]	83	Cj0497	putative lipoprotein (308/1278)
1D1	15	n/d	100	Cj0497	putative lipoprotein (563/1278)
10H1	2	3	88	Cj1245c	putative membrane protein
Signal transduction					
3E9	4	n/d	100	*flg*S	signal transduction histidine kinase
Surface polysaccharides, lipopolysaccharides and antigens					
3A10	1	2	88	Cj0685c	invasion protein *cip*A
9D2	2	26	92	JJD26997_1801	capsule polysaccharide biosynthesis protein of *C. jejuni* subsp. *doylei* 269.97
10E9	3	6	88	Cj1136	putative

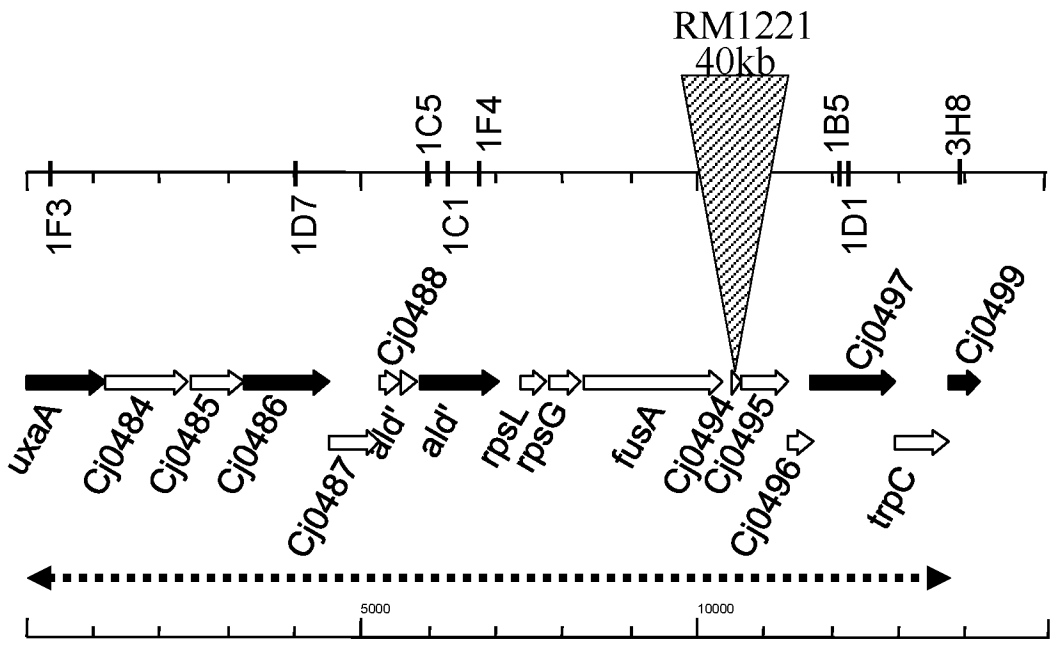
					glycosyltransferase
Transport- Anions					
6A7	11	4	86	Cj1539c	putative anion-uptake ABC-transport system permease
Transport- Carbohydrates, organic acids and alcohols					
1D7	10	3	100	Cj0486	putative sugar transporter
Transport- Peptides					
1H10	0	2	92	*dtp*T	di-tripeptide transporter protein not present in NCTC11168
Unknown function/miscellaneous					
1H6	0	n/d	75	Cj0519	putative rhodanese-like domain protein; molybdopterin biosynthesis protein
3H8	3	3	82	Cj0499	putative HIT family protein
10D12	5	4	76	*rlo*E	putative hypothetical protein, *C. jejuni* 260.94
10H12	1	2	96	Cj1305c	hypothetical protein
Location not yet determined					
10A4	8	4	98		not determined
10E11	2	3	94		not determined
10H2	<1	<1	112		not determined

a: Invasiveness is presented as a percentage of the invasiveness of the parent strain 01/51. The figure given is the mean of three replicates from within a single assay. Each assay was repeated at least once more.

b: Motility is presented as a percentage of the motility of the wild-type. The figure given is the mean of three replicates

c: Nucleotide position of the transposons in Cj0490 and Cj0497/the length of both genes (bp)

Figure 1



Cj0483-Cj0499 (14240 bp)

Figure 2A

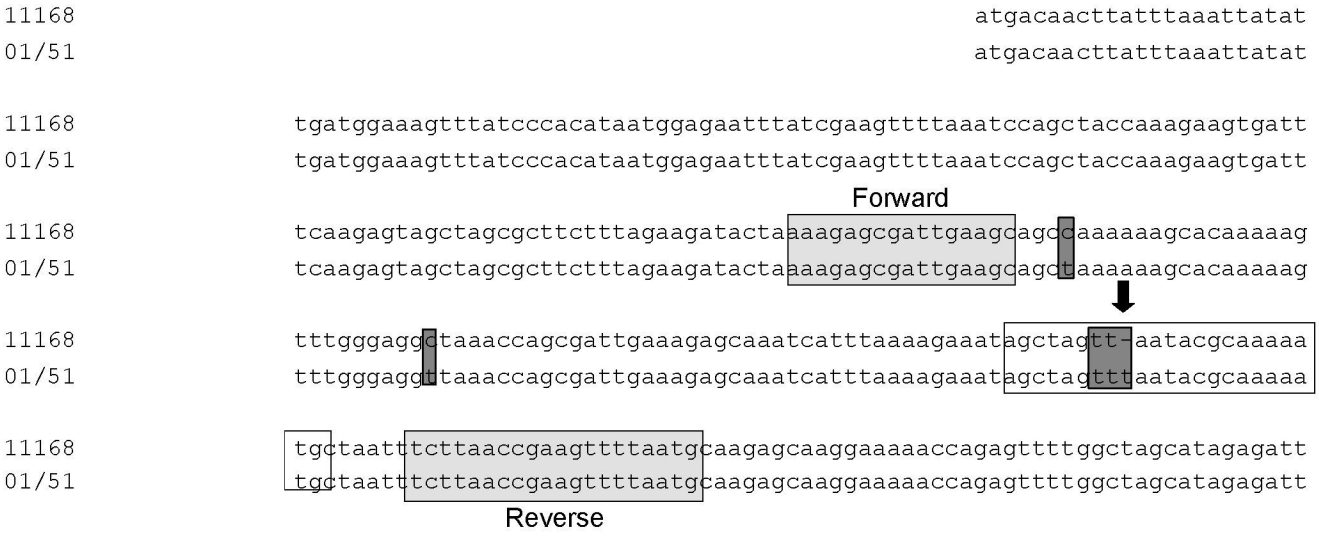


Figure 2B