# Re-annotation of the *Campylobacter jejuni* NCTC11168 genome and functional characterisation of selected genes involved in strain pathogenesis

**Thesis submitted for the degree of Doctor of Philosophy**

**By**

**Ozan Gundogdu B.Sc (Hons) M.Sc**

**Faculty of Infectious and Tropical Diseases**
**Department of Pathogen Molecular Biology**
**London School of Hygiene & Tropical Medicine**
**University of London**

**January 2011**

## Abstract

*Campylobacter jejuni* is the leading bacterial cause of foodborne human gastroenteritis worldwide. The first *C. jejuni* genome (strain NCTC11168) was sequenced in 2000. This original annotation was a milestone in *Campylobacter* research, but soon became outdated. A re-annotation and re-analysis of this genome sequence was performed resulting in updates to over 90% of coding sequences (CDSs) and modification of 18.2% of CDS product functions (Gundogdu *et al*., 2007). Following this re-annotation, 15 uncharacterised CDSs with revised functions relating to virulence, signal transduction or regulation of gene expression were selected for further investigation. Defined isogenic *C. jejuni* 11168H mutants were constructed and after preliminary analysis, the *Cj1556* and *Cj0248* mutants were selected for further characterisation. Cj1556 was originally annotated as a hypothetical protein and was updated to a MarR family transcriptional regulator. Further bioinformatic analysis indicated a putative role in regulating the oxidative stress response. A *C. jejuni* 11168H *Cj1556* mutant exhibited increased sensitivity to oxidative and aerobic ($O_2$) stress, decreased ability for intracellular survival in both Caco-2 intestinal epithelial cells (IECs) and J774A.1 mouse macrophages and a reduction in virulence in the *Galleria mellonella* infection model. Microarray analysis of gene expression changes in the *Cj1556* mutant compared to the wild-type strain indicated negative autoregulation of *Cj1556* expression and down-regulation of genes associated with oxidative and aerobic ($O_2$) stress responses. Cj0248 was originally annotated as a hypothetical protein however the re-annotation identified a HD domain linked to a superfamily of metal-dependent phosphohydrolases with roles in signal transduction in bacteria. Previously a *C. jejuni* 81-176 *Cj0248* mutant was shown to be deficient for motility and chick colonisation, however the exact function of Cj0248 was not investigated. The *C. jejuni* 11168H *Cj0248* mutant also possessed a reduced motility phenotype and exhibited reduced interaction and invasion when co-cultured with Caco-2 IECs compared to the wild-type strain. However the *Cj0248* mutant showed no difference in autoagglutination compared to the wild-type strain and TEM analysis indicated the mutant possessed intact flagella. Higher magnification TEM indicated the possibility of an altered flagella basal body region in the *Cj0248* mutant. Secretion profile analysis identified no differences in the protein profile of the *Cj0248* mutant compared to the wild-type strain. The exact function of Cj0248 remains unclear.

**Contents** **Page**

## Declaration

I declare that all of the work included in this thesis is my own unless otherwise stated. All experiments were performed at the London School of Hygiene & Tropical Medicine (LSHTM). All of the re-annotation work was performed at the Wellcome Trust Sanger Institute (WTSI) with assistance from Stephen Bentley, Matt Holden and Julian Parkhill in accordance with WTSI annotation guidelines. All TEM images were captured and analysed by Maria McCrossan and David Ellis in the LSHTM TEM unit.

**Acknowledgements**

There are a number of people I would like to thank for helping me assemble this thesis. First and foremost I would like to thank my supervisor Dr. Nick Dorrell. I will always be grateful for the guidance, support and time you have given me during this project, from the phone calls I used to make asking endless questions, to the "do you have two minutes Nick?". You have continuously challenged me to ensure good research comes to the fore. It has been a privilege to work with you. Thank you Nick!

I would like to thank Prof. Brendan Wren for allowing me to carry out this PhD and for all his guidance throughout the research project. You have always pushed me to challenge myself further and to never stand still. I have been extremely lucky to have worked in your group. Thank you Brendan!

I would like to give special thanks to Dr Dorrell's team; Abdi Elmi, Dominic Mills, Nevida Naz and Baraa Alhaj-Hussein. Thank you for all your questions, help and support!

I would like to thank everyone past and present in Prof. Wren's group for all their support and help throughout the years: Dr Jon Cuccui, Dr Emily Kay, Dr Rebecca Langdon, Melissa Martin, Dr Richard Stabler, Prof. Andrey Karlyshev, Dr. Olivia Champion, Dr Philippa Strong, Dr George Joshua, Gillian Thacker, Laura Yates, Elizabeth Donahue, Dr Esmeralda Valiente, Dr Vanessa Terra and Dr Madeleine Moule; you have all been fantastic!

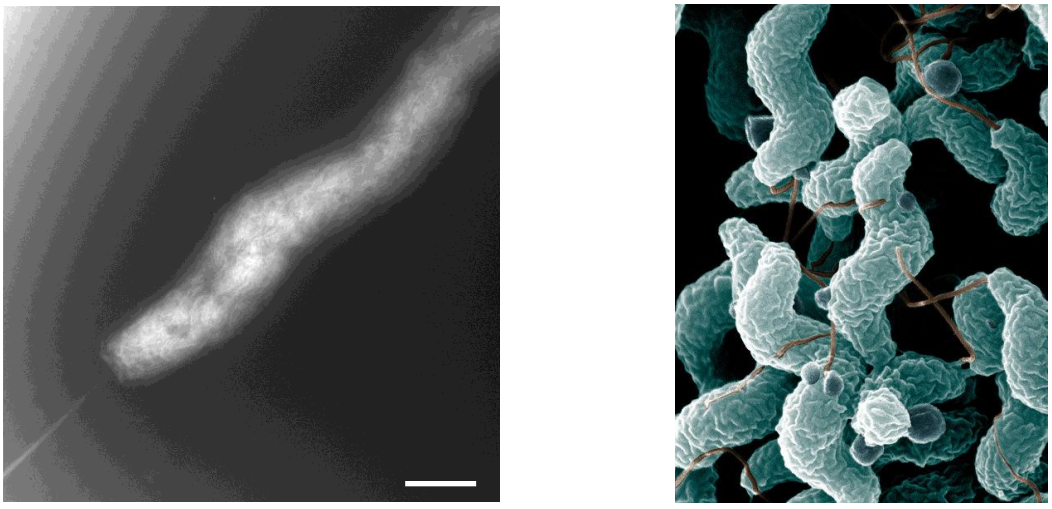
I would like to thank Prof. Julian Parkhill, Dr. Stephen Bentley, Dr. Nick Thompson and Dr. Matt Holden at the WTSI for allowing me to carry out the re-annotation work, teaching and supporting me along the way.

Finally, I would like to thank Dina Vara for her support throughout the years. I would like to give a special thanks to my family for supporting me throughout the years. Volkan, Mum and Dad, you have all been amazing and put up with me for all these years! I could not have done this without you and won't ever be able to repay you for all the support you have given me. Thank you!

# Chapter 1: Introduction

## 1.1 *Campylobacter jejuni*

The human intestinal pathogen *C. jejuni* is a microaerophilic bacterium which grows optimally in a 12% $CO_2$ and 5% $O_2$ atmosphere between 37°C and 42°C (Garenaux *et al*., 2008, Park, 2002). *C. jejuni* has the general characteristics of being Gram-negative, curved rod-shaped or spiral, with an approximate size of 0.2 – 0.8 μm wide and 0.5 – 5.0 μm long. The bacterium is motile via uni- or bi-polar flagella (Figure 1.1).



(A) (B)

Figure 1.1. TEM of a *C. jejuni* cell from individual colonies grown on blood plates resuspended in a 1.5 ml microcentrifuge tube containing 2.5% (v/v) glutaraldehyde, 2.5% (v/v) paraformaldehyde and 0.1 M sodium cacodylate buffer (pH 7.4) (A). Scanning electron micrograph of *C. jejuni* (B). Image obtained from the Agricultural Research Service – United States Department of Agriculture (http://www.ars.usda.gov/main/main.htm). Magnification = 250,000x. Size bar = 100 nm.

*C. jejuni* is capable of growth at temperatures ranging from 30°C to 47°C and therefore is capable of growth at the body temperatures of human and avian hosts, 37°C and 42°C respectively (Blaser *et al.*, 1983, Garenaux *et al.*, 2008). Traditionally, several problems have hindered the study of *Campylobacter* species compared to other enteric bacteria such as *E. coli* and *Salmonella* species. In particular, for decade's scientists did not have

the appropriate conditions to culture the bacterium and were unknown until microaerobic growth conditions for bacteria were established. In the last two decades microaerobic cabinets have greatly assisted in the routine growth and study of *Campylobacter* species. The bacterium is oxidase-positive and unable to ferment or oxidise sugars. Energy is obtained from amino acids or tricarboxylic acid cycle intermediates rather than utilisation of carbohydrates (Debruyne *et al*., 2008). *C. jejuni* has two sub-species, *C. jejuni* subsp. *jejuni* and *C. jejuni* subsp. *doylei* which are distinguished by biochemical testing with the latter characterised by the absence of nitrate reduction and susceptibility to cephalothin (Debruyne *et al*., 2008).

### 1.2 *Campylobacter* species

It is believed that the earliest description of *Campylobacter* was by Theodor Escherich in 1886, followed by the 1913 publication describing the isolation of *Vibrio*-like organisms from aborted ovine foetuses by McFadyean and Stockman (Escherich, 1886, McFadyean and Stockman, 1913). In later years, several novel "Vibrio" species in a range of animals with and without clinical symptoms were identified. In 1963, Sebald and Véron transferred two of these *Vibrio* species into the new genus *Campylobacter* due to low DNA base composition, microaerophilic growth requirements and non-fermentative metabolism (Sebald and Veron, 1963). In 1973 Veron and Chatelain outlined four distinct species in the genus *Campylobacter* (Véron and Chatelain, 1973, Debruyne *et al*., 2008). Improved isolation, growth and incubation methods have allowed further species to be identified (Figure 1.2). There are currently 23 *Campylobacter* species and 21 different *C. jejuni* strains (July 2010). *Campylobacter* species are generally motile, typically microaerophilic, have a low G + C content and do not form spores. Strict anaerobiosis is a stress condition for *C. jejuni*, however alternative respiratory pathways can contribute significantly to energy conservation under oxygen-limited conditions (Sellars *et al*., 2002). Anaerobic growth can occur with fumarate or nitrate as an electron acceptor, whereas microaerobic growth will occur with hydrogen, formate or succinate used as the electron source (Debruyne *et al*., 2008).

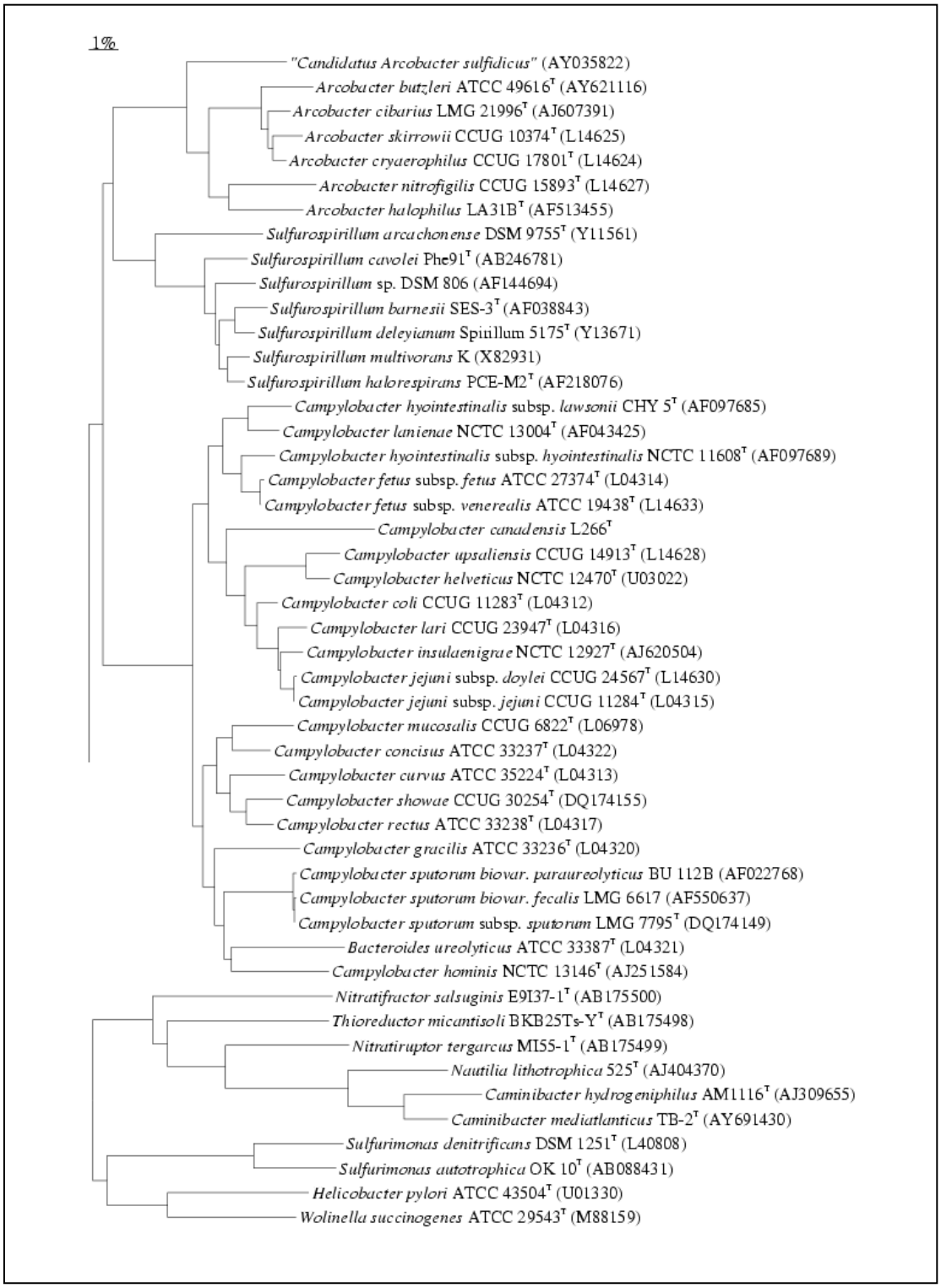



Figure 1.2. Phylogenetic tree of the relatedness of *Campylobacter* species to other bacteria as inferred by comparison of 16S rRNA gene sequences. The scale bar indicates 1% distance for the dendogram. Image obtained from Debruyne, *et al*., 2008.

### 1.3 Pathology of *C. jejuni* infection

*C. jejuni* infection is one of the leading bacterial causes of gastroenteritis in humans and is the leading foodborne cause of gastroenteritis worldwide (Allos, 2001). *Campylobacter* species such as *C. fetus* have long been recognised as a cause of diarrhoea in cattle and of septic abortion in both cattle and sheep (Jerrett *et al.*, 1984, Sauerwein *et al.*, 1993). The identification of *C. jejuni* as a causative agent of human gastroenteritis became apparent in 1977 due to improved selective and transport media for the isolation of the organism from clinical specimens (Skirrow, 1977). Historically, more than 95% of *Campylobacter* strains isolated and identified in cases of human disease have been *C. jejuni* and *C. coli* (Lastovica and Allos, 2008). Thus, the clinical aspects described in this section are mainly attributed to *C. jejuni* and *C. coli*. The essential lesion in *Campylobacter* enteritis is an acute inflammatory enteritis which commonly extends down the intestine to affect the colon and rectum (Blaser and Engberg, 2008). Terminal ileitis and cecitis with inflammation of the mesentery is also common (Blaser and Engberg, 2008). In terms of disease symptoms, there does not seem to be any clear difference between infections caused by *C. jejuni* and *C. coli* when assessed by frequency of diarrhoea, blood in stool, abdominal pain, fever, vomiting, mean duration of illness or admission to hospitals (Blaser and Engberg, 2008). One other *Campylobacter* species of significance is *C. fetus* which is of particular economic importance as this species can cause sterility and abortion in cattle and sheep (Bergen *et al*., 2008).

#### 1.3.1 Gastroenteritis

The most common clinical symptom observed in humans infected with *C. jejuni* is gastroenteritis (Figure 1.3) (Blaser, 1997). *C. jejuni* causes acute inflammatory enteritis, affecting the intestine down to the colon and rectum (Blaser and Engberg, 2008). Symptoms vary based on the individual infected, however following an incubation period of 24-72 hours, acute diarrhoea usually occurs (Young and Mansfield, 2005). This may be followed or preceded by non-specific syndrome of fever, chills, myalgia, headache and can lead to abdominal cramps. The illness typically lasts for 5 - 10 days and is self-limiting (Blaser, 1997). However extended illness can also occur, specifically in immunocompromised individuals such as elderly or human immunodeficiency syndrome (HIV) patients (Sorvillo *et al*., 1991). Presentation of clinical symptoms in humans from *C. jejuni* infection has been noted to vary between developed and developing countries. Clinical symptoms of humans in developed countries infected by

*C. jejuni* typically results in inflammatory bloody diarrhoea and abdominal pain (Black *et al*., 1988, Karmali and Fleming, 1979, Ketley *et al*., 1996, Richard and David, 2000). Occasionally, intestinal complications, extra-intestinal infections and late onset complications can follow initial gastroenteritis (Figure 1.3).

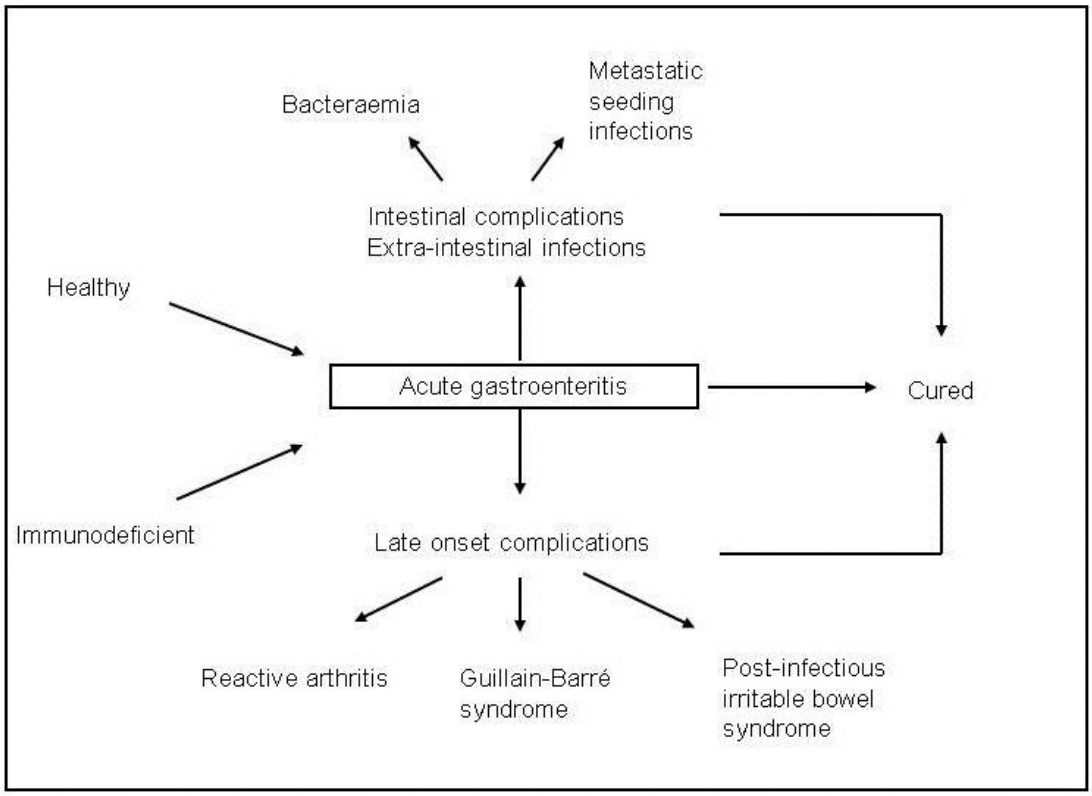



Figure 1.3. Overview of illnesses caused by *C. jejuni*. Image adapted from Blaser and Engberg., 2008.

*Campylobacter* enteritis in developing countries is low frequency in adults (Blaser and Engberg, 2008) yet it has been noted from a study in Egypt that *C. jejuni* shedding can occur from adults (Pazzaglia *et al*., 1991). In developing countries, infection rates decline with age with fewer infections being associated with diarrhoea (Blaser and Engberg, 2008). The clinical symptoms for humans in developing countries infected by *C. jejuni* are generally non-inflammatory watery diarrhoea (Richard and David, 2000). The differences observed in clinical symptoms after *C. jejuni* infection between patients from developed and developing countries have been linked to pre-exposure of the developing country population to the bacterium at not only an earlier stage in life, which allows immunity to develop, but from probable multiple re-exposure to the bacterium

(Calva *et al*., 1988, Richardson *et al*., 1981, Taylor *et al*., 1988). In addition, factors such as diet, lifestyle, common exposure to parasitic infections and even strain differences may explain the reported differences in clinical outcomes between *C. jejuni* infection in developed and developing countries.

Bacteraemia has been observed rarely following *C. jejuni* infection (Crushell *et al*., 2004) and is believed to occur as a transient event in the early stages of infection, especially in patients with high fever (Blaser and Engberg, 2008). *C. jejuni* bacteraemia is difficult to detect because of the lack of blood samples taken in the early stages of disease presentation and not all methods of detecting bacteraemia are equally sensitive for *Campylobacter* species (Wang and Blaser, 1986). In a recent study in Denmark, the highest frequency of bacteraemia associated with *Campylobacter* species was 2.9 per 1,000,000 persons (Nielsen *et al*., 2010).

### 1.3.2 Post-infectious sequelae

Even though gastroenteritis is the most common clinical symptom observed in humans infected with *C. jejuni*, on rare occasions a number of post-infectious sequelae such as Guillain-Barré syndrome (GBS), Miller Fisher syndrome (MFS), endocarditis, meningitis, post-dysenteric irritable bowel syndrome (PD-IBS) and reactive arthritis (ReA) can develop (Kuroki *et al*., 1993, Skirrow, 1991). Current evidence suggests *C. jejuni* as the predominant preceding infection in GBS and that this infection triggers the production of cross-reactive antibodies to human gangliosides that damage peripheral nerve tissue (Jacobs *et al*., 2008). GBS is characterised by an acute progressive and symmetrical motor weakness of limbs with loss of tendon reflexes (Asbury and Cornblath, 1990). GBS is believed to occur in between 0.6-1.9 cases for every 100,000 people infected with *Campylobacter* (Schonberger *et al*., 1981). The mortality rate is less than 5-10% and is mainly the result of respiratory or autonomic nervous system complications (Ropper *et al*., 1991).

MFS is a sub-form of GBS that is characterised by areflexia, ataxia, and ophthalmoplegia (Fisher, 1956). *C. jejuni* is also the predominant type of preceding infection in patients with MFS (Koga *et al*., 2005). PD-IBS causes continued diarrhoeal complications and the prevalence of such cases are believed to be far higher than observed (5-20%) (Thornley *et al*., 2001, Marshall *et al*., 2006, Dunlop *et al*., 2003). ReA typically affects ankles, knees, wrists and the small joints of the hands and feet (Schaad, 1982, Peterson,

1994). The prevalence of ReA is believed to be 1-5% of those previously infected with *Campylobacter* species (Pope *et al*., 2007). Little is known about the mechanism(s) of how *Campylobacter* infection causes ReA.

### 1.3.3 Epidemiology

Most illnesses caused by *Campylobacter* species occur sporadically, however the bacterium can still be associated with large outbreaks affecting thousands of people (Miller and Mandrell, 2005, Clark *et al*., 2003). Infection from *Campylobacter* species is estimated to cause 5–14% of diarrhoea worldwide (Coker *et al*., 2002). Each year *Campylobacter* is responsible for an estimated 400 million human cases of gastroenteritis worldwide, making this the leading cause of bacterial food borne disease and a major causative agent of traveller's disease (Walker, 2005, Friedman *et al*., 2000). Herein, lies a peculiarity often referred to as the "*Campylobacter* conundrum" (Jones, 2001) – how can a bacterium that rarely causes outbreaks be responsible for such large numbers of infection? This may be because the organism is omnipresent in the environment and probably survives better in the oxidative stress of the environment than traditionally thought. The number of cases of *Campylobacter* species reported in England and Wales over the last 20 years is illustrated in Figure 1.4. However, the actual numbers are estimated to be nearer half a million per year due to lack of reporting by patients with less severe clinical symptoms (Wheeler *et al*., 1999).

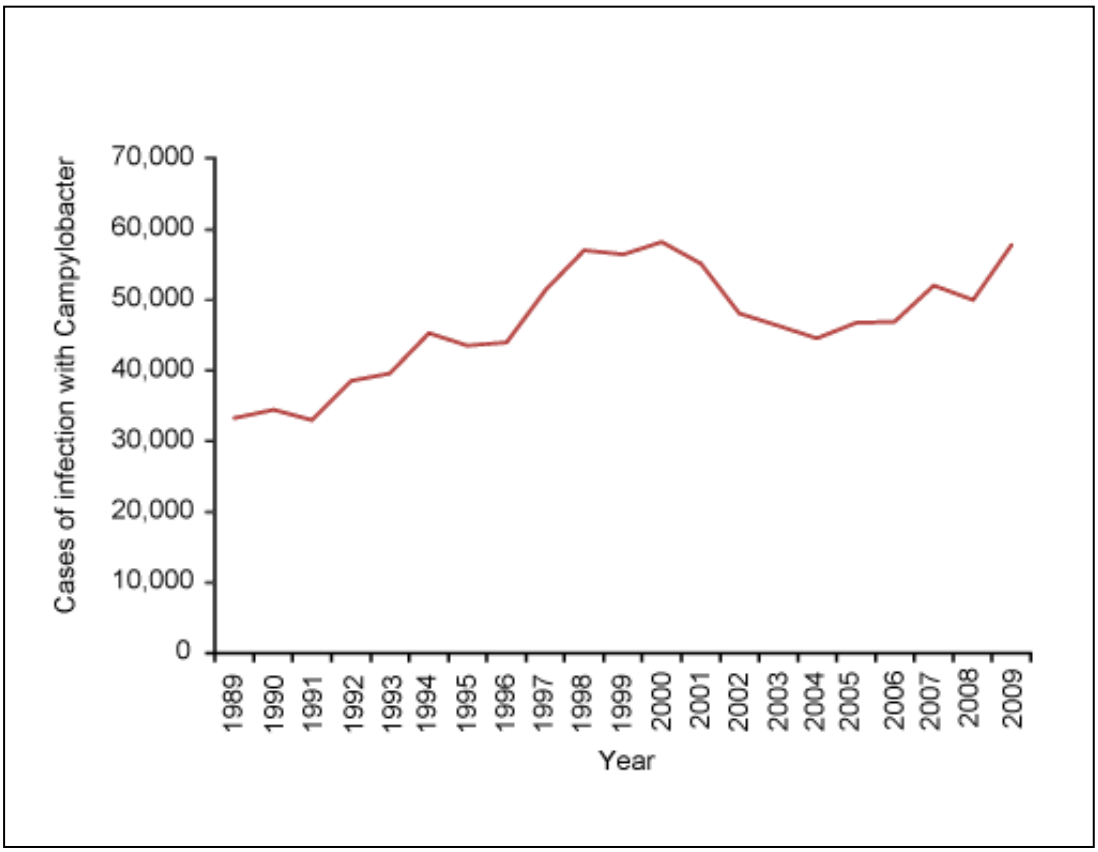



Figure 1.4. Laboratory cases of *Campylobacter* reported to the Health Protection Agency Centre for Infections England and Wales, 1989-2009. Data shown is human *Campylobacter* cases reported from diagnostic laboratories to the HPA's local and national surveillance. It includes patients with enteric and non-enteric infections. Data obtained from http://www.hpa.org.uk/infections/topics_az/campy/data_ew.htm.

Many countries in temperate zones observe a sharp increase in the isolation of *Campylobacter* species infections in the spring, a well-defined summer peak and a gradual decrease in the autumn or winter (Skirrow, 1987, Kovats *et al*., 2005, Miller *et al*., 2004, Olson *et al*., 2008). Human behaviour has been suggested as a possible reason for the summer peak with the increased barbecuing of meat and water-associated recreational activities (Olson *et al*., 2008). Investigation into the age distribution of humans infected with *Campylobacter* species in the U.S.A has identified children younger than 12 months of having the highest rate of *Campylobacter* species infection at 27 per 100,000 cases. Other notable peaks occurred at 1-4 and 40-49 years of age in males, and 1-4 and 20-29 years of age in females (Tauxe *et al*., 1988). This bimodal age distribution has also been described in other European countries (Olson *et al*., 2008).

### 1.3.4 Sources of *C. jejuni* infection

*Campylobacter* infection is most commonly associated with the consumption of contaminated poultry or meat products (MacKichan *et al*., 2004). Potential sources and routes of human infection by *C. jejuni* are shown in Figure 1.5. *Campylobacter* species are a typical commensal microorganism of the gastrointestinal tract of many birds and animals (Altekruse *et al*., 1999). Studies have shown the bacterium to be maintained asymptomatically at levels of $10^6$ to $10^8$ CFU per gram of caecal contents within poultry (Stern, 2008). Other reported sources of infection include untreated water, raw or unpasteurized milk, vegetables and transmission from pets (Crushell *et al*., 2004, Olson *et al*., 2008, Harris *et al*., 1987, Friedman *et al*., 2004, Miller and Mandrell, 2005).

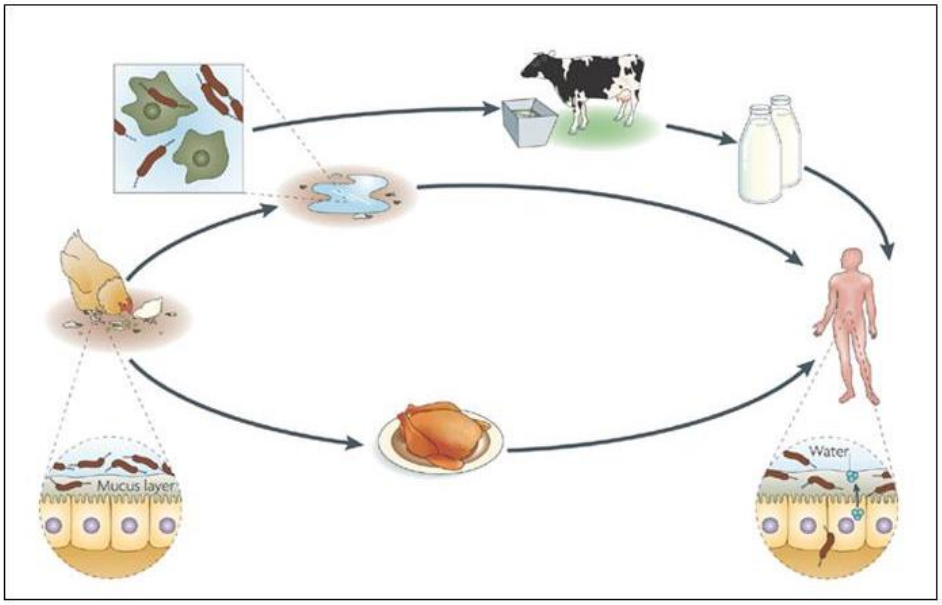



Figure 1.5. Potential environmental reservoirs of *C. jejuni* leading to human infection. The bacterium colonises the chicken gastrointestinal tract and is passed between chicks within a flock through the faecal-oral route. *C. jejuni* can enter the water supply, where it can associate with protozoans, such as freshwater amoebae. *C. jejuni* can also infect humans directly through the drinking water or through the consumption of animal products. Image from Young *et al*. 2007.

### 1.3.5 Diagnosis and Treatment

Accurate diagnosis requires the isolation of bacteria from patient stools and performing growth-dependent tests, immunological assays or a Gram-stain. Stool antigen tests for *Campylobacter* species have also been developed (Fitzgerald *et al*., 2008). Sensitivity is

believed to vary between 80-96% compared with culture and has specificity greater than 97% (Dediste *et al*., 2003, Hindiyeh *et al*., 2000, Tolcin *et al*., 2000). Currently there is a lack of reliable molecular-based assays to detect *Campylobacter* species in stool samples (Dediste *et al*., 2003, Hindiyeh *et al*., 2000, Tolcin *et al*., 2000). Molecular based techniques have advantages over culturing including same-day detection, additional data regarding mixed infections, automation and identification of uncommon species (Dediste *et al*., 2003, Hindiyeh *et al*., 2000, Tolcin *et al*., 2000).

Antibiotics are not typically administered as *Campylobacter* infections are generally self-limiting. Where antibiotics have been administered, notable differences in resistance rates have been identified between countries (Blaser and Engberg, 2008). Treatment consists principally of oral replacement of fluid and electrolytes lost through diarrhoea and vomiting (Taylor and Tracz, 2005). However for severe cases, antibiotics from the fluoroquinolone class such as ciprofloxacin are administered which function by inhibiting bacterial DNA gyrase activity and thus inhibiting DNA unwinding and replication. However resistance to ciprofloxacin and other fluoroquinolones have been shown for *Campylobacter* strains isolated from humans (Blaser and Engberg, 2008). Up to 20% of patients treated with a fluoroquinolone for a *Campylobacter* infection will develop resistance (Wistrom and Norrby, 1995, Fitzgerald *et al*., 2008). Gentamicin has also been recommended for the treatment of campylobacteriosis, particularly in patients with systemic infection (Skirrow and Blaser, 2002). Resistance to antibiotics for *Campylobacter* strains is believed to have developed as a result of fluoroquinolones being used in veterinary medicine and the resulting resistance developing in livestock animals (Blaser and Engberg, 2008).

### 1.4 Virulence determinants

The four major virulence determinants in *C. jejuni* were identified before, during and after the publication of the original genome sequenced strain NCTC11168 (Parkhill *et al*., 2000) and are shown in Figure 1.6. A virulence determinant is a feature of the bacterium (in this case sugar structures) that allows for colonisation (by adhesion and/or invasion), immune evasion, immune suppression and/or generally assisting the persistence of the bacterium in a host. In *C. jejuni* these include flagella, lipooligosaccharide (LOS), *N*-linked glycoproteins, *O*-linked glycosylated flagella, capsular polysaccharide (CPS), cytolethal distending toxin (CDT) and also adhesins.

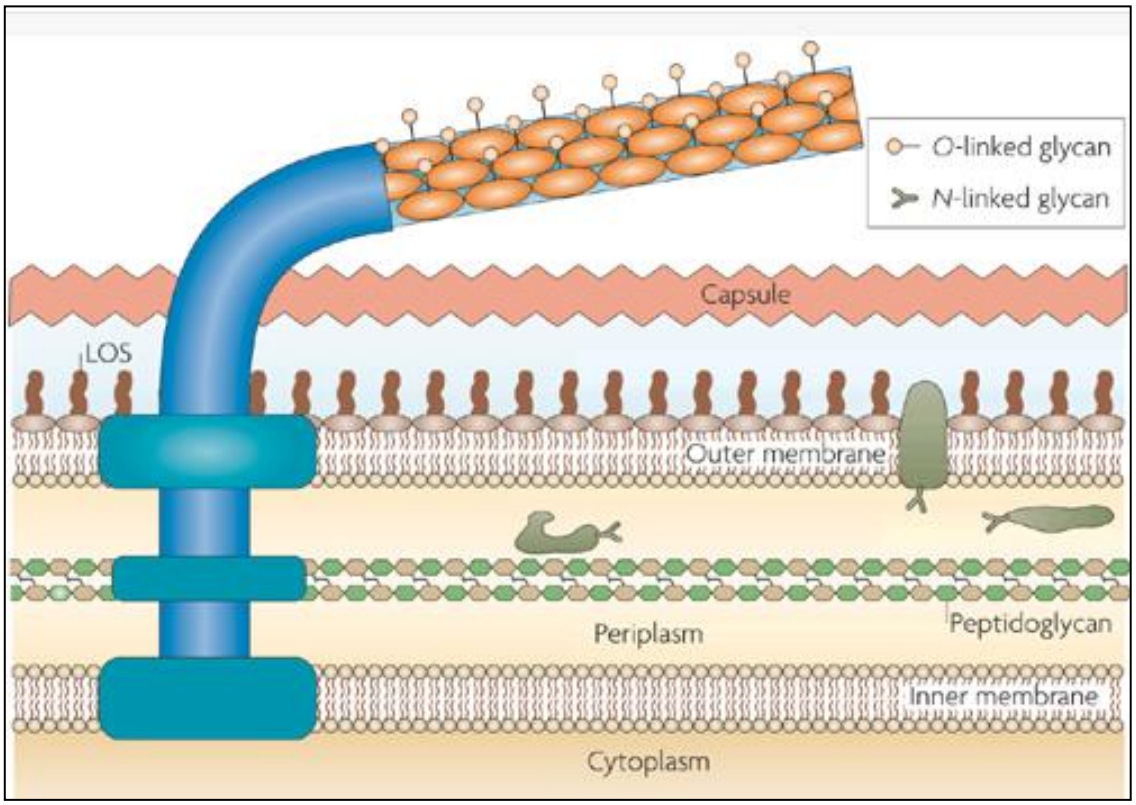



Figure 1.6. Potential glyco-based virulence determinants in *C. jejuni*. These include LOS, *N*-linked glycoproteins, *O*-linked glycosylated flagella and CPS. Image obtained from Young *et al*. 2007.

### 1.4.1 Flagella

The flagellum is a key virulence factor for *C. jejuni* enabling motility and intestinal colonisation (Hendrixson, 2006, Hendrixson and DiRita, 2004). Motility via functional flagella is not only important for colonisation, but studies have demonstrated that the flagella can also secrete important virulence factors that enable the bacterium to invade human intestinal cells (Black *et al*., 1988, Grant *et al*., 1993, Hendrixson, 2008). Flagella allow the bacterium to penetrate, colonize and persist in the thick mucus lining the intestinal surface and crypts (Guerry, 2007, Hendrixson, 2008).

#### 1.4.1.1 Biosynthesis of flagella

Up to 50 genes have been implicated in the biosynthesis of the *C. jejuni* flagella requiring a finely tuned regulatory system to control the biosynthesis of the organelle (Hendrixson, 2008). The flagella rod links the flagella export apparatus to the flagella hook, which connects to the major and minor flagellins where FlaA is the predominant flagellin, whereas FlaB appears to be sparse, contributing to less than 20% of the flagella

filament (Guerry *et al*., 1991) (Figure 1.7). The flagella basal body proteins (FlgABC, FlgFGHIJ and FliE), the ring structures in the peptidoglycan and outer membrane (FliFHI, FlhAB and FliOPQR) and rod proteins (FlgDE and FlgKL) are all present in *C. jejuni*. Key components of the flagella include cytoplasmic flagella motor switch complex proteins, inner-membrane-localised export apparatus and a cytoplasmic ATPase which together form a secretion system for flagella proteins. Regulation of the expression of these genes is controlled by RpoN ($\sigma^{54}$), FliA ($\sigma^{28}$) and the two-component regulatory system FlgSR (Hendrixson, 2008). Studies have shown that site directed mutagenesis of *rpoN* and *fliA* result in non-motile mutants (Hendrixson *et al*., 2001, Jagannathan *et al*., 2001). RpoN regulates expression of *flaB* along with the hook and basal-body genes, whereas FliA regulates expression of *flaA*. Recent studies have also identified FlhF to be necessary for flagella organelle development and also required for $\sigma^{54}$-dependent flagella expression (Balaban *et al*., 2009). In addition, until the flagella secretory system has formed, FlgM can inhibit the activity of $\sigma^{28}$. Once the secretory system is formed, FlgM is transported out of the cytoplasm and $\sigma^{28}$ can initiate the expression of genes such as *flaA* encoding the major flagellin (Hendrixson and DiRita, 2003, Wosten *et al*., 2004).

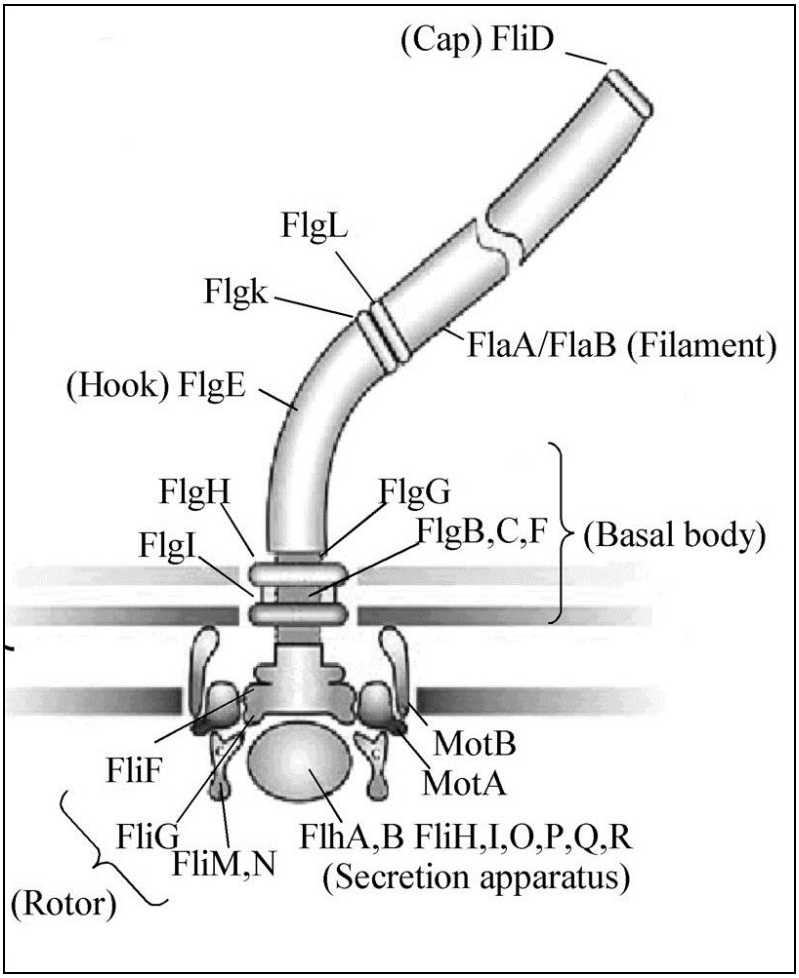



Figure 1.7. Biosynthesis of *C. jejuni* flagella. Image obtained from Wosten, *et al*. 2004.

Much of the early understanding of flagella biosynthesis and regulation has developed from mutagenesis studies on *Salmonella* and *Vibrio* species (Hughes *et al*., 1993, Karlinsey *et al*., 2000, Klose and Mekalanos, 1998). Many of the orthologous genes have been identified in *C. jejuni*, however important differences exist between the flagella biosynthesis mechanisms in these microorganisms when compared to *C. jejuni*. Analysis of different *Campylobacter* species has shown that flagella genes are not organised into one big operon, but are dispersed throughout the genome (Fouts *et al*., 2005, Chilcott and Hughes, 2000, Liu and Ochman, 2007, Parkhill *et al*., 2000). This is in contrast to other bacteria which typically have one operon for flagella genes allowing expression to be controlled from a single promoter. Formation of flagella within *Campylobacter* species is considered to be a three stage process. Initial steps include the formation of export apparatus, the expression of regulatory genes encoding RpoN, FliA, FlgM, FlgS, FlgR and FlhF. Unlike *C. jejuni*, other motile bacteria contain a master regulator for controlling flagella formation; FlhDC in *Salmonella* species, FlrA for *Vibrio* species and

FleQ for *Pseudomonas* species (Klose and Mekalanos, 1998, Arora *et al*., 1997, Karlinsey *et al*., 2000, Kutsukake *et al*., 1990). After formation of the flagella export apparatus, FlgS may sense an undetermined signal to autophosphorylate and begin a signal transduction cascade, terminating in activation of FlgR and expressing $\sigma^{54}$-dependent flagella genes (Hendrixson, 2008). Differences exist in the flagella biosynthesis process when compared to *Salmonella* species where three classes of genes exist; class I encodes the master regulator FlhDC, class II genes encode FliA, FlgM, export apparatus, basal body and hook components and class III genes encode flagellins (Hughes *et al*., 1993, Karlinsey *et al*., 2000, Kutsukake *et al*., 1990). In the case of *Vibrio* species there exists four classes of genes for flagella biosynthesis; class I encodes the master regulator FlrA and in conjunction with RpoD activates class II genes. These activate FliA, export genes, a two component regulatory system and a $\sigma^{54}$-dependent response regulator (Arora *et al*., 1997, Dasgupta *et al*., 2000). The response regulator FlrB for *Vibrio* species allows RpoD to initiate expression of hook and basal genes. Only one class III gene encodes the flagellin and this is followed by class IV genes encoding further flagellins and motor proteins (Dasgupta *et al*., 2003, Prouty *et al*., 2001, Klose and Mekalanos, 1998).

#### 1.4.1.2 Flagella and chemotaxis

A key element in flagella motility is chemotaxis which enables bacteria to move towards nutrient and away from noxious agents (Hendrixson, 2008). Extracellular signals, often sugars or amino acids, are sensed by chemoreceptors called methyl-accepting chemotaxis proteins (MCPs) which typically contain a periplasmic domain that binds to the signal. *C. jejuni* contains approximately 10 chemoreceptor proteins involved in this process of sensing. CheAY is a two-component regulatory system which controls chemotaxis in response to environmental stimuli. The binding of the signal ligand is relayed by the MCP to CheA, a histidine kinase that forms a complex with the MCP in conjunction with CheW. CheA autophosphorylates and subsequently phosphorylates CheY, the response regulator (Young *et al*., 2007). Phosphorylated CheY interacts with the FliM of the flagella motor to initiate movement (Yao *et al*., 1997). *C. jejuni* lacks a homologue of CheZ, a phosphatase which dephosphorylates CheY, but does possess a homologue of the poorly understood protein CheV (Marchant *et al*., 2002, Young *et al*., 2007, Parkhill *et al*., 2000). CheV has an amino terminal CheW-like domain and a carboxyl terminal CheY-like domain, so has been hypothesised to act as a phosphate sink for the chemotaxis signal-transduction machinery (Young *et al*., 2007, Marchant *et al*., 2002,

Pittman *et al*., 2001). It is believed this may ameliorate the effect of the absence of a CheZ phosphatase on phosphate flow through this signal-transduction pathway (Young *et al*., 2007, Marchant *et al*., 2002). Another type of taxis, called energy taxis, is a response to an intracellular signal, such as the proton motive force or the redox state of the electron-transport system (Young *et al*., 2007). Sequence and genetic analyses indicate that *C. jejuni* transduces an energy taxis (or aerotaxis) signal using two proteins, CetA and CetB, in place of the single protein (Aer) that is used by *E. coli* and other species (Hendrixson *et al*., 2001, Young *et al*., 2007).

Based on *E. coli* studies, the bacterial chemotaxis signal transduction pathway is considered a model for bacterial two-component regulatory systems and specifically for chemotaxis signal transduction (Eisenbach, 1996, Falke *et al*., 1997). Movement is either based on a tumbling mode induced by clockwise-rotating dissociated flagella or a smooth, straight swimming mode induced by anti-clockwise flagella rotation (Korolik and Ketley, 2008). Tumbling motion is performed to allow reorientation, whereas the straight motion is used by the bacterium to swim to a specific concentration gradient (Korolik and Ketley, 2008). Importantly, in the absence of chemoattractants, autophosphorylation of CheA is inhibited and CheY is not phosphorylated (Korolik and Ketley, 2008). The signal transduction pathway is initiated by the MCP sensory receptors described above. The interaction between CheY and FliM determines the type of movement, as studies in *E. coli* have shown phosphorylated CheY need to occupy at least 70% of the available FliM molecules in the basal body for a change of rotational direction to occur (Bren and Eisenbach, 2001). It is predicted that when CheY binds to FliM in the flagella motor complex, this leads to a clockwise rotation, resulting in tumbling of the cell (Falke *et al*., 1997, Korolik and Ketley, 2008, Bren and Eisenbach, 2001, Spohn and Scarlato, 2001). In contrast, if a chemoattractant is bound to the MCP, this leads to suppression of CheA activity which in turn leads to less CheY binding to FliM and hence a return to anti-clockwise rotation leading to swimming (Falke *et al*., 1997, Korolik and Ketley, 2008, Bren and Eisenbach, 2001).

#### 1.4.1.3 Flagella and secretion

In addition to providing motility, it has been proposed that flagella can act as a secretory channel in *C. jejuni*. *C. jejuni* lack classical type III and IV secretion systems used by many bacteria to secrete a range of determinants important in survival and pathogenesis (Parkhill *et al*., 2000). However, *C. jejuni* 81-176 has been noted to contain a pVIR

plasmid containing homologues of CDSs matching components of the type IV secretion system (Bacon *et al.*, 2000). Studies have described the identification of at least eight proteins from the culture supernatant when *C. jejuni* was grown in the presence of INT 407 IECs (Konkel *et al*., 2004, Konkel *et al*., 1999b, Young *et al*., 1999). Protein secretion was not detected when *C. jejuni* was incubated in the absence of IECs. These proteins, termed Cia (Campylobacter Invasion Antigen) proteins are further stimulated in the presence of other factors such as serum and bile salts (Guerry, 2007, Konkel *et al*., 1999b, Rivera-Amill *et al*., 2001, Malik-Kale *et al*., 2008). A *ciaB* mutant was shown to be 50- fold less invasive compared to the *C. jejuni* F38011 wild-type strain (Konkel *et al*., 1999b). The secretion of Cia proteins is dependent on a functional flagellum, indicating that this organelle has a dual function in motility and as a type III secretion system (Konkel *et al*., 2004). Secretion of Cia proteins was not detected in a *flhB* mutant (defective in export), or in *flgB, flgC* and *flgE* mutants (defective for the basal body and hook) (Larson *et al*., 2008). Secretion of Cia proteins occurred in both *flaA* and *flaB* single mutants, but not a double *flaA flaB* mutant that lacked the filament (Konkel *et al*., 2004).

The putative role of flagella as secretory tools is controversial. Watson and Galan suggest that there is little evidence that any component of the flagella apparatus is directly secreting proteins involved in invasion when in contact with IECs (Watson and Galán, 2008). The authors argue that previous research shows that gene mutations causing a reduced motility phenotype, but not affecting the structure of the flagella apparatus show reduced invasion into host IECs (Yao *et al*., 1994). The argument is that these mutants, though non-motile, still harbour secretion positive flagella and so should be able to secrete the proposed virulence proteins. Watson and Galan argue that it is motility that is essential for invasion (Watson and Galán, 2008). Recent studies by this group have also shown that a *flaA* mutant (lacking the major flagellin filament subunit) and a *motA* mutant (lacking rotation), both have severely reduced ability to invade IECs even though both mutants should still harbour a functional secreting flagella structure (Novik *et al*., 2010).

### 1.4.2 Lipooligosaccharide

Gram-negative bacteria contain lipopolysaccharide (LPS) or LOS in the outer membrane which are important variable cell surface structures playing a key role in virulence (Moran *et al*., 1996). LPS is made up of an O-polysaccharide chain, core oligosaccharide

and a lipid A component in the outer membrane. LOS lacks an O-polysaccharide repeating structure and exhibits greater structural diversity in the outer core than LPS (Gilbert *et al*., 2008, Moran and Penner, 1999). LPSs and LOSs are families of phosphorylated lipoglycans and glycolipids that are considered toxic with potent immunomodulating and immune stimulating properties.

Analysis of the LOS from different *C. jejuni* strains have shown lipid A to have an unusual mixed lipid A containing D-glucosamine (GlcN) and 2,3-diamino-2,3-dideoxy-D-glucose (GlcN3N) (Gilbert *et al*., 2008). The LOS outer core of many species is characterised by the presence of sugars such as D-glucose (Glc), D-galactose (Gal), *N*-acetyl-D-glucosamine (GlcNAc), and *N*-acetyl-D-galactosamine (GalNAc), whereas the inner core contains 2-keto-3-deosyoctulosonic acid (Kdo) and the heptose sugars, L-*glycero*-D-*manno*-heptose (L,D-Hep) and D-*glycero*-D-*manno*-heptose (D,D-Hep) (Rietschel *et al*., 1993). *C. jejuni* LOS inner cores contain the trisaccharide L,D-Hep-α(1,3)-L,D-hep-α(1,5)-Kdo (Moran, 1997). Another feature is that the heptose adjacent to Kdo (HepI) is substituted by D-glucose in a β(1→4) linkage, hence producing a common tetrasaccharide (Moran, 1997) (Figure 1.8). The genetic locus responsible for the biosynthesis of LOS in *C. jejuni* NCTC11168 is *Cj1131c - Cj1152c.*

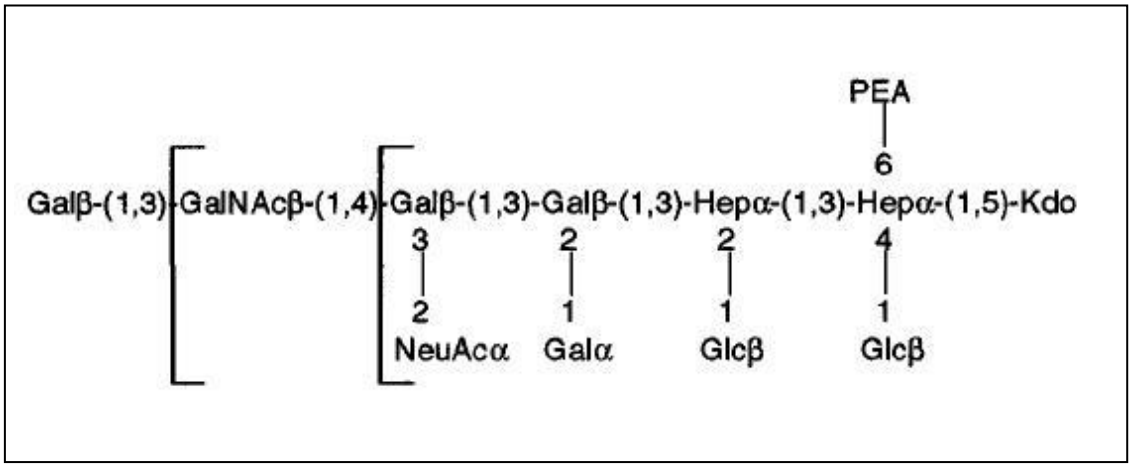



Figure 1.8. The proposed complete LOS core structure of *C. jejuni* NCTC11168. Kdo, 2-keto-3-deosyoctulosonic acid, Hep, L-*glycero*-D-*manno*-heptose. glc, Glucose. Gal, galactose. NeuAc, *N*-acetylneuraminic acid. GalNAc, *N*-acetyl-D-galactosamine. PEA, phosphoethanolamine. Image obtained from Gilbert *et al*. 2005.

Much of the interest in the structure and biosynthesis of the core LOS of *C. jejuni* resulted from the ganglioside-like structures in the outer core of *C. jejuni* LOSs that may be responsible for inducing autoimmune neurological disorders such as GBS and MFS. Outer core structures of many *C. jejuni* strains mimic the saccharide portion of certain human gangliosides. LOS locus classes A, B, C, M and R encode genes responsible for the production of sialylated LOS that are human ganglioside mimics, while LOS locus classes such as D, F and G lack a *cst* gene that encodes a sialic acid transferase (Parker *et al*., 2008). Different *C. jejuni* serotypes have shown varying LOS structures (Gilbert *et al*., 2008). *C. jejuni* HS:19 serostrain expresses a heterogeneous outer core composed of a mixture of GM1a and GD1a ganglioside mimics (Aspinall *et al*., 1994). In parallel, studies have shown that *C. jejuni* HS:19 isolates from GBS patients and enteritis patients produce LOS that contain terminal tetra- and pentasaccharide moieties identical to those of GM1a and GD1a gangliosides (Yuki *et al*., 1993, Moran and O'Malley, 1995). *C. jejuni* NCTC11168 (HS:2) exhibits GM2 and GM1a mimicry over tri- and tetrasaccharides (Oldfield *et al*., 2002, Prendergast and Moran, 2000, St Michael *et al*., 2002).

There exists great variation in LOS between *Campylobacter* strains and species. A number of genetic mechanisms exist to induce this diversity, namely genetic differences in gene content and organisation. Studies have shown that introduction of a complete new class of LOS biosynthesis loci can occur between strains by horizontal transfer (Phongsisay *et al*., 2006, Gilbert *et al*., 2004). There are at least four genetic mechanisms that affect glycosyltransferase activity and hence vary the LOS outer core structures; i) phase variation of certain genes due to homopolymeric tracts, ii) gene inactivation by deletion of insertion of single or multiple bases, but without phase variation, iii) amino acid substitution resulting in an inactive variant of the glycosyltransferase and, iv) single or multiple mutations leading to variant glycosyltransferases with different acceptor specificities. The huge variation in LOS structures generated by the mechanisms outlined above allow *C. jejuni* to modulate the structure of this cell-surface carbohydrate (Gilbert *et al*., 2008).

### 1.4.3 Capsular polysaccharide

The belated discovery that *C. jejuni* has a CPS was an important step in our understanding of *C. jejuni* pathogenesis (Karlyshev *et al*., 2005b). CPS consists of repeating oligosaccharide units attached to a phospholipid and is not chemically linked to LOS. The variability in CPS composition is the main determinant exploited by the Penner serotyping system (Penner and Hennessy, 1980, Moran and Penner, 1999). Great variety exists in the make-up of CPS due to the phospholipid chain containing different numbers of repeating sugar units (Karlyshev *et al*., 2005a). Diversity in CPS between *C. jejuni* strains can occur by multiple mechanisms such as exchange of capsular genes and gene clusters by horizontal transfer, gene duplication, phase, variation, deletion and fusion (Karlyshev *et al*., 2005a). CPS is believed to be an important factor protecting the bacterium against environmental conditions and also plays a role in pathogenesis (Karlyshev *et al*., 2008). CPS has been shown to be required for biofilm formation which contributes to virulence of the bacteria infecting the gastrointestinal tract and survival within environments that are not ideal. However, some studies have demonstrated that the presence of CPS may actually hinder the formation of biofilms in certain strains, and also mask bacterial adhesins required for interaction with host cells (Karlyshev *et al*., 2008). In addition, a recent study demonstrated that *C. jejuni* alters its surface polysaccharides when co-cultured with IECs suggesting the existence of cross talk mechanism(s) that modulate CPS expression during infection (Corcionivoschi *et al*., 2009). *C. jejuni* grown in conditioned medium taken from co-culture experiments after a single passage showed no loss of CPS. However, *C. jejuni* grown in conditioned medium taken from co-culture experiments after two rounds of passaging showed reduced CPS. In contrast, there was no change in CPS profile for bacteria grown in conditioned medium taken from HCT-8 cells alone or in medium conditioned by bacteria alone. These results suggested that a soluble factor present in the conditioned medium from HCT-8 cells co-cultured with *C. jejuni* was responsible for CPS alteration. The alteration in CPS appears to be dependent on a soluble factor that is both heat labile and proteinase K sensitive. CPS depletion occurred when *C. jejuni* organisms were exposed to conditioned media from a different *C. jejuni* strain, but not when exposed to conditioned media from other bacterial species. Thus, *C. jejuni* alters its surface polysaccharides when co-cultured with IECs. The genetic locus responsible for the biosynthesis of CPS in *C. jejuni* NCTC11168 is between *Cj1413c - Cj1448c*.

There are currently four different classifications of CPS (I-IV) (Whitfield, 2006, Whitfield and Roberts, 1999). These classifications are not based on antigenic or structural differences of the phospholipid chain, but on mechanisms of biosynthesis, assembly, genetic regulation, and sequence similarity (Whitfield, 2006, Whitfield and Roberts, 1999). *C. jejuni* CPSs fall in groups II and III which are typically characterised by organisation of the respective gene clusters containing a major internal biosynthetic region flanked by two groups of genes involved in CPS transport and assembly (*kps* genes). These CPSs contain a phospholipid moiety attached to a repeating unit either with or without a relatively labile linkage involving the sugar Kdo.

There remains unanswered questions as to the specific role of *C. jejuni* CPS structures in bacteria-host interactions and gastroenteritis development and there is a significant gap in our knowledge as to the regulation of CPS expression in *C. jejuni* (Karlyshev *et al*., 2008).

### 1.4.4 Glycosylation

Glycosylation of proteins is important for biological activity (Nothaft *et al*., 2008). Although it has been predicted that bacteria synthesize at least six-fold more sugar building blocks compared to eukaryotes, it was previously assumed prokaryotes were incapable of modifying proteins with sugars (Nothaft *et al*., 2008). This was due to lack of reagents and sensitive techniques capable of detecting these unusual sugars.

#### 1.4.4.1 *O*-linked glycosylation

Bacteria can attach sugars via the hydroxyl groups of Ser and Thr in an *O*-linkage (Thibault *et al*., 2001, Logan *et al*., 2002). *O*-linked glycosylation of *C. jejuni* appears to be a specific process where only the flagellin subunits of the flagella become glycosylated (Logan *et al*., 2008). *O*-linked glycosylation is integral to the flagella assembly process in *Campylobacter* and prevention of glycosylation leads to an inability to assemble the flagella filament leading to a non-motile phenotype (Logan *et al*., 2008). Unlike other prokaryotic glycosylation systems which occur in association with the cell membrane, flagellin monomers are secreted through the flagella apparatus and are therefore not found in the periplasm (Aas *et al*., 2007, Chaban *et al*., 2006). Hence, glycosylation takes place either in the cytoplasm, possibly in close proximity to the flagella machinery, or in the basal body compartment (Logan *et al*., 2008). There is no specific consensus sequence for the addition of *O*-linked glycosylated proteins on

flagellin (Logan *et al*., 2008). The *C. jejuni O*-linked glycosylation locus encompasses *Cj1293 - Cj1342c*.

Structural and metabolomics studies have elucidated the *O*-linked glycosylation pathway to produce Pse5Ac7Ac (pseudaminic acid) (Thibault *et al*., 2001, Logan *et al*., 2002, Schoenhofen *et al*., 2006a, Schoenhofen *et al*., 2006b). The pathway commences from UDP-GlcNAc and evolves to UDP-2 acetamido-2,6-dideoxy-β-L-*arabino*-hexos-4-ulose via PseB. PseC further converts this to UDP-4-amino-4,6-dideoxy-β-L-AltNAc. PseH converts this to UDP-2,4-diacetamido-2,3,6-trideoxy-β-L-altropyranose. PseG acts as a sugar hydrolase resulting in the release of 2,4-diacetamido-2,4,6-trideoxy-L-altro-pyranose. This is then converted by PseI to the nine-carbon structure Pse5Ac7Ac (pseudaminic acid). PseF allows CMP activation of Pse5Ac7Ac (Logan *et al*., 2008) (Figure 1.9).

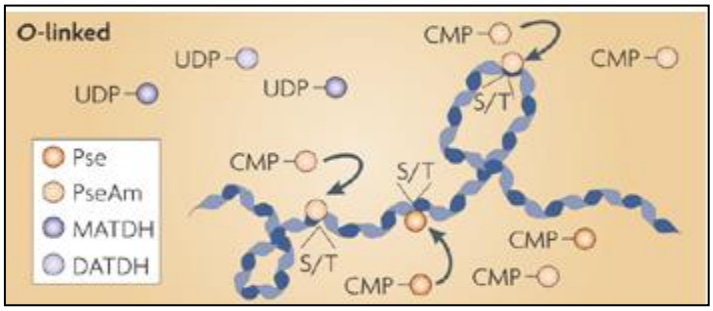



Figure 1.9. *O*-linked glycosylation of flagellin monomers is proposed to occur in the cytoplasmic inner membrane where nucleotide-activated sugars are individually added to serine or threonine residues that are surface exposed. Pse, 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acid ($Pse_5Ac_7Ac$); PseAm, 5-acetamidino analogue of Pse ($Pse_5Am_7Ac$); MATDH, monoacetamido-trideoxyhexose; NeuAc, N-acetylneuraminic acid; DATDH, diacetamido-trideoxyhexose; UDP, uridine diphosphate; CMP, cytidine monophosphate. Image obtained from Young *et al*. 2007.

The requirement that flagellin proteins are glycosylated for filament assembly illustrates the biological significance of glycosylation (Logan *et al*., 2008). The function of modifications to flagella structure by *O*-linked glycosylation is still not fully understood. Mutation of genes involved in *O*-linked glycosylation generally do not result in loss of

motility (Logan, 2006). *C. jejuni* 81-176 has been shown to require glycosylation to assemble a flagella filament (Ewing *et al*., 2009, Goon *et al*., 2003). However *O*-linked glycosylation can occur in the absence of a functional flagella (Ewing *et al*., 2009). Loss of flagella and changes in sugars attached to flagellin have been shown to affect autoagglutination in *C. jejuni* (Ewing *et al*., 2009).

#### 1.4.4.2 *N*-linked glycosylation

*C. jejuni* was the first bacterium identified to be capable of attaching sugars to proteins via the amide group of Asn in an *N*-linkage (Szymanski *et al*., 1999). The identification of *N*-linked glycosylation within bacteria was a breakthrough in glycobiology as it was initially considered that this glycosylation pathway was only present in eukaryotes. The field of functional glycomics has received a large amount of interest due to the fact that in eukaryotes, up to 90% of all proteins are modified with sugars. Disruption of these sugar pathways has been implicated in multiple disease states ranging from congenital disorders of glycosylation to autoimmune disease and cancer (Nothaft *et al*., 2008). The *N*-linked glycosylation system in *C. jejuni* has been exploited as an ubiquitous toolbox for understanding the pathway, developing new techniques for glycobiology and to exploit novel diagnostics and therapeutics strategies (Nothaft *et al*., 2008, Karlyshev *et al*., 2004, Langdon *et al*., 2009).

Disruption of the *N*-glycosylation system within *C. jejuni* is believed to have pleiotropic effects affecting iron transport (Palyada *et al*., 2004), amino acid transport (Nothaft *et al*., 2008), induction of the host-immune response (Nothaft *et al*., 2008), chicken colonisation (Jones *et al*., 2004, Szymanski and Wren, 2005, Kelly *et al*., 2006), adhesion/invasion (Szymanski *et al*., 2002), mouse colonisation (Szymanski *et al*., 2002) and even the *O*-linked glycosylation pathway (Schoenhofen *et al*., 2006b). Previously, it has been shown that glycans play an important role in immunity during infection by masking the primary amino acid epitopes (Szymanski *et al*., 1999). *N*-linked glycosylation also plays an important role in *C. jejuni*-host cell interactions where *C. jejuni* 81-176 mutants in the *pgl* pathway showed reduced levels of adherence to and invasion of INT 407 cells and colonisation of the intestinal tracts of mice (Szymanski *et al*., 2002, Karlyshev *et al*., 2004). A number of glycoproteins have been identified with a virulence related role e.g. PEB3, VirB10 and *Cj1496c* (Young *et al*., 2002, Kowarik *et al*., 2006, Nothaft *et al*., 2008, Scott *et al*., 2010). For example, *Cj1496c* has been shown

to be required for efficient cell invasion and chick gastrointestinal colonisation (Kakuda and DiRita, 2006).

The process of *N*-linked glycosylation starts with the synthesis of a heptasaccharide as the final product on the lipid bactoprenylpyrophosphate at the cytoplasmic side of the inner membrane. Initially UDP-GlcNAc is converted to UDP-bacillosamine by PglD, PglE and PglF proteins. The glycosyltransferases PglC, PglA, PglI, PglH and PglJ are involved in the formation of the heptasaccharide (Figure 1.10). The PglK flippase allows the translocation of the lipid-linked heptasaccharide across the inner membrane. The oligosaccharyltransferase PglB catalyses the transfer of the heptasaccharide from the lipid carrier to selected asparagine residues on nascent polypeptide chains. *N*-linked protein glycosylation takes place in the periplasmic space with the PglB requiring the extended consensus sequence Asp/Glu-X-Asn-Z-Ser/Thr (D/E-X-N-Z-S/T; X, Z ≠ P) (Kowarik *et al*., 2006) for transfer of oligosaccharides to protein. The *C. jejuni N*-linked glycosylation locus is between *Cj1119c* - *Cj1130c*.

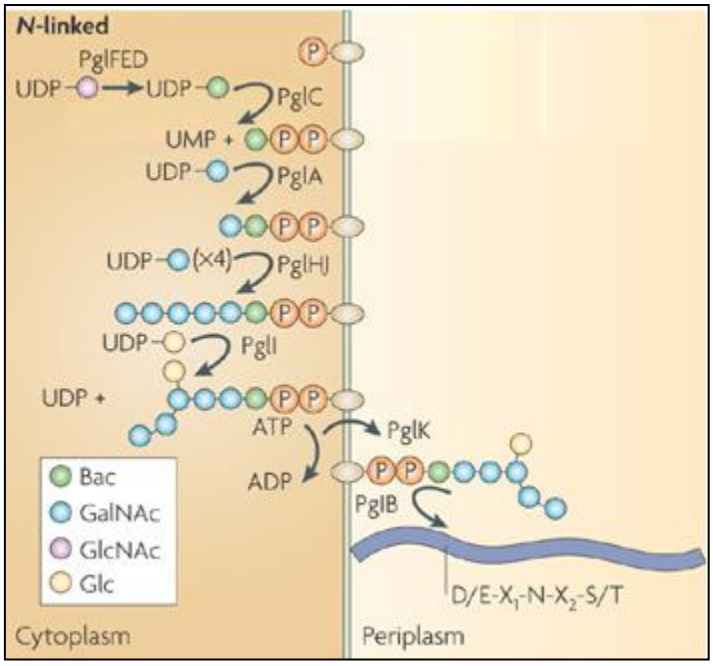



Figure 1.10. *N*-linked glycosylation is proposed to proceed through the sequential edition of nucleotide-activated sugars onto a lipid carrier, resulting in the formation of a branched heptasaccharide. The glycan is then 'flipped' across the inner membrane into the periplasm by a ATP-binding cassette (ABC) transporter. Bac, 2,4-diacetamido-2,4,6-

trideoxyglucose; GalNAc, *N*-acetylgalactosamine; HexNAc, *N*-acetylhexosamine; Glc, glucose; UDP, uridine diphosphate. Image obtained from Young *et al*. 2007.

Recently, studies have demonstrated that *C. jejuni* also produces free heptasaccharides derived from the *N*-glycan pathway reminiscent of the free oligosaccharides (fOS) produced by eukaryotes (Nothaft *et al*., 2009). It has been shown that *C. jejuni* produces fOS in response to changes in the osmolarity of the environment and bacterial growth phase (Nothaft *et al*., 2009). This study showed the conserved WWDYG motif of the oligosaccharyltransferase PglB is necessary for fOS release into the periplasm and the work demonstrates that fOS from an *N*-glycosylation pathway in bacteria are potentially equivalent to osmoregulated periplasmic glucans in other Gram-negative bacteria (Nothaft *et al*., 2009, Nothaft *et al*., 2010). Thus, *C. jejuni* heptasaccharides play a dual role in the periplasm being involved in the post-translational modification of proteins as well as potentially existing in the free with a role in osmotic control (Nothaft *et al*., 2010).

### 1.4.5 Cytolethal distending toxin

*C. jejuni* has been reported to produce CDT encoded by the genes *cdtA*, *cdtB* and *cdtC* (Whitehouse *et al*., 1998). Not all strains of *C. jejuni* contain the complete *cdt* operon (Engberg *et al*., 2005). Studies have shown CDT to irreversibly block eukaryotic cells in the G2 phase of the cell cycle (Whitehouse *et al*., 1998, Lara-Tejero and Galan, 2001). CDT has also been shown to cause progressive cellular distension, chromatic fragmentation and apoptotic cell death (Frisan *et al*., 2002, Lara-Tejero and Galan, 2001). The active subunit CdtB, a type I deoxyribonuclease, enters the nucleus resulting in DNA double-strand breaks causing cell arrest (Lara-Tejero and Galan, 2001). Studies using *E. coli* have shown the eukaryotic cell responds to the DNA double-strand breaks by initiating a regulatory cascade that results in cell cycle arrest, cellular distension and cell death (Smith and Bayles, 2006, Bielaszewska *et al*., 2005). CDT contributes to pathogenesis by inhibiting both cellular and humoral immunity via apoptosis of immune response cells and also by generating necrosis of epithelial-type cells and fibroblasts involved in the repair of lesions produced by pathogens (Smith and Bayles, 2006, Bielaszewska *et al*., 2005). This results in slow healing and production of disease symptoms (Smith and Bayles, 2006, Bielaszewska *et al*., 2005). *C. jejuni* co-cultured with INT 407 IECs has been shown to induce IL-8 secretion (Hickey *et al*., 2000), however CDT has not been shown to play a role in colonisation of chickens (Hu and

Kopecko, 2008). *cdtB* mutants have been shown to exhibit impaired ability to cause disease in immunodeficient mice (Fox *et al*., 2004, Purdy *et al*., 2000). Studies have also shown *cdt* mutants to have normal levels of adhesion to HeLa and HD-11 cells, but to have a 10- fold decrease in the level of invasion (Biswas *et al*., 2006). The role of cytotoxin or enterotoxin during *C. jejuni* infection in initiating inflammation or diarrhoea remains unclear (Florin and Antillon, 1992, Ketley, 1997, Wassenaar and Blaser, 1999). It has been hypothesised that CDT may promote epithelial cell death at a quicker rate and thus allow further *C. jejuni* to translocate across the mucosa (Hu and Kopecko, 2008). Alternatively, diarrhoeal symptoms may result from *C. jejuni* induced epithelial cell death or translocation across the epithelium which has been noted to loosening of tight junctions leading to compromised mucosal barrier function.

### 1.4.6 Adhesins

One of the major adhesins of *C. jejuni* to IECs is the flagella (Hu *et al*., 2008). Disruption of flagella functionality results in a reduction of *C. jejuni* binding to host cells (Konkel and Joens, 1989, Konkel *et al*., 2004, Carrillo *et al*., 2004). Major adhesins involved in *C. jejuni* adhering to IECs include PEB1, a conserved antigen proposed to be an adhesin (Pei and Blaser, 1993, Pei *et al*., 1991). PEB1 shares homology with a periplasmic binding protein involved in nutrient acquisition (Pei and Blaser, 1993, Garvis *et al*., 1996). A *peb1A* mutant demonstrates a 50- to 100- fold reduction in adherence to HeLa cells (Pei and Blaser, 1993, Pei *et al*., 1998). *C. jejuni* has been noted to bind to components of the extracellular matrix such as fibronectin, laminin, vitronectin and collagen (Hu and Kopecko, 2008). CadF (*Campylobacter* adhesion to fibronectin) has been identified as an adhesin to fibronectin and is conserved among *C. jejuni* strains (Konkel *et al*., 1999a). Studies have shown that anti-CadF antibodies have reduced the binding of *C. jejuni* clinical isolates to immobilised fibronectin by greater than 50% (Monteville *et al*., 2003). Another adhesin identified for *C. jejuni* was the surface-exposed lipoprotein JlpA (Jin *et al*., 2001) which acts by binding to Hsp90-alpha on the surface of Hep-2 epithelial cells (Jin *et al*., 2003). Both LOS and CPS have important roles to play in adherence. It has also been reported that *N*-linked glycosylation proteins on the cell surface have a role to play in adhesion and even invasion of Caco-2 cells (Karlyshev *et al*., 2004, Hu and Kopecko, 2008).

### 1.5 Host-pathogen interactions

*C. jejuni* is a pathogen that has established a close association with a variety of vertebrate hosts and thus has evolved specific adaptations to modulate cellular functions (Watson and Galán, 2008). Studies have shown that *C. jejuni* can gain intracellular access to nonphagocytic IECs (Konkel and Joens, 1989, Ketley, 1997, De Melo *et al*., 1989). In humans, *C. jejuni* typically colonises the ileum and colon, where bacteria interfere with the absorptive capacity of the intestine (MacCallum *et al.*, 2005b). *C. jejuni* has evolved mechanisms such as attachment to and invasion of IECs to combat the defensive mechanisms posed by the host such as the mucous layer, epithelial barrier and innate immune response. *C. jejuni*, like other pathogenic bacteria, adhere to colonic epithelial cells and trigger signal transduction events that induce host cytoskeletal rearrangements, bacterial internalisation and translocation across IECs (Hu and Kopecko, 2008). During mucosal penetration, different bacteria-host cell interactions induce interleukin and chemokine production, the recruitment of inflammatory cells, host cell death and cause diarrhoea, colitis, or both (Hu and Kopecko, 2008). *C. jejuni* does not contain homologues of classical bacterial enterotoxins, adhesins, invasins or type III protein secretion systems or even pathogenicity islands (Fouts *et al*., 2005, Parkhill *et al*., 2000). The presence of a type IV secretion system has been noted in certain *C. jejuni* strains; in particular on the pVIR plasmid of *C. jejuni* 81-176 (Bacon *et al*., 2000, Bacon *et al*., 2002), however the role in disease is uncertain. A recent study identified pVIR in 17 of 104 (16.3%) clinical *C. jejuni* isolates and found that these particular isolates were associated with the presence of a tetracycline resistance plasmid (Tracz *et al*., 2005).

#### 1.5.1 *C. jejuni* adhesion to host intestinal epithelial cells

*C. jejuni* adherence, as in other bacteria, is typically due to specific interactions between molecules on the surface (adhesins) and molecules on the host surface (receptors). Putative adhesins for *C. jejuni* attachment to IECs have been described in Section 1.4.6. The mechanisms of binding and attachment are still largely uncharacterised, though it is clear that attachment is a prerequisite to adhesion (Hu and Kopecko, 2008).

#### 1.5.2 *C. jejuni* invasion of host intestinal epithelial cells

*C. jejuni* transcytose host IECs to emerge in the lamina propria resulting in an inflammatory response (Everest *et al*., 1992, Konkel *et al*., 1992b, Harvey *et al*., 1999, MacCallum *et al*., 2005a) (Figure 1.11). The two main mechanisms by which *C. jejuni* passes through the impermeable epithelial layer is via transcellular methods whereby

bacteria pass through the absorptive enterocytes or M cells, or paracellular where bacteria pass between adjacent epithelial cells (Hu and Kopecko, 2008). Polarized cell lines such as Caco-2 (human colonic carcinoma) have provided a simple and alternative method to study host-pathogen interactions without using animal models (Hu and Kopecko, 2008). *C. jejuni* invasion into the epithelial mucosa is an essential process to forming colitis (Allos, 1997, Russell and Blake, 1994). *C. jejuni* interacts with the intestinal mucosal surface, triggering host signal transduction events that lead to host cytoskeletal rearrangements, resulting in bacterial uptake (Hu and Kopecko, 2008). There remains confusion as to the exact mechanism for *Campylobacter* entering and crossing the intestinal mucosa. Bacterial internalisation into IECs has typically been linked to involve pathogen-induced rearrangement of host cytoskeletal structures, resulting in endocytosis of the pathogen (Hu and Kopecko, 2008). These structures comprise microfilaments (MFs) and microtubules (MTs), which are made up of actin and tubulin respectively. The structures are involved in both cellular and subcellular movements and along with intermediate filaments help determine cell shape (Hu and Kopecko, 2008). Internalisation of *C. jejuni* has been reported to be MF-dependent (De Melo *et al*., 1989, Konkel and Joens, 1989), MT-dependent (Hu and Kopecko, 1999, Oelschlaeger *et al*., 1993), both MF- and MT-dependent (Biswas *et al*., 2003, Monteville *et al*., 2003), or neither (Russell and Blake, 1994). There is evidence that suggests *C. jejuni* may encode separate MF-dependent (Konkel *et al*., 1992a, Konkel and Joens, 1989) and MT-dependent (Hu and Kopecko, 1999, Oelschlaeger *et al*., 1993) pathways for host invasion. A number of explanations have been put forward as to why such differences in host cell cytoskeletal requirements for *C. jejuni* invasion occur. These including experimental techniques, cell lines used and strains used (Hu and Kopecko, 2008). It is believed that the MT-dependent pathway is observed at the apical host cell surface, whereas MF-dependent invasion occurs mainly at the basolateral surface (Hu and Kopecko, 2008).

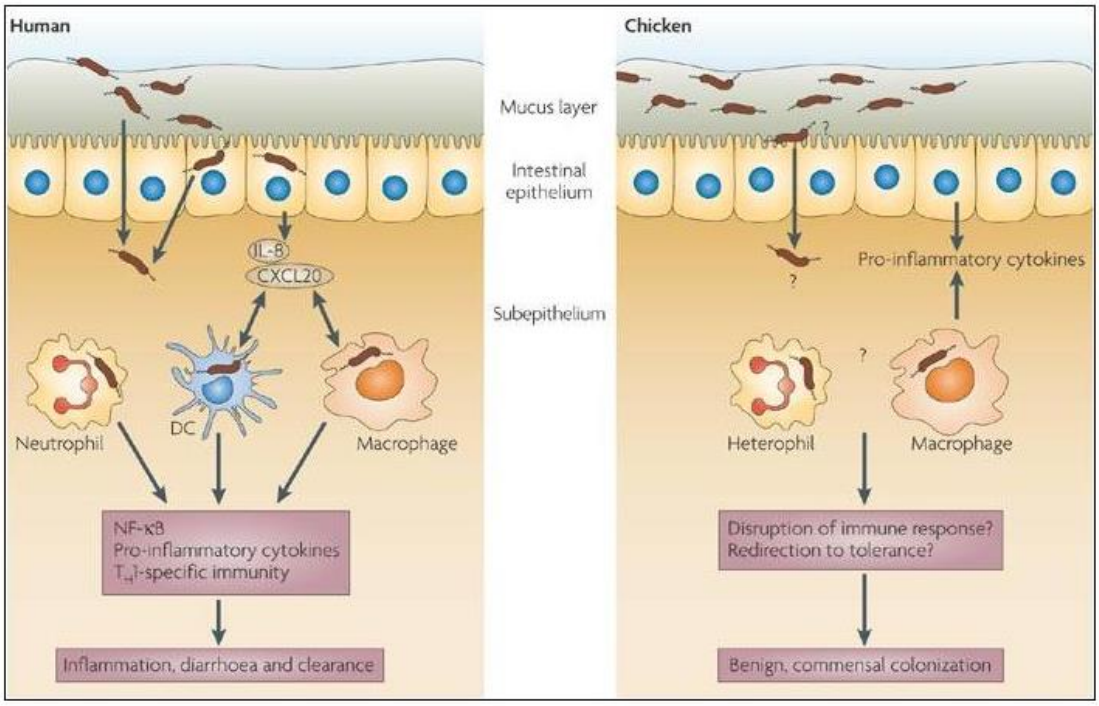



Figure 1.11. *C. jejuni* evades the mucus layer in humans and interacts with the IECs causing IL-8 production. *C. jejuni* binds to, and is internalised by epithelial cells. This leads to interaction with dendritic cells, macrophages and neutrophils, which further leads to a pro-inflammatory response and increase in cytokines production. *C. jejuni* resides primarily in the mucosal layer in chicken intestines. *In vitro* studies have shown that *C. jejuni* can stimulate the production of IL-1, IL-6 and intracellular nitric oxide synthase from IECs and macrophages, but importantly the resulting host response does not typically lead to inflammatory diarrhoea in chickens. Image obtained from Young *et al*. 2007.

Invasion and translocation are important steps in the survival and persistence of *C. jejuni*. An important factor for *C. jejuni* invasion is the disruption of tight junctions as a route for translocation, which was shown using Caco-2 cell monolayers (MacCallum *et al*., 2005b). Polarized Caco-2 cells with differentiated apical and basolateral surfaces are separated by tight junctions, express several markers characteristic of normal small intestinal cells and have a well-defined brush border (Everest *et al*., 1992, Finlay and Falkow, 1990). Studies have demonstrated that *C. jejuni* translocate tight polarized epithelial monolayers (Bras and Ketley, 1999, Everest *et al*., 1992, Grant *et al*., 1993). A putative paracellular pathway of *C. jejuni* mucosal translocation termed subvasion has also been described (van Alphen *et al*., 2008). The findings suggest that subvasion does

not occur in polarized IECs, and is only observed in cultured IECs maintained under nutrient-limiting conditions, raising the relevance of this process is for *C. jejuni* pathogenesis (Hu and Kopecko, 2008). In addition, infection of polarised Caco-2 cells with high multiplicities of infection (MOI 10,000:1) were found to cause loss of transepithelial electrical resistance (TEER) by 24 h after infection (MacCallum *et al*., 2005b). This loss of TEER was accompanied by a rearrangement of the tight junction protein occludin (MacCallum *et al*., 2005b). Loss of tight junctions, electrolyte and fluid absorption are likely to be compromised and this may likely contribute to the clinical manifestations of diarrhoea (MacCallum *et al*., 2005b)

In addition to host cytoskeletal requirements during *C. jejuni* adhesion and invasion, signal transduction events are also important where pathogens modulate host signalling systems both to allow for invasion and to trigger disease pathogenesis (Hu and Kopecko, 2008). The divalent calcium cation $Ca^{2+}$ plays a role in host signal transduction and other cellular processes (Clapham, 1995, Jacob, 1990, Marks and Maxfield, 1990a, Marks and Maxfield, 1990b). Increased free intracellular $Ca^{2+}$ has been demonstrated to link cell surface receptor stimulation via signalling pathways with intracellular effectors and to modulate cytoskeletal structure, chemotaxis, membrane fluidity, chromosome segregation, cell cycle transition, enzyme activity, transmembrane ion fluxes, proteolysis and other functions (Clapham, 1995, Jacob, 1990, Marks and Maxfield, 1990a, Marks and Maxfield, 1990b). Although many signalling events have been found to affect *C. jejuni* invasion there still remains gaps in our knowledge of the signal transduction pathway involvement in bacterial internalisation (Hu and Kopecko, 2008).

A key bacterial factor for invasion is motility as non-motile mutants are non-invasive (Watson and Galán, 2008, Yao *et al*., 1994). Non-flagella proteins secreted through the flagella (e.g. Cia proteins) have been shown to trigger internalisation (Konkel *et al*., 2004). However, this hypothesis is controversial, as mutants defective in motility but with an intact flagella structure should still be invasive as secretion is possible. However, these mutants were non-invasive (Yao *et al*., 1994). Mutations affecting protein glycosylation or capsular synthesis also affect bacterial entry, thus there exists mixed data as to what is exactly required for invasion (Bachtiar *et al*., 2007, Bacon *et al*., 2001, Guerry *et al*., 2006, Kakuda and DiRita, 2006, Karlyshev *et al*., 2004, Szymanski *et al*., 2002, Vijayakumar *et al*., 2006).

### 1.5.3 *C. jejuni* internalisation and survival

An important factor involved in *C. jejuni* internalisation are lipid rafts or caveolae which are microdomains within plasma membranes enriched for cholesterol, glycolipids, sphingolipids and signalling molecules such as receptor tyrosine kinases (Hu *et al*., 2006, Wooldridge *et al*., 1996). *C. jejuni* are believed to exploit these areas to gain entry into the cell and importantly avoid delivery to lysosomes (Watson and Galan, 2008). It is therefore possible that in order to survive within cells, *C. jejuni* has evolved specific adaptations to survive within lysosomes or to modulate host cellular trafficking events to avoid fusion with lysosomes (Watson and Galan, 2008). Once internalised into IECs, *C. jejuni* is believed to reside within a vacuole (Campylobacter Containing Vacuole - CCV). Characterisation of this compartment indicates that it is functionally distinct from lysosomes (Watson and Galan, 2008). Studies have shown that the CCV acquires lipid raft or caveolae markers at early time points during infection and that a functional caveolin-1 is required for efficient *C. jejuni* entry (Watson and Galan, 2008). Even though *C. jejuni* internalisation into IECs requires caveolae, bacterial internalisation has been shown to not require dynamin, which is an essential component of the endocytic machinery associated with caveolae (Watson and Galán, 2008). Thus, it is possible that caveolae may be required for the assembly of signalling molecules used for the signal transduction pathways that leads to *C. jejuni* entry rather than to allow bacterial entry through its associated endocytic machinery (Watson and Galán, 2008). The CCV is not part of the canonical endocytic pathway and studies have demonstrated *C. jejuni* does not survive within lysosomes (Watson and Galan, 2008). *C. jejuni* has been identified in macrophages, but long term survival is not feasible within this cell type due to oxidative stress (Watson and Galan, 2008, Sikic Pogacar *et al*., 2009). There has also been some data demonstrating the potential importance of cellular processes in *C. jejuni* internalisation e.g. tyrosine kinase inhibitors have been noted to inhibit bacterial internalisation into cultured IECs (Woolcock *et al*., 1996, Biswas *et al*., 2003). The identity of the tyrosine kinase or kinases required for bacterial entry or the specific role that they may play in the entry process is currently unknown (Watson and Galán, 2008).

*C. jejuni* utilises a number of different mechanisms to survive and replicate within host cells. In general, intracellular *C. jejuni* lose viability within IECs over 24 hours, with no evidence of replication (Day *et al*., 2000, Candon *et al*., 2007, Konkel *et al*., 1992a). Recent studies have demonstrated not only that intracellular *C. jejuni* remains viable for up to 24 hours, but also that it shifts to a physiological state where it can only be initially

cultured under anaerobic conditions (Watson and Galan, 2008). This is believed to be a mechanism of survivability in low oxygen environments.

It is noteworthy that not only different species, but different strains of *C. jejuni* have varying abilities to adhere, invade and survive within IECs. Even though a great deal has been elucidated in the past decade, there are still major gaps in our knowledge of the cellular mechanisms that lead to internalisation (Hu and Kopecko, 2008).

### 1.6 *C. jejuni* genome sequences

Given the socioeconomic importance of this pathogen, it is surprising that the ecology, epidemiology and, in particular, the pathogenesis are still so poorly understood. The lack of information on this problematic pathogen was the driving force for the original *C. jejuni* NCTC11168 genome sequencing project (Parkhill *et al*., 2000). *C. jejuni* NCTC11168 was selected for sequencing as it is genetically tractable and widely available. This strain was isolated in 1977 from a U.K. patient with severe gastroenteritis (Skirrow, 1977). The sequencing process involved the construction of seven libraries in pUC18 using size fractions ranging from 1.0 kb to 2.2 kb (Parkhill *et al*., 2000). Approximately 19,400 pUC clones were sequenced using Dye-terminator chemistry on ABI 373 and 377 sequencing machines. The sequencing project was initiated in 1997 and completed in 1999 (Parkhill *et al*., 2000). Initial annotation was performed at the WTSI and the genome was published in 2000 (Parkhill *et al*., 2000).

#### 1.6.1 Key features of *C. jejuni* genome sequences

Key features identified in the original annotation of the *C. jejuni* NCTC11168 genome were CPS (surface polysaccharide at the time), flagella modification and LOS loci (Figure 1.12). There was a notable lack of insertion or phage-associated sequences and very few repeat sequences within the genome. Hypervariable sequences allowing phase variation were identified in genes encoding proteins involved in the biosynthesis or modification of surface structures (Parkhill *et al*., 2000). Initially these were considered sequencing errors, but closer inspection revealed that these were genuine and were on surface polysaccharide structures that had been previously shown to phase vary (De Bolle *et al*., 2000, Saunders *et al*., 1998). A number of known and potential pathogenicity factors were identified including genes encoding for CDT, haemolysin-like toxins, putative type II export and chemotaxis genes. The *C. jejuni* genome was shown to

be highly compact (1654 CDSs representing 94.3% of the genome), with a relatively low G/C percentage at 30.55%.

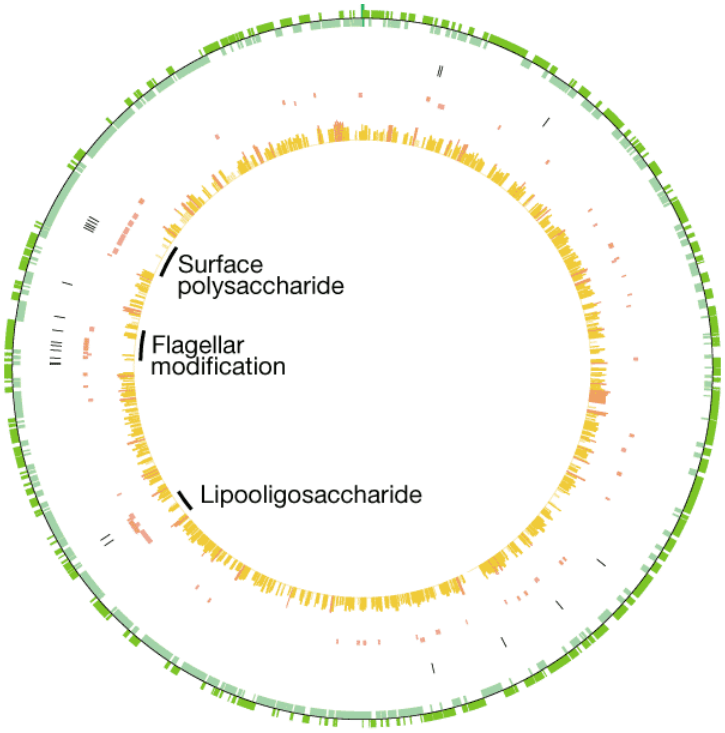



Figure 1.12. Schematic diagram of the *C. jejuni* NCTC11168 genome. Green circular rings show CDSs transcribed in forward (dark green) and reverse (light green) orientation. Black lines indicate locations of hypervariable sequences. Pink lines indicate locations of surface structures. The innermost histogram shows the similarity of each CDS to its *H. pylori* orthologue where present; height of the bar and the intensity of the colour are proportional to the degree of similarity. Figure reproduced from NCTC11168 original genome sequence project where 'surface polysaccharide' was later noted as CPS. Image obtained from Parkhill *et al*. 2000.

Since this work was published in 2000, there has been continued interest in this important human pathogen. A consequence of this has been significant revisions of the genetic loci that code for important surface structures on *C. jejuni* strains. The surface polysaccharide region (*Cj1413c – Cj1448c*) has been renamed as the capsule locus (Karlyshev *et al*., 2005a, Karlyshev *et al*., 2000, Karlyshev *et al*., 2001). The flagella modification locus (*Cj1293 – Cj1342c*) has been confirmed as an *O*-linked glycosylation

pathway which is responsible for glycosylation of the flagellin structural proteins FlaA and FlaB (Thibault *et al*., 2001, Szymanski *et al*., 2003a, Liu and Tanner, 2006, Karlyshev *et al*., 2002). Progress has also been made in our understanding of the LOS locus (*Cj1131c* - *Cj1152c*) (Guerry *et al*., 2000, Valvano *et al*., 2002, Gilbert *et al*., 2002, Linton *et al*., 2000). In addition, the *N*-linked glycosylation pathway (*Cj1119c* – *Cj1130c*) was identified before the original genome annotation (though was not mentioned) and this locus was further characterised in due course (Linton *et al*., 2005, Glover *et al*., 2005, Kelly *et al*., 2006, Szymanski *et al*., 2003a). To date, up to 81 *C. jejuni* proteins modified with the same *N*-linked heptasaccharide glycan structure have been identified (Schoenhofen *et al*., 2006b, Nothaft *et al*., 2008, Scott *et al*., 2010).

### 1.6.2 Further *C. jejuni* genome sequences

Further strains and species isolated from different environmental sources have since been sequenced (Tables 1.1 and 1.2). These include *C. jejuni* RM1221, isolated in 2000 from a chicken carcass in the U.S.A (Miller *et al*., 2000), *C. jejuni* 81-176, a highly pathogenic clinical strain originally isolated from a patient during an outbreak of *C. jejuni* campylobacteriosis (Korlath *et al*., 1985) isolated from a milk outbreak in the US (Hofreuter *et al*., 2006), a genetically stable *C. jejuni* 81116 strain originally isolated from a case of campylobacteriosis associated with a human waterborne outbreak (Pearson *et al*., 2007) and CG8486 which was isolated from a patient with inflammatory diarrhoea in Thailand (Poly *et al*., 2007). In addition, a number of different *Campylobacter* species have been sequenced and annotated such as *C. lari*, *C. upsaliensis* and *C. coli* (Fouts *et al*., 2005).

Table 1.1. *C. jejuni* sequencing projects at July 2010 with sequence characteristics.

| ***C. jejuni* strain** | **Source [a]** | **State of genome sequence project** | **Genome size (Mb)[b]** | **GC content (%)** | **Number of plasmids** | **GenBank Accession number** | **Sequencing centre [c]** | **Publication** |
|---|---|---|---|---|---|---|---|---|
| *C. jejuni subsp. jejuni* NCTC11168 | Clinical isolate / faeces of a diarrheic patient in 1977 | Completed | 1.60 | 30.6 | 0 | AL111168.1 | Sanger | Parkhill *et al*., 2000 |
| *Campylobacter jejuni* RM1221 | Skin of a retail chicken in 2000 | Completed | 1.80 | 30.3 | 0 | CP000025.1 | JCVI | Fouts *et al*. 2005 |
| *Campylobacter jejuni* subsp. *jejuni* 81116 (NCTC11828) | Clinical isolate / waterborne outbreak in 1983 | Completed | 1.60 | 30.3 | 0 | CP000814.1 | BBSRC IFR | Pearson *et al*., 2007 |
| *Campylobacter jejuni* subsp. *jejuni* 81-176 | Clinical isolate / Campylobacteriosis Outbreak in 1985 | Completed | 1.68 | 30.5 | 2 | CP000538.1 | JCVI | Hofreuter *et al*., 2006 |
| *Campylobacter jejuni* subsp. *doylei* 269.97 | Isolated from human bacteraemia | Completed | 1.8 | 30.6 | 0 | CP000768 | TIGR | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* IA3902 | Causes abortion in sheep | Completed | 1.64 | 30.5 | 0 | CP001876 | Iowa State University | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* CG8486 | Clinical isolate / faeces of a diarrheic patient in 1999 | In progress | ** | 30.4 | ** | AASY00000000 | NMRC | Poly *et al*., 2007 |
| *Campylobacter jejuni* subsp. *jejuni* 260.94 | Clinical isolate (GBS associated) | In progress | ** | 30.5 | ** | AANK00000000 | JCVI | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* 84-25 | Child cerebrospinal fluid | In progress | ** | 30.4 | ** | AANT00000000 | JCVI | Unpublished |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Campylobacter jejuni* subsp. *jejuni* CF93-6 | Clinical isolate (MFS) | In progress | ** | 30.5 | ** | AANJ00000000 | JCVI | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* HB93-13 | Child faeces (GBS) | In progress | ** | 30.6 | ** | AANQ00000000 | JCVI | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* 1336 | ** | In progress | 1.70 | 29.7 | ** | ADGL00000000 | University of Liverpool | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* 305 | Turkey isolate | In progress | ** | ** | ** | ** | University of Copenhagen | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* 327 | Turkey isolate | In progress | ** | ** | ** | ** | University of Copenhagen | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* 414 | ** | In progress | 1.70 | 29.6 | ** | ADGM00000000 | University of Liverpool | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* BH-01-0142 | Clinical isolate from patient with gastroenteritis | In progress | ** | 27.7 | ** | ABKD00000000 | Naval Medical Research Center | Poly, F *et al*, 2008 |
| *Campylobacter jejuni* subsp. *jejuni* CG8421 | Clinical isolate from patient with gastroenteritis | In progress | ** | 30.4 | ** | ABGQ00000000 | Naval Medical Research Center | Poly, F *et al*, 2008 |
| *Campylobacter jejuni* subsp. *jejuni* DFVF1099 | ** | In progress | ** | ** | ** | ** | University of Copenhagen | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* ICDCCJ07001 | ** | In progress | ** | ** | ** | ** | Dept. of Diagnosis for Infectious Disease | Unpublished |
| *Campylobacter jejuni* subsp. *jejuni* M1 | ** | In progress | ** | ** | ** | ** | Danish Technical University | Unpublished |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Campylobacter jejuni* subsp. *jejuni* S3 | Poultry isolate | In progress | 1.70 | ** | ** | ** | University of Arizona | Unpublished |

Data obtained from NCBI website (http://www.ncbi.nlm.nih.gov/).** Denotes information not yet available. [a] Source abbreviations: GBS, Guillain Barré syndrome; MFS, Miller Fisher syndrome. [b] Denotes genome is estimated. [c] Institute abbreviations: JCVI, J. Craig Venter Institute; BBSRC IFR. BBSRC Institute of Food Research; NMRC, Naval Medical Research Centre.

Table 1.2. Features of five selected *C. jejuni* genomes.

| **Trait** | **Strains** | | | | |
|---|---|---|---|---|---|
| Strain | NCTC11168 | RM1221 | 81-176 | 81116 | CG8486 |
| Serotype | HS:2 | HS:53 | HS:23/36 | HS:6 | HS:4 |
| MLST[a] | ST-21 (43) | ST-354 (354) | ST-42 (913) | ST-283 (267) | unknown |
| Origin | Clinical | Chicken | Clinical | Clinical | Clinical |
| Genome size (Mb) [b] | 1.64 | 1.77 | 1.62 | 1.63 | 1.60 |
| GC content (%) | 30.55 | 30.31 | 30.62 | 30.54 | 30.43 |
| Predicted CDS numbers | 1654 (1643)[e] | 1835 | 1568 | 1626 | 1588 |
| Pseudogenes | 20(19)[e] | 47 | 0[c] | 1[c] | 3[c] |
| Poly G/C tracts[d] | 29 (22) | 25 (8) | 19 | 17 | 23 |
| Plasmids | 0 | 0 | 2 | 0 | 0 |
| References | (Parkhill *et al*., 2000) | (Fouts *et al*., 2005) | (Hofreuter *et al*., 2006) | (Pearson *et al*., 2007) | (Poly *et al*., 2007) |

Legend. [a] ST represents clonal complex. () indicates sequence type. [b] indicates genome size made by approximation. CG8486 genome sequence is currently in 19 contigs. [c] indicates approximate number of pseudogenes. [d] Poly G/C tracts represent total found. () indicate tracts greater than seven or more nucleotides in length and have been shown to vary during sequencing project. [e] () indicates number after NCTC11168 re-annotation.

The first *Campylobacter* genome sequence was completed in 2000 and allowed researchers to partially identify the genetic makeup of this microaerophilic organism. This genomic data provided renewed impetus for *Campylobacter* research. Multiple

strains and species have since been sequenced and have led to greater insights into the organism. However, somewhat surprisingly, the ecology, epidemiology and in particular, the pathogenesis of *C. jejuni* are still poorly understood. Even a basic question, such as how this microaerophilic organism can persist in the natural environment yet for many years prove difficult to grow in laboratories, remains unanswered. Genome sequence data from all sequenced *Campylobacter* strains have identified a large number of CDSs still denoted as hypothetical proteins with no designated function. Characterising such CDSs should help further our understanding of this unique human pathogen.

### 1.7 Aims of study

The initial aim of this study was to undertake a comprehensive, up-to-date re-annotation of the *C. jejuni* NCTC11168 genome sequence. From this re-annotation, CDSs with a putative role in virulence, signal transduction or regulation of gene expression were selected for further investigation.

# Chapter 2: Materials and Methods

## 2.1 Re-annotation of *C. jejuni* NCTC11168 genome sequence

A manual re-annotation of the *C. jejuni* NCTC11168 genome sequence was performed at the WTSI using the methodology outlined below.

### 2.1.1 Sequence searches for re-annotation

A manual re-annotation of the *C. jejuni* NCTC11168 genome sequence was performed using Artemis software release 8 (Rutherford *et al*., 2000). BLASTP (Altschul *et al*., 1990) and FASTA (Pearson and Lipman, 1988) sequence comparisons from non-redundant databases were performed for every CDS. Re-annotation was based, wherever possible on characterised genes/proteins (Parkhill *et al*., 2000). Re-annotation also used information generated from PFAM (Sonnhammer *et al*., 1997), PROSITE (Falquet *et al*., 2002), RFAM (Griffiths-Jones *et al*., 2003), TMHMM (Sonnhammer *et al*., 1998b) and SIGNALP (Nielsen *et al*., 1997) databases search results. These databases are described in Table 2.1.

Table 2.1. Databases with corresponding search methodology used in this re-annotation.

| **Database Name** | **Database description** |
|---|---|
| BlastP | Database of protein sequences used to identify similar matches to a query sequence via a heuristic algorithm |
| FASTA | Database of protein sequences used to identify similar matches to a query sequence via a local sequence alignment algorithm |
| PFAM | Database of protein families that includes their annotations and multiple sequence alignments generated using Hidden Markov Models |
| PROSITE | Database of protein domains, families and functional sites as well as amino acid patterns, signatures, and profiles |
| RFAM | Database containing non-coding RNA (ncRNA) families |
| TMHMM | Database predicting transmembrane helices in proteins |
| SIGNALP | Database predicting the presence and location of signal peptide cleavage sites in amino acid sequences |

### 2.1.2 Literature and additional searches for re-annotation

Literature searches with CDS numbers and gene names was performed using PubMed (Entrez Pubmed), HighWire Press (HighWire Press), Scirus (Scirus - for scientific information only) and Google Scholar (Scholar). Updated ‘note’, ‘product’ and ‘gene’ qualifiers were added to each CDS in the EMBL (European Molecular Biology Laboratory) file. The original ‘note’ qualifier was retained for reference. Gene Ontology

(GO) annotation of the NCTC11168 genome sequence was performed automatically on submission to EMBL (accessed via the Gene Ontology Annotations link at EBI (European Bioinformatics Institute) (http://www.ebi.ac.uk/GOA/proteomes.html). GO annotation was also performed within GeneDB by performing a reciprocal FASTA comparison with *C. jejuni* RM1221 and adopting the GO annotation from orthologous CDSs.

### 2.1.3 Product designation

Results from the methods detailed in 2.1.1 and 2.1.2 were added to the 'updated' note qualifier. This data was used to evaluate whether the product function was to be updated. In this study, a sequence similarity of 35% was used as a benchmark for updating the product function. The decision to update product function was greatly aided by using supporting motif and characterisation data from the searches performed in 2.1.1 and 2.1.2. Importantly, any update based on alignment data was assessed in terms of coverage, as partial matches without key motifs may provide inaccurate designations.

### 2.1.4 Re-designation of pseudogenes

Pseudogene updates were performed with the aid of TBLASTX results. A complete re-analysis of all pseudogenes was performed using two alternative techniques. Two or more adjacent, in frame CDSs (previously annotated as separate pseudogene CDSs) were merged to create a single pseudogene containing internal stop codons. Alternatively, pseudogenes were modified by inserting multiple CDSs on different frames representing one or more frameshift in the full CDS.

### 2.1.5 Re-designation of CDSs with an intersecting homopolymeric tract

CDSs containing an intersecting homopolymeric tract (phase-variable CDSs) were merged to reflect the complete amino acid sequence, irrespective of whether the genome sequence indicated the CDS was in or out of phase. The result of this modification was to better illustrate a complete encoding of the CDS as opposed to a partial encoding represented by two separate CDSs.

### 2.1.6 Genome submission

The re-annotated genome was submitted to EMBL which acts as a database for genome sequences. The re-annotation can be accessed through - http://www.ncbi.nlm.nih.gov/nuccore/AL111168.

## 2.2 Microbiological and molecular biology techniques

### 2.2.1 Chemicals and reagents

Chemicals were purchased from Sigma-Aldrich (Poole, U.K), VWR/BDH/Merck (Poole, U.K), Fisher Scientific (Loughborough, U.K) or Invitrogen/Gibco (Paisley, U.K), unless stated otherwise (Appendix 1). All buffers and solutions were prepared as indicated in Sambrook and Russell unless otherwise stated (Sambrook and Russell, 2001). All media and reagent compositions not defined in the text are listed in Appendix 2.

### 2.2.2 Sterilisation

Solutions and buffers were sterilised either by autoclaving (Touchclave System PL, LTE Scientific, Oldham, U.K) or filtration. Wet steam sterilisation was performed at 121°C for 20 minutes. Dry steam sterilisation was performed at 134°C for 15 minutes. Filter sterilisation was performed using a 10 ml syringe (BD Plastipak, Oxford, U.K) and a 0.2 µm (32 mm) Acrodisc Syringe Filter with Supor Membrane (Pall Life Sciences, Portsmouth, U.K). All molecular reactions used filter sterilised Milli-Q grade water (Millipore, Billerica, U.S.A). All other solutions and buffers were prepared using Milli-RO (Reverse Osmosis) grade water (Millipore).

### 2.2.3 Bacterial strains

*C. jejuni* strains used in this study are described in Table 2.2.

Table 2.2. *C. jejuni* strains used in this study.

| ***C. jejuni* strains** | **Description** | **References** |
|---|---|---|
| 11168H | A hypermotile derivative of the original sequence strain NCTC11168 that shows higher levels of caecal colonisation in a chick colonisation model | (Karlyshev *et al*., 2002, Jones *et al*., 2004) |
| 11168H *Cj1556* mutant | Isogenic mutant of *Cj1556* with the insertion of a 1.4 kb $Km^R$ cassette at position 202 of the *Cj1556* nucleotide sequence | This study |
| 11168H *Cj1556* complement | *Cj1556* complement constructed with the insertion of *Cj1556* (0.33 kb) into the *Cj0233* pseudogene, in the *Cj1556* mutant chromosome | This study |
| 11168H *Cj0248* mutant | Isogenic mutant of *Cj0248* with the insertion of a 1.4 kb $Km^R$ cassette at position 454 of the *Cj0248* nucleotide sequence | This study |

*E. coli* strains used in this study are described in Table 2.3.

Table 2.3. *E. coli* strains used in this study.

| ***E. coli* strains** | **Description** | **Reference** |
|---|---|---|
| XL2-Blue MRF strain | Epicurian coli XL2 Blue MRF' Competent cells | Stratagene |
| SCS110 | *E. coli* competent cells - DAM methylase negative. | Stratagene |

### 2.2.4 Plasmids

Plasmids used in this study are described in Table 2.4.

Table 2.4. Plasmids used in this study.

| **Plasmids** | **Description** | **Reference** |
|---|---|---|
| pGEM-T Easy | PCR cloning vector, ampicillin resistant | Promega |
| pUC-*Cj1556* | pUC library clone (cam25a2 (1489074..1490567)) used for the *C. jejuni* NCTC11168 genome sequencing project. A 1.49 kb insert containing the CDSs *Cj1555c* – *Cj1560* including *Cj1556* (0.33 kb). | (Parkhill *et al*., 2000) |
| pGEM-*Cj0248* | pGEM-T Easy vector containing *Cj0248* CDS fragment (0.84 kb) | This study |
| pUC-*Cj1556*-$Km^R$ | pUC-*Cj1556* containing *Cj1556* with a 1.4 kb $Km^R$ inserted at position 202 of the *Cj1556* nucleotide sequence | This study |
| pGEM-*Cj0248*-$Km^R$ | pGEM-T Easy containing *Cj0248* with a 1.4 kb $Km^R$ inserted at position 454 of the *Cj0248* nucleotide sequence | This study |

### 2.2.5 Primers

Primers used in this study are described in Table 2.5.

Table 2.5. Primers used in this study.

| **Primer Name** | **Sequence** |
|---|---|
| *Cj1556*-F | ATCATTCTCTTTGTCCTAT |
| *Cj1556*-R | TAAGATGGATTCTAAACTATTG |
| Comp-*Cj1556*-F | CCCCCATGGATAAGGATTTATAATGAAAAAATATCATTCTCT |
| Comp-*Cj1556*-R | CCCGCTAGCTTAAACGATATTTTTATAGCTAT |
| *Cj0248*-F | ATGATTGGAGATATGAATGA |
| *Cj0248*-R | TTTCCTTAGCTTTGTTAGGA |
| *Cj0248*-IPCR-F | GGGAGATCTTCTTATACAAAATC |
| *Cj0248*-IPCR-R | GGAGATCTTAACAATACCAAGCC |
| Comp-*Cj0248*-F | CCCCCATGGATTAAGGATAAATTAAATATGATTGGAGATATGAATG |
| Comp-*Cj0248*-R | CCCGCTAGCTTAATCTTCTTTGTTTAAATTT |
| $Km^R$ forward-out | TGGGTTTCAAGCATTAGTCCATGCAAG |
| $Km^R$ reverse-out | GTGGTATGACATTGCCTTCTGCG |

### 2.2.6 Statistical analysis

All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Statistical analyses were performed using GraphPad Prism software (v4.02). Variables

were compared for significance using a student's *t*-test with one star (*) indicating $p$ = 0.01 - 0.05, two stars (**) indicating $p$ = 0.001- 0.01 and three stars (***) $p < 0.001$.

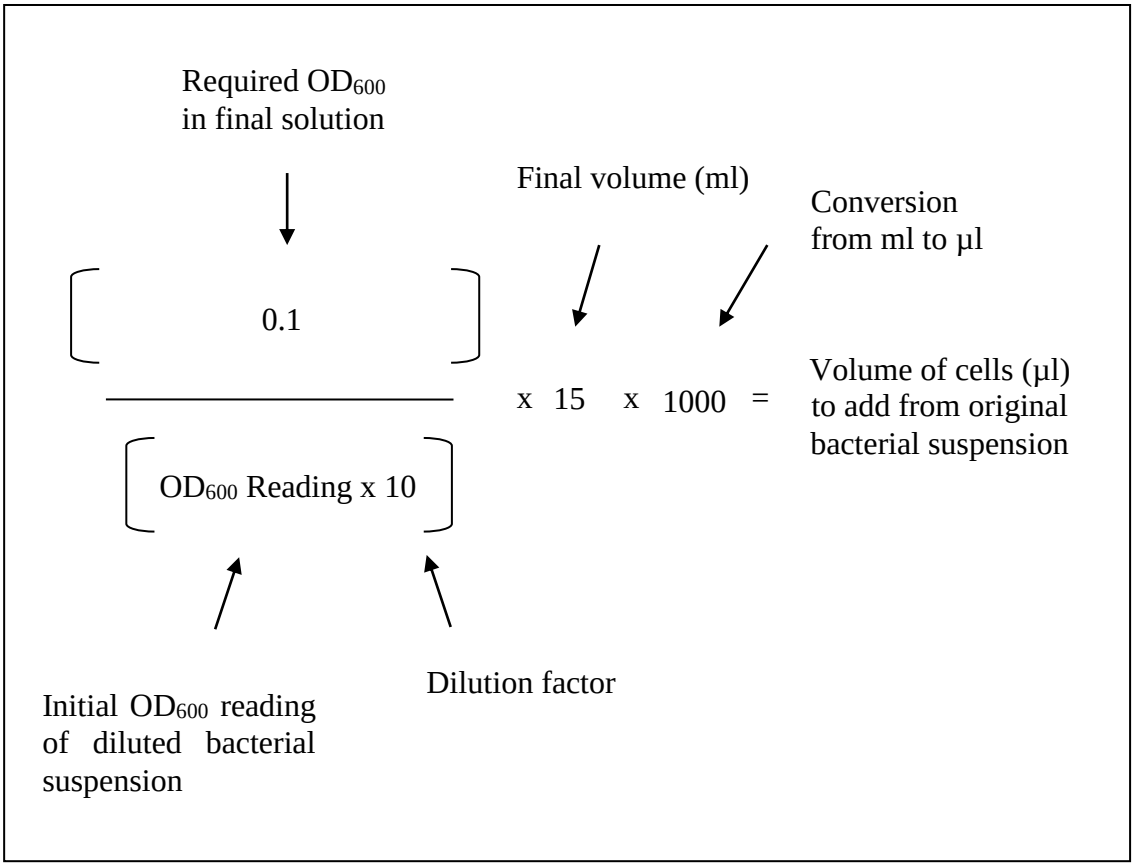
### 2.2.7 Bacterial growth and storage

All media and reagents not defined in this text are described in Appendix 2. *C. jejuni* cells were resuscitated from glycerol stocks by plating onto blood agar plates and incubated at 37°C under microaerobic conditions (85% Nitrogen, 10% Carbon dioxide and 5% Oxygen) in a variable atmospheric incubator (VAIN) (Don Whitley Scientific, Shipley, U.K). *C. jejuni* restreaks were performed every 3 or 4 days up to a maximum of 10 passages. Overnight cultures of *C. jejuni* were prepared by adding an appropriate quantity of a bacterial suspension (see Section 2.2.8) to 10 ml Brucella broth (Sigma-Aldrich) in a 50 ml flask (Thermo Scientific, Massachusetts, U.S.A). Cultures were incubated at 37°C under microaerobic conditions on a shaker (75 rpm) (Platform Shaker STR6, VWR-Jencons, East Grinstead, U.K). *Escherichia coli* cells were resuscitated from glycerol stocks and plated onto Luria-Bertani (LB) agar plates and were incubated at 37°C under aerobic conditions in an incubator (Sanyo, Loughborough, U.K). LB plates with *E. coli* grown overnight were sealed with Parafilm M (Neenah, U.S.A) and stored at 4°C for 2-3 weeks. Overnight cultures of *E. coli* were prepared by adding a single colony to 10 ml LB broth in a 30 ml universal. Cultures were grown at 37°C under aerobic conditions in an incubator shaking at 200 rpm (Weiss Gallenkamp, Loughborough, U.K). Appropriate antibiotics were added at the following concentrations; ampicillin (100 μg/ml), kanamycin (50 μg/ml), chloramphenicol for *E. coli* growth (50 μg/ml) or for *C. jejuni* growth (10 μg/ml). Glycerol stocks were prepared for *C. jejuni* (from an overnight plate) and *E. coli* (from an overnight liquid culture) cells using 10% (v/v) glycerol, 10% (v/v) Foetal Calf Serum (FCS) (Sigma-Aldrich) and 80% (v/v) Mueller-Hinton broth mixture (*C. jejuni*) or 15% (v/v) glycerol in LB broth (*E. coli*). 500 μl aliquots were snap-frozen using dry ice in 100% (v/v) ethanol and stored in a -80°C freezer (New Brunswick Scientific, St. Albans, U.K).

### 2.2.8 Preparation of a specific $OD_{600}$ *C. jejuni* suspension

15 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h. *C. jejuni* were grown on blood agar plates for 24 h. *C. jejuni* were collected using a sterile swab and resuspended in 1 ml Brucella broth. 100 μl of this suspension was resuspended in 900 μl Brucella broth in a 1.5 ml microcentrifuge tube. This diluted suspension was vortexed, transferred to a 1 ml spectrophotometer cuvette

and the $OD_{600}$ recorded using a spectrophotometer (S2000 UV/Vis Spectrophotometer, VWR-Jencons, Leighton Buzzard, U.K). 1 ml Brucella broth was used as a blank. The volume of the original *C. jejuni* suspension to be added to the 15 ml of equilibrated Brucella broth to produce an initial inoculum with an $OD_{600}$ of 0.1 was calculated as follows:-




**2.2.9 Isolation of genomic DNA**

Bacterial genomic DNA was isolated using the Gentra – Puregene DNA Purification Kit (Gentra Systems Inc, Flowgene, Lichfield, U.K). *C. jejuni* cells were collected from a 24 h blood agar plate using a sterile swab and resuspended in 0.75 ml Phosphate Buffered Saline (PBS) in a 1.5 ml microcentrifuge tube and centrifuged at 13,000 rpm for 5 seconds. The supernatant was removed and 300 µl Cell Lysis Solution was added to the sample tube and mixed until the bacterial pellet was thoroughly resuspended. The suspension was incubated at 80°C for 5 minutes. 1.5 µl RNase A Solution (80 U/ml) was then added to the cell lysate. The tube was mixed by inverting 25 times and incubated at 37°C for 45 minutes. Once cooled to room temperature, 100 µl Protein Precipitation Solution was added to the RNase A-treated cell lysate, vortexed for 20 seconds and then

centrifuged at 13,000 rpm for 3 minutes. The process of vortexing and centrifugation was repeated to achieve greater separation of DNA from protein. The supernatant containing the DNA was transferred into a 1.5 ml microcentrifuge tube containing 300 µl 100% (v/v) isopropanol and mixed by inverting 50 times. The resulting suspension was centrifuged at 13,000 rpm for 1 minute at which point the DNA was usually visible as a small white pellet. The supernatant was carefully removed and the microcentrifuge tube inverted and drained on absorbent paper for 20 minutes. 300 µl of 70% (v/v) ethanol was then added and the DNA pellet was rinsed by inverting the tube several times. This tube was centrifuged at 13,000 rpm for 1 minute. The remaining ethanol was poured off and the tube was air dried for 15 minutes. 50 µl DNA Hydration Solution was added to the sample tube and the DNA rehydrated by incubating the sample for 1 h at 65°C. Genomic DNA concentrations were quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, U.S.A). Genomic DNA was stored at -20°C.

### 2.2.10 RNA Extraction

10 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h. A *C. jejuni* suspension was prepared and used to inoculate this broth to a starting $OD_{600}$ of 0.1 (see Section 2.2.8). *C. jejuni* cultures were grown for 16 h. RNA extraction was performed using the Qiagen RNeasy kit and RNAprotect Bacteria Reagent (Qiagen) as follows. 8 ml RNAprotect Bacteria Reagent was added to a 30 ml universal. 4 ml of the 16 h *C. jejuni* culture was added to this universal and briefly vortexed. This mixture was incubated for 5 minutes at room temperature, then centrifuged at 4,000 rpm for 10 minutes followed by the removal of the resulting supernatant. 200 µl TE buffer/lysozyme mix (950 µl TE (pH 8.0) + 50 µl lysozyme (1 mg/ml)) was added and mixed to resuspend the pellet using a P1000 Gilson pipette for 1 minute. This suspension was then incubated for 10 minutes at room temperature, vortexing every 2 minutes. 700 µl buffer RLT mix (10 µl β-mercaptoethanol (14.3 M) (Stratagene, Amsterdam, Holland) + 1 ml buffer RLT) was added to the mix and vortexed vigorously. 500 µl 100% (v/v) ethanol was then added and mixed by pipetting. 700 µl of this suspension was added to an RNeasy Mini spin column and centrifuged for 15 seconds at 10,000 rpm. The flow through was discarded. The above procedure was repeated for any remaining suspension to further increase RNA yield. 350 µl of buffer RW1 was added to the spin column and centrifuged for 15 seconds at 10,000 rpm. The resulting flow through was discarded. 80 µl DNaseI/buffer RDD mix (Qiagen) (10 µl DNaseI stock solution + 70 µl Buffer RDD) was added directly onto the membrane and

incubated at room temperature for 15 minutes. 350 µl Buffer RW1 was added to the spin column and centrifuged for 15 seconds at 10,000 rpm. The spin column was then transferred to a fresh 2 ml collection tube. 500 µl Buffer RPE was added to the spin column and centrifuged for 15 seconds at 10,000 rpm. The flow through was discarded. Another 500 µl Buffer RPE was added to the spin column and centrifuged for 2 minutes at 10,000 rpm. The column was then placed in a new 2 ml collection tube and centrifuged for a further 1 minute at 13,000 rpm to allow removal of any Buffer RPE carryover. The spin column was then removed from the collection tube and placed into a new 1.5 ml microcentrifuge tube. 50 µl RNase-free water was added directly onto the RNeasy silica membrane in the spin column and then centrifuged for 1 minute at 10,000 rpm to allow elution. The elution was reapplied to the membrane and centrifugation was repeated for 1 minute at 10,000 rpm. RNA concentrations were quantified using a NanoDrop ND-1000 spectrophotometer. RNA was stored at -80°C.

### 2.2.11 Polymerase Chain Reaction

Oligonucleotide primers (Sigma-Aldrich) were designed manually or using the web program Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Primers were checked for 'self-dimers' and 'hairpins' using the web program IDT SciTools – OligoAnalyzer 3.0 (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx). Primers were resuspended in sterilised Milli-Q water and diluted to a final concentration of 100 pmol/µl.

A dNTP stock solution containing 1.25 mM of dATP, dTTP, dCTP and dCTP was prepared as follows:-

dNTP 1.25 mM Stock

| | |
|---|---|
| dATP (100mM) | - 12.5 µl |
| dTTP (100mM) | - 12.5 µl |
| dGTP (100mM) | - 12.5 µl |
| dCTP (100mM) | - 12.5 µl |
| Sterilised Milli-Q water | - to 1000 µl |

Standard polymerase chain reaction (PCR) pre-mix was prepared as follows:-

| | |
|---|---|
| 10X Buffer | - 100 µl |
| dNTPs (1.25mM) | - 168 µl |
| $MgCl_2$ (25 mM) | - 60 µl |
| *Taq* polymerase (5 U/µl) | - 6 µl |
| Sterilised Milli-Q water | - to 1000 µl |

Standard PCR pre-mix was mixed using a P100 Gilson pipette and stored at -20°C.

Typical PCR reactions were performed as follows:-

| | |
|---|---|
| Forward primer (100 pmol/µl) | - 0.075 µl* |
| Reverse primer (100 pmol/µl) | - 0.075 µl |
| Genomic DNA (10 – 100 ng/ /µl) | - 1 µl |
| Standard PCR pre-mix | - to 20 µl |

* A 10 reaction volume was typically used to allow greater accuracy during pipetting e.g. 0.75 µl forward and reverse primer.

Standard PCR programme

Step 1 – Denature at 94°C for 15 seconds

Step 2 – Anneal at 50°C for 1 minute

Step 3 – Extension at 72°C for 1 minute

Repeat steps 1-3 34 times

Step 4 – End cycle with 72°C for 7 minutes

Steps 2 or 3 were varied to optimise PCR amplification or to amplify larger products where each minute for extension amplified approximately 1.5 kb. PCR amplification was performed using an MJ Research PTC-225 Peltier Thermal Cycler (GRI, Braintree, U.K).

### 2.2.12 Agarose gel electrophoresis

0.7% (w/v) TAE agarose gels were prepared as described in Appendix 2. The agarose suspension was heated in a microwave until molten and allowed to cool to approximately 50°C. Ethidium bromide (10 mg/ml / Promega) was added to the agarose to a final concentration of 0.5 µg/µl. The agarose was poured into a sealed gel tray (Scie-Plas, Southam, U.K) with appropriate combs in place and allowed to gel. The well combs

were then removed. The gel was then placed into a gel tank (Scie-Plas) and fully immersed in 1X TAE buffer. 20 µl of PCR reactions were mixed with 5 µl gel loading buffer (Appendix 2) then loaded into individual wells. A 1-10 kb ladder (HyperLadder 1 – Bioline, London, U.K) was run parallel to the samples to allow PCR product sizes to be ascertained. Electrophoresis was performed using a Hybaid PS250 (Hybaid Ltd, Middlesex, U.K) with settings of 120V and 500mA. Running time was approximately 45 minutes. Gels were visualised using a GeneGenius Gel-Documentation System (SynGene, Cambridge, U.K).

### 2.2.13 PCR product purification

PCR products and also restriction endonuclease reactions (see Section 2.2.11 and 2.2.19) were purified using the QIAquick PCR Purification kit (Qiagen). 5 volumes of buffer PB was added to 1 volume of PCR reaction. This mixture was transferred onto a QIAquick spin column in a 2 ml collection tube and centrifuged at 13,000 rpm for 1 minute. The flow-through was discarded and the QIAquick column was placed back into the same 2 ml collection tube. 0.75 ml buffer PE was applied to the QIAquick column and centrifuged for 1 minute at 13,000 rpm. The flow-through was discarded and the QIAquick column was placed back into the 2 ml collection tube and centrifuged at 13,000 rpm for an additional minute. The QIAquick column was transferred to a new 1.5 ml microcentrifuge tube and 50 µl sterilised Milli-Q water was applied to the centre of the QIAquick membrane. This was allowed to fully absorb for 1 minute at room temperature and then centrifuged at 13,000 rpm for an additional minute. Sample concentrations were recorded using a NanoDrop ND-1000 spectrophotometer. Purified PCR products were stored at -20°C.

### 2.2.14 Gel fragment purification

Purification of PCR products from agarose gels was performed using the GenElute Gel Extraction Kit (Sigma-Aldrich). Samples to be purified were separated on a 0.7% (w/v) TAE agarose gel (see Section 2.2.12). Gels were visualised using near-UV light (365 nm) on a TM-20 transilluminator (UVP, Cambridge, U.K) in a dark room under red light. The appropriate band was excised from the agarose gel with a sterile scalpel and placed into a 1.5 ml microcentrifuge tube. The excised agarose sample was weighed and 3 gel volumes of Gel Solubilization Solution was added to the 1.5 ml microcentrifuge tube containing the gel fragment then heated at 55°C for 10 minutes with brief vortexing every 2-3 minutes to ensure the gel was fully dissolved. A GenElute Binding Column G

was placed into a 2 ml collection tube. 500 µl of Column Preparation Solution was applied to the column and centrifuged at 13,000 rpm for 1 minute. The flow through was discarded. 1 gel volume of 100% (v/v) isopropanol was added to the solubilized gel and mixed until homogenous. The solubilized gel solution was then loaded onto the prepared binding column and centrifuged at 13,000 rpm for 1 minute. The flow-through was discarded. 700 µl of Wash Solution was applied to the binding column and centrifuged for 1 minute at 13,000 rpm. The flow through was discarded and the column was placed back into the same 2 ml collection tube and centrifuged for an additional minute at 13,000 rpm. DNA elution was performed by transferring the binding column to a fresh collection tube and applying 50 µl Elution Solution to the centre of the binding column. This was allowed to absorb for 1 minute before centrifuging for 1 minute at 13,000 rpm. Sample concentrations were recorded using a NanoDrop ND-1000 spectrophotometer. Purified DNA fragments were stored at -20°C.

### 2.2.15 DNA ligation

Ligation reactions were performed using the pGEM-T Easy kit (Promega). The amount of PCR product to be ligated into the pGEM-T Easy vector was calculated using the equation below.

$$\frac{[\text{ng of vector} \times \text{kb size of insert}]}{[\text{kb size of vector}]} \times \text{insert:vector molar ratio} = \text{ng of insert}$$

e.g.

$$\frac{[50 \text{ ng} \times 1.3 \text{ kb}]}{[3.0 \text{ kb}]} \times 1{:}8 = 173.3 \text{ ng}$$

Molar ratios of 3:1 for insert:vector were recommended by the manufacturer, however ratios down to 1:8 were also successfully used. The ligation reaction was set up as follows:-

| | |
|---|---|
| 2X Rapid Ligation Buffer | - 5 µl |
| pGEM-T Easy Vector (50 ng/µl) | - 1 µl |
| PCR product | - X µl |
| T4 DNA Ligase (3 Weiss units/µl) | - 1 µl |
| Sterilised Milli-Q water | - to 10 µl |

The reaction mixture was mixed gently by pipetting and incubated overnight at 4°C.

### 2.2.16 Transformation of *E. coli* cells

The DNA ligation mixture (see Section 2.2.15) was briefly centrifuged. 2 µl ligation mix was added to a 1.5 ml microcentrifuge tube and incubated on ice. A separate 1.5 ml tube

was incubated on ice and 2 µl of β-mercaptoethanol (14.3 M) added. To this tube, 100 µl Epicurian coli XL2 Blue MRF' Competent cells (Stratagene) was added, mixed and incubated on ice for 5-10 minutes. The cells were mixed by gently tapping the tubes every 2/3 minutes. 50 µl of these competent cells were transferred into the tube containing the 2 µl ligation mix. This tube was gently tapped to mix and then incubated on ice for 20 minutes. The cells were heat-shocked in a 42°C water bath for 45-50 seconds and then incubated on ice for 2 minutes. 950 µl SOC media (at room temperature) was then added to the transformation reaction and incubated for 1.5 h at 37°C with shaking at 200 rpm. 100 µl of the transformation culture was plated onto LB/ampicillin/IPTG/X-Gal plates (Appendix 2). The plates were incubated for 12 - 16 h at 37°C. White colony transformants were selected for further study.

### 2.2.17 Screening by vector primer PCR / insert specific PCR

White colony transformants were restreaked onto LB plates and incubated overnight at 37°C. A single colony from the restreaked plate was selected and resuspended in 100 µl sterilised Milli-Q water, briefly vortexed and heated at 95°C for 10 minutes followed by centrifuging at 13,000 rpm for 5 minutes. Vector primer PCR (VPPCR) reactions were set up as follows:-

| | |
|---|---|
| Forward T7 primer (100 pmol/µl) | - 0.075 µl |
| Reverse SP6 primer (100 pmol/µl) | - 0.075 µl |
| Boilate supernatant | - 1 µl |
| VPPCR pre-mix | - to 25 µl |

Standard PCR programmes were performed (see Section 2.2.11). PCR reactions were analysed on a 0.7% (w/v) TAE agarose gel (see Section 2.2.12). A variation of VPPCR was to use insert specific primers (ISPCR) using the original primers used for amplification of the CDS of interest (see Section 2.2.11) to test for positive transformants.

### 2.2.18 Isolation of plasmid DNA

Plasmid minipreps were performed using QIAprep Miniprep Kits (Qiagen). An overnight culture of *E. coli* was prepared by adding a single colony to 10 ml LB broth in a 30 ml universal, then incubated at 37°C in an incubator shaking at 200 rpm. The overnight bacterial suspension was centrifuged at 4,000 rpm for 10 minutes. The supernatant was

discarded. The pellet of bacterial cells was resuspended in 250 µl buffer P1 and transferred to a 1.5 ml microcentrifuge tube. 250 µl buffer P2 was added and mixed by inverting 4-6 times. This was followed by adding 350 µl buffer N3 and mixing immediately by inverting the tube 4-6 times. The mixture was then centrifuged at 13,000 rpm for 10 minutes. The supernatant was transferred to a QIAprep spin column and centrifuged for 1 minute. The resulting flow-through was discarded. 0.5 ml buffer PB was added to the column and this was centrifuged for 1 minute. The flow-through was discarded and 0.75 ml buffer PE was added to the column. The column was then centrifuged at 13,000 rpm for 1 minute. The flow-through was discarded and centrifuged again for 1 minute at 13,000 rpm to remove residual wash buffer. Finally, the QIAprep column was transferred to a clean 1.5 ml microcentrifuge tube and 50 µl buffer EB was applied to the centre of the column. This was allowed to absorb for 1 minute before centrifuging for 1 minute at 13,000 rpm. Plasmid DNA concentrations were recorded using a NanoDrop ND-1000 spectrophotometer. Plasmid DNA samples were stored at -20°C.

### 2.2.19 Restriction endonuclease digests

Standard restriction digests were prepared using a final volume of 20 µl. 0.5 – 1.0 µg plasmid DNA was digested with 10 U of the appropriate restriction endonuclease and buffer (New England Biolabs, Hitchin, U.K). Reactions were prepared in 1.5 ml microcentrifuge tubes and incubated at the appropriate temperature and time duration recommended by the manufacturer. Digests were purified using the QIAquick PCR Purification kit (see Section 2.2.13).

### 2.2.20 DNA sequencing

Sequencing of plasmid DNA was performed using ABI Prism Terminator Ready Reaction Mix (Applied Biosystems, Warrington, U.K). Sequencing reactions were also performed on genomic DNA (see Section 2.2.9) using similar concentrations as for plasmid DNA.

| | |
|---|---|
| ABI Reaction Mix | - 8 µl |
| Plasmid or genomic DNA (200 – 500 ng) | - 1 µl |
| Forward or reverse CDS specific primer* (1 pmol/µl) | - 4 µl |
| Sterilised Milli-Q water | - to 20 µl |

*Sequencing reactions were also performed using a vector specific primer.

PCR cycle

Step 1 – Denature at 96°C for 10 seconds

Step 2 – Anneal at 50°C for 5 seconds

Step 3 – Extension at 60°C for 4 minutes

Repeat steps 1-3 25 times

Sequencing reactions were performed on a GeneAmp 9600 thermal cycler (Perkin-Elmer, Beaconsfield, U.K). The entire sequencing reaction was then transferred to a 1.5 ml microcentrifuge tube containing 80 µl 75% (v/v) isopropanol. The sample was briefly vortexed and incubated at -20°C for 1 h. The sample was then centrifuged at 13,000 rpm for 30 minutes and the supernatant discarded. 400 µl of 75% (v/v) isopropanol was added to the pellet and briefly vortexed before centrifuging at 13,000 rpm for 5 minutes. The supernatant was removed using a pipette. The pellet was dried at room temperature for approximately 20 minutes. 10 µl HiDi solution (Applied Biosystems) was applied to the pellet and the reaction mixture sequenced using the ITD Faculty facilities. Analysis of sequence data was performed using Chromas v1.61 software (Technelysium Pty. Ltd).

### 2.2.21 Mutagenesis strategy

Defined isogenic *C. jejuni* mutants were constructed in this study as follows. Before designing amplifying primers, the CDS of interest was checked for the presence of *Bam*HI, *Bcl*I or *Bgl*II restriction sites. If any of these restriction sites were present in the CDS of interest, the mutagenesis strategy would not require inverse PCR mutagenesis (IPCRM) (see Section 2.2.23). Primers for the CDS of interest were designed (see Section 2.2.11) and the CDS/CDS fragment was PCR amplified (see Section 2.2.11). After confirmation of PCR product size and purification (see Section 2.2.13), the PCR product was cloned into the pGEM-T Easy vector followed by transformation into *E.coli* XL2 cells (or *E. coli* SCS110 cells if the unique restriction site was *Bcl*I) (see Sections 2.2.15 and 2.2.16). Isolation of plasmid DNA was followed by checking for positive transformants and glycerol stock preparation (see Sections 2.2.7, 2.2.17, 2.2.19 and 2.2.20).

The second stage was to incorporate a $Km^R$ cassette into the CDS of interest. IPCRM was performed at this stage if required (see Section 2.2.23). Restriction digests were performed to allow the ligation of *Bam*HI digested $Km^R$ cassette into the centre of the

cloned CDS/CDS fragment (see Sections 2.2.19) followed by transformation into *E.coli* XL2 cells (see Sections 2.2.15 and 2.2.16). Isolation of plasmid DNA was followed by checking for correct sized insert, correct orientation of the $Km^R$ cassette and glycerol stock preparation (see Sections 2.2.7, 2.2.17, 2.2.19 and 2.2.20).

The final stage was to transform this plasmid construct by electroporation into *C. jejuni* 11168H wild-type strain (see Section 2.2.24) followed by screening for positive mutants (see Section 2.2.25).

### 2.2.22 Mutation of individual *C. jejuni* CDSs

A $Km^R$ cassette (*aphA-3* (Trieu-Cuot *et al*., 1985)) was isolated from pJMK30 (van Vliet *et al*., 1998) following a *Bam*HI restriction digest. The *C. jejuni* CDS to be mutated was analysed for the presence of a single *Bam*HI, *Bcl*I or *Bgl*II site within the sequence using NEBcutter v2.0 (http://tools.neb.com/NEBcutter2/index.php). The unique restriction site should ideally be located near the centre of the amplified CDS/CDS fragment to improve efficiency of recombination. All three restriction sites are complementary to *Bam*HI. When using *Bcl*I, the original plasmid and CDS of interest transformation was performed using SCS110 *E. coli* competent cells (Stratagene), rather than Epicurian coli XL2 Blue MRF' Competent cells (Stratagene), due to the *Bcl*I restriction site exhibiting DAM methylase sensitivity. It would not be possible to *Bcl*I digest plasmid DNA from XL2 cells which are DAM methylase positive. SCS110 cells are DAM methylase negative.

To allow insertion of the $Km^R$ cassette, 2 µg plasmid DNA containing the cloned CDS of interest was digested with approximately 20 units restriction enzyme (*Bam*HI, *Bcl*I or *Bgl*II*)* (see Section 2.2.19). The digest were purified using the QIAquick PCR Purification kit (see Section 2.2.13) and sample concentrations were recorded using a NanoDrop ND-1000 spectrophotometer. The following ligation reaction was then set up:-

| | |
|---|---|
| Plasmid DNA digested with unique restriction enzyme (≈ 250 ng/µl) | - 2 µl |
| $Km^R$ cassette digested with *Bam*HI (≈ 20 ng/µl) | - 5 µl |
| 10X Ligase buffer | - 1 µl |
| T4 DNA Ligase (3 Weiss units/µl) | - 2 µl |

Reactions were incubated overnight at 4°C. Transformation of the reaction mixture into XL2 or SCS110 cells was performed as described in Section 2.2.16. Transformants were selected on LB/ampicillin/kanamycin plates and incubated overnight at 37°C. Colonies were restreaked onto LB/ampicillin/kanamycin plates and checked for the insertion of a $Km^R$ cassette by carrying out ISPCR (see Section 2.2.17). Plasmid DNA was isolated from positive clones as described in Section 2.2.18. The orientation of the $Km^R$ cassette was determined by performing multiple combinations of CDS specific and kanamycin specific primer PCR reactions (Appendix 3 and 4).

### 2.2.23 Inverse PCR mutagenesis

IPCRM was performed when the CDS to be mutated did not contain a *Bam*HI, *Bcl*I or *Bgl*II restriction site. IPCRM primers were designed by selecting 15-20 nucleotides in the centre of the CDS to be mutated with an interspacing region of 10-15 nucleotides (Figure 2.1).

ATGATTGGAGATATGAATGAGCTTTTATTAAAAAGCGTTGAAGTATTGCCACCTTTACCTGATAC
TGTAAGTAAGTTAAGAAAATATGTGAGCGAGGCTAATTCAAATATAGAAACTATGAAAGTTGCTG
AAATCATTTCAAGCGATCCGTTGATGACGGCTAAGCTTTTGCAATTAGCAAATTCTCCTTATTAT
GGTTTTACAAGAGAAATTACAACCATAAATCAAGTGATTACTTTATTAGGCGTTGGTAATATCAT
CAATATAGTTATGGCTGACTCCATTAGAGATAATTTTAAAATAGACGTTTCACCTTATGGTTTAA
ATACTCAAAATTTTTTAAAAACGTGCAATGAAGAGGCAACTTTTATCGCAAATTGGCTTAATGAT
GAAGATAAAAAACTTTCTCATCTTTTAGTTCCTTGTGCAATGCTTTTAAGGCTTGGTATTGTTAT
TTTTTCAAATTTTCTTATACAAAATCATAAGGATAAGGATTTTTTAGCTTTTTTAAATAAAAATG
AAAATCTTGCTTTAGCGGAGAATGAATTTTTAGGCGTAGATCATATTTCTTTCTTGGGATTTTTG
TTACATCGTTGGAATTTTGATGATGTTTTGATTGAAAGTATATGTTTTGTTCGCACTCCTCATGC
TGCTCGCGAAAAAGTGAAAAAATCCGCTTATGCTTTAGCAATAACAGATCATCTTTTTGCTCCGC
ATGATGGTTCTTCTCCATTTAACGCAAAAGCTGCAGTTGCTTTACTTAAAGAGGCAAAAACTCAA
GGAATTAATTTTGATTTAAACAATCTTTTATCTAAGCTTCCTAACAAAGCTAAGGAAAATTTAAA
CAAAGAAGATTAA

*Cj0248*-IPCRM–F = 5’ GGGAGATCTTCTTATACAAAATC 3’

*Cj0248*-IPCRM–R = 5’ GGGAGATCTTAACAATACCAAGCC 3’

Figure 2.1. IPCRM primers designed for *Cj0248*. The 5’-end of the IPCRM primers contained three guanine residues to allow efficient functionality of the *Bam*HI restriction site. This was followed by the *Bam*HI complementary sequence – AGATCT.

IPCRM reactions were performed using cloned CDS fragments at a diluted concentration of approximately 0.1-10 ng. The reduced concentration of template DNA was used to minimise the number of false positives after transformation into XL2 cells (described below).

The following reactions were prepared:-

IPCRM Pre-Mix

| | |
|---|---|
| Buffer I (10X) | - 10 µl |
| AccuPrime Taq DNA high fidelity polymerase* (5 U/µl) (Invitrogen) | - 0.4 µl |
| Milli-Q water | - 89.6 µl |

*dNTPs are present in Buffer I at a concentration of 400 µM.

IPCRM reaction

IPCRM Pre-mix - 98.4 µl

DNA (0.1 – 10 ng) - 1.0 µl

IPCRM primers* - 0.6 µl

*IPCRM primers were resuspended in sterilised Milli-Q water and diluted to a final concentration of 100 pmol/µl. A 1:1 mix of both primers was prepared giving a working concentration of each primer of 50 pmol/µl.

IPCRM reaction conditions

Step 1 – 94°C for 2 minutes

Step 2 – Denature at 94°C for 1 minute

Step 3 – Anneal at 45°C for 1 minute

Step 4 – Extension at 72°C for X minutes (X = size of plasmid in kb x 1.5 minutes)

Repeat steps 2-4 40 times

Step 5 – 72°C for 7 minutes

15 µl of the PCR reaction was analyzed on an agarose gel (see Section 2.2.12). The amplified IPCRM product was purified using the QIAquick PCR Purification kit (see Section 2.2.13). Sample concentrations were recorded using a NanoDrop ND-1000 spectrophotometer. The amplified IPCRM product was then digested with *Bgl*II and *Dpn*I for 3 h at 37°C. Digestion using *Dpn*I was performed to ensure the reduction/elimination of methylated template DNA that would lead to false positive transformation (Shenoy and Visweswariah, 2003). Digestion volumes were as follows:-

| | |
|---|---|
| Amplified IPCRM product (10-500 ng/µl) | - 20 µl |
| *Bgl*II (20 units) | - 2 µl |
| *Dpn*I (20 units) | - 2 µl |
| NEBuffer 2 | - 10 µl |
| Milli-Q water | - to 100 µl |

Restriction endonuclease reactions were purified using the QIAquick PCR Purification kit as described in Section 2.2.13. Sample concentrations were again recorded using a NanoDrop ND-1000 spectrophotometer. To perform the insertion of the $Km^R$ cassette and re-ligation of the digested IPCRM product, the following reaction was set up:-

| | |
|---|---|
| Amplified IPCRM product digested with *Bgl*II and *Dpn*I (≈ 250 ng/µl) | - 2 µl |
| $Km^R$ cassette digested with *Bam*HI (≈ 20 ng/µl) | - 5 µl |
| 10X Ligase buffer | - 1 µl |
| T4 DNA Ligase (3 Weiss units/µl) | - 2 µl |

The reaction was incubated overnight at 4°C. The ligation reactions were then transformed into XL2 *E. coli* cells (see Section 2.2.16). Transformants were selected on LB/ampicillin/kanamycin plates. Colonies were screened using ISPCR (see Section 2.2.17). Positive colonies were re-streaked onto LB/ampicillin/kanamycin plates and plasmid DNA isolated (see Section 2.2.18).

### 2.2.24 Transformation of *C. jejuni* cells by electroporation

Bacterial cells from a 24 h plate of *C. jejuni* 11168H were collected using a sterile loop and resuspended in 10 ml ice cold EBF buffer (Appendix 2) in a universal container. The suspension was centrifuged at 4,000 rpm for 10 minutes and the supernatant discarded. The pellet was then resuspended in 1 ml ice cold EBF buffer, transferred to a 1.5 ml microcentrifuge tube and centrifuged at 13,000 rpm for 2 minutes. This process was repeated before resuspending the pellet in 250 µl ice cold EBF buffer. 5 µl ice-cold plasmid DNA (1-5 µg) was then added to 50 µl *C. jejuni* 11168H suspension and mixed by pipetting. This reaction was incubated on ice for 5-10 minutes. The mixture was then transferred to a cold electroporation 2 mm gap cuvette (Bio-Rad Laboratories, Hemel Hempstead, U.K). The cuvette was gently tapped to ensure no bubbles were present. Electroporation was performed using a GenePulser Xcell (Bio-Rad) using 2.5 kV, 25 µFD and 200 Ω settings. Immediately after electroporation, 100 µl pre-warmed SOC media was added to the cuvette to aid bacterial recovery. The bacterial cells were then plated onto blood agar plates and incubated at 37°C under microaerobic conditions for 2 days. After 2 days, bacterial cells were harvested using a sterile swab and resuspended in 0.5 ml PBS solution. 100 - 200 µl aliquots were then pipetted onto blood agar plates containing 50 µg/ml kanamycin. The bacterial cells were spread across the plate surface using a sterile spreader. Plates were incubated for up to 5-7 days at 37°C under microaerobic conditions.

### 2.2.25 Screening *C. jejuni* clones for positive mutants

Kanamycin resistant colonies were restreaked onto blood agar plates containing kanamycin. A single loop of this putative kanamycin resistant colony was resuspended in

100 µl sterilised Milli-Q water, briefly vortexed and heated at 95°C for 10 minutes followed by centrifuging at 13,000 rpm for 5 minutes. PCR screening was performed using CDS specific primers as described in Section 2.2.17. Glycerol stocks were prepared and genomic DNA isolation was performed on positive mutants.

### 2.2.26 Complementation of *C. jejuni* 11168H mutants

Complementation constructs were produced using the *Campylobacter* complementation vector (named pDENNIS in this study) (Hitchen *et al*., 2010). The forward complement primer was designed by incorporating three cytosine residues, an *Nco*I restriction site, approximately 15 bps upstream of the intended CDS start codon (allowing incorporation of the ribosome binding site), followed by 15-20 bps of the 5' end of the CDS sequence (Figure 2.2).

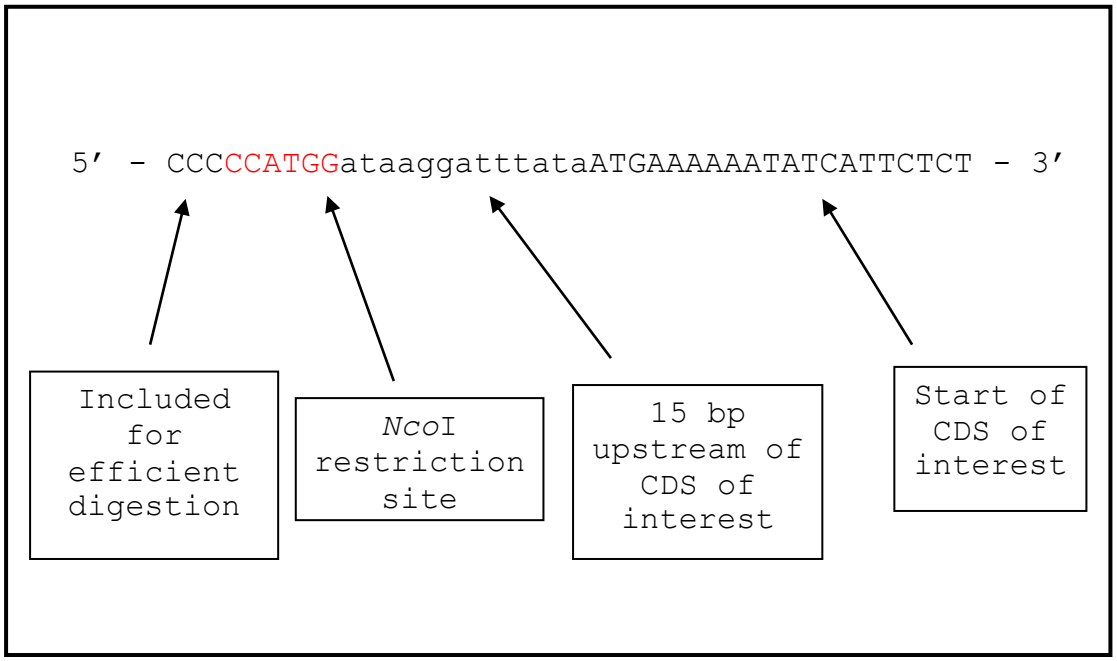



Figure 2.2. Forward complement primer incorporating three cytosine residues, an *Nco*I restriction site, approximately 15 bps upstream of the intended CDS (allowing incorporation of the ribosome binding site) followed by 15-20 bps of the 5' end of the CDS sequence.

The reverse complement primer was designed by incorporating three cytosine residues, an *Nhe*I restriction site, a stop codon (TTA), followed by 15-20 bps of the 3' end of the CDS sequence (Figure 2.3).

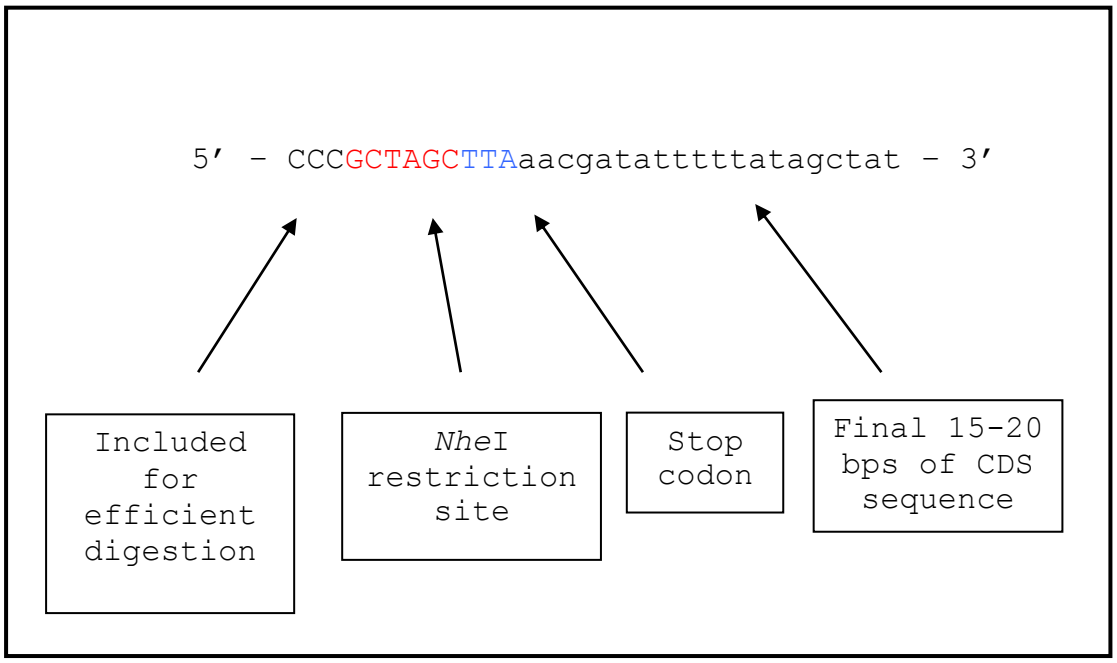



Figure 2.3. Reverse complement primer incorporating three cytosine residues, an *Nhe*I restriction site, a stop codon (TTA), followed by 15-20 bps of the 3' end of the CDS sequence.

The CDS of interest was amplified using AccuPrime™ *Taq* DNA Polymerase High Fidelity (Invitrogen) to ensure nucleotide modifications did not occur during PCR amplification (see Section 2.2.11). PCR products were analysed on an agarose gel (see Section 2.2.12). Both the amplified PCR product and the pDENNIS vector were digested with *Nhe*I and *Nco*I (see Section 2.2.19), followed by ligation and subsequent transformation into XL2 cells and selection on LB agar plates containing ampicillin and chloramphenicol (see Sections 2.2.15 and 2.2.16). Transformants were screened using ISPCR (see Section 2.2.17) and positive clones sequenced (see Section 2.2.20). Electroporation of constructs into *C. jejuni* mutant strains was performed and screened as described previously with kanamycin and chloramphenicol respectively (see Sections 2.2.24 and 2.2.25). Complemented *C. jejuni* mutants were checked for the presence of the non-disrupted CDS and the CDS containing the $Km^R$ cassette using ISPCR and by sequencing (see Sections 2.2.17 and 2.2.20).

**2.2.27 RT-PCR**

RNA was isolated as described in Section 2.2.10. DNA contamination of the RNA sample was removed using the TURBO DNA-*free* kit (Ambion, Austin, U.S.A) using the following protocol:-

| | |
|---|---|
| 10X TURBO DNase buffer | - 5 µl |
| TURBO DNase (2 U/µl) | - 1 µl |
| RNA (0.2 – 1.0 µg/µl) | - 44 µl |

The reaction mixture was gently vortexed and incubated at 37°C for 30 minutes. 5 µl DNase Inactivation Reagent was added and mixed. The reaction was kept at room temperature for 5 minutes with occasional mixing followed by centrifugation for 90 seconds at 10,000 rpm. The supernatant containing the RNA was transferred to a fresh tube and used for cDNA synthesis by RT-PCR.

cDNA synthesis was performed using SuperScript III first-strand synthesis system for RT-PCR (Invitrogen). The following reaction was initially set up:-

| | |
|---|---|
| Random hexamers (50 ng/µl) | - 1 µl |
| dNTP mix (10 mM) | - 1 µl |
| RNA (1 pg - 5 µg) | - 7 µl |

The reaction was incubated at 65°C for 5 minutes and immediately transformed onto ice for a further minute. 10 µl of cDNA synthesis mix (see below) was added to each RNA/primer mixture and incubated for 10 minutes at 25°C, followed by 50 minutes at 50°C.

<u>cDNA synthesis mixture</u>

| | |
|---|---|
| 10X RT buffer | - 2 µl |
| $MgCl_2$ (25 mM) | - 4 µl |
| DTT (0.1 M) | - 2 µl |
| RNaseOUT (40 U/µl) | - 1 µl |
| SuperScript III RT (200 U/µl) | - 1 µl |

The reaction was terminated by incubating at 85°C for 5 minutes, followed by incubating on ice for 1 minute. The reaction was briefly centrifuged and 1 µl RNase H (2 U/µl) (Invitrogen) was added and incubated for 20 minutes at 37°C. Introduction of RNase H after first strand synthesis can increase the sensitivity of the PCR step by removing the RNA template from the cDNA-RNA hybrid molecule by digestion (Invitrogen). PCR reactions were performed and analysed as described in Section 2.2.11.

### 2.2.28 Designing 6XHis-tagged complementation primer

To allow protein purification and further studies of the CDS of interest, a 6XHis-tagged protein sequence was cloned into the complementation construct for detection purposes. Primers were designed as in Section 2.2.26, however the reverse primer contained a 6XHis-tag in between the CDS sequence and the stop codon (Figure 2.4).

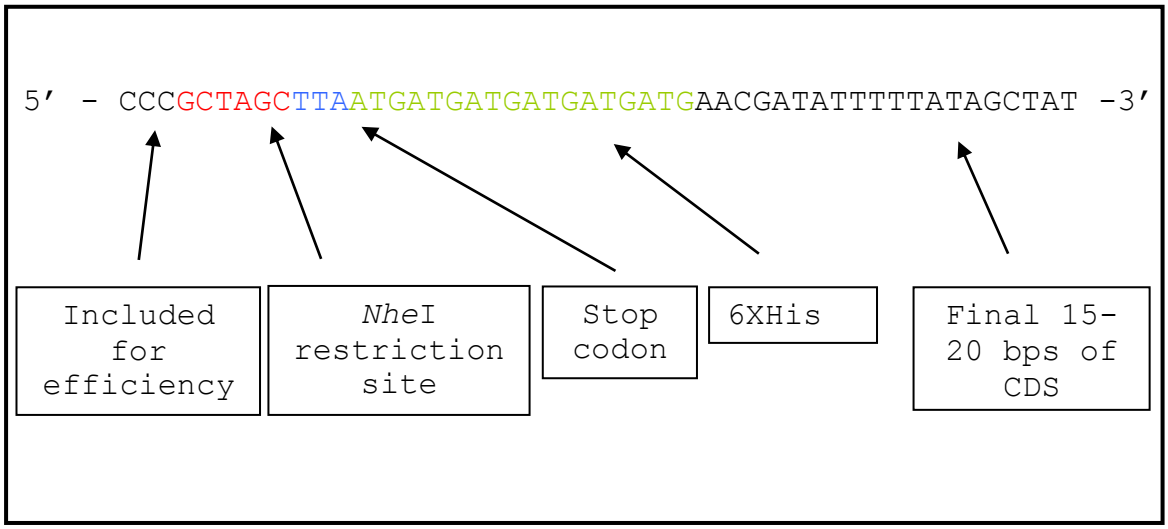



Figure 2.4. Reverse complement primer incorporating three cytosine residues, an *Nhe*I restriction site, a stop codon (TTA), 6XHis-tag, followed by 15-20 bps of the 3' end of the CDS sequence.

### 2.2.29 Protein sample preparation

#### 2.2.29.1 Whole-cell lysate preparation

Whole-cell lysates of *C. jejuni* or *E. coli* strains were prepared from either a plate grown for 24 h or a 20 ml culture grown for 16 h (see Section 2.2.7 for *C. jejuni* and *E. coli*

respectively). Bacteria from a plate were resuspended in 1 ml PBS. This suspension (or culture) was then spun at 4,000 rpm for 10 minutes at 4°C. The supernatant was discarded and the bacterial pellet was resuspended with 100 µl 1X Laemmli buffer (Sigma-Aldrich), then boiled for 10 minutes and spun at 13,000 rpm for 5 minutes. For experiments requiring normalisation of protein content between different strains, bacteria were resuspended in 1X Laemmli buffer at a ratio of 2 $OD_{600}$/100 µl. Protein samples were stored at -20°C.

#### 2.2.29.2 Whole-cell lysate preparation using sonication

Bacterial pellets obtained from the method described in Section 2.2.29.1 and were resuspended in 1 ml PBS. Bacterial samples were sonicated for 10 minutes (1 min on / 30 seconds off) using setting H (High) (Diagenode, Liege, Belgium). Samples were centrifuged after sonication for 5 minutes at 13,000 rpm. The supernatant containing lysed cell content was poured into a new 1.5 ml microcentrifuge tube. The samples were concentrated using a centrifugal filter column with a 10 kDa cut-off. Equal volume of 2X Laemmli buffer was added and samples were boiled for 10 minutes to ensure denaturing of proteins. Samples were centrifuged after sonication for 5 minutes at 13,000 rpm and were stored at -20°C.

#### 2.2.29.3 Purification of 6XHis-tagged proteins

Bacterial pellets were obtained from the method described in 2.2.29.1 and were resuspended in 1 ml equilibration buffer (Appendix 2). Lysed cells were obtained using the sonication method described in Section 2.2.29.2 were incubated with Ni-NTA (binding capacity up to 50 mg/ml) (Qiagen) for 1 h at 4°C on a rotator. Liquid chromatography (Pierce Biotechnology, Inc, Rockport, U.S.A) was performed using a His-Select spin column (Sigma-Aldrich). Equilibration, Wash and Elution buffers (Appendix 2) were used to elute the His-tagged protein. Three 2 ml aliquots were prepared and samples were concentrated using a centrifugal filter column with a 10 kDa cut-off. Equal volume of 2X Laemmli buffer was added and samples were boiled for 10 minutes to ensure denaturing of proteins. Samples were centrifuged for 5 minutes at 13,000 rpm and were stored at -20°C.

### 2.2.30 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

Protein samples were loaded onto 16% (w/v) Tris-Glycine gels (Invitrogen) and set-up in a Novex EI9001-XCell II Mini Cell (Invitrogen) and loaded with 1X running buffer

(Appendix 2). 20 µl of protein sample (prepared as described in 2.2.29) was loaded into each well. For elucidation of band sizes during sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent transfer to nitrocellulose membranes, a Spectra BR marker (Fermentas, Ontario, Canada) was used. Gels were run at 150 V for approximately 1 h.

### 2.2.31 Coomassie staining of SDS-PAGE gels

Coomassie staining was performed when analysing protein sample content. For staining, 0.125% (w/v) coomassie stain (see Appendix 2) (Amresco, Ohio, U.S.A) was used to stain the gel for 1 h shaking at 75 rpm, followed by destaining overnight. Destain solution (Appendix 2) was re-applied three times. Coomassie gels were visualised on a GeneGenius Gel-Documentation System (SynGene).

### 2.2.32 Western blots

Protein samples were loaded onto SDS-PAGE gels and were separated as described in 2.3.30. For elucidation of band sizes after transfer onto Hyperfilm ECL (GE Healthcare, Buckinghamshire, U.K), MagicMark XP (Invitrogen) was used. Whatman paper 3MM (Scientific Lab Supplies, Nottingham, U.K) and nitrocellulose Hybond-C extra (GE Healthcare) were cut into gel shapes and incubated in transfer buffer (Appendix 2). Completion of the SDS-PAGE was followed by transfer of the proteins from the gel to the nitrocellulose membrane, performed using a Hoefer SemiPhor System (GE Healthcare) set at 45 mA for 1.5 h (or 90 mA for two gels). A 1.5 kg weight was placed onto the transfer apparatus to ensure efficient transfer.

#### 2.2.33.1 Western blot 6XHis-tag detection

The following methodology was used for detection of 6XHis-tagged proteins. All stages were performed at room temperature unless stated otherwise. Whilst performing the transfer stage, blocking reagent was prepared. Blocking reagent was provided as part of the Penta-His HRP Conjugate kit (Qiagen) used for the detection of HIS-tagged proteins. Blocking reagent was prepared by adding 0.1 g of blocking reagent to 20 ml PBS containing 0.05% (v/v) Tween-20 and 0.2% (v/v) Triton X-100 (Sigma-Aldrich). Blocking was performed for either 1 h shaking at 75 rpm (Platform Shaker STR6, VWR-Jencons) or overnight at 4°C without shaking. Once blocking was completed, the Penta-His HRP conjugate antibody (Qiagen) was added at a dilution of 1:5000 (1 µl in 5 ml blocking reagent solution). The nitrocellulose membrane was incubated for 1 h with

shaking with the Penta-His HRP conjugate antibody in a square petri dish. The nitrocellulose membrane was then washed twice with PBS containing 0.05% (v/v) Tween-20 and 0.2% (v/v) Triton X-100 for 10 minutes, followed by PBS alone for 10 minutes. Detection was performed by chemoluminescence using ECL Plus Western blotting reagent pack (GE Healthcare) in a Hypercassette case (GE Healthcare). Chemiluminescent detection was performed using Hyperfilm ECL (GE Healthcare). Detection was performed after placing the Hyperfilm ECL on top of the nitrocellulose membrane inside the Hypercassette case for 1 minute and also for longer time points as required.

#### 2.2.33.2 Western blot polyclonal antibody detection

The following methodology was used for detection of proteins using a primary antibody. All stages were performed at room temperature unless stated otherwise. Whilst performing the transfer stage, blocking reagent was prepared. Blocking reagent was made up of 10% (w/v) dried skimmed milk (Tesco, Cheshunt, U.K) in PBS containing 0.05% (v/v) Tween-20. Blocking was performed for either 1 h shaking at 75 rpm or overnight at 4°C without shaking. Once blocking was completed, the primary antibody was added at a dilution of 1:5000 (1 µl in 5 ml blocking reagent solution). The dilution was prepared using PBS with 5% (w/v) skimmed milk and 0.05% (v/v) Tween-20. The nitrocellulose membrane was incubated with the primary antibody for 1 h with shaking in a square petri dish. The nitrocellulose membrane was then washed twice with PBS containing 0.05% (v/v) Tween-20 and 0.2% (v/v) Triton X-100 for 10 minutes. The secondary antibody (anti-rabbit IgG - whole molecule / peroxidase antibody produced in goat (Sigma-Aldrich)) was added at a dilution of 1:5000 (1 µl in 5 ml blocking reagent solution). The dilution was prepared using PBS with 5% (w/v) skimmed milk and 0.05% (v/v) Tween-20. The nitrocellulose membrane was incubated with the secondary antibody for 1 h shaking in a square petri dish. The nitrocellulose membrane was then washed twice with PBS containing 0.05% (v/v) Tween-20 and 0.2% (v/v) Triton X-100 for 10 minutes. The nitrocellulose membrane was finally washed with PBS for 10 minutes. Detection was performed by chemoluminescence using ECL Plus Western blotting reagent pack (GE Healthcare) in a Hypercassette case (GE Healthcare). Developing and detection was as described in Section 2.2.33.1.

### 2.2.35 Protein quantification

Protein concentrations were analyzed using a BCA assay (Thermo Scientific) following the manufacturers guidelines.

### 2.2.36 Gram stains

Gram stains were performed using a Diagnostics Kit (Sigma-Aldrich). A blank glass slide was spotted with Milli-Q water. A single colony was selected from an agar plate and smeared onto the spotted Milli-Q water on the slide and left to dry. The smear was heat fixed by passing the slide over a blue flame. The slide was then immersed in Crystal Violet for 45 seconds then rinsed with tap water. The slide was immersed in iodine for 45 seconds then rinsed again with tap water. 3 drops of decolouriser were added onto the stained section of the slide and then rinsed with water. Finally, Safranin O solution (counter-wash) was added onto the slide for 45 seconds. This was rinsed off with tap water and the slide dried. The slide was viewed using a Laborlux K microscope (Leica, Milton Keynes, U.K).

## 2.3 Assays

### 2.3.1 Primary culture growth kinetics

15 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h, shaking at 75 rpm. A bacterial suspension was prepared and the volume to inoculate the 15 ml Brucella broth to a starting $OD_{600}$ of 0.1 was calculated as described in Section 2.2.8. The calculated volume was removed from the pre-incubated Brucella broth. This volume of cells from the original bacterial cell suspension was then added to produce a final volume of 15 ml. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were performed at selected time points.

### 2.3.2 Secondary culture growth kinetics

15 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h, shaking at 75 rpm. A bacterial suspension was prepared and the volume to inoculate the 15 ml Brucella broth to a starting $OD_{600}$ of 0.1 was calculated as described in Section 2.2.8. This primary culture was grown for 16 h at 37°C under microaerobic conditions, shaking at 75 rpm. 15 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h. The primary culture was

used to inoculate the preincubated broth to a starting $OD_{600}$ of 0.1. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were performed at selected time points.

### 2.3.3 Motility assay

Brucella motility plates were prepared as described in Appendix 2. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared as described in Section 2.2.8. The suspension was briefly vortexed and 5 µl of this suspension pipetted into the centre of the motility plate using a P10 tip. Plates were incubated at 37°C under microaerobic conditions. Plate images were recorded at 24, 48 and 72 h using a GeneGenius Gel-Documentation System (SynGene). Motility was quantified by measuring the diameter of the halo on the plate.

### 2.3.4 Oxidative stress assay

A stock of 100 mM $H_2O_2$ was prepared by performing a 1/10 dilution with Milli-Q water from 1 M $H_2O_2$ (Sigma-Aldrich). Serial dilution tubes were also prepared up to $10^{-6}$ with 900 µl PBS in each tube. *C. jejuni* were grown on blood agar plates for 24 h. Bacteria were collected using a sterile swab and resuspended in 1 ml PBS. A bacterial suspension was prepared and the volume to inoculate into 900 µl broth to produce a starting $OD_{600}$ of 1.0 was calculated as described in Section 2.2.8. The suspension was briefly vortexed. 100 µl of 100 mM $H_2O_2$ was added to the test sample to produce a final concentration of 10 mM $H_2O_2$ and 100 µl Brucella broth added to the control sample. Both test and control tubes were incubated at 37°C under microaerobic conditions for 15 minutes with the caps open. Tubes were briefly vortexed and serial dilutions were performed down to $10^{-6}$. A 10 µl volume of each dilution was pipetted onto blood agar plate(s) and incubated uninverted at 37°C under microaerobic conditions for 48 h. Plate images were taken after 48 h using a GeneGenius Gel-Documentation System. The level of sensitivity was measured as log survival based on the serial dilutions.

### 2.3.5 Acidified nitrite assay

A stock of 100 mM sodium nitrite ($NaNo_2$) at pH 5 was prepared by performing 1/10 dilution with Milli-Q water from 1 M $NaNo_2$ (Sigma-Aldrich). The assay method used was as described in Section 2.3.4, with a final concentration of 10 mM $NaNo_2$ used. Both test and control tubes were incubated at 37°C under microaerobic conditions for 75 minutes. Further assays using a final concentration of 100 mM $NaNo_2$ with a 30 minute

incubation time at 37°C under microaerobic conditions were performed. The level of sensitivity was measured as log survival based on the serial dilutions.

### 2.3.6 Oxidative growth inhibition assay

10 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h. A bacterial suspension was prepared as described in Section 2.2.8 and used to inoculate the 10 ml Brucella broth to a starting $OD_{600}$ of 0.1. Control flasks had the calculated volume of Brucella broth removed and the same volume of *C. jejuni* added. Test flasks had the calculated volume of Brucella broth removed, in addition to another 500 µl Brucella broth. 500 µl of 20 mM $H_2O_2$ was added to the flask followed by the calculated cell volume. This gave a final concentration of 1 mM $H_2O_2$. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h.

### 2.3.7 Acidified nitrite growth inhibition assay

Acidified nitrite growth inhibition assays were performed basically as described in Section 2.3.6 with the modification that a 2 mM stock of $NaNo_2$ at pH 5 was prepared and added to the test flask containing Brucella broth with *C. jejuni* at a starting $OD_{600}$ of 0.1 to produce a final concentration of 0.1 mM $NaNo_2$.

### 2.3.8 Sodium deoxycholate growth inhibition assay

Sodium deoxycholate (sDOC) growth inhibition assays were performed basically as described in Section 2.3.6 with the modification that a 2% (w/v) sDOC stock was prepared and added to the test flask containing Brucella broth with *C. jejuni* at a starting $OD_{600}$ of 0.1 to produce a final concentration of 0.1% (w/v) sDOC.

### 2.3.9 Iron chelating (deferoxamine) growth inhibition assay

Iron chelating (deferoxamine) growth inhibition assays were performed basically as described in Section 2.3.6 with the following modifications. Stocks of deferoxamine at 1M, 100mM and 10mM were prepared and added to the flask containing Brucella broth with *C. jejuni* at a starting $OD_{600}$ of 0.1 to produce a final concentration of 0.01, 0.1 and 1 mM deferoxamine respectively.

### 2.3.10 Microarray experimental design, template labelling and hybridisation

Gene expression profiling of *C. jejuni* 11168H during the late log phase of growth (16 h) was performed using an indirect comparison method or type 2 experimental design (Kamal *et al*., 2007). In this experimental design, replicate test sets of Cy5-labelled *C. jejuni* 11168H total RNA samples were combined with a common reference sample (Cy3-labelled *C. jejuni* 11168H genomic DNA). The microarrays used in this study were whole genome *C. jejuni* NCTC11168 arrays printed on Ultragaps glass slides (Corning, NY, U.S.A), produced by the BμG@S Group (www.bugs.sghms.ac.uk).

The procedures for the Cy3-labelling of 11168H genomic DNA and Cy5-labelling of total RNA samples are detailed below. Initially, 50 ml pre-hybridisation solution was prepared and placed into a Coplin jar (Fisher). Reagents were added in the following order:-

| | |
|---|---|
| SDS 20% (v/v) | - 250 μl |
| Milli-Q water | - 36.0 ml |
| 20X saline sodium citrate (SSC) | - 8.75 ml |
| Bovine serum albumin (BSA) (100 mg/ml) | - 5.0 ml |

The pre-hybridisation solution was incubated at 65°C for 1 h before commencing the labelling reactions.

The following reaction was set up for labelling of control genomic DNA:-

| | |
|---|---|
| Genomic DNA | ≈ 2 μg |
| Random primers (3 μg/μl) (Invitrogen) | - 1 μl |
| Milli-Q water | to 41.5 μl |

Reactions were heated at 95°C for 5 minutes followed by snap cooling on ice and brief pulse centrifugation. Control Cy3 DNA labelling reactions were prepared as follows, adding reagents in the following order:-

| | |
|---|---|
| 10xREact 2 buffer (Invitrogen) | - 5 μl |
| dNTP's (5 mM each dATP, dGTP and dTTP, 2 mM dCTP) | - 1 μl |
| Cy3-labelled dCTP (25 mmol) (GE healthcare) | - 1.5 μl |
| Klenow fragment (10 U/μl) (Invitrogen) | - 1 μl |

The control reaction was incubated at 37°C for 90 minutes.

The following reaction was set up for labelling of test Cy5 total RNA:-

| | |
|---|---|
| Total RNA | ≈ 10 μg |
| Random primers (3 μg/μl) (Invitrogen) | - 1 μl |
| Milli-Q water | - to 11 μl |

Reactions were heated at 95°C for 5 minutes followed by snap cooling on ice and brief pulse centrifugation. Test reactions were prepared, adding reagents in the following order:-

| | |
|---|---|
| 5X first strand buffer (Invitrogen) | - 5 μl |
| DTT (100mM) | - 2.5 μl |
| dNTP's (5 mM each dATP, dGTP and dTTP, 2 mM dCTP) | - 2.3 μl |
| Cy5-labelled dCTP (25 mmol) (GE Healthcare) | - 1.7 μl |
| Reverse transcriptase (Superscript II 200 U/μl) (Invitrogen) | - 2.5 μl |

The Test Cy5-labelled reaction was incubated at 25°C for 10 minutes, followed by incubating at 42°C for 90 minutes. 20 minutes prior to completion of this incubation period, microarray slide(s) were placed into the pre-hybridisation solution in the Coplin jar and incubated at 65°C. Both control and test reactions were then combined in a 1.5 ml microcentrifuge tube. The reaction was purified using a MinElute PCR purification kit (Qiagen). 500 μl buffer PB was added to the combined reaction and centrifuged using a 2 ml MinElute column for 1 minute at 13,000 rpm. 250 μl buffer PE was added to the column and centrifuged for 1 minute at 13,000 rpm. The cDNA was eluted from the column by adding 14 μl Milli-Q water. Samples were incubated on ice. Pre-hybridised microarray slides were rinsed in water for 1 minute using a trough and slide holder (Raymond A. Lamb, Eastbourne, U.K). Slides were then rinsed in 100% isopropanol for

1 minute. Slides were dried by placing into 50 ml Falcon tubes (Fisher) and centrifuging at 1500 rpm for 5 minutes. The Hybridisation solution was prepared as follows in a 1.5 ml microcentrifuge:-

| | |
|---|---|
| Eluted sample | - 13 µl |
| Milli-Q water | - 26 µl |
| 20X SSC | - 12 µl |
| 2% (v/v) SDS | - 9 µl |

This hybridisation solution was then heated at 95ºC for 2 minutes and allowed to cool at room temperature. The reaction was briefly pulse centrifuged.

Microarray slide(s) were placed in a humidified hybridisation cassette (Telechem International, Sunnyvale, U.S.A) and covered with a LifterSlip™ (Erie Scientific, Portsmouth, U.S.A) glass coverslip (22 mm x 25 mm) over the array section. The hybridisation solution was carefully pipetted under the coverslip. The microarray chamber was sealed and incubated in a water bath at 65°C overnight without shaking. Wash A and B solutions were prepared as below:-

Wash A

| | |
|---|---|
| 20X SSC | - 20 ml |
| 20% (v/v) SDS | - 1 ml |
| Milli-Q water | - 379 ml |

Wash B

| | |
|---|---|
| 20X SSC | - 2.4 ml |
| Milli-Q water | - 797.6 ml |

Wash A solution was incubated overnight at 65°C. Prior to performing the wash protocol, Wash A solution was dispensed into a pre-rinsed single trough. Wash B was dispensed into two troughs. The microarray slides were removed from the chambers and placed into the slide holder in Wash A. This was gently agitated for 5 minutes. The slides were transferred to a clean slide holder and rinsed with gentle agitation in Wash B for 2 minutes. This was followed by another 2 minutes in the second Wash B trough. Microarray slides were placed into 50 ml Falcon tubes (label away from lid), and

centrifuged at 1500 rpm for 5 minutes to dry. Microarray experiments were performed with three biological replicates, each with one technical replicate.

### 2.3.11 Data acquisition and microarray data analysis

The microarray slides were scanned with an Affymetrix 418 array scanner (MWG Biotech) according to the manufacturer's guidelines. Signal and local background intensity readings for each spot were quantified using ImaGene software v8.0 (BioDiscovery, El Segundo, U.S.A). Quantified data were analysed using GeneSpring GX software v7.2 (Agilent, Santa Clara, U.S.A). Expression analysis was performed using a DNA versus RNA experimental set-up. Statistically significant up- and down-regulated genes were selected when comparing gene expression from mutant test arrays against 11168H control arrays using ANOVA (ANalysis Of VAriance) (Bacon *et al.*, 2004, Corcionivoschi *et al.*, 2009). ANOVA was performed using a Benjamini and Hochberg False Discovery Rate as the Multiple Testing Correction within the GeneSpring software. All microarray data was deposited into ArrayExpress.

### 2.3.12 *C. jejuni* co-culture studies

Co-culture experiments were set-up in a sterile hood (Envair, Haslingden, U.K). The Caco-2 human intestinal epithelial and J774A.1 mouse macrophage cell lines were maintained using the following media:-

| | |
|---|---|
| Dulbecco's Modified Eagle's Medium (DMEM) | - 500 ml |
| FCS | - 50 ml |
| Non-essential amino acids | - 5 ml |
| Penicillin (100 U/ml) / Streptomycin (100 µg/ml) solution | - 5 ml |

The T84 human colonic epithelial cell line was maintained using the following media:-

| | |
|---|---|
| 1:1 mixture of DMEM and Ham's F-12 medium containing Glutamax®, 2.5 mM l-glutamine, 15 mM HEPES, and 0.5 mM sodium pyruvate (Gibco) | - 500 ml |
| FCS | - 50 ml |
| Penicillin (100 U/ml) / Streptomycin (100 µg/ml) solution | - 5 ml |

Cells were maintained in a $CO_2$ incubator (Sanyo). For Caco-2 cell co-culture experiments, cells were seeded at ≈ 1 x $10^5$ cells/ml and grown in 24-well plates to >90%

confluence (≈ 1 x $10^6$ cells/ml). For T84 cell co-culture experiments, cells were seeded at ≈ 5 x $10^5$ cells/ml and grown in 24-well plates to >90% confluence (≈ 5 x $10^6$ cells/ml). For co-culture experiments involving J774A.1 mouse macrophage, cells were seeded at ≈ 5 x $10^5$ cells/ml and grown in 24-well plates for 24 h. For ELISA experiments, T84 cells were maintained in low serum 1% (v/v) and antibiotic-free media overnight prior to co-culture. Supernatants for ELISA experiments were stored at -80°C until required.

Caco-2, T84 and J774A.1 cells reaching a confluent state were split and/or seeded for future experiments. Cells were washed three times with 10 ml PBS. 5 ml 0.25% (v/v) Trypsin-EDTA solution was added to the flask containing cells and incubated in a 37°C $CO_2$ incubator for approximately 7 minutes to allow removal of cells (20 minutes for J774A.1 cells). 20 ml culture media was added to the cells and pipetted vigorously. Media containing cells was added to a 50 ml Falcon with culture media added to make a final volume of 50 mls. Cells were centrifuged at 1500 rpm for 10 minutes. The supernatant was decanted and cells were resuspended in 1 ml culture media. 9 ml culture media was added to complete a 10- fold dilution. 3 ml of this suspension was mixed with 27 ml culture media in a new 50 ml Falcon, giving ≈ 1 x $10^5$ cells/ml. All 30 mls were pipetted into a new 75 $cm^2$ flask (Corning) and incubated at 37°C in a $CO_2$ incubator. Dilutions were adjusted for different cell lines.

To calculate the required cells for interaction, invasion, intracellular survival and macrophage survival assays, seeding was prepared at ≈ 1 x $10^5$ cells/ml (e.g. for Caco-2 cells). 800 µl DMEM, 100 µl Trypan Blue (Sigma-Aldrich) solution and 100 µl cells from the cell suspension (≈ 1 x $10^6$ cells/ml) were briefly mixed in a 1.5 ml microcentrifuge tube. This suspension was added to a Neubauer cell counter (Weber, Hamilton New Jersey, US) and covered with a slide (VWR). Cells were counted on a light microscope (DM1000 Leica, Milton Keynes, U.K) and calculated as follows:-

Number of cells in total area within cell counter x dilution factor (1000) x dilution (10) = approximate number of cells

Appropriate dilutions were performed to obtain ≈ 1 x $10^5$ cells/ml. 1 ml volumes were seeded onto 24 well plates (Corning) and were incubated at 37°C in a $CO_2$ incubator. All reagents were obtained from Invitrogen unless otherwise stated.

### 2.3.13 Interaction, invasion and intracellular survival assays

Interaction (adhesion and invasion), invasion and intracellular survival assays were performed using Caco-2 cells. All assays used the following media:-

| | |
|---|---|
| DMEM | - 500 ml |
| FCS | - 5 ml |
| Non-essential amino acids | - 5 ml |

*C. jejuni* cells from a 24 h grown bacteria plate were collected using a sterile swab and resuspended in a 1.5 ml microcentrifuge tube containing 1 ml DMEM. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared as described in Section 2.2.8 using DMEM as the blank. The volume was made up with DMEM to 2 ml. 1 ml of this mixture was added to a 30 ml universal and 9 ml DMEM was added, to produce a final $OD_{600}$ of 0.1 (≈ 1 x $10^8$ cells). 100 µl of this suspension was taken and serial dilutions performed by resuspending in a 1.5 ml microcentrifuge tube containing 900 µl PBS. Serial dilutions were performed down to $10^{-6}$. 200 µl volume from the final tube was pipetted and spread onto three dried blood agar plates. Plates were incubated at 37°C under microaerobic conditions for 72 h. Plates were inverted after 24 h. Colonies were counted on each plate and this data was used to calculate the initial inoculum (CFU/ml). *C. jejuni* cells (≈ 1 x $10^8$ cells) in DMEM were added to a monolayer of ≈ 1 x $10^6$ Caco-2 cells (MOI 100:1) and incubated for 3, 6 or 24 h. The number of interacting bacteria was determined by washing the monolayers three times with PBS then lysing the cells by addition of 0.2% (v/v) Triton X-100 in PBS. The number of intracellular bacteria was determined by further incubating the monolayers after the initial interaction time point with 150 µg/ml gentamicin in DMEM for 2 h at 37°C to allow killing of extracellular bacteria. Monolayers were then washed three times in PBS and the epithelial cells lysed as above. For intracellular survival assays, bacterial cells were co-cultured with a monolayer of Caco-2 cells for 3 h following by washing the monolayers three times with PBS. The monolayers were then incubated in DMEM containing gentamicin (150 µg/ml) for 2 h and then incubated in DMEM containing a reduced concentration of gentamicin (10 µg/ml) for 19 h (Naito *et al.*, 2010). Monolayers were then washed three times in PBS and the epithelial cells lysed as above. Bacterial survival in the tissue culture medium during co-culture experiments was also investigated after 24 h co-culture, the tissue culture medium alone was removed followed by plating of serial dilutions to

determine the CFU/ml. In all cases, serial dilutions, plating and enumeration of bacterial numbers were performed as stated above.

### 2.3.14 Macrophage survival assays

Macrophage survival assays were performed using J774A.1 mouse macrophages. A bacterial suspension with an $OD_{600}$ of 0.1 was prepared in DMEM as described in Section 2.3.13. *C. jejuni* cells (≈ 1 x $10^8$ cells) were added to a culture of ≈ 5 x $10^5$ J774A.1 mouse macrophage cells (MOI 200:1) and incubated for 3 h. Cells were washed three times in PBS followed by incubation in DMEM containing gentamicin (150 µg/ml) for 2 h to allow killing of extracellular bacteria. The macrophages were then incubated in DMEM containing a reduced concentration of gentamicin (10 µg/ml) and bacterial survival determined at 0, 4 and 16 h. At each time point, the macrophages were washed three times with PBS and lysed by adding 0.2% (v/v) Triton X-100. Serial dilutions, plating and enumeration of bacterial numbers were performed as stated in Section 2.3.13.

### 2.3.15 ELISA

Levels of IL-8 and IL-6 secreted from T84 cells during co-culture with *C. jejuni* were quantified using a human IL-8 and IL-6 ELISA development kit (Peprotech, New Jersey, U.K). For ELISA experiments, T84 cells were maintained in antibiotic-free media with 1% (v/v) FCS overnight prior to co-culture. A bacterial suspension with an $OD_{600}$ of 2 was prepared as described in Section 2.3.13. 50 µl of this suspension was added to each well of the prepared T84 cells and incubated in a 37°C $CO_2$ incubator (Sanyo). This ensured a dilution of 1:20 and produced a working $OD_{600}$ of 0.1 (≈ 1 x $10^8$ CFU/ml). Supernatants from uninfected T84 cells and T84 cells infected with *C. jejuni* (MOI of 20:1) for 24 h were removed and placed into 1.5 ml microcentrifuge tubes and spun for 1 minute at 13,000 rpm. The prepared samples were stored at -80°C and used for ELISA.

ELISA 96 well microplates were used during the assay (Nunc, Roskilde, Denmark). ELISA assays were performed following the manufacturer's instructions. Briefly, 100 µl capture antibody (0.5 µg/ml) was added to each well, plates were sealed, then incubated overnight at room temperature. The wells were aspirated four times with 300 µl wash buffer (0.05% (v/v) Tween-20 in PBS). 300 µl block buffer (1% (w/v) BSA in PBS) was added to each well and incubated for 1 h at room temperature. Wells were aspirated four times with 300 µl wash buffer. Human IL-8 Standard (1 µg/ml) was diluted to 2 ng/ml using diluent (0.05% (v/v) Tween-20, 0.1% (w/v) BSA in PBS). Serial dilutions of this 2

ng/ml standard were performed using a 1:2 dilution to 0 using diluent. 100 µl of each standard and samples were added to duplicate wells. The plates were incubated at room temperature for 2 h. Wells were then aspirated four times with wash buffer. 100 µl detection antibody (0.25 µg/ml) was added to each well. The plate was incubated at room temperature for 2 h. Wells were then aspirated four times with wash buffer. 100 µl avidin-HRP conjugate was added at a dilution of 1:2000 using diluent and incubated at room temperature for 30 minutes. Wells were aspirated four times with PBS and 100 µl 2,2'-azino-bis-(3-benzthiazoline-6-sulfonic acid) (ABTS) substrate solution (Sigma) was added to each well. Plates were incubated at room temperature for 10 minutes. Detection was performed using a Dynex MRX II 96 well plate reader (Dynex, Chantilly, U.S.A) at an absorbance of 405 nm (*A*405) and analysed on Revelation software (Dynex, Chantilly, U.S.A).

### 2.3.16 *Galleria mellonella* model of infection

*G. mellonella* larvae (LiveFoods Direct, Sheffield, U.K) were kept on wood chips at room temperature. A 16 h secondary culture was prepared as described in Section 2.2.8, and adjusted to $OD_{600}$ of 0.1. Larvae were injected with a 10 µl inoculum of the *C. jejuni* culture by micro-injection (Hamilton, Switzerland) in the right foremost leg, giving an infectious dose of $\approx 10^6$ CFU. Larvae were incubated at 37ºC with survival and appearance recorded at 24 h intervals. Brucella broth injection and no injection controls were also prepared. For each experiment, 10 *G. mellonella* larvae were infected and experiments were repeated three times. *G. mellonella* larvae were scored as alive if able to rotate fully onto front when placed onto their back and also when demonstrating sufficient twitching in the presence of contact.

### 2.3.17 Biofilm formation assay

10 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h. A bacterial suspension with an $OD_{600}$ of 0.1 was prepared as described in Section 2.2.8. 1 ml of this suspension was added to 24 well plates (Corning) and plates were incubated at 37°C under microaerobic conditions with shaking at 75 rpm for 3 days. Wells were washed three times with PBS, followed by addition of 1 ml 0.2% (w/v) crystal violet for 10 minutes. Wells were then washed with PBS three times followed by dissolving the crystal violet stain with 1:4 acetone/ethanol (v/v) followed by recording the absorbance at 595 nm using a ELx800 Absorbance Microplate Reader (BioTek, Winooski, U.S.A).

### 2.3.18 Secretion profile analysis of *C. jejuni*

30 ml Brucella broth was preincubated in a 150 ml flask at 37°C under microaerobic conditions for 24 h. *C. jejuni* were grown on blood agar plates for 24 h. A bacterial suspension was prepared as described in Section 2.2.8 and used to inoculate the 30 ml Brucella broth to a starting $OD_{600}$ of 0.1. The culture was incubated for 16 h with shaking at 75 rpm, followed by centrifuged for 1 h at 6,000 rpm. Filter sterilisation was performed using 0.2 µm Supor Acrodisc, 32 mm filter (Pall Life Sciences). Samples were concentrated using a centrifugal filter column with a 10 kDa cut-off. 0.1% (v/v) FCS serum and/or 0.1% (w/v) sDOC were also used to supplement the culture media at the point of inoculation. Negative controls contained no bacteria. Protein samples were mixed with equal volume of 2X Laemmli buffer, then boiled for 10 minutes and spun at 13,000 rpm for 5 minutes. In order to semi-quantify any differences in protein secretion between strains, the starting $OD_{600}$ and growth conditions were standardized. 30 µl of final sample were analysed by SDS-PAGE as described in Sections 2.2.30.

### 2.3.19 Transmission Electron Microscopy

Samples for TEM were prepared from both solid and liquid media. *C. jejuni* were grown on blood agar plates for 24 h at 37°C under microaerobic conditions. Alternatively, 10 ml Brucella broth was preincubated in a 50 ml flask at 37°C under microaerobic conditions for 24 h. *C. jejuni* grown on blood agar plates for 24 h at 37°C under microaerobic conditions was used to prepare a bacterial suspension with an $OD_{600}$ of 0.1. This bacterial culture was prepared as described in Section 2.2.8. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm for 24 h. Individual colonies of *C. jejuni* from blood agar plates or 10 µl of *C. jejuni* culture were resuspended in a 1.5 ml microcentrifuge tube containing 2.5% (v/v) glutaraldehyde / 2.5% (v/v) paraformaldehyde / 0.1 M sodium cacodylate buffer pH 7.4. TEM was performed by an in-house technician (Maria McCrossan).

### 2.3.20 Autoagglutination ability of *C. jejuni*

*C. jejuni* from a 24 h grown blood agar plate were collected using a sterile swab and resuspended in a 1.5 ml microcentrifuge tube containing 1 ml PBS. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared as described in Section 2.2.8. The appropriate volume was made up with PBS to 10 ml in glass tubes. Cultures were incubated for 24 h at 37°C under microaerobic conditions. 1 ml of the suspension (taken

from approximately the 5 ml mark of the glass tube) was measured using $OD_{600}$ after 24 h.

# Chapter 3: Re-annotation and re-analysis of the *C. jejuni* NCTC11168 genome sequence

## 3.1 Introduction

### 3.1.1 Aims

Since the publication of the *Haemophilus influenzae* strain Rd KW20 genome sequence in 1995 (Fleischmann *et al*., 1995), the availability of genome sequence data has increased at an exponential rate. To date (July 2010), there have been 12,567 genomes sequenced, however only a handful of these have had re-annotations performed (http://www.ncbi.nlm.nih.gov/sites/genome). Genomes which are not re-annotated can start to hinder research as scientists may rely on out of date information. Re-annotation is required as new methods for gene analysis become available and as our knowledge of gene function increases over time. The *C. jejuni* strain NCTC11168 was the second genome to be completely sequenced and annotated at the WTSI after *Mycobacterium tuberculosis* strain H37Rv. The aims of this work were to perform a complete re-annotation of the *C. jejuni* NCTC11168 genome sequence and to use this new information to select and characterise CDSs with an uncertain function, but a possible role in virulence, signal transduction or regulation of gene expression. The original genome sequencing project of *C. jejuni* strain NCTC11168 was published in 2000 (Parkhill *et al*., 2000) and since this time considerable research has been undertaken on *C. jejuni* and other bacteria. Re-annotation of *C. jejuni* NCTC11168 was undertaken in collaboration with Stephen Bentley, Matt Holden and Julian Parkhill in accordance with WTSI annotation guidelines.

### 3.1.2 Genome annotations and re-annotations

Once a genome sequence has been generated, the primary aim is to identify potential open reading frames (ORFs) and annotate each ORF accordingly as a predicted CDS. The annotation process has traditionally been performed manually whereby similarity searches, motif searches and also literature searches are performed by dedicated scientists (annotators). However the recent scale of genome sequence generation has now made it near impossible to undertake manual annotation as the time required to generate a genome sequence has decreased so rapidly that there is insufficient resources to perform manual annotations. The problem has become even more acute with the advent

of next-generation sequencing technology and is likely to continue for the foreseeable future. As a result, automated annotation procedures are increasingly used (Friedberg, 2006, Stothard and Wishart, 2006, Brent, 2008).

Re-annotation is defined as the process of re-analysis of a previously annotated genome (Ouzounis and Karp, 2002). A re-annotation is performed to allow incorporation of data generated since the original annotation was produced. Examples of re-annotated genomes are rare compared to the number of sequenced genomes (Camus *et al*., 2002, Dandekar *et al*., 2000). Re-annotation projects are typically performed as a result of large amounts of new information becoming available since the original annotation was produced and as such rendering the original annotation a less useful resource. An example is the *Mycobacterium tuberculosis* H37Rv genome (the first genome to be completed at the WTSI) where a re-annotation was performed four years after the original annotation, leading to updates on 300 CDSs. Currently, new sequenced genomes receive priority when annotation resources are available. However due to the number of genome sequences currently being generated, the resources simply do not exist for all these new genome sequences to be annotated manually. Manual annotation is time consuming and costly, however automated techniques do not yet provide the level of detail and accuracy that manual techniques provide. Thus, it is no surprise that re-annotations of existing genomes are so rare. Re-annotation of genome sequences can significantly reduce the time spent on misdirected research. For example, attempts at characterising a specific CDS may be wasteful on time and resources if the CDS had already been studied, but not updated in terms of annotation, or if the original annotation was incorrect. An outdated annotation can lead to gaps in our knowledge, particularly if genome information is already available. Hence, there is a need for a research review and regular update of the genome annotation of any given organism.

### 3.1.3 Methods available for annotating genomes

Newly sequenced genomes have ORF prediction performed using automated programs such as ORPHEUS (Frishman 1998) or Glimmer (Delcher, 2007), followed by sequence and additional motif searches. Most sequenced genomes will have an automated annotation performed that uses genome wide comparison searches such as BLAST (Altschul *et al*., 1990) or FASTA (Pearson and Lipman, 1988) to find characterised orthologues. This can be a quick and efficient way of providing a basic annotation for a genome sequence. The incorporation of additional specialist searches and specific tools

can further enhance the annotation process, such as running motif databases like PFAM (Sonnhammer *et al*., 1998a) or PROSITE (Falquet *et al*., 2002). The more up-to-date the genome, the more useful the annotation is for researchers utilising this resource.

### 3.1.4 Comparison of manual and automated annotation methods

The process of manual annotation takes far longer to perform than automated annotation, however there are a number of advantages in a manual approach. Manual annotation allows the incorporation of knowledge from experts on the organism of interest. In particular, it is often the case that annotators have actually worked on the organism. The annotator will be able to use multiple sources of information and resources for the annotation process that will not be possible with an automated process. One such example demonstrating the importance of manual annotation is from the original *C. jejuni* NCTC11168 genome sequencing project. The original annotation identified a number of hypervariable sequences that encoded phase-variable CDSs (Parkhill *et al*., 2000). Until then, such homopolymeric tracts were assumed to be a result of sequencing errors (Saunders *et al*., 1998). This manual identification of phase-variable CDSs has also been performed for a number of *Helicobacter pylori* genome sequences whereby a re-analysis of the genomes altered the number of phase-variable CDSs from 36 to 46 (Salaun *et al*., 2004).

The majority of genome annotations are currently performed by automated methods, however the accuracy of these approaches has been questioned since the beginning of the sequencing era (Poptsova and Gogarten, 2010). Automated annotation methods also rely on genome-wide searches. However, based on information from additional searches and programs, a more informative decision can be made in the designation of product function during a manual annotation. Indeed, it is exactly this point which is often debated during the annotation process. When does a CDS become a fully characterised gene considering all the annotation information? In some cases, CDSs that have been characterised as genes will have a definitive product function. However this is often not the case, so many CDSs have a ‘putative’ designation as the annotator cannot be 100% sure of the product function. The annotator must rely on a wide range of sources to finally make an informed decision on the designation of the product function, especially if they have research experience with the particular organism. In addition, manual annotators can take advantage of additional resources such as recent characterisation

papers. This advanced level of data integration will have far greater benefits for downstream research (Smith and Zhang, 1997).

Annotation can be performed at different levels of specificity. Different annotators will use different thresholds for deciding whether the acquired information is worthy of designating a product function to a CDS and also whether a 'putative' designation is more appropriate. The 'putative' designation applies to a CDS when evidence is present to support the product function, but this evidence is inconclusive. The majority of product function designations are made using similarity search results. An accepted criteria is to use 35% sequence similarity as a benchmark for naming a putative product function and then to use supplementary data to ascertain an unanimous designation (Sander and Schneider, 1991). Errors exist in the assignment of CDS function mainly due to use of databases that are still of insufficient quality (Bork, 2000). Problems can also exist with product function assignment where low similarity search results are used to incorrectly assign function (Bork, 2000).

In summary, automated annotation methods can save time and resources, but will not incorporate the maximum information available from experienced manual annotators, leading to incomplete or even false designations. The main drawbacks of manual annotation or re-annotation are that they are subjective, costly (in terms of curator wages) and time consuming.

### 3.1.5 Sequencing of the *C. jejuni* NCTC11168 genome

*C. jejuni* NCTC11168 was selected for sequencing as it is genetically tractable and widely available. This strain was isolated in 1977 from a U.K patient with severe gastroenteritis (Skirrow, 1977). The sequencing process involved construction of seven libraries in pUC18 using size fractions ranging from 1.0 kb to 2.2 kb (Parkhill *et al*., 2000). Approximately 19,400 pUC clones were sequenced using Dye-terminator chemistry on ABI 373 and 377 sequencing machines. The sequencing project was initiated in 1997 and completed in 1999 (Parkhill *et al*., 2000). Annotation was performed at the WTSI and the genome was published in 2000 (Parkhill *et al*., 2000).

## 3.2 Results

### 3.2.1 Re-annotation overview of the *C. jejuni* NCTC11168 genome sequence

A complete re-annotation of the *C. jejuni* NCTC11168 genome was performed according to WTSI guidelines and submitted to EBI. The updated EMBL file (AL111168) is publicly available (http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?val=AL111168.1). The re-annotation has been acknowledged on the WTSI website (http://www.sanger.ac.uk/Projects/C_jejuni/). This re-annotation resulted in a reduction in the total number of CDSs in the NCTC11168 genome sequence from 1654 to 1643. The reduction in the number of CDSs was due to either the merging of adjacent CDSs or the removal of CDSs. Based on the annotation criteria, a re-analysis of pseudogenes was performed on the *C. jejuni* NCTC11168 genome (see Section 3.2.3). CDSs designated as pseudogenes were updated to reflect the complete amino acid sequence for the encoded protein regardless of the number of multiple reading frames represented in the genome sequence. The implication for this type of designation is that CDSs from multiple reading frames can be joined to represent one single pseudogene or CDS. For example, *Cj0290c*, *Cj0291c* and *Cj0292c* were merged to represent a single CDS on multiple frames denoting a pseudogene and re-annotated as *Cj0292c* (putative glycerol-3-phosphate transporter). *Cj0968* and *Cj0969* were also merged to represent a single CDS on multiple frames denoting a pseudogene and re-annotated as *Cj0969* (putative periplasmic protein). All merged CDSs representing pseudogenes were named after the most downstream CDS (e.g. the new name for *Cj0290c*, *Cj0291c* and *Cj0292c* is now *Cj0292c*). Phase-variable CDSs that contained an intersecting homopolymeric region between adjacent CDSs on separate frames were also merged (see Section 3.2.4). This allowed the complete amino acid sequence for newly merged adjacent CDSs to be represented, regardless of possible variation in the size of the homopolymeric region. Re-interpretation of phase-variable CDSs resulted in removal of seven CDSs. All merged CDSs representing phase-variable CDSs linked by a homopolymeric region were named after the upstream CDS. For example, the new name for *Cj0031* and *Cj0032* is *Cj0031* (see Section 3.2.4). The *Cj1520* CDS was removed because of the identification of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) structural moieties (Schouls *et al*., 2003) (see Section 3.2.5). In total, 11 CDSs were removed from the re-annotated sequence and are listed in Table 3.1.

Table 3.1. CDSs from the original *C. jejuni* NCTC11168 annotation that were merged or removed during re-annotation.

| **Original CDS Number** | **Original Annotation** | **Type/Description** | **New CDS Number** | **Updated Annotation** |
|---|---|---|---|---|
| *Cj0031* | probable DNA restriction/modification enzyme, N-terminal half | *Cj0031*/*Cj0032* - Phase-variable | *Cj0031* | putative type IIS restriction/modification enzyme |
| *Cj0032* | probable DNA restriction/modification enzyme, C-terminal half | *Cj0031*/*Cj0032* - Phase-variable | *Cj0031* | putative type IIS restriction/modification enzyme |
| *Cj0170* | unknown | *Cj0170*/*Cj0171* - Phase-variable | *Cj0170* | hypothetical protein *Cj0170* |
| *Cj0171* | unknown | *Cj0170*/*Cj0171* - Phase-variable | *Cj0170* | hypothetical protein *Cj0170* |
| *Cj0290c* | probable pseudogene representing the C-terminus of Cj0291c (glpT') | *Cj0290c*/*Cj0291c*/*Cj0292c* - Pseudogenes | *Cj0292c* | pseudogene (putative glycerol-3-phosphate transporter) |
| *Cj0291c* | glycerol-3-phosphate transporter, possible pseudogene | *Cj0290c*/*Cj0291c*/*Cj0292c* - Pseudogenes | *Cj0292c* | pseudogene (putative glycerol-3-phosphate transporter) |

| | | | | |
|---|---|---|---|---|
| *Cj0292c* | probable pseudogene representing the N-terminus of Cj0291c (glpT') | *Cj0290c*/*Cj0291c*/*Cj0292c* - Pseudogenes | *Cj0292c* | pseudogene (putative glycerol-3-phosphate transporter) |
| *Cj0628* | probable lipoprotein | *Cj0628*/*Cj0629* - Phase-variable | *Cj0628* | putative lipoprotein |
| *Cj0629* | possible lipoprotein | *Cj0628*/*Cj0629* - Phase-variable | *Cj0628* | putative lipoprotein |
| *Cj0968* | probable periplasmic protein | *Cj0968*/*Cj0969* - Pseudogenes | *Cj0969* | pseudogene (putative periplasmic protein) |
| *Cj0969* | probable pseudogene | *Cj0968*/*Cj0969* - Pseudogenes | *Cj0969* | pseudogene (putative periplasmic protein) |
| *Cj1144c* | unknown | *Cj1144c*/*Cj1145c* - Phase-variable | *Cj1144c* | hypothetical protein *Cj1144c* |
| *Cj1145c* | unknown | *Cj1144c*/*Cj1145c* - Phase-variable | *Cj1144c* | hypothetical protein *Cj1144c* |
| *Cj1325* | unknown | *Cj1325*/*Cj1326* - Phase-variable | *Cj1325* | putative methyltransferase |
| *Cj1326* | unknown | *Cj1325*/*Cj1326* - Phase-variable | *Cj1325* | putative methyltransferase |

| | | | | |
|---|---|---|---|---|
| *Cj1335* | unknown | *Cj1335*/*Cj1336* - Phase-variable | *Cj1335* | motility accessory factor (function unknown) |
| *Cj1336* | unknown | *Cj1335*/*Cj1336* - Phase-variable | *Cj1335* | motility accessory factor (function unknown) |
| *Cj1520* | hypothetical protein | *Cj1520* - removed | CRISPR region identified | CRISPR |
| *Cj1677* | probable lipoprotein | *Cj1677*/*Cj1678* - Phase-variable | *Cj1677* | putative lipoprotein |
| *Cj1678* | possible lipoprotein | *Cj1677*/*Cj1678* - Phase-variable | *Cj1677* | putative lipoprotein |

### 3.2.2 Re-annotation of CDS function

A systematic re-annotation of all 1654 CDSs in the original annotation of the NCTC11168 genome sequence was performed. As a result of this re-annotation, all CDSs with additional new information have had an 'updated' note qualifier attached to the published EMBL file (Figure 3.1). Qualifiers are the different categories for describing each CDS. The 'colour' and 'literature' qualifiers were only displayed in the EMBL file deposited at the WTSI and GeneDB (analysis website of WTSI genomes - http://www.genedb.org) due to EMBL restrictions at EBI. A 'colour' qualifier refers to the colouring scheme adopted by the WTSI – Pathogen Genomics Department (http://www.sanger.ac.uk/Projects/Microbes/) to allow classification of functionality for each CDS. The colour qualifier is used in the genome viewing software Artemis produced by the WTSI (Rutherford *et al*., 2000). Importantly, the 'colour' qualifier adopted by the WTSI does not correlate with the more detailed functional classification system defined by the 'class' qualifier (Appendix 5 – Additional File 1). The six broad category headings of this 'class' qualifier are:-

1. Small molecule metabolism
2. Broad regulatory
3. Macromolecule metabolism
4. Cell processes
5. Other
6. Miscellaneous

The 'literature' qualifier refers to a new qualifier adopted in this particular re-annotation whereby relevant literature was attached to each individual CDS (Figure 3.1). The 'updated' note qualifier contains consistent free-hand descriptions on recently identified motifs, relevant similarity search results and any published characterisation work performed in *Campylobacter* species/strains or any orthologues in related organisms. Additionally, the 'updated' note qualifier also contains reasoning for designating a product function for a CDS as putative or definitive. Putative designations infer an accepted product function without definitive evidence. For each CDS, a full literature search was performed. Detailed statistics on genome modifications are given in Table 3.2. The product functions of 299 (18.2% CDSs) were updated (Appendix 5 – Additional File 5). The classification of each CDS into a 'functional classification' was adopted using WTSI classification guidelines (Appendix 5 – Additional File 4). Out of the 299

CDSs with an updated product function, 181 (60.5% CDSs) now have a different functional classification. One example is *Cj1546*, where this CDS was originally annotated as a 'hypothetical protein'. The re-annotation changed the product function to a 'putative transcriptional regulator'. This change has also altered the functional classification from '5.H - Conserved hypothetical protein' to '2 - Broad regulatory functions'. A different 'functional classification' may or may not be within the same broad category of functional classifications (outlined as 1-6 above). Of the 181 CDSs with new product functions AND a new functional classification, 177 were moved into a new broad category of functional classification (as was the case for *Cj1546*). This was originally in the functional classification '5.I - Unknown', which is in broad category 5. The re-annotation has moved the functional classification to '2 - Broad regulatory functions', which is in broad category 2. One example of a CDS that had a new functional classification, but remained in the same broad category was *Cj1317*, where the CDS was originally annotated as a '*N*-acetylneuraminic acid synthetase' in the functional classification '3.C.2 - Surface polysaccharides, lipopolysaccharide and antigens'. The re-annotation changed the product function to 'Pse synthetase'. This CDS is part of the pseudaminic acid biosynthesis pathway of the *O*-linked glycosylation pathway that modifies the flagellin protein. The functional classification was altered to '3.C.3 - Surface structures'. As such, *Cj1317* has a new product name, a different functional classification, but still within the same broad category of functional classification (in this case, category 3).

Appendix 5 – Additional Files 2 & 3 display in-depth data on the change and distribution of CDSs within these functional categories. Importantly, the number of CDSs in the 'Unknown and other' category was reduced from 389 to 267 (a reduction of 122). This category includes functional classifications such as 'IS elements', 'plasmid related functions', 'antibiotic resistance', 'conserved hypothetical proteins' and 'unknown' (where no motifs were identified). The reduction in this category demonstrates the significant number of CDSs with new information relating to product function. However it is also clear that there remain a large number of CDSs yet to be characterised. Also the number of CDSs in the 'Miscellaneous' category increased from 75 to 152 (an increase of 77). CDSs were designated as miscellaneous when a motif was present, but there was no clear indication as to what functional classification the CDS belonged to. The increase in 'miscellaneous' numbers was due to the identification of new motifs not present during the original annotation, such as an ATP/GTP-binding site motif A (P-loop) in

*Cj0096*. Sequence comparisons and structural data have shown that proteins containing this motif have a role in ATP or GTP binding (Saraste *et al*., 1990). The fact that a number of CDSs have new information relating to a product function from such motifs, has led to an increase in CDSs within this particular category. This motif identification may not provide the full functional characterisation, however the designation of such motifs will aid in the future characterisation of the CDS in question.

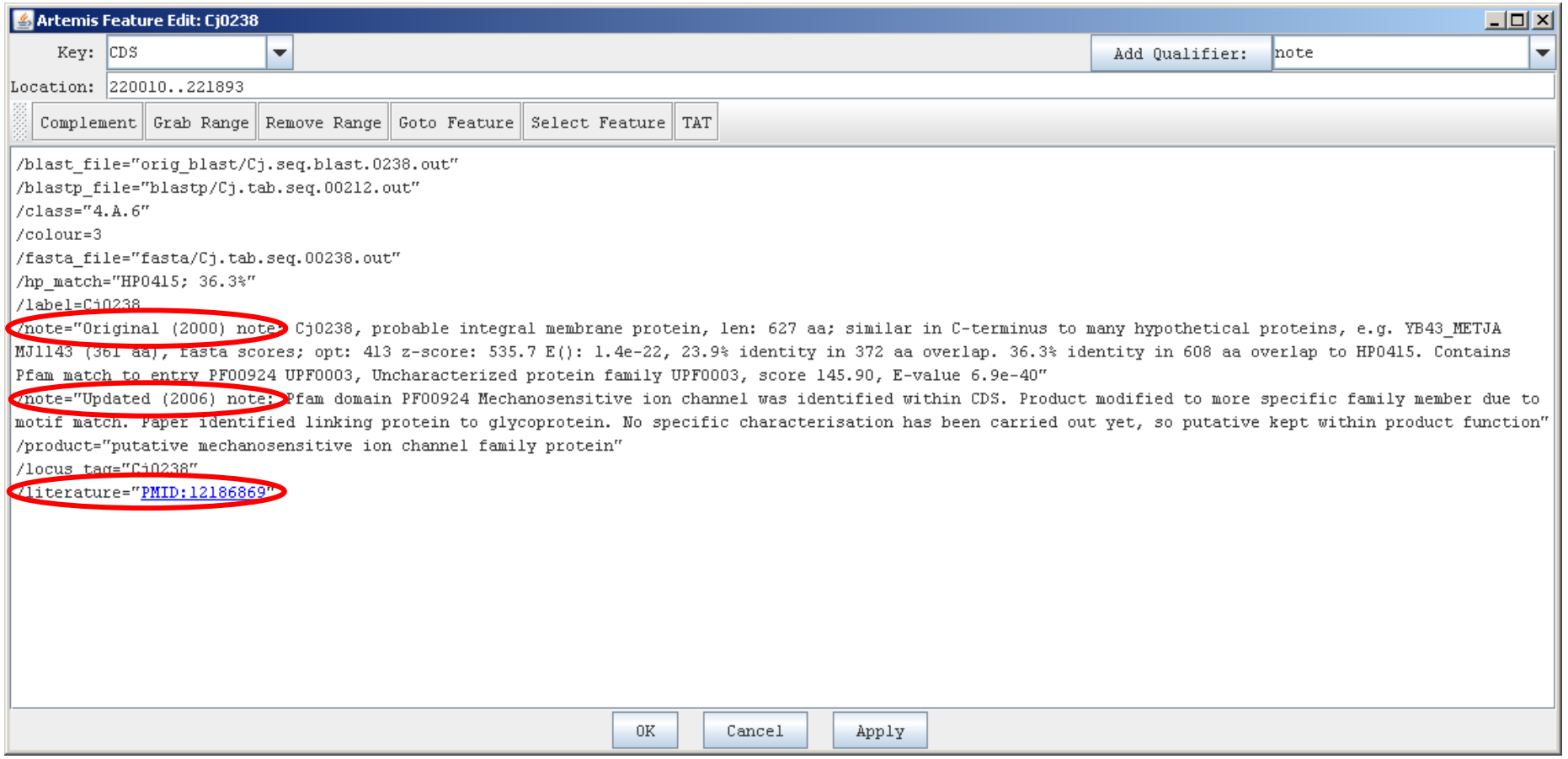



Figure 3.1. An example of a re-annotated CDS (*Cj0238*) displayed in Artemis. Highlighted is the 'original' and 'updated' annotations along with a new 'literature' qualifier, which was added in the re-annotation and details relevant publications related to the CDS. This CDS shows the different qualifiers used in the re-annotation.

Table 3.2. Statistics on CDSs from the *C. jejuni* NCTC11168 re-annotation.

| **Categories** | **Number of CDS** | **Percentage change** |
|---|---|---|
| CDSs with new motifs identified | 743 | +44.9% |
| CDSs with new literature qualifier | 1067 | +64.5% |
| CDSs with new updated note qualifier | 1489 | +90.0% |
| CDSs with hypothetical designations in product function | 124 | -7.5% |
| CDSs with conserved hypothetical in product function | 62 | +3.7% |
| CDSs with putative designation in product function | 98 | +5.9% |
| CDSs with new gene qualifier | 105 | +6.3% |
| CDSs with new product function | 299 | +18.2% |
| CDSs with new product function and with a new functional classification | 181 | +60.5% |
| CDSs with new product function and a new functional classification with a different broad category | 177 | +97.8% |

Since the original annotation was published in 2000, significant new information has been generated about the genetic loci encoding the four main *C. jejuni* surface structures. The *C. jejuni N*-linked glycosylation pathway (*Cj1119c – Cj1130c*) was not described in the original annotation, but has been studied extensively since the original annotation (Linton *et al*., 2005, Glover *et al*., 2005, Kelly *et al*., 2006, Szymanski *et al*., 2003a). This re-annotation includes the nomenclature for the *pglA-K* (*p*rotein *gl*ycosylation) genes and has updated all the product functions for each gene within this locus. The LOS locus (*Cj1131c* - *Cj1152c*) described in the original annotation was updated to include product functions for genes identified since the original annotation such as *neuA1*, *neuB1*, *neuC1* and *hldDE* (Guerry *et al*., 2000, Valvano *et al*., 2002, Linton *et al*., 2000, Gilbert *et al*., 2002). The *O*-linked glycosylation loci (*Cj1293* - *Cj1342*) which is responsible for glycosylation of the flagellin structural proteins, has been updated to include the *neu*, *pse* and *maf* genes (Thibault *et al*., 2001, Szymanski *et al*., 2003a, Liu and Tanner, 2006, Karlyshev *et al*., 2002). Finally, the CPS locus (*Cj1413c* - *Cj1448c*) has now been updated to include *kps* and *hdd* genes (Karlyshev *et al*., 2000, Karlyshev *et al*., 2001, Karlyshev *et al*., 2005a). The new product designations made to these four loci in the *C. jejuni* NCTC11168 genome are shown in Tables 3.3-3.6.

Table 3.3. Re-annotation updates for the *C. jejuni* NCTC11168 *N*-glycosylation locus.

| **CDS Number** | **Gene Name** | **Original Annotation** | **Re-annotation** |
|---|---|---|---|
| *Cj1119c* | *pglG* | probable integral membrane protein | putative integral membrane protein |
| *Cj1120c* | *pglF* | possible sugar epimerase/dehydratase | UDP-GlcNAc C4,6 dehydratase |
| *Cj1121c* | *pglE* | possible aminotransferase | UDP-4-keto-6-deoxy-GlcNAc C4 aminotransferase |
| *Cj1122c* | | possible integral membrane protein | putative integral membrane protein |
| *Cj1123c* | *pglD* | possible transferase | acetyltransferase |
| *Cj1124c* | *pglC* | probable galactosyltransferase | galactosyltransferase |
| *Cj1125c* | *pglA* | probable galactosyltransferase | GalNAc transferase |
| *Cj1126c* | *pglB* | probable integral membrane protein | oligosaccharide transferase to *N*-glycosylate proteins |
| *Cj1127c* | *pglJ* | probable glycosyltransferase | GalNAc transferase |
| *Cj1128c* | *pglI* | probable glycosyltransferase | glucosyl transferase |
| *Cj1129c* | *pglH* | probable glycosyltransferase | GalNAc transferase/polymerase |
| *Cj1130c* | *pglK* | ABC-type transport protein | flippase |

Table 3.4. Re-annotation updates for the *C. jejuni* NCTC11168 LOS locus.

| **CDS Number** | **Gene Name** | **Original Annotation** | **Re-annotation** |
|---|---|---|---|
| *Cj1131c* | *gne* | probable UDP-glucose 4-epimerase | UDP-GlcNAc/Glc 4-epimerase |
| *Cj1132c* | | unknown | conserved hypothetical protein |
| *Cj1133* | *waaC* | probable lipopolysaccharide heptosyltransferase | heptosyltransferase I |
| *Cj1134* | *htrB* | probable lipid A biosynthesis lauroyl acyltransferase | putative lipid A biosynthesis lauroyl acyltransferase |
| *Cj1135* | | probable two-domain glycosyltransferase | putative two-domain glucosyltransferase |
| *Cj1136* | | probable galactosyltransferase | putative glycosyltransferase |
| *Cj1137c* | | unknown | putative glycosyltransferase |
| *Cj1138* | | probable galactosyltransferase | putative glycosyltransferase |
| *Cj1139c* | *wlaN* | probable galactosyltransferase | beta-1,3 galactosyltransferase |
| *Cj1140* | *cstIII* | unknown | alpha-2,3 sialyltransferase |
| *Cj1141* | *neuB1* | probable *N*-acetylneuraminic acid synthetase | sialic acid synthase (*N*-acetylneuraminic acid synthetase) |
| *Cj1142* | *neuC1* | probable *N*-acetylglucosamine-6-phosphate 2- | putative UDP-*N*-acetylglucosamine 2-epimerase |

| | | | |
|---|---|---|---|
| | | epimerase/*N*-acetylglucosamine-6-phosphatase | |
| *Cj1143* | *neuA1*/*cgtA* | probable acylneuraminate cytidylyltransferase (CMP-*N*-acetylneuraminic acid synthetase) | two-domain bifunctional protein (beta-1,4-*N*-acetylgalactosaminyltransferase/CMP-Neu5Ac synthase) |
| *Cj1144c* | | unknown | hypothetical protein |
| *Cj1146c* | *waaV* | possible glucosyltransferase | putative glucosyltransferase |
| *Cj1148* | *waaF* | probable ADP-heptose--LPS heptosyltransferase | heptosyltransferase II |
| *Cj1149c* | *gmhA* | probable phosphoheptose isomerase | sedoheptulose 7-phosphate isomerase |
| *Cj1150c* | *hldE* | possible ADP-heptose synthase | D-beta-D-heptose 7-phosphate kinase/D-beta-D-heptose 1-phosphate adenylyltransferase |
| *Cj1151c* | *hldD* | probable ADP-L-*glycero*-D-*manno*-heptose-6-epimerase | ADP-*glyceromanno*-heptose 6-epimerase |
| *Cj1152c* | *gmhB* | possible phosphatase | D,D-heptose 1,7-bisphosphate phosphatase |

Table 3.5. Re-annotation updates for the *C. jejuni* NCTC11168 *O*-glycosylation locus.

| **CDS Number** | **Gene Name** | **Original Annotation** | **Re-annotation** |
|---|---|---|---|
| *Cj1293* | *pseB* | possible sugar nucleotide epimerase/dehydratase | UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase |
| *Cj1294* | *pseC* | probable aminotransferase | C4 aminotransferase specific for PseB product |
| *Cj1295* | | unknown | conserved hypothetical protein |
| *Cj1296* | | unknown | hypothetical protein |
| *Cj1297* | | unknown | hypothetical protein |
| *Cj1298* | | unknown | putative *N*-acetyltransferase |
| *Cj1299* | *acpP2* | probable acyl carrier protein | putative acyl carrier protein |
| *Cj1300* | | unknown | putative SAM domain containing methyltransferase |
| *Cj1301* | | unknown | hypothetical protein |
| *Cj1302* | | unknown | putative HAD-superfamily phosphatase, subfamily IIIC |
| *Cj1303* | *fabH2* | probable 3-oxoacyl-[acyl-carrier-protein] synthase | putative 3-oxoacyl-[acyl-carrier-protein] synthase |
| *Cj1304* | *acpP3* | probable acyl carrier protein | putative acyl carrier protein |

| | | | |
|---|---|---|---|
| *Cj1305c* | | unknown | hypothetical protein |
| *Cj1306c* | | unknown | hypothetical protein |
| *Cj1307* | | possible amino acid activating enzyme | putative amino acid activating enzyme |
| *Cj1308* | | possible acyl carrier protein | putative acyl carrier protein |
| *Cj1309c* | | unknown | hypothetical protein |
| *Cj1310c* | | unknown | hypothetical protein |
| *Cj1311* | *pseF* | probable acylneuraminate cytidylyltransferase (CMP-*N*-acetylneuraminic acid synthetase) | putative acylneuraminate cytidylyltransferase |
| *Cj1312* | *pseG* | possible flagella protein | nucleotidase specific for PseC product, UDP-4-amino-4,6-dideoxy-beta-L-AltNAc |
| *Cj1313* | *pseH* | possible flagella protein | *N*-acetyltransferase specific for PseC product, UDP-4-amino-4,6-dideoxy-beta-L-AltNAc |
| *Cj1314c* | *hisF* | probable cyclase | imidazole glycerol phosphate synthase subunit |
| *Cj1315c* | *hisH* | probable amidotransferase | imidazole glycerol phosphate synthase subunit |
| *Cj1316c* | *pseA* | unknown | pseudaminic acid biosynthesis PseA protein |
| *Cj1317* | *pseI* | *N*-acetylneuraminic acid synthetase | Pse synthetase |

| | | | |
|---|---|---|---|
| *Cj1318* | *maf1* | unknown | motility accessory factor (function unknown) |
| *Cj1319* | | probable nucleotide sugar dehydratase | putative nucleotide sugar dehydratase |
| *Cj1320* | | probable aminotransferase | putative aminotransferase (*degT* family) |
| *Cj1321* | | probable transferase | putative transferase |
| *Cj1322* | | unknown | hypothetical protein |
| *Cj1323* | | unknown | hypothetical protein |
| *Cj1324* | | unknown | hypothetical protein |
| *Cj1325* | | unknown | putative methyltransferase |
| *Cj1327* | *neuB2* | *N*-acetylneuraminic acid synthetase | *N*-acetylneuraminic acid synthetase |
| *Cj1328* | *neuC2* | probable *N*-acetylglucosamine-6-phosphate 2-epimerase/*N*-acetylglucosamine-6-phosphatase | putative UDP-*N*-acetylglucosamine 2-epimerase |
| *Cj1329* | | probable sugar-phosphate nucleotide transferase | putative sugar-phosphate nucleotide transferase |
| *Cj1330* | | unknown | hypothetical protein |
| *Cj1331* | *ptmB* | probable acylneuraminate cytidylyltransferase (CMP-*N*-acetylneuraminic acid synthetase) | acylneuraminate cytidylyltransferase (flagellin modification) |
| *Cj1332* | *ptmA* | probable oxidoreductase | putative oxidoreductase (flagellin modification) |

| | | | |
|---|---|---|---|
| *Cj1333* | *pseD* | unknown | PseD protein |
| *Cj1334* | *maf3* | unknown | motility accessory factor (function unknown) |
| *Cj1335* | *maf4* | unknown | motility accessory factor (function unknown) |
| *Cj1337* | *pseE* | unknown | PseE protein |
| *Cj1338c* | *flaB* | flagellin B | flagellin |
| *Cj1339c* | *flaA* | flagellin A | flagellin |
| *Cj1340c* | | unknown | conserved hypothetical protein |
| *Cj1341c* | *maf6* | unknown | motility accessory factor (function unknown) |
| *Cj1342c* | *maf7* | unknown | motility accessory factor (function unknown) |

Table 3.6. Re-annotation updates for the *C. jejuni* NCTC11168 CPS biosynthesis locus.

| **CDS Number** | **Gene Name** | **Original Annotation** | **Re-annotation** |
|---|---|---|---|
| *Cj1413c* | *kpsS* | possible polysaccharide modification protein | capsule polysaccharide modification protein |
| *Cj1414c* | *kpsC* | possible polysaccharide modification protein | capsule polysaccharide modification protein |
| *Cj1415c* | *cysC* | possible adenylylsulfate kinase | putative adenylylsulfate kinase |
| *Cj1416c* | | probable sugar nucleotidyltransferase | putative sugar nucleotidyltransferase |
| *Cj1417c* | | unknown | putative amidotransferase |
| *Cj1418c* | | unknown | putative transferase |
| *Cj1419c* | | possible methyltransferase | putative methyltransferase |
| *Cj1420c* | | unknown | putative methyltransferase |
| *Cj1421c* | | possible sugar transferase | putative sugar transferase |
| *Cj1422c* | | possible sugar transferase | putative sugar transferase |
| *Cj1423c* | *hddC* | possible sugar-phosphate nucleotidyltransferase | putative D-glycero-D-manno-heptose 1-phosphate guanosyltransferase |
| *Cj1424c* | *gmhA2* | probable phosphoheptose isomerase | phosphoheptose isomerase |

| | | | |
|---|---|---|---|
| *Cj1425c* | *hddA* | possible sugar kinase | putative D-glycero-D-manno-heptose 7-phosphate kinase |
| *Cj1426c* | | unknown | putative methyltransferase family protein |
| *Cj1427c* | | probable sugar-nucleotide epimerase/dehydratease | putative sugar-nucleotide epimerase/dehydratease |
| *Cj1428c* | *fcl* | probable fucose synthetase | GDP-L-fucose synthetase |
| *Cj1429c* | | unknown | hypothetical protein |
| *Cj1430c* | *rfbC* | probable nucleotide-sugar epimerase/dehydratase | putative dTDP-4-dehydrorhamnose 3,5-epimerase |
| *Cj1431c* | *hddC* | unknown | capsular polysaccharide heptosyltransferase |
| *Cj1432c* | | possible sugar transferase | putative sugar transferase |
| *Cj1433c* | | unknown | hypothetical protein |
| *Cj1434c* | | probable sugar transferase | putative sugar transferase |
| *Cj1435c* | | unknown | putative phosphatase |
| *Cj1436c* | | probable aminotransferase | aminotransferase |
| *Cj1437c* | | probable aminotransferase | aminotransferase |
| *Cj1438c* | | probable sugar transferase | putative sugar transferase |
| *Cj1439c* | *glf* | probable UDP-galactopyranose mutase | UDP-galactopyranose mutase |

| | | | |
|---|---|---|---|
| *Cj1440c* | | probable sugar transferase | putative sugar transferase |
| *Cj1441c* | *kfiD* | probable UDP-glucose 6-dehydrogenase | UDP-glucose 6-dehydrogenase |
| *Cj1442c* | | unknown | putative sugar transferase |
| *Cj1443c* | *kpsF* | unknown | D-arabinose 5-phosphate isomerase |
| *Cj1444c* | *kpsD* | probable capsule polysaccharide export system periplasmic protein | capsule polysaccharide export system periplasmic protein |
| *Cj1445c* | *kpsE* | probable capsule polysaccharide export system inner membrane protein | capsule polysaccharide export system inner membrane protein |
| *Cj1447c* | *kpsT* | probable capsule polysaccharide export ATP-binding protein | capsule polysaccharide export ATP-binding protein |
| *Cj1448c* | *kpsM* | probable capsule polysaccharide export system inner membrane protein | capsule polysaccharide export system inner membrane protein |

Additional genome-wide updates were performed of which a large proportion entailed adding specificity to existing product function. For example, the identification of a new PFAM or PROSITE motif for the CDS *Cj1556* has allowed the product function to become further specified. In this example, a hypothetical protein was re-annotated as a putative transcriptional regulator. A complete list of changes throughout the *C. jejuni* NCTC11168 genome is provided in Appendix 5 – Additional Files 4 & 5.

### 3.2.3 Modifications to pseudogene annotations

Pseudogene identification is a challenging process with discrepancies existing between the methods used for assignment of pseudogenes (Lerat and Ochman, 2005). Methods include detection of ORFs belonging to a single CDS on multiple frames, the presence of one or more stop codons within a CDS and extra information from the biology of the organism. More recently, comparative genomics have been used as a technique for pseudogene identification where alignment to a fully functional version of the CDS from a different strain/species has assisted in the annotation of pseudogenes (Lerat and Ochman, 2004). The driving force for these studies is that the identification of such mutations aids in determining the phylogenetic relatedness and evolution of bacterial strains. A full re-analysis of all pseudogenes in the NCTC11168 genome was performed as part of this re-annotation (Table 3.7). The original annotation had only displayed a pseudogene on a single frame so did not reflect the true position of the mutated CDS. This process of identifying pseudogenes is difficult and support from homologues identified using FASTA and TBLASTX search results were used. The majority of revisions in this re-annotation incorporated multiple features created from different co-ordinates on more than one reading frame. This full re-analysis of all pseudogenes resulted in the adjustment of 19 out of 20 pseudogenes by introducing multiple CDSs on different reading frames representing the hypothetical fully functional original single CDS. The final pseudogene number was reduced to 19 due to the merging of two CDSs (*Cj0290c* and *Cj0292c*) designated as pseudogenes along with an intermediate CDS *Cj0291c*, which had been originally annotated as a glycerol-3-phosphate transporter and had been noted as a possible pseudogene (Figure 3.2).

Table 3.7. Re-annotated pseudogenes in *C. jejuni* NCTC11168 with modifications.

| **CDS Number** | **Product function** | **Modification** |
|---|---|---|
| *Cj0046* | pseudogene (putative sodium:sulfate transmembrane transport protein) | CDS split into multiple frames |
| *Cj0072c* | pseudogene (putative iron-binding protein) | CDS split into multiple frames |
| *Cj0223* | pseudogene (putative IgA protease family protein) | CDS split into multiple frames |
| *Cj0292c* | pseudogene (putative glycerol-3-phosphate transporter) | Merging of multiple CDS |
| *Cj0444* | pseudogene (putative TonB-denpendent outer membrane receptor) | CDS split into multiple frames |
| *Cj0501* | pseudogene (ammonium transporter) | CDS split into multiple frames |
| *Cj0565* | pseudogene (conserved hypothetical protein) | CDS split into multiple frames |
| *Cj0654c* | pseudogene (putative transmembrane transport protein) | CDS split into multiple frames |
| *Cj0676* | pseudogene (potassium-transporting ATPase A chain) | CDS split into multiple frames |
| *Cj0678* | pseudogene (potassium-transporting ATPase C chain) | CDS split into multiple frames |
| *Cj0742* | pseudogene (putative outer membrane protein) | CDS split into multiple frames |
| *Cj0752* | pseudogene (IS element transposase) | CDS split into multiple frames |
| *Cj0866* | pseudogene (arylsulfatase) | CDS split into multiple frames |
| *Cj0969* | pseudogene (putative periplasmic protein) | Merging of multiple CDS |

| | | |
|---|---|---|
| *Cj1064* | pseudogene (nitroreductase) | CDS split into multiple frames |
| *Cj1389* | pseudogene (putative C4-dicarboxylate anaerobic carrier) | CDS split into multiple frames |
| *Cj1395* | pseudogene (putative MmgE/PrpD family protein) | Unmodified |
| *Cj1395* | pseudogene (putative MmgE/PrpD family protein) | Unmodified |
| *Cj1470c* | pseudogene (type II protein secretion system F protein) | CDS split into multiple frames |
| *Cj1528* | pseudogene (putative C4-dicarboxylate anaerobic carrier) | CDS split into multiple frames |


Figure 3.2. Artemis genome viewer illustrating CDSs *Cj0290c*, *Cj0291c* and *Cj0292c*. Upper panel illustrates CDS layout from the original annotation. Lower panel illustrates CDS layout from the re-annotation. The two pseudogenes along with *Cj0291c*, which had originally been annotated as a 'glycerol-3-phosphate transporter – putative pseudogene' were all merged together. Pseudogenes are illustrated in brown. The result of this modification was the removal of one of the two original pseudogenes *Cj0290c* and *Cj0292c*).

### 3.2.4 Modifications to phase-variable CDS annotations

Predicted phase-variable CDSs containing an intersecting homopolymeric tract were merged to reflect the complete amino acid sequence for appropriate genes regardless of phase variation (Table 3.8). The merging of such CDSs was not undertaken in the original annotation. The merging of CDSs with an intersecting homopolymeric tract allows a better representation of phase-variable CDSs within the *C. jejuni* NCTC11168 genome.

Table 3.8. Merged phase-variable CDSs in *C. jejuni* NCTC11168 re-annotation.

| **CDS Number** | **Product** | **Re-annotation CDS** | **Effect** |
|---|---|---|---|
| *Cj0031/Cj0032* | putative type IIS restriction/modification enzyme | *Cj0031* | Merged |
| *Cj0170/Cj0171* | hypothetical protein Cj0170 | *Cj0170* | Merged |
| *Cj0628/Cj0629* | putative lipoprotein | *Cj0628* | Merged |
| *Cj1144/Cj1145* | hypothetical protein Cj1144c | *Cj1144* | Merged |
| *Cj1325/Cj1326* | putative methyltransferase | *Cj1325* | Merged |
| *Cj1335/Cj1336* | motility accessory factor (function unknown) | *Cj1335* | Merged |
| *Cj1677/Cj1678* | putative lipoprotein | *Cj1677* | Merged |

### 3.2.5 Additional genome data

In addition to CDS updates, novel features were also added to the re-annotation. CRISPR regions have been shown to provide acquired immunity against viruses and plasmids by targeting nucleic acid in a sequence in a sequence-specific manner (Schouls *et al*., 2003, Godde and Bickerton, 2006, Jansen *et al*., 2002, Bolotin *et al*., 2005, Horvath and

Barrangou, 2010, Marraffini and Sontheimer, 2010). To date, there has only been one identified CRISPR region within *C. jejuni* NCTC11168 and this has now been incorporated within this re-annotation (PMID: 12517820). As a result one CDS (*Cj1520)* has been removed. This CDS was previously annotated as having five repeat regions. In addition, three CDSs (*Cj1521c*, *Cj1522c* and *Cj1523c*) adjacent to the CRISPR repeats were re-annotated as CRISPR associated proteins and this concurs with existing CRISPR structures (Schouls *et al*., 2003, Godde and Bickerton, 2006, Jansen *et al*., 2002).

RFAM databases investigating non-coding RNAs (Griffiths-Jones *et al*., 2003) were performed and identified two new non-coding RNA structures. RFAM RF00169, a bacterial signal recognition particle (SRP) RNA, was identified upstream of *Cj0046*. The SRP is a universally conserved ribonucleoprotein involved in the co-translational targeting of proteins to membranes (Rosenblad *et al*., 2003, Regalia *et al*., 2002). RFAM RF00059, a thiamin pyrophosphate riboswitch, was identified upstream of *Cj0453* (thiamin biosynthesis protein ThiC). This RFAM motif is a conserved structure involved in thiamin-regulation (Rodionov *et al*., 2002).

One of the major advantages of performing a manual annotation or re-annotation is the potential inclusion of the advice from experts on the organism being annotated. This was the case for this re-annotation whereby experts contributed product functions and descriptive information for the four main glyco-structure loci in the *C. jejuni* NCTC11168 genome. The *N*-linked glycosylation locus was reviewed by Dr. Christine Szymanski. The *O*-linked glycosylation locus was reviewed by Dr. Susan Logan. The LOS locus was reviewed by Dr. Michel Gilbert and the capsule locus was reviewed by Prof. Andrey Karlyshev.

The final step of the re-annotation process was the incorporation of Gene Ontology (GO) annotation. GO annotation attempts to link three structured, controlled vocabularies (ontologies) that describe gene products in terms of their associated biological processes, cellular components and molecular functions in a species-independent manner (Ashburner *et al*., 2000). There are currently two different versions of the GO annotation. The EMBL genome file created by EBI carries a GOA link that lists the GO annotation created automatically by EBI. GeneDB maintains an alternative list of GO annotations created automatically by performing a reciprocal FASTA of the *C. jejuni* NCTC11168 genome against *C. jejuni* RM1221. GO annotation is a valuable feature in current

annotation techniques that can expedite systems biology approaches to genome analysis (Ashburner *et al*., 2000).

### 3.3 Further *C. jejuni* genome analysis since re-annotation

Further studies have been performed since the re-annotation was published in 2007. The number and type of predicted pseudogenes within *C. jejuni* NCTC11168 and *C. jejuni* RM1221 were analysed (Table 3.9). The table displays the genomes in synteny. Hence, identical pseudogenes are on the same row. This was to ensure that variation in product function naming did not exclude identical pseudogenes.

Table 3.9. Comparison of predicted pseudogenes in *C. jejuni* NCTC11168 and *C. jejuni* RM1221. Table represents the genomes in synteny. A cell which is grey in colour denotes a pseudogene. Hence, identical pseudogenes are on the same row. A parallel italicised white cell can have the pseudogene represented as an intact CDS (present), not have the CDS (absent) or have the CDS but mis-annotated (should be a pseudogene).

| **NCTC11168** | | **RM1221*** | |
|---|---|---|---|
| **CDS Number** | **Product** | **CDS Number** | **Product** |
| | *mis-annotated as a CDS (should be a pseudogene) in NCTC11168* | ***CJE0013*** | **membrane protein, putative** |
| ***Cj0046*** | **pseudogene (putative sodium:sulfate transmembrane transport protein)** | ***CJE0046*** | **C4-dicarboxylate transporter** |
| ***Cj0072c*** | **pseudogene (putative iron-binding protein)** | ***CJE0068*** | **conserved hypothetical protein** |
| | *present as intact CDS* | ***CJE0196*** | **citrate transporter** |
| ***Cj0223*** | **pseudogene (putative IgA protease family protein)** | | *mis-annotated as a CDS (should be a pseudogene) in RM1221* |
| | *absent* | ***CJE0274*** | **pathogenicity protein** |
| | *absent* | ***CJE0292*** | **conserved hypothetical protein** |
| | *present as intact CDS* | ***CJE0297*** | **conserved hypothetical protein** |

| | | | |
|---|---|---|---|
| ***Cj0292c*** | **pseudogene (putative glycerol-3-phosphate transporter)** | ***CJE0338*** | **glycerol-3-phosphate transporter** |
| ***Cj0444*** | **pseudogene (putative TonB-denpendent outer membrane receptor)** | ***CJE0496*** | **TonB-dependent receptor** |
| | *mis-annotated as a CDS (should be a pseudogene) in NCTC11168* | ***CJE0533*** | **altronate hydrolase** |
| ***Cj0501*** | **pseudogene (ammonium transporter)** | ***CJE0609*** | **ammonium transporter** |
| **Cj0565** | **pseudogene (conserved hypothetical protein)** | | *absent* |
| | *present as intact CDS* | ***CJE0665*** | **MATE efflux family protein** |
| | *present as intact CDS* | ***CJE0673*** | **conserved hypothetical protein** |
| | *mis-annotated as a CDS (should be a pseudogene) in NCTC11168* | ***CJE0720*** | **hypothetical protein** |
| ***Cj0654c*** | **pseudogene (putative transmembrane transport protein)** | ***CJE0758*** | **peptide transporter** |
| ***Cj0676*** | **pseudogene (potassium-transporting ATPase A chain)** | ***CJE0774*** | **potassium-transporting ATPase, A subunit** |

| | | | |
|---|---|---|---|
| | *present as intact CDS* | ***CJE0775*** | **potassium-transporting ATPase, B subunit** |
| ***Cj0678*** | **pseudogene (potassium-transporting ATPase C chain)** | | *present as intact CDS* |
| ***Cj0742*** | **pseudogene (putative outer membrane protein)** | | *absent* |
| | *present as intact CDS* | ***CJE0777*** | **sensor histidine kinase KdpD** |
| | *present as intact CDS* | ***CJE0803*** | **conserved hypothetical protein** |
| | *present as intact CDS* | ***CJE0836*** | **conserved hypothetical protein** |
| ***Cj0752*** | **pseudogene (IS element transposase)** | ***CJE0844*** | **transposase** |
| ***Cj0866*** | **pseudogene (arylsulfatase)** | ***CJE0953*** | **arylsulfate sulfotransferase** |
| | *present as intact CDS* | ***CJE0955*** | **cytochrome c family protein** |
| | *present as intact CDS* | ***CJE1032*** | **membrane protein, putative** |
| | *present as intact CDS* | ***CJE1047*** | **conserved hypothetical protein** |
| | *present as intact CDS* | ***CJE1056*** | **conserved hypothetical protein** |
| | *absent* | ***CJE1121*** | **DNA primase TraC** |
| | *absent* | ***CJE1140*** | **hypothetical protein** |

| | | | |
|---|---|---|---|
| | *absent* | ***CJE1175*** | **conserved hypothetical protein** |
| ***Cj0969*** | **pseudogene (putative periplasmic protein)** | | *present as intact CDS* |
| ***Cj1064*** | **pseudogene (nitroreductase)** | ***CJE1208*** | **nitroreducatase family protein** |
| | *mis-annotated as a CDS (should be a pseudogene) in NCTC11168* | ***CJE1294*** | **conserved hypothetical protein** |
| | *present as intact CDS* | ***CJE1393*** | **multidrug resistance efflux transporter** |
| | *present as intact CDS* | ***CJE1544*** | **enterochelin ABC transporter** |
| | *absent* | ***CJE1549*** | **vacuolating cytotoxin, putative** |
| ***Cj1389*** | **pseudogene (putative C4-dicarboxylate anaerobic carrier)** | ***CJE1580*** | **cryptic C4-dicarboxylate transporter** |
| ***Cj1395*** | **pseudogene (putative MmgE/PrpD family protein)** | | *present as intact CDS* |
| ***Cj1470c*** | **pseudogene (type II protein secretion system F protein)** | | *present as intact CDS* |
| | *present as intact CDS* | ***CJE1585*** | **ferrous iron transport protein B** |
| | *present as intact CDS* | ***CJE1697*** | **CRISPR-associated protein** |

| | | | |
|---|---|---|---|
| ***Cj1528*** | **pseudogene (putative C4-dicarboxylate anaerobic carrier)** | ***CJE1699*** | **cryptic C4-dicarboxylate transporter** |
| | *absent* | ***CJE1718*** | **outer membrane lipoprotein Blc** |
| | *absent* | ***CJE1720*** | **type I restriction-modification system, R subunit** |
| | *absent* | ***CJE1722*** | **type I restriction-modification system, S subunit** |
| | *present as intact CDS* | ***CJE1729*** | **conserved hypothetical protein** |
| | *present as intact CDS* | ***CJE1734*** | **transcriptional regulator, putative** |
| | *present as intact CDS* | ***CJE1759*** | **proline/betaine transporter, putative** |
| | *present as intact CDS* | ***CJE1790*** | **conserved hypothetical protein** |
| | *present as intact CDS* | ***CJE1802*** | **TonB-dependent receptor, putative** |

* *C. jejuni* RM1221 was predicted to have 47 pseudogenes and 35 of these were noted as genes in *C. jejuni* NCTC11168 (Fouts *et al*., 2005). However, the RM1221 genome sequence obtained from the WTSI only contained 46 pseudogenes. Thus, the above 35 "non-pseudogenes" in *C. jejuni* NCTC11168 are 34 in number. RM1221 genes are represented as 'CJE'.

Publication of further *C. jejuni* genome sequences has allowed analysis between different genomes. To further investigate the number of pseudogenes in different *C. jejuni* strains

and also to analyse genome features in greater detail, a further four *C. jejuni* strains were analysed (Table 3.10).

Table 3.10. Genome features of five *C. jejuni* genomes.

| **Trait** | **Strains** | | | | |
|---|---|---|---|---|---|
| Strain | NCTC11168 | RM1221 | 81-176 | 81116 | CG8486 |
| Serotype | HS:2 | HS:53 | HS:23/36 | HS:6 | HS:4 |
| MLST[a] | ST-21 (43) | ST-354 (354) | ST-42 (913) | ST-283 (267) | unknown |
| Origin | Clinical | Chicken | Clinical | Clinical | Clinical |
| Genome size (Mb) [b] | 1.64 | 1.77 | 1.62 | 1.63 | 1.60 |
| GC content (%) | 30.55 | 30.31 | 30.62 | 30.54 | 30.43 |
| Predicted CDS numbers | 1654 (1643)[e] | 1835 | 1568 | 1626 | 1588 |
| Pseudogenes | 20(19)[e] | 47 | 0[c] | 1[c] | 3[c] |
| Poly G/C tracts[d] | 29 (22) | 25 (8) | 19 | 17 | 23 |
| Plasmids | 0 | 0 | 2 | 0 | 0 |
| References | (Parkhill *et al*., 2000) | (Fouts *et al*., 2005) | (Hofreuter *et al*., 2006) | (Pearson *et al*., 2007) | (Poly *et al*., 2007) |

Legend. [a] ST represents clonal complex. () indicates sequence type. [b] indicates genome size made by approximation. CG8486 genome sequence is currently in 19 contigs. [c] indicates approximate number of pseudogenes. [d] Poly G/C tracts represent total found. () indicate tracts greater than seven or more nucleotides in length and have been shown to vary during sequencing project. [e] () indicates number after NCTC11168 re-annotation.

### 3.4 Rationale for the selection of predicted CDSs for further characterisation

Following the re-annotation of the *C. jejuni* NCTC11168 genome, 15 CDSs were selected for further characterisation (Table 3.11). Selection was primarily based on new motif and literature information not available at the time of the original annotation.

Selected CDSs were grouped into either putative virulence-associated, signal transduction or regulation of gene expression (Table 3.11). Defined isogenic *C. jejuni* 11168H mutants were successfully constructed for 8 of the selected CDSs. Following a number of preliminary phenotypic assays such as growth kinetics, motility, interaction and invasion assays on all successfully constructed mutants, the *Cj0248* and *Cj1556* mutants displayed the most interesting phenotypic results and were selected for further investigation to characterise each CDS function in detail. The characterisation of the Cj1556 and Cj0248 mutants are described in Chapters 4 and 5 respectively.

Table 3.11. CDSs selected for further study after re-annotation of the NCTC11168 genome sequence. CDSs were selected in three main categories; regulation of gene expression, signal transduction and virulence-associated.

| **CDS number** | **Product function** | **Mutant created** | **Function category** |
|---|---|---|---|
| *Cj0394c* | putative transcriptional activator | No | Broad regulatory function |
| *Cj0883c* | putative transcriptional regulator | No | Broad regulatory function |
| *Cj1546* | putative transcriptional regulator | No | Broad regulatory function |
| *Cj1556* | putative transcriptional regulator | Yes | Broad regulatory function |
| *Cj0144** | putative methyl-accepting chemotaxis signal transduction protein | No | Signal transduction |
| *Cj0145* | putative TAT (Twin-Arginine Translocation) pathway signal sequence domain protein | Yes | Signal transduction |
| *Cj0262c** | putative methyl-accepting chemotaxis signal transduction protein | Yes | Signal transduction |
| *Cj1505c* | putative two-component response regulator (SirA-like protein) | No | Signal transduction |
| *Cj1506c** | putative MCP-type signal transduction protein | Yes | Signal transduction |
| *Cj1507c* | putative regulatory protein | No | Signal transduction |
| *Cj0248* | hypothetical protein Cj0248 | Yes | Virulence-associated |
| *Cj0619* | putative MATE family transport protein | Yes | Virulence-associated |

| | | | |
|---|---|---|---|
| *Cj0737* | putative haemagglutination activity domain containing protein | Yes | Virulence-associated |
| *Cj0952* | putative HAMP containing membrane protein | Yes | Virulence-associated |
| *Cj0977* | hypothetical protein Cj0977 | No | Virulence-associated |

* Denotes product function was identified in 2000.

### 3.5 Discussion

Existing genome sequences from model strains should ideally be re-annotated continuously to incorporate the maximum information available and to maintain the genome sequence as a useful resource. This is important as many aspects of research in the 21st century depend on genome sequence data. Global access to genome information has made it vital that up-to-date information is provided for each CDS. Using outdated CDS data could result in a waste of resources and affect strategic planning of research projects. Examples where up-to-date sequence data is important are microarray and proteomic studies where throughout the analysis pipeline, genomic data is used in conjunction with appropriate software. In addition, the use of up-to-date sequence data for next-generation sequencing techniques is also important. Whether the application is *de novo* sequencing, transcriptomics, interactomics or any other related application, annotated genome data is the baseline information that can have a strong influence on the interpretation of data analysis.

The re-annotation of the *C. jejuni* NCTC11168 genome sequence led to 299 of product functions being updated and 1489 of CDSs having new information added to the note qualifier. The number of CDSs in the 'Unknown and other' category was reduced from 389 to 267 (a reduction of 122). This category includes functional classifications such as 'IS elements', 'plasmid related functions', 'antibiotic resistance', 'conserved hypothetical proteins' and 'unknown' (where no motifs were identified). This decrease is predominantly due to novel motifs being identified within a given CDS allowing the CDS to be placed into different functional category. There remains 267 CDSs in this 'Unknown and other' category that required further investigation to pinpoint the functionality of these CDSs. The number of CDSs in the 'Miscellaneous' category has risen from 75 to 152 (an increase of 77). CDSs in this category contained a motif,

however there were no clear indications as to what functional classification the CDS belongs to. The fact that a number of CDSs have new information relating to a product function from uncharacterised motifs has led to the increase in CDSs within this particular category. This motif identification may not provide the ultimate product characterisation, however the designation of such motifs will assist in the future characterisation of the CDS in question, especially when a function is associated with the currently undefined motif.

The re-annotation of the *C. jejuni* NCTC11168 genome was performed manually in collaboration with WTSI staff. Most genome sequencing projects now use automated annotation due to time and cost restraints. So it is rare for a re-annotation to be performed by a manual annotator. It is clear that researchers performing a re-annotation project can still offer a greater level of insight compared to automated approaches. However it is also clear that better automated annotation methods are needed to cope with the deluge of genome sequence data generated by next-generation sequencing. Automated tools need to be able to incorporate information from wider sources, such as literature searches. Searches should include a wide range of motif databases and be able to use this information to describe a product function. This of course is extremely challenging and requires a high level of programming and computing power. Currently automated searches rely primarily on results from similarity searches. This in itself can be highly misleading as an incorrectly assigned product function could result in the propagation of false designations across genome annotations for many different bacterial species and beyond.

This re-annotation has attempted to incorporate data from a number of additional sources to allow for a more accurate product function designation. In addition to the classical motif databases of PFAM and Prosite, membrane and ribosomal motif databases were also searched. One of the most advantageous aspects of manual annotation is the incorporation of additional data from publicly available literature. Researchers may also have contacts or even information within their own laboratory that can be used to add further information to the file. Indeed this was the case with this re-annotation of the *C. jejuni* NCTC11168 genome sequence where a number of experts contributed to the updated annotation.

An important feature of the current re-annotation project (and indeed annotation projects in general) is the improved designation of pseudogenes and phase-variable genes. Pseudogenes are difficult to pinpoint due to sections of the full CDS residing on different reading frames of the forward or reverse strands. An example of the difficulty and complexity associated with pseudogene designation is observed when analysing the CDSs *Cj0522*, *Cj0523* and *Cj0524* in the *C. jejuni* NCTC11168 genome (Figure 3.3). These three CDSs are represented as one whole CDS on a single frame within *C. jejuni* RM1221 (*Cje0628*). The three CDSs are large enough to be represented as individual CDSs and were represented on more than one reading frame in the original *C. jejuni* NCTC11168 annotation. Do these CDSs represent a pseudogene in *C. jejuni* NCTC11168, which are intact in *C. jejuni* RM1221? Given the fact that in *C. jejuni* RM1221 *Cje0628* does code for a putative product (a putative Na/Pi-co-transporter), it is most likely that these three CDSs represent a pseudogene in *C. jejuni* NCTC11168. In this re-annotation, a re-analysis of annotated pseudogenes was performed. However on the advice of WTSI staff, the elucidation of new pseudogenes was not undertaken. In this particular case, the potential for a pseudogene was noted in the 'note' qualifier.


Figure 3.3. Artemis genome viewer with upper panel illustrating CDSs *Cj0522*, *Cj0523* and *Cj0524* from the original NCTC11168 genome annotation. Lower panel illustrates CDS layout from the RM1221 genome annotation illustrating *Cje0628*. This figure illustrates the likelihood that the three NCTC11168 CDSs (upper panel) should actually be represented as a pseudogene due to the presence of stop codons between *Cj0523* and *Cj0524*, and also the presence of a frame-shift for *Cj0522*. In this re-annotation, a re-analysis of pseudogenes was performed. However the elucidation of new pseudogenes was not undertaken. In this particular example, the potential for a pseudogene was noted in the 'note' qualifier.

The frequency and importance of pseudogene formation in microorganisms has attained increased significance in recent years with the development of genome reduction theories and the demonstration of enhanced virulence through pathoadaptive mutations (Harrison and Gerstein, 2002, Ochman and Davalos, 2006). Recent studies for example in *Yersinia* species have suggested that ever increasing non-functional CDSs are being identified within microorganisms and in particular are more common in genomes of recently evolved pathogens than in their benign relatives (Tong *et al.*, 2005, Lerat and Ochman, 2005). It should be noted that sequencing errors are also a factor in the pseudogene designation process and that re-sequencing is common to ensure mis-annotation has not occurred. Based on guidelines from the WTSI, this re-annotation did not attempt to identify novel pseudogenes, but did edit them to better illustrate the non-functional CDS. The number and type of predicted pseudogenes within *C. jejuni* NCTC11168 and *C. jejuni* RM1221 are compared in Table 3.9. After this re-annotation, *C. jejuni* NCTC11168 is predicted to contain at least 19 pseudogenes. 12/19 (63%) of the pseudogenes from *C. jejuni* NCTC11168 were also present in RM1221 as pseudogenes. Of the remaining seven pseudogenes, four were present as intact CDSs in RM1221, two were absent and one was mis-annotated and should have been designated a pseudogene in RM1221 (Table 3.9). *C. jejuni* RM1221 was predicted to contain 47 pseudogenes (Fouts *et al.*, 2005). 12 pseudogenes are shared with NCTC11168 as noted above. Of the remaining 35 pseudogenes, 21 were present as intact CDSs in NCTC11168, 9 were absent and 4 were mis-annotated and should have been designated a pseudogene in NCTC11168 (Table 3.9). The mis-annotated CDSs are represented as two or more CDSs in *C. jejuni* NCTC11168 and were also noted to be putative pseudogenes.

Assuming these are genuine pseudogenes, there is a possibility that *C. jejuni* NCTC11168 (isolated in 1977 from a UK patient with severe gastroenteritis and sequenced in 1999 (Parkhill *et al.*, 2000, Skirrow, 1977)) and *C. jejuni* RM1221 (isolated in 2000 from a chicken isolate in the U.S and sequenced in 2004 (Miller *et al.*, 2000, Fouts *et al.*, 2005)) share a set of identical pseudogenes. However the sequenced strains only represent a single isolated clone and there is scope for significant adaptation from the original isolation date and sequence date. Both *C. jejuni* strains were likely to be passaged a number of times and may be laboratory variants with significant modifications in genome content compared to the original isolated strain. In addition, there is no evidence to date that shows *C. jejuni* RM1221 to be pathogenic in humans (see Section 1.12 – Table 1.1). The number of pseudogenes identified within these two

strains can also be attributed to putative pathoadaptive mutations, where a greater number of pseudogenes are being identified in recently evolved pathogens. The *C. jejuni* strain RM1221 was isolated from the skin of a supermarket chicken and contains 47 pseudogenes, whereas *C. jejuni* NCTC11168 was a human isolate and contains 19 pseudogenes. This may indicate a more rapidly evolving genome for RM1221 especially given the number of *C. jejuni* present, for example in a poultry farming environment. The increased number of pseudogenes identified within *C. jejuni* RM1221 may be due to these CDSs not being necessary as functional copies, as the strain was isolated from chicken. Clearly analysing only two genomes is insufficient to make such conclusions and a greater sample pool would be required to investigate this question further.

Pseudogenes numbers were analysed in *C. jejuni* 81-176, RM1221 and NCTC11168 (Table 3.10) (Hofreuter *et al*., 2006, Fouts *et al*., 2005). The annotation of *C. jejuni* 81-176 identified no designated pseudogenes (Hofreuter *et al*., 2006). A total of 37 CDSs in 81-176 that are absent or are pseudogenes in NCTC11168 and RM1221 were identified, such as *kdpA*, *kdpB*, and *kdpC*, which encode a potassium-transporting ATPase and are pseudogenes in the two latter strains (Hofreuter *et al*., 2006). A total of 51 CDSs were identified that were present in NCTC11168, but absent in 81-176 (Hofreuter *et al*., 2006). This kind of analysis allows the identification of non-functional CDSs within specific strains and thus may lead to identification of potential reasons for phenotypic strain variations.

The significance of pseudogenes in early genome annotations were frequently ignored, as these were considered sequencing artefacts. However, given the recent realisation of the importance of pseudogenes in pathoadaptive mutations, a greater significance is placed on their identification (Lerat and Ochman, 2004, Homma *et al*., 2002). An example of this is the re-analysis of the *Escherichia coli* strain K-12 genome sequence, which predicted an additional 160 pseudogenes in comparison to the single pseudogene identified in the original annotation (Ochman and Davalos, 2006). The same study also indicated pseudogenes are continually generated, with existing pseudogenes being eliminated over a period of time (Ochman and Davalos, 2006). Pseudogenes can accumulate in the genomes of some bacterial species, especially those undergoing processes like niche adaptation or host specialisation (Mira and Pushker, 2005). Analysis of further *Campylobacter* strains and species along with additional epsilon proteobacteria species will aid our understanding on this emerging research area. Also, greater

understanding of pseudogene dynamics and in particular innovative pseudogene identification techniques will yield more information about the actual number and purpose of these entities within microorganisms.

Phase-variable CDSs containing hypervariable regions were also analysed. The initial annotation correctly identified for the first time a number of hypervariable sequences found within the *C. jejuni* genomic shotgun sequence (Parkhill *et al*., 2000). These hypervariable sequences are scattered throughout the genome, however there are notable clusters present within the *O*-linked glycosylation, CPS and LOS loci. Further research on these loci has illustrated the impact of phase-variation on *C. jejuni* pathogenicity (Linton *et al*., 2001, Karlyshev *et al*., 2002, Szymanski *et al*., 2003b). A comparison of the number of phase-variable CDSs (i.e. those with an intersecting homopolymeric tract) was performed between *C. jejuni* NCTC11168, RM1221, 81-176, 81116 and CG8486 strains and all strains were identified as possessing a large number of homopolymeric tracts (polyG/C), thus denoting the importance of phase-variable CDSs within *C. jejuni* genomes (Table 3.10). It should also be noted that sequencing errors are also a factor in not only the identification of homopolymeric tracts, but also to show whether the phase-variable CDS are in frame or not (Karlyshev *et al*., 2002). This is even more pertinent with next-generation sequencing as certain methodologies (i.e. 454/pyrosequencing) have been noted to have difficulties identifying homopolymeric regions.

The selection of 15 CDSs for further characterisation was based predominantly on identification of novel motifs and literature searches. These CDSs were grouped into either putative virulence-associated, signal transduction or regulation of gene expression (Table 3.11). Table 3.11 lists the selected mutants and indicates where defined isogenic *C. jejuni* 11168H mutants were successfully prepared. Although cloned CDS fragments with $Km^R$ cassettes were constructed for all 15 CDSs, it is interesting to note that seven of these had no $Km^R$ transformants identified suggesting that some or all of these mutations were lethal to the host. This suggests the selection strategy for the identification of CDSs with putative functions linked to roles in virulence, signal transduction or regulation of gene expression was accurate as it is possible a number of the mutants which could not be constructed were important to the host.

### 3.6 Conclusion

The re-annotation of the *C. jejuni* NCTC11168 genome sequence was completed in 2007 and is arguably already out of date. Novel studies such as *C. jejuni* RNA-Seq data should be incorporated into future re-annotations (van Vliet, 2010). In this re-annotation, 18.2% of product functions were updated and 90.0% of CDSs had their note qualifier updated. The re-annotated genome sequence has been made publicly available with an updated EMBL file and a supporting publication (Gundogdu. *et al*, 2007/Appendix 5). The manuscript has allowed detailed genomic information to be made available regarding the evolution of our understanding of this genome. This re-annotation has been a useful resource for the *Campylobacter* research community, allowing up-to-date information to be accessible for not only this particular strain, but also related species. This updated genome sequence has also been used in third-party software such as microarray and proteomic genome annotations. The re-annotation process identified 20 additional CDSs linked to 'small molecule metabolism', six additional CDSs linked to 'broad regulatory functions', 26 additional CDSs linked to 'cell process' and 77 CDSs linked to 'miscellaneous' functions. The latter function relates to CDSs with new motifs that have not yet been linked to a specific function. Interestingly, 122 CDSs were removed from the 'unknown and other' category mainly due to the assignment of CDSs with new motifs. Significant updates were made within the four loci encoding the *N*- and *O*-linked glycosylation systems, CPS and LOS bioysthesis. Over the last decade, the CDSs within these loci have been characterised and this has greatly enhanced our understanding of the pathogenesis for this bacterium. In addition, novel CDSs such as regulators and the further refinement of many CDS product functions have allowed a greater understanding of the genome, mainly due to the availability of new motifs. All of these updates have, and will continue to be used in *Campylobacter* research.

The manual re-annotation of the genome identified a set of CDSs yet to be fully characterised, that may have putative roles in virulence, signal transduction or regulation of gene expression. As a result of this re-annotation, the functions of Cj1556 and Cj0248 were investigated as reported in Chapters 4 and 5 respectively.

# Chapter 4: *Cj1556* encodes a putative transcriptional regulator which has a role in oxidative and aerobic ($O_2$) stress response along with bacterial survival *in vivo*

## 4.1 Introduction

Re-annotation of the NCTC11168 genome sequence resulted in the identification of 15 CDSs for further study (see Section 3.4). The selection of these 15 CDSs was based on a number of different criteria such as newly identified motifs along with literature searches. The function of each CDS was unknown, however putative functions were linked to roles in virulence, signal transduction or regulation of gene expression. Following the successful construction of eight defined isogenic *C. jejuni* 11168H mutants, a number of preliminary phenotypic assays such as stress tests, motility, interaction and invasion assays were performed. Two mutants were chosen for further characterisation and this chapter details the characterisation of the 11168H *Cj1556* mutant.

### 4.1.1 Aims

- Construct a defined isogenic *C. jejuni* 11168H *Cj1556* mutant
- Construct a 11168H *Cj1556* complement
- Characterise the 11168H *Cj1556* mutant

## 4.2 Construction and characterisation of the 11168H *Cj1556* mutant

### 4.2.1 Bioinformatic analysis of *Cj1556*

The 333 nucleotide predicted CDS of *Cj1556* was originally annotated as a hypothetical protein in the genome sequence of *C. jejuni* NCTC11168. Following the re-annotation of the *C. jejuni* NCTC11168 genome sequence performed in this study (see Chapter 3), the updated product function indicated Cj1556 as a putative transcriptional regulator based on the identification of a new Pfam motif (PF01638), defined as an HxlR-like helix-turn-helix motif. The HxlR-like helix-turn-helix motif is located 45 nucleotides into the CDS and encompasses the remainder of the CDS. The HxlR-like helix-turn-helix motif is part of the MarR family of transcriptional regulators that includes proteins that control virulence factor production, bacterial responses to both antibiotics and oxidative stress and also catabolism of environmental aromatic compounds (Wösten *et al*., 2008,

Wilkinson and Grove, 2004). In addition to Cj1556, the *C. jejuni* NCTC11168 genome contained one other CDS (Cj1546) with the MarR family motif. Cj1546 was also re-annotated as a putative transcriptional regulator with 43.6% identity and 58.4% similarity to Cj1556. (Figure 4.1).

```
1556    ATGAAAAAATATCATTCTCTTTGTCCTATTGAAAC---CACGCTTAATTTGA---TAGGA 54
1546    ATGACTAAAGAGAATTCTCCGTGCAATTTCGAAGAATGTGGATTTAACTATACTCTAGCA 60
        ****  *** *  ******  **   * * ***          **** *  *   *** *

1556    --AATAAATGGAAAATTTTAATTATTCGAGATTTGTTGCAAGGGACTAAAAGATTTGGAG 112
1546    TTAATTAATGGAAAATATAAAATGAGTATTTTATACTGCTTATTTCGCTATGAAATCGTG 120
          *** ********** * ** *        * *  ***      *   * **  * * *

1556    AGTTAAGAAAAAGCATTTCATTTACAAAAAATCAAAACATCTCTCAAAATGTCCTTACAC 172
1546    CGTTATAATGAACTCAAACGCTTTTTAAGCTCCA---TATCTTTTAAAA--CTTTAACCA 175
         ****  *  **      *  **   **    **    **** * ****     * **

1556    AAAATTTACGAGAATTAGAAGAAGCAAAACTGATCAAACGCAAAGTTTATGCAGAAGTTC 232
1546    ATACCTTAAGAGAACTAGAAAACGATGGTTTAATTATCAGAAAAGAATACGCTCAAATTC 235
        * *  *** ***** ***** * *      * ** *   * ****  ** **  ** ***

1556    CTCCAAAGGTTGAGTACTCACTTA--CATCTTTAGGCAATAGTTTAGAATCCATCTTAAA 290
1546    CACCCAAGGTAGAATA-TAGCTTATCCAAACGTGGACAA-AGTTTAATTCCCATTTTACA 293
        * ** ***** ** ** *  ****  **    * * *** ******    **** *** *

1556    AAGTCTTGAGAATTGGGGAAATAGCTATAAAAATATCGTTTAA------------ 333
1546    AGCTATGAGCAAATGGGGCAAGAAAGACAAAAAAGGGAAAAAATGCTTGAATTAA 348
        *  * *    ** ***** ** *   * *****        **
```

Figure 4.1. ClustalW alignment comparison of *Cj1556* and *Cj1546* nucleotide sequences. Star indicates identical match. Dash indicates insertion site. ClustalW tool was used at - http://www.ebi.ac.uk/Tools/clustalw2/help.html. Underline indicates location of HxlR-like helix-turn-helix motif.

The predicted function of Cj1556 was investigated further using the *Campylobacter* Protein Interaction Database (Parrish *et al*., 2007) and putative interactions with Ctb (Cj0465c) were identified. Ctb is a group III truncated haemoglobin and characterisation studies in *C. jejuni* have already shown Ctb to be part of the nitrosative stress response regulon (Monk *et al*., 2008) (Appendix 6). Ctb has also been linked with moderating oxygen metabolism within *C. jejuni* (Monk *et al*., 2008). The *Campylobacter* Protein Interaction Database was also used to investigate Cj1546. A number of common interactions e.g. CheA, were identified for both Cj1546 and Cj1556 (Appendix 7), however interactions relating to nitrosative stress were only identified for Cj1556.

Protein Blast search results surprisingly identified this CDS to be present in relatively few strains/species of *Campylobacter* (Table 4.1). Homologues with relatively high identity and expect scores were present in *C. jejuni* strains CF93-6, 84-25 and 81-176, along with the species *C. fetus* 82-40. The significance of this CDS not being present in all *C. jejuni* strains is unclear, although one of the most commonly used laboratory strains, *C. jejuni* 81-176, a highly pathogenic strain does have the CDS (Hofreuter *et al*., 2006). Apart from *C. jejuni*, other related epsilon proteobacteria with *Cj1556* homologues included *Helicobacter bilis* ATCC 43879, *H. pullorum* MIT 98-5489 and *H. hepaticus* ATCC 51449. The homologue was also found in unrelated bacterium such as *Clostridium leptum*. Search results identified *Campylobacter* species and related organisms harbouring homologues named as either a 'hypothetical protein' or a regulatory/transcription designation based on the presence of the motif. No homologue has been characterised as yet. Considering all the bioinformatic analysis, Cj1556 was hypothesised to have a role as a stress response regulator.

Table 4.1. BlastP results for *Cj1556*. Campylobacter strains/species indicated below.

| Species / Strain | Identities score | Expect score |
|---|---|---|
| *Campylobacter jejuni* subsp. *jejuni* NCTC11168 | 110/110 | 6e-56 |
| *Campylobacter jejuni* subsp. *jejuni* CF93-6 | 110/110 | 6e-56 |
| *Campylobacter jejuni* subsp. *jejuni* 84-25 | 110/110 | 6e-56 |
| *Campylobacter jejuni* subsp. *jejuni* 81-176 | 110/110 | 6e-56 |
| *Helicobacter bilis* ATCC 43879 | 85/109 | 1e-43 |
| *Campylobacter fetus* subsp. *fetus* 82-40 | 84/108 | 4e-43 |
| *Campylobacter fetus* subsp. *Venerealis* str. Azul-94 | 84/108 | 4e-43 |
| *Helicobacter pullorum* MIT 98-5489 | 86/108 | 7e-43 |
| *Helicobacter hepaticus* ATCC 51449 | 84/107 | 2e-42 |
| *Anaerofustis stercorihominis* DSM 17244 | 72/110 | 1e-33 |
| *Clostridium leptum* DSM 753 | 8e-31 | 64/100 |

### 4.2.2 Construction of 11168H *Cj1556* mutant and complement

The mutagenesis strategy and all techniques used to create a defined isogenic 11168H *Cj1556* mutant and a 11168H *Cj1556* complement are described in Chapter 2. Briefly, primers were designed for PCR amplification and detection of a *Cj1556* CDS fragment were performed using *Cj1556*-F and *Cj1556*-R (see Section 2.2.5). A vector construct (cam25a2 - 1489074..1490567) containing the *Cj1556* CDS was selected from the *C.*

*jejuni* NCTC11168 genome sequence pUC library. This vector construct is 1.49 kb in size and includes the CDSs *Cj1555c – Cj1560*. The vector construct was named pUC-*Cj1556*. Digestion with *Bcl*I was followed by ligation with the *aph-3* (aminoglycoside 3'-phosphotransferase) kanamycin resistance cassette ($Km^R$) (Trieu-Cuot *et al*., 1985) and transformed into XL-2 cells. The $Km^R$ cassette was derived from pJMK30 which does not contain a transcriptional terminator. The allows the formation of multicistronic mRNA as the $Km^R$ can be inserted in the same orientation as the CDS of interest, hence removing the potential to form polar effects (van Vliet *et al*., 1998). Transformants were screened by PCR using *Cj1556* specific and $Km^R$ specific primers; $Km^R$ forward-out and $Km^R$ reverse-out (see Section 2.2.5). pUC-*Cj1556*-$Km^R$ plasmids with the $Km^R$ cassette in the same orientation as the *Cj1556* CDS were selected and electroporated into 11168H wild-type as described previously (Jones *et al*., 2004, Karlyshev *et al*., 2002). Electroporated *C. jejuni* were plated onto blood agar plates and incubated at 37°C under microaerobic conditions for 2 days. Cells were harvested and resuspended in 0.5 ml PBS. 200 μl of this suspension was spread onto blood agar plates containing kanamycin. Putative *Cj1556* mutants were screened using PCR and sequencing (Figure 4.2).

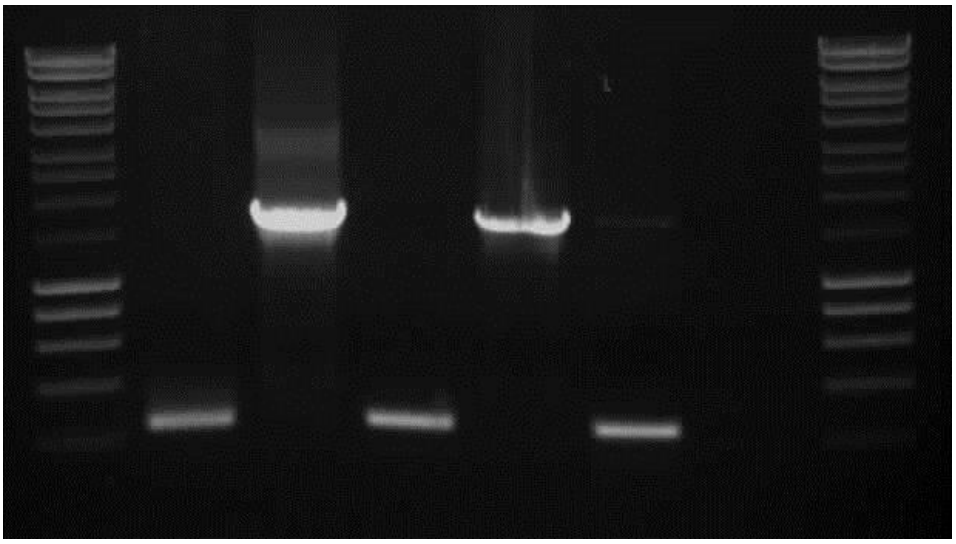



Figure 4.2. Confirmation of *Cj1556* constructs and 11168H *Cj1556* mutant using PCR with *Cj1556* specific primers.

Lane 1 – 1 kb ladder / Lane 2 – pUC-*Cj1556* plasmid DNA amplified with *Cj1556* specific primers (0.33 kb) / Lane 3 – pUC-*Cj1556*-$Km^R$ plasmid DNA amplified with *Cj1556* specific primers (1.76 kb) / Lane 4 – *C. jejuni* 11168H genomic DNA amplified with *Cj1556* specific primer (0.33 kb) / Lane 5 – *C. jejuni Cj1556* mutant genomic DNA amplified with *Cj1556* specific primer (1.76 kb) / Lane 6 – *C. jejuni Cj1556* complement genomic DNA amplified with *Cj1556* specific primer (0.33 kb and 1.76 kb) / Lane 7 – Negative control

Insertion of a copy of the *Cj1556* CDS into the *Cj1556* mutant chromosome for complementation purposes was achieved using pDENNIS, a *C. jejuni* 11168H complementation vector allowing the insertion of a functional *Cj1556* CDS into the pseudogene *Cj0223* (Hitchen *et al*., 2010). The coding region for *Cj1556* was amplified by PCR using primers; Comp-*Cj1556*-F and Comp-*Cj1556*-R, which introduced an *Nco*I site at the 5' end and an *Nhe*I site at the 3' end as well as the native ribosome binding site of *Cj1556* (Svensson *et al*., 2008, Wosten *et al*., 1998). Following digestion with *Nhe*I and *Nco*I, this PCR product was ligated into the pDENNIS vector. The construct was sequenced to ensure there were no mutations in the *Cj1556* nucleotide sequence. This construct was electroporated into the *Cj1556* mutant strain and grown on blood agar plates. Putative mutants were restreaked onto blood agar plates containing kanamycin

and chloramphenicol. Positive clones were confirmed by PCR checking for the correct distance between *Cj1556* and the inserted $Km^R$ cassette (*Cj1556*-F and $Km^R$ forward-out). This confirmed the presence of the mutated CDS with the $Km^R$ cassette in the correct orientation. PCR was also used to confirm the presence of an intact CDS (*Cj1556*-F and *Cj1556*-R). In addition, sequencing was also used to further confirm positive mutants

### 4.2.3 The 11168H *Cj1556* mutant exhibits increased sensitivity to oxidative but not nitrosative stress

Data from the *Campylobacter* Protein Interaction Database identified putative interactions between Cj1556 and Ctb. Ctb, a truncated haemoglobin encoded by *Cj0465c*, along with Cgb, a single domain haemoglobin encoded by *Cj1586*, are part of the *C. jejuni* nitrosative stress regulon (Wainwright *et al*., 2006, Wainwright *et al*., 2005, Elvers *et al*., 2005, Elvers *et al*., 2004). The regulon is under the control of NssR (encoded by *Cj0466*). Previous studies have implicated Ctb with a role in oxygen metabolism (detoxification, sequestration or transfer of oxygen) within *C. jejuni* (Wainwright *et al*., 2005, Wainwright *et al*., 2006). Identification of Cj1556 as a member of the MarR family of transcriptional regulators and the link to Ctb through the Protein Interaction Database searches suggested a role in the *C. jejuni* nitrosative stress response. Nitrosative stress assays were performed using acidified $NaNO_2$ (10 mM / 75 mins). No differences between the survival of the *Cj1556* mutant and the 11168H wild-type strain were observed under nitrosative stress (Figure 4.4). However, MarR family proteins have been shown to be involved in oxidative stress (Wilkinson and Grove, 2004) and even though Ctb is part of the nitrosative stress regulon, it has also been shown to be involved in oxygen delivery (Wainwright *et al*., 2005, Wainwright *et al*., 2006). Oxidative stress assays were performed using $H_2O_2$. The *Cj1556* mutant exhibited increased sensitivity to $H_2O_2$ (10 mM / 15 mins) compared to the 11168H wild-type strain (Figures 4.5 and 4.6). In addition, the *Cj1556* complement restored the 11168H wild-type strain $H_2O_2$ sensitivity phenotype. The increased resistance to $H_2O_2$ observed for the *Cj1556* complement compared to the 11168H wild-type strain may be due to the *Cj1556* expression being controlled by the constitutively expressed chloramphenicol promoter on the complementation vector instead of the native *Cj1556* promoter.

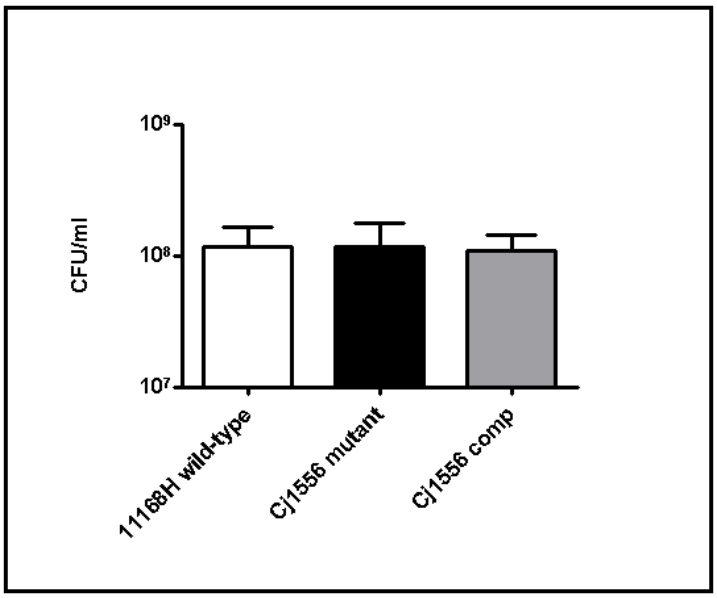



Figure 4.4. Nitrosative stress assays on *C. jejuni* 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp). *C. jejuni* were grown on blood agar plates for 24 h at 37°C under microaerobic conditions. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared and incubated with 10 mM $NaNO_2$ for 75 minutes under microaerobic conditions. Serial dilutions ($10^{-1}$ - $10^{-6}$) were performed and 10 µl volumes were spotted onto blood agar plates. The plates were incubated at 37°C under microaerobic conditions for 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

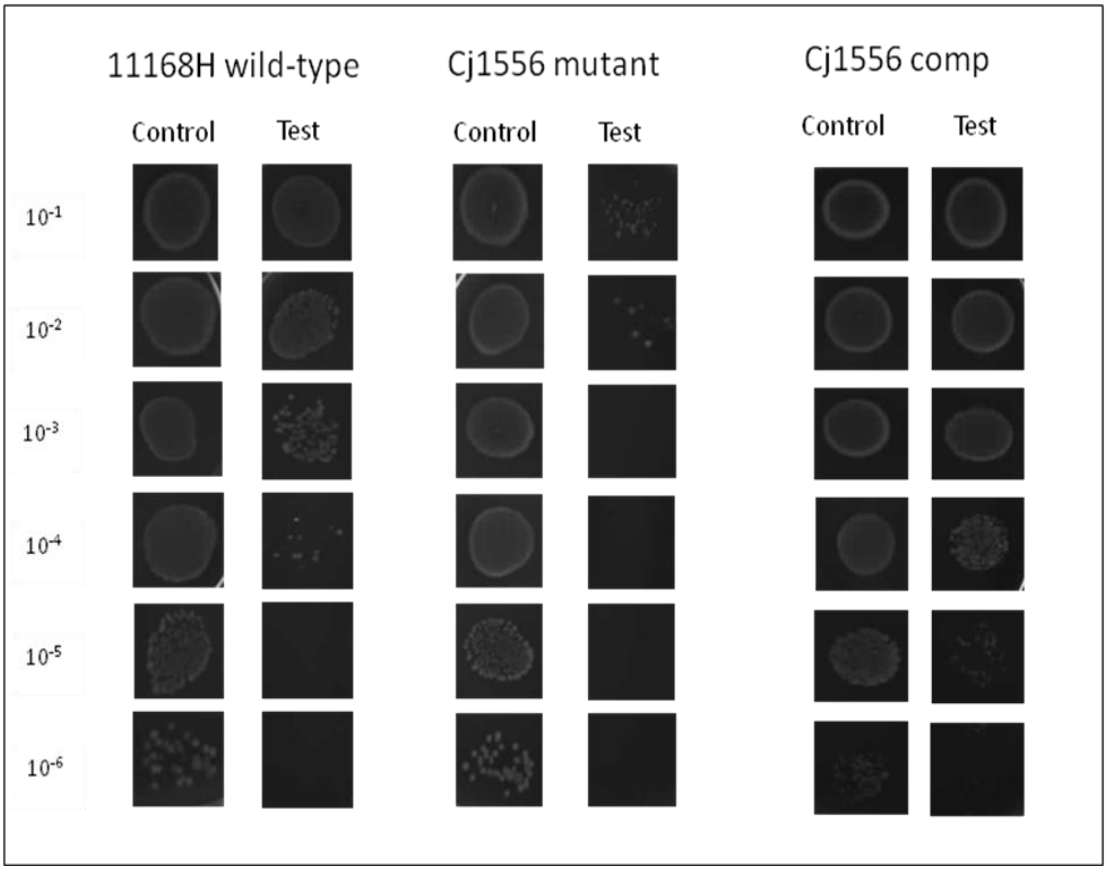



Figure 4.5. Oxidative stress assays on *C. jejuni* 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp). *C. jejuni* were grown on blood agar plates for 24 h at 37°C under microaerobic conditions. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared and incubated with 10 mM $H_2O_2$ for 15 minutes under microaerobic conditions. Serial dilutions ($10^{-1}$ - $10^{-6}$) were performed and 10 µl volumes were spotted onto blood agar plates. The plates were incubated at 37°C under microaerobic conditions for 48 h. An example is shown of one of three biological replicates performed in triplicate.

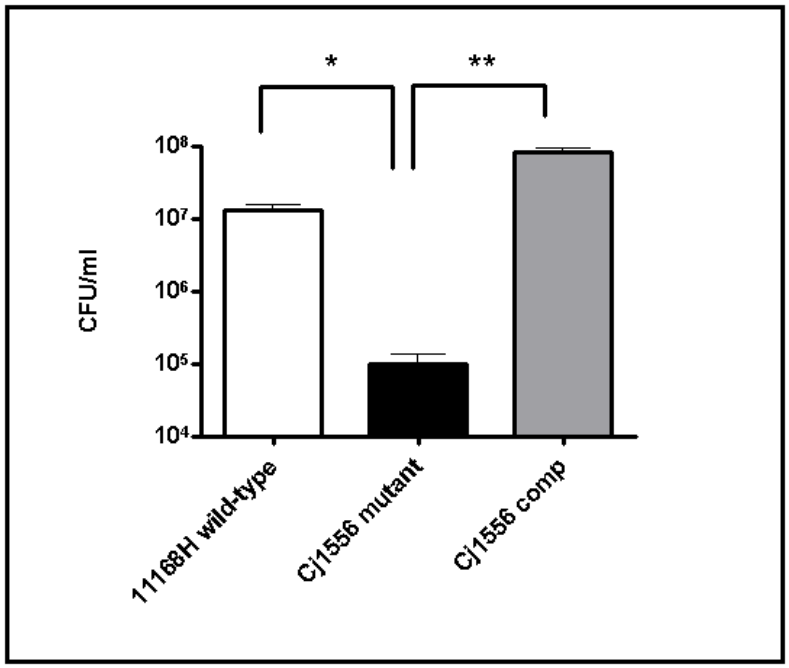



Figure 4.6. Oxidative stress assays on *C. jejuni* 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp). *C. jejuni* were grown on blood agar plates for 24 h at 37°C under microaerobic conditions. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared and incubated with 10 mM $H_2O_2$ for 15 minutes under microaerobic conditions. Serial dilutions ($10^{-1}$ - $10^{-6}$) were performed and 10 µl volumes were spotted onto blood agar plates. The plates were incubated at 37°C under microaerobic conditions for 48 h. Figure shows the viable counts in terms of CFU per ml. The asterisk denotes a statistically significant difference (* = $p$ <0.05) for *Cj1556* mutant compared to the 11168H wild-type strain. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p$ = 0.01 - 0.05 and two stars (**) indicating $p$ = 0.001- 0.01.

### 4.2.4 The 11168H *Cj1556* mutant reveals unaltered growth kinetics during growth inhibition studies

As a result of the oxidative stress assay, a number of growth inhibition studies were performed to ascertain whether any modification in growth kinetics was observed due to the addition of specific inhibitory compounds. Growth inhibition studies were performed using conditions to represent both oxidative and nitrosative inhibition. No differences were observed between the *Cj1556* mutant and the 11168H wild-type strain for oxidative (Figure 4.7) and nitrosative (Figure 4.8) inhibition. Recently *Cj1556* was shown to be 2.8

fold up-regulated in the presence of 0.1% (w/v) sDOC (Malik-Kale *et al*., 2008). Growth analysis of the *Cj1556* mutant in broth supplemented with 0.1% (w/v) sDOC showed no difference compared to the 11168H wild-type strain (Figure 4.9). In addition, no differences were observed between *Cj1556* mutant and the 11168H wild-type strain during growth in broth supplemented with the iron chelator deferoxamine (Figure 4.10).

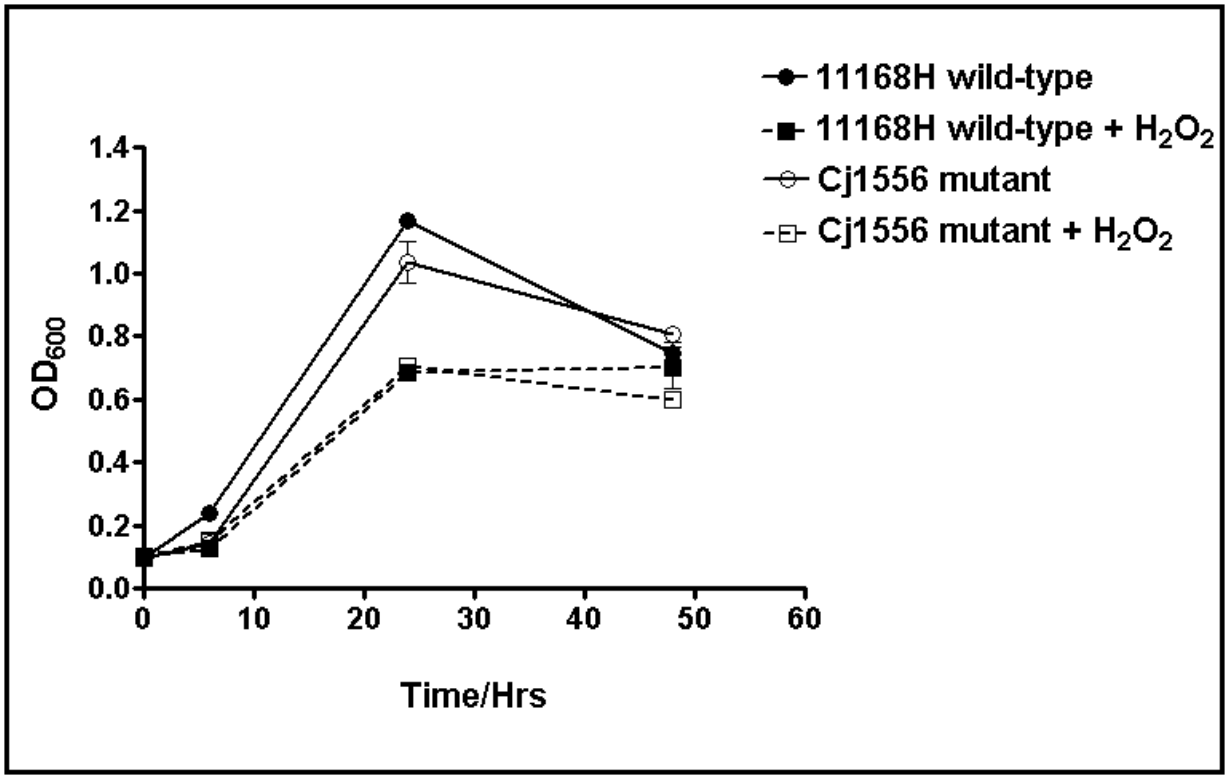



Figure 4.7. Oxidative stress growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj1556* mutant. *C. jejuni* were grown in the presence of 1 mM $H_2O_2$. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

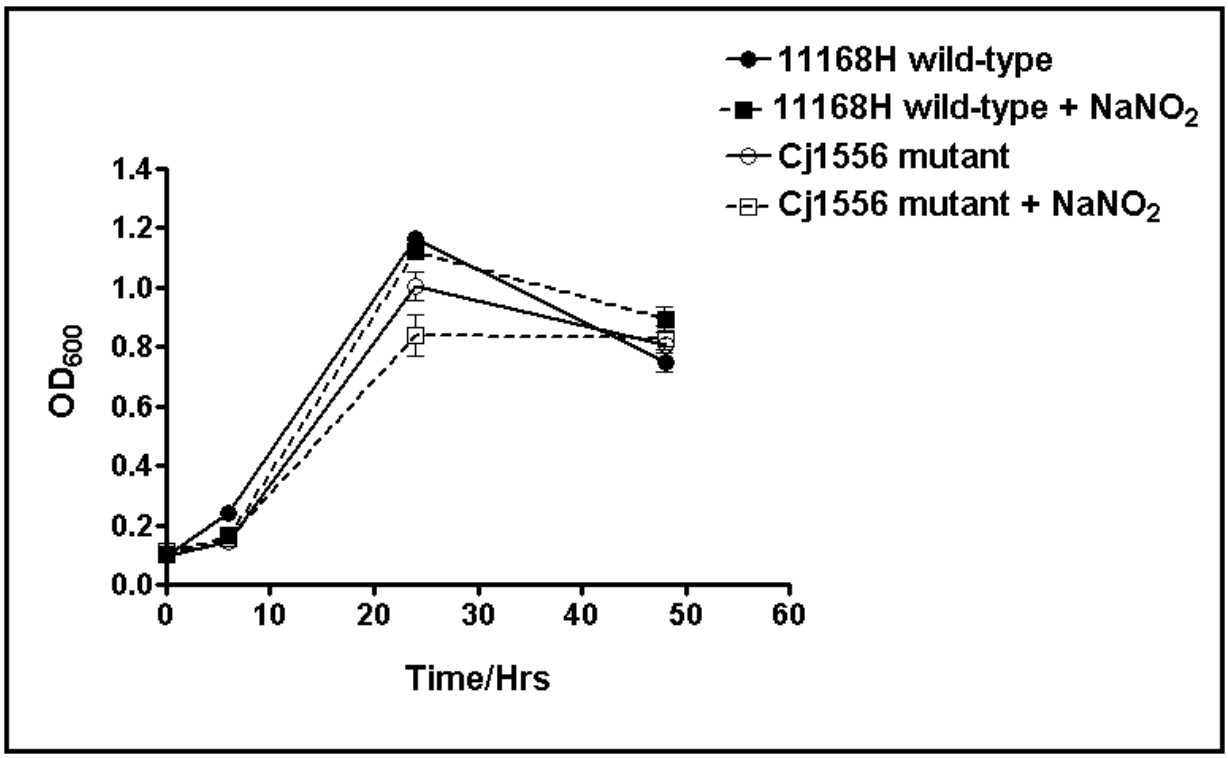



Figure 4.8. Nitrosative stress growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj1556* mutant. *C. jejuni* were grown in the presence of 0.1 mM $NaNO_2$ at pH 5. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

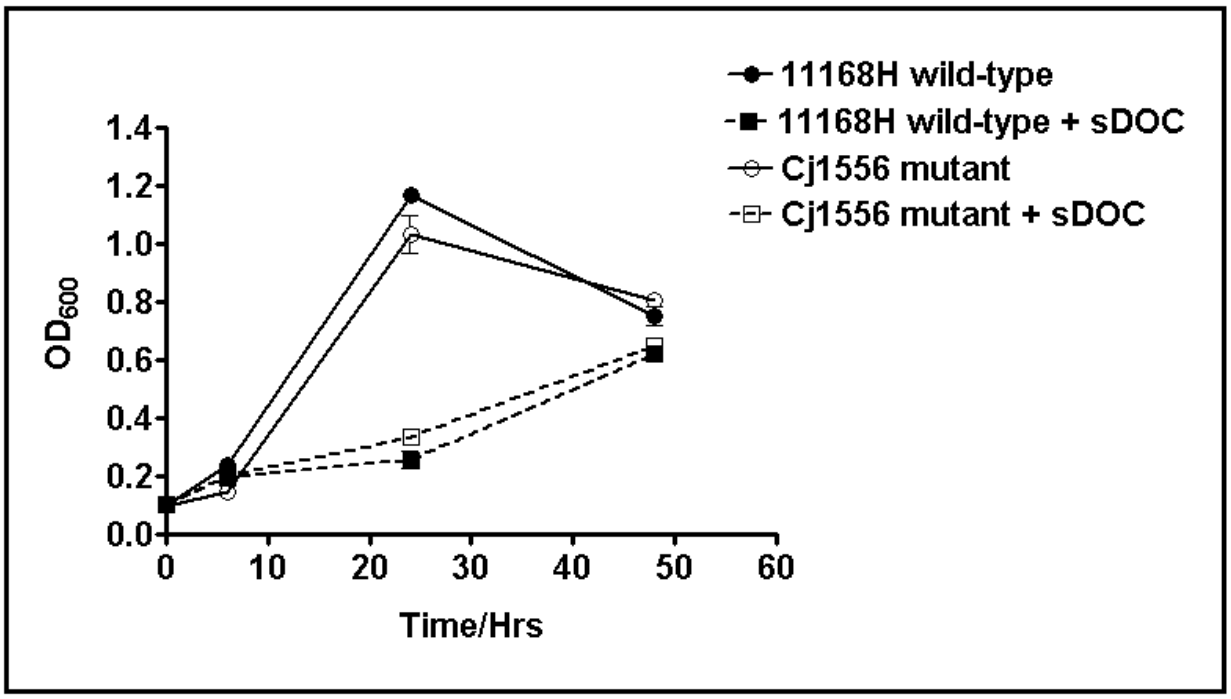



Figure 4.9. Bile stress growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj1556* mutant. *C. jejuni* were grown in the presence of 0.1% sDOC. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

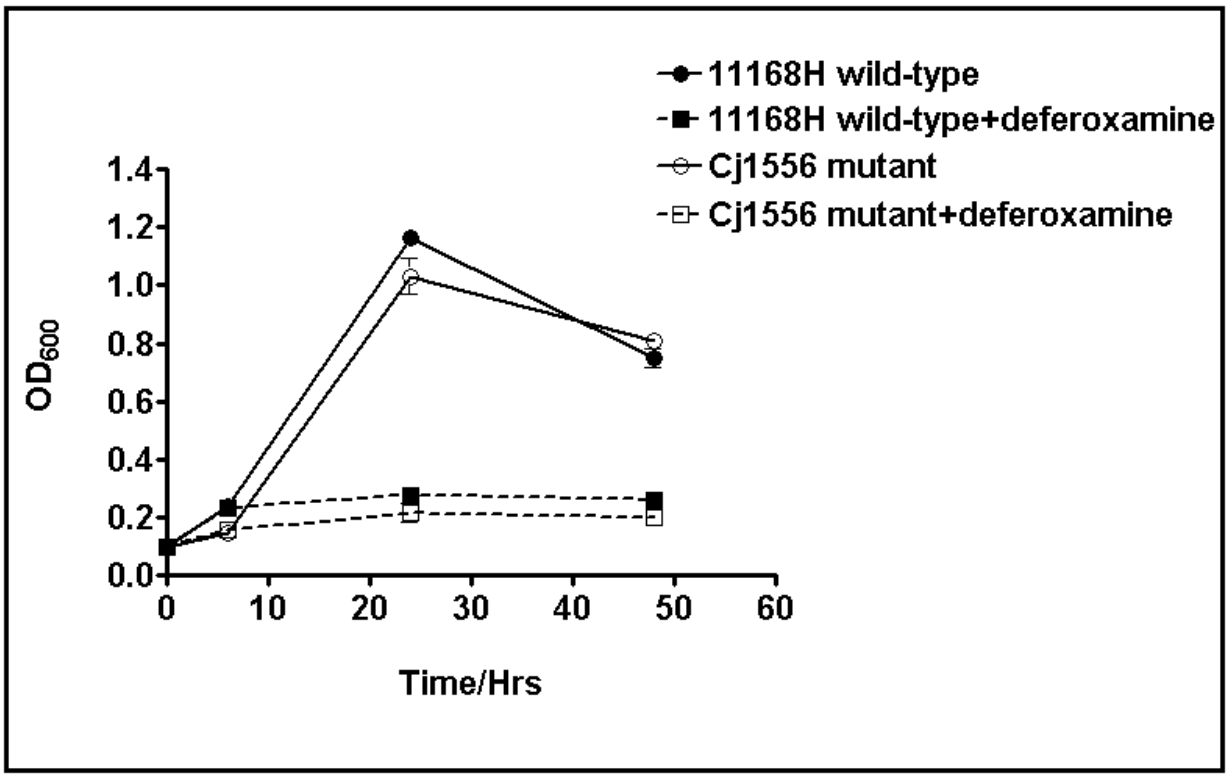



Figure 4.10. Iron limitation growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj1556* mutant. *C. jejuni* were grown in the presence of 1 mM deferoxamine. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

### 4.2.5 The 11168H *Cj1556* mutant exhibits increased sensitivity to heat stress

Previous research has suggested a link between aerobic and heat stress (Lin *et al*., 2009, Andersen *et al*., 2005, Brondsted *et al*., 2005, Phongsisay *et al*., 2007). The HtrA protease and HspR regulator have already been described as playing a role in aerobic ($O_2$) stress, however their primary role is related to heat tolerance (Brondsted *et al*., 2005). The microarray data identified *hspR* as being 2.07 fold down-regulated in the *Cj1556* mutant compared to the 11168H wild-type strain. In order to investigate this further, a range of heat stress experiments were performed. No significant differences in survival were observed at 42°C / 60 mins or 55°C / 15 mins. However the *Cj1556* mutant displayed increased sensitivity compared to the 11168H wild-type strain at 60°C / 5 minutes and the *Cj1556* complement restored the 11168H wild-type strain phenotype

(Figure 4.3). Heat stress above 55°C has been noted to accelerate the spiral-to-coccoid transition and result in cell death (Nguyen *et al*., 2006).

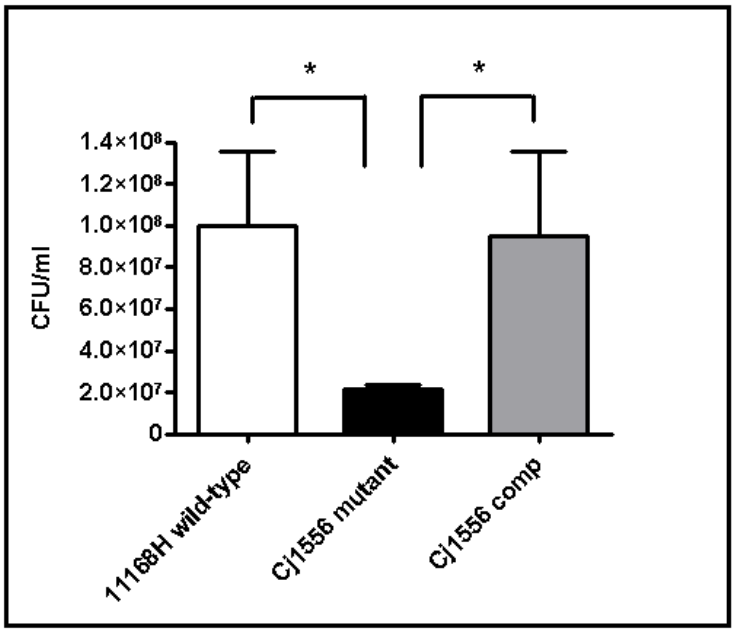


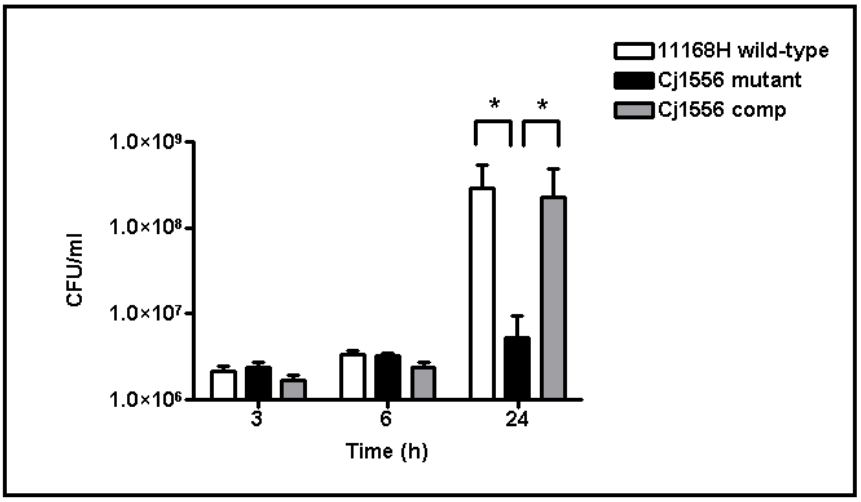

Figure 4.3. Heat stress assays on *C. jejuni* 11168H wild-type strain, *Cj1556* mutant and *Cj1556* complement. *C. jejuni* grown overnight had $OD_{600}$ adjusted to 1.0 and were exposed to 60°C for 5 minutes. Serial dilutions were performed and 200 µl volumes were spotted onto blood agar plates. The plates were incubated at 37°C under microaerobic conditions for 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$.

### 4.2.6 The 11168H *Cj1556* mutant displays a reduced ability to interact with and invade Caco-2 intestinal epithelial cells

In order to investigate the colonisation potential of the *Cj1556* mutant, interaction (adhesion and invasion) and invasion assays were performed using co-culture experiments with Caco-2 cells and the 11168H wild-type strain, *Cj1556* mutant and *Cj1556* complement. No significant differences were observed when comparing the levels of interaction at either 3 h or 6 h, however the *Cj1556* mutant displayed a reduced ability to interact with Caco-2 cells after 24 h co-culture compared with the 11168H wild-type strain and *Cj1556* complement (Figure 4.11A). The *Cj1556* mutant also displayed a reduced ability to invade Caco-2 cells after 24 h co-culture compared with

the 11168H wild-type strain and *Cj1556* complement (Figure 4.11B). No significant differences were observed when comparing the levels of invasion at either 3 h or 6 h.

A)




B)

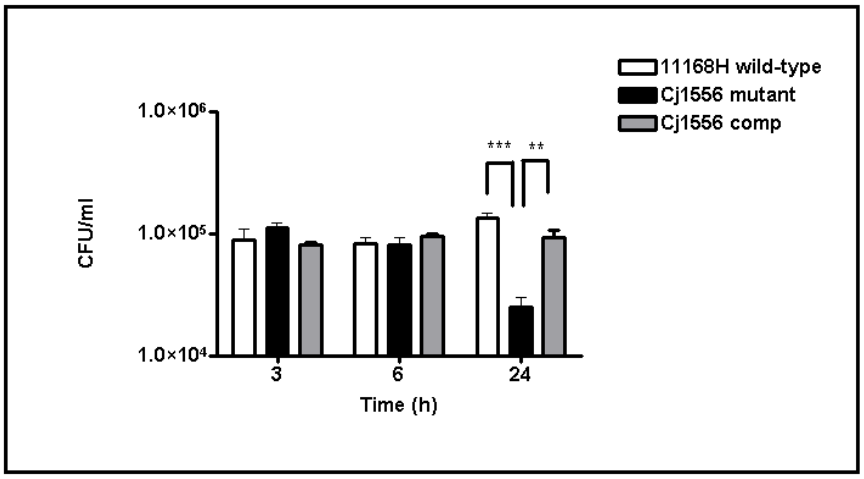



Figure 4.11. Interaction (adhesion and invasion) and invasion assays on 11168H wild-type strain, *Cj1556* mutant and *Cj1556* complement (*Cj1556* comp). *C. jejuni* were co-cultured with Caco-2 IECs for 3, 6 or 24 h. Caco-2 cells were either lysed and numbers of interacting bacteria assessed (A) or incubated with gentamicin (150 μg/ml) for 2 h to kill extracellular bacteria, then lysed and numbers of intracellular bacteria assessed (B). All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$, two stars (**) indicating $p = 0.001 - 0.01$ and three stars (***) $p < 0.001$.

### 4.2.7 The 11168H *Cj1556* mutant exhibits reduced intracellular survival in Caco-2 intestinal epithelial cells

In order to investigate the ability of *C. jejuni* to survive when exposed to intracellular stress such as ROS, a modification of the interaction and invasion assays was used to analyse the level of intracellular survival in Caco-2 IECs (Naito *et al*., 2010). Here, *C. jejuni* were co-cultured with Caco-2 cells for 3 h, extracellular *C. jejuni* were removed and intracellular bacteria were maintained for 19 h. The results indicated a decrease in the level of intracellular survival of the *Cj1556* mutant compared to the 11168H wild-type strain and *Cj1556* complement (Figure 4.12).

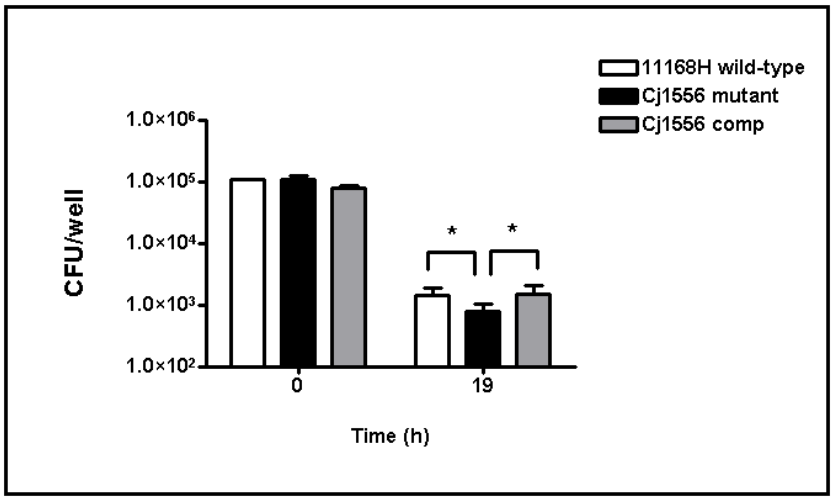



Figure 4.12. Intracellular survival assays on 11168H wild-type strain, *Cj1556* mutant and *Cj1556* complement (*Cj1556* comp). *C. jejuni* were co-cultured with Caco-2 IECs for 3 h, then incubated with gentamicin (150 μg/ml) for 2 h to kill extracellular bacteria, followed by further incubation with gentamicin (10 μg/ml) for 19 h. Cells were lysed and numbers of intracellular bacteria assessed. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$.

To ascertain whether the above results were due to a genuine *Cj1556* mutant phenotype and not to increased sensitivity to Triton X-100, stress assays were performed on all three strains with 0.2% (v/v) Triton X-100. No difference in the level of sensitivity to

Triton X-100 was observed between 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (Figure 4.13).

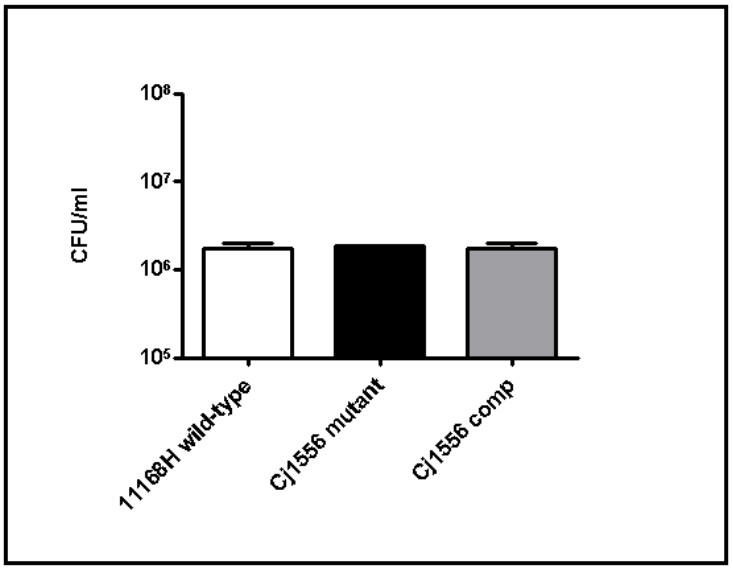



Figure 4.13. 0.2% (v/v) Triton X-100 stress assay on *C. jejuni* 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp). *C. jejuni* were grown on blood agar plates for 24 h at 37°C under microaerobic conditions. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared and incubated with 0.2% (v/v) Triton X-100 for 15 minutes under microaerobic conditions. Serial dilutions ($10^{-1}$ - $10^{-6}$) were performed and 10 µl volumes were spotted onto blood agar plates. The plates were incubated at 37°C under microaerobic conditions for 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

### 4.2.8 The 11168H *Cj1556* mutant exhibits reduced survival in co-culture media

A further experiment was performed to compare the number of *C. jejuni* present in the tissue culture medium from supernatant after 24 hours co-culture with Caco-2 cells. There was a statistically significant decrease in the number of viable bacterial cells obtained from the supernatant after 24 hours co-culture with Caco-2 cells when comparing the *Cj1556* mutant to the 11168H wild-type strain and *Cj1556* complement (Figure 4.14).

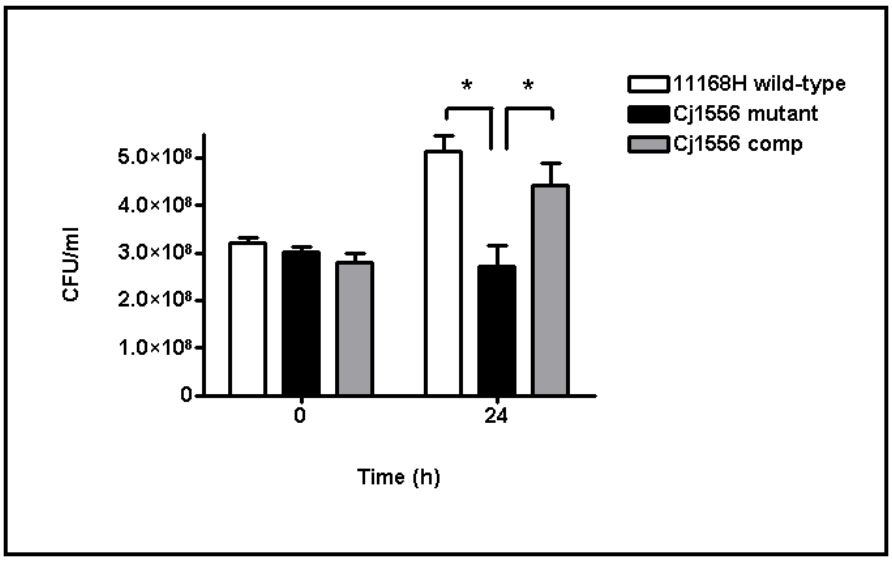



Figure 4.14. Survival assays on 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp). *C. jejuni* were co-cultured with Caco-2 IECs for 24 h followed by assessing the number of bacteria in the co-culture media. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$.

### 4.2.9 The 11168H *Cj1556* mutant exhibits reduced intracellular survival in J774A.1 macrophage cells

Intracellular survival assays using J774A.1 mouse macrophages cells were also performed to further investigate the survival rates of the 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement strains. Macrophages internalise and destroy *C. jejuni* with previous studies showing that *C. jejuni* are killed within 24 h of internalisation (Watson and Galan, 2008). There was a statistically significant reduction in the level of intracellular survival of the *Cj1556* mutant compared to the 11168H wild-type strain and *Cj1556* complement (Figure 4.15).

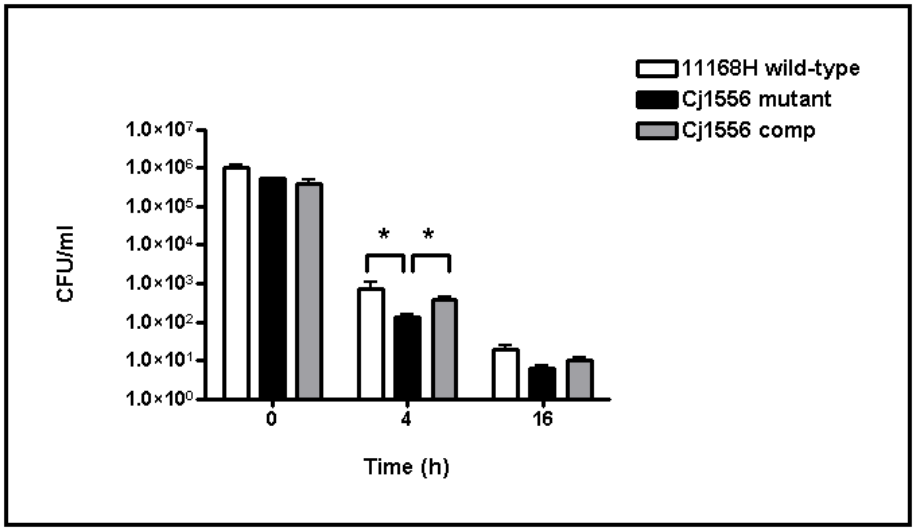



Figure 4.15. Intracellular survival assays on 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp). *C. jejuni* were co-cultured with J774A.1 mouse macrophages for 3 h, then incubated with gentamicin (150 μg/ml) for 2 h to kill extracellular bacteria, followed by further incubation with gentamicin (10 μg/ml) for 4 and 16 h. Cells were lysed and numbers of intracellular bacteria assessed. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p$ = 0.01 - 0.05.

### 4.2.10 The 11168H *Cj1556* mutant exhibits reduced survival in an aerobic environment

Following the identification of significant differences between the 11168H wild-type strain and *Cj1556* mutant in response to oxidative stress and intracellular survival, further investigations on the ability of these strains to survive aerobic ($O_2$) stress were performed. The difference in the level of *Cj1556* mutant survival between the interaction and intracellular assays suggested that additional stresses might affect *C. jejuni* during these assays. Survival assays using the 11168H wild-type strain, *Cj1556* mutant, and *Cj1556* complement were performed under either microaerobic or aerobic conditions in either Brucella broth or tissue culture media with no shaking to replicate the conditions for the co-culture assays. A reduction in the number of viable bacterial cells with the *Cj1556* mutant compared to the 11168H wild-type strain in both types of media were observed after 24 h incubation under aerobic conditions (Figure 4.16C and D), but not under microaerobic conditions (Figure 4.16A and B).

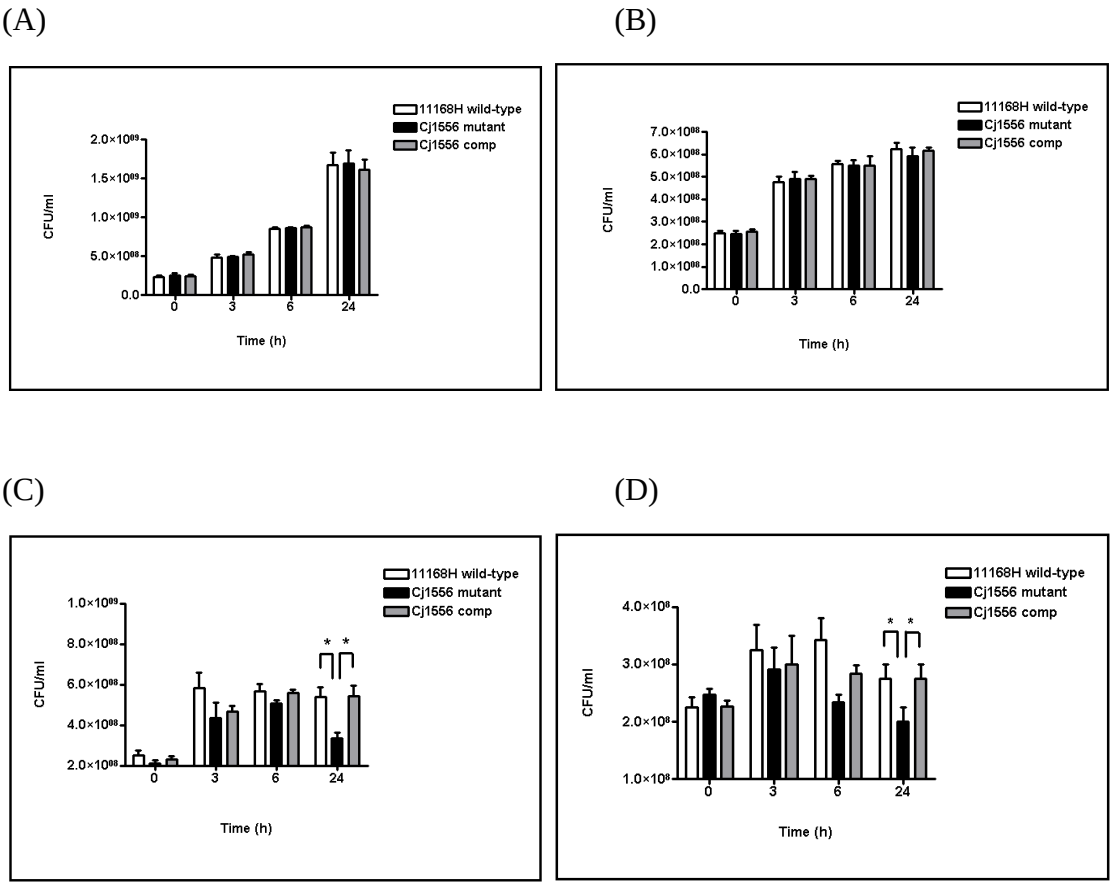


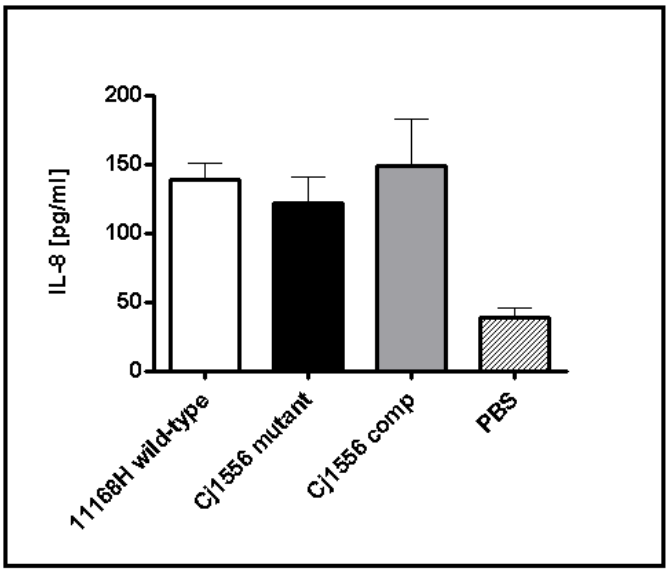

Figure 4.16. Aerobic ($O_2$) stress assays on 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp). *C. jejuni* were grown under microaerobic (A and B) and aerobic conditions (C and D) in Brucella broth (A and C) or tissue culture media (B and D), then numbers of viable bacteria assessed. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$.

### 4.2.11 The 11168H *Cj1556* mutant induces a reduced IL-6 response from T84 cells

Following on from the co-culture studies investigating the interaction and invasion properties of the *Cj1556* mutant, further co-culture investigations were performed analysing the host innate immune response. To analyse the host innate immune response, IL6 and IL8 were selected. IL-6 is an interleukin that acts as both a pro-inflammatory and anti-inflammatory cytokine and has been shown to be important for epithelial cell integrity (Friis *et al*., 2009). IL-6 is typically associated with an acute phase response (Heinrich *et al.*, 1990). Interleukin-8 is a chemokine produced by macrophages and other cell types such as epithelial cells. Both are well-characterised markers denoting a host

innate immune response against pathogens (Wolff *et al*., 1998, Oppenheim *et al*., 1991). To investigate any differences between the level of IL-6 and IL-8 induced by the *Cj1556* mutant compared to the 11168H wild-type strain, co-culture experiments with two different cell lines were performed. Only minimal secretion of IL-6 and IL-8 was detected when the 11168H wild-type strain and *Cj1556* mutant were co-cultured with Caco-2 cells (data not shown). However using the T84 cell line, significant levels of induction of both IL-6 and IL-8 by the 11168H wild-type strain and *Cj1556* mutant were observed (Figure 4.17). There was no significant difference in the level of IL-8 induction by the *Cj1556* mutant compared to that by the 11168H wild-type strain (Figure 4.17A), however a significant reduction in the level of IL-6 induction by the *Cj1556* mutant was observed (Figure 4.17B).

A)




B)

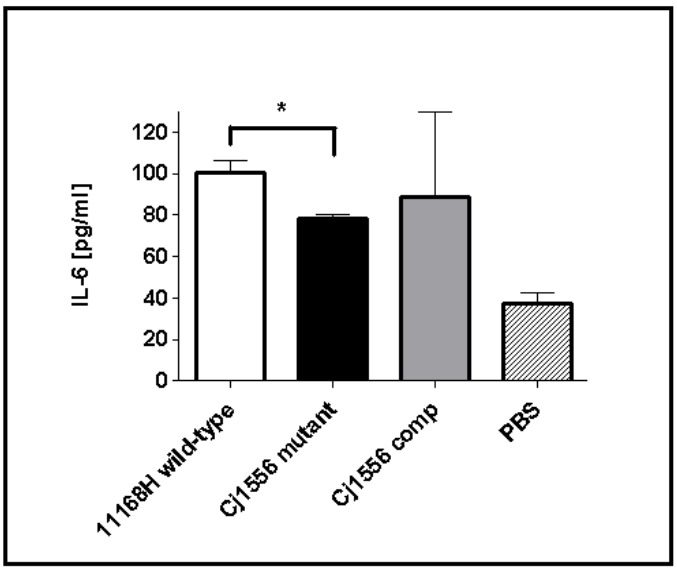



Figure 4.17. Analysis of the host innate immune response during 11168H wild-type strain, *Cj1556* mutant and the *Cj1556* complement (*Cj1556* comp) infection. *C. jejuni* were co-cultured with T84 IECs for 24 h and the levels of IL-8 and IL-6 secreted were quantified using either a human IL-8 ELISA (A) or IL-6 ELISA (B). All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$.

### 4.2.12 Microarray analysis of gene expression differences between *Cj1556* mutant and 11168H wild-type strain suggest negative autoregulation of Cj1556

To analyse the gene expression profile of the *Cj1556* mutant compared to the 11168H wild-type strain, microarray experiments were performed using total RNA samples isolated from *C. jejuni* grown to late-log phase (16 h). Microarray experiments were performed with three biological replicates, each with one technical replicate. Statistically significant up- and down-regulated genes were selected when comparing gene expression from mutant test arrays against 11168H control arrays using ANOVA (ANalysis Of VAriance) (Bacon *et al.*, 2004, Corcionivoschi *et al.*, 2009). ANOVA was performed using a Benjamini and Hochberg False Discovery Rate as the Multiple Testing Correction within the GeneSpring software (v7.3). A total of 91 genes were differentially expressed in the *Cj1556* mutant compared to the 11168H wild-type strain based on an ANOVA selection methodology. A total of 73 genes, including *Cj1556* were up-regulated in the *Cj1556* mutant, whilst 18 genes were down-regulated compared to the 11168H wild-type strain (Appendix 8 and 9). As most of the genes were up-regulated (73), this suggests a putative role for Cj1556 as a transcriptional repressor and thus the inactivation of this CDS results in the transcription of many different genes that would be repressed under these conditions in the 11168H wild-type strain. Changes in expression of genes linked to the *C. jejuni* oxidative and aerobic ($O_2$) stress responses in the *Cj1556* mutant compared to the 11168H wild-type strain are shown in Table 4.2. The genes listed in Table 4.2 do not appear in the significant ANOVA selection (Appendix 8 and 9), however some genes would be significant if a more classical selection method was adopted e.g. 2-fold up/down. In this scenario, genes such as *katA* (5.13), *perR* (5.05) and *hspR* (2.07) would be identified as down-regulated. Table 4.2 is only used to display expression changes of important oxidative and aerobic ($O_2$) stress genes for further discussion. *Cj1556* was identified with the highest fold up-regulation (10.4 fold). Sequence analysis of the *Cj1556* reporter element used on the arrays showed that this particular sequence was present upstream of the $Km^R$ cassette in the *Cj1556* mutant (data not shown). Typically the mutated gene in a defined mutant would be expected to be significantly down-regulated, however the microarray data indicates that in the absence of the Cj1556 protein, *Cj1556* gene expression is dramatically increased. This suggests that Cj1556 represses further expression of the *Cj1556* gene, acting as a negative autoregulation system.

Table 4.2. Changes in expression of genes linked to the *C. jejuni* oxidative and aerobic ($O_2$) stress responses in the *Cj1556* mutant compared to the 11168H wild-type strain.

| Gene Name | Fold Change | Product function |
|---|---|---|
| *spoT* | +1.26 | putative guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase |
| *sodB* | +1.24 | superoxide dismutase (Fe) |
| *htrA* | +1.21 | serine protease (protease DO) |
| *fdxA* | +1.07 | ferredoxin |
| *dcuA* | -1.17 | anaerobic C4-dicarboxylate transporter |
| *ahpC* | -1.27 | alkyl hydroperoxide reductase |
| *dps* | -1.36 | putative bacterioferritin |
| *hspR* | -2.07 | heat shock transcriptional regulator |
| *perR* | -5.05 | peroxide stress regulator |
| *katA* | -5.13 | catalase |
| *htrB* | No hybridisation | putative lipid A biosynthesis lauroyl acyltransferase |

### 4.2.13 *Cj1556* promoter analysis reveals similar promoter sequences to *ctb*

*In silico* analysis of *Cj1556* was performed using BPROM to predict the presence of bacterial $\sigma^{70}$ promoters (http://linux1.softberry.com/berry.phtml?topic=bprom). 300 nucleotides upstream and 50 nucleotides downstream of *Cj1556* were analysed. Both a -10 (ataaat) and -35 (tgttataat) promoter site were identified. In addition, a number of other transcription factor binding sites were identified (Appendix 10). To investigate links between *Cj1556* and *ctb*, promoter sequences upstream of *ctb* were also searched in the database. Two -35 boxes, three *lrp* (attttttt) sequences and a *lexA* (ttttttta) sequence

was identified. Both *lrp* and *lexA* sequences were also present in the original *Cj1556* search result and indicate possible common promoters for both *Cj1556* and *ctb*. In addition, analysis of *C. jejuni katA*, *perR* and *hspR* promoter sequences identified the presence of these promoter sequences upstream of *Cj1556* (Appendix 10). This suggests Cj1556 has the potential to bind to the promoter regions of *katA*, *perR* and *hspR*.

### 4.2.14 The 11168H *Cj1556* mutant reveals unaltered motility

In response to microarray data described in Section 4.2.12 where a number of flagella related genes (*Cj0041*, *Cj0195*, *Cj0335*, *Cj0320*) were up-regulated in the *Cj1556* mutant, motility assays were performed on the *Cj1556* mutant and compared to the 11168H wild-type strain. The results identified no significant difference in the level of motility between the strains (Figure 4.18).

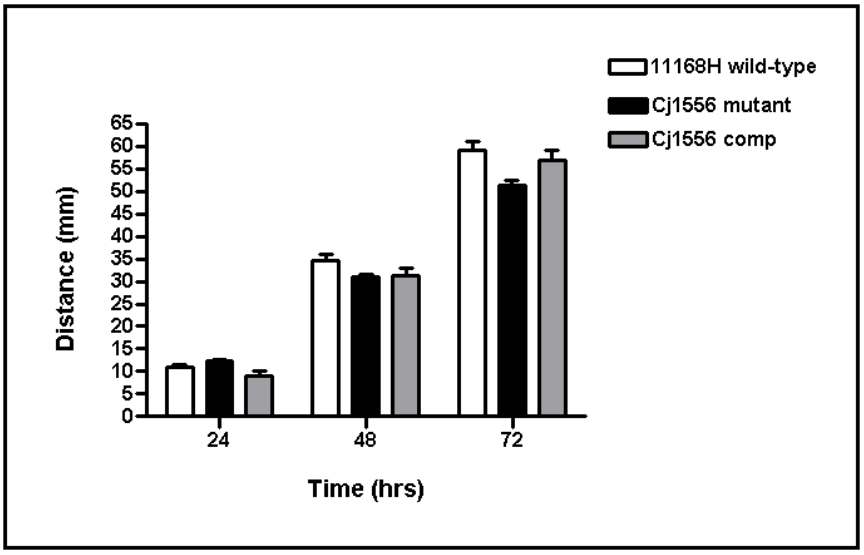



Figure 4.18. Motility assay comparing the *Cj1556* mutant to the 11168H wild-type strain and *Cj1556* complement (*Cj1556* comp) on semi-solid surfaces. *C. jejuni* from an overnight culture was adjusted to an $OD_{600}$ of 1.0. 5 µl of this suspension was pipetted into the centre of the motility plate using a Gilson pipette (P10). Plates were incubated at 37°C using microaerobic conditions. Plate images were recorded at 24, 48 and 72 h. Motility was assessed by measuring the diameter of the halo on the plate. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

### 4.2.15 The 11168H *Cj1556* mutant reveals unaltered autoagglutination levels

Autoagglutination or self-association has been linked with virulence in some Gram-negative bacteria (Roggenkamp *et al*., 1995, Bieber *et al*., 1998, Brondsted *et al*., 2005). Here, the autoagglutination ability of the *Cj1556* mutant was compared to the *C. jejuni* 11168H wild-type strain. No difference in the level of autoagglutination ability was observed between the strains (Figure 4.19). An *rpoN* mutant was used as a autoagglutination negative control. Inactivation of RpoN ($\sigma^{54}$) renders the bacterium non-motile and without flagella (Jagannathan *et al*., 2001).

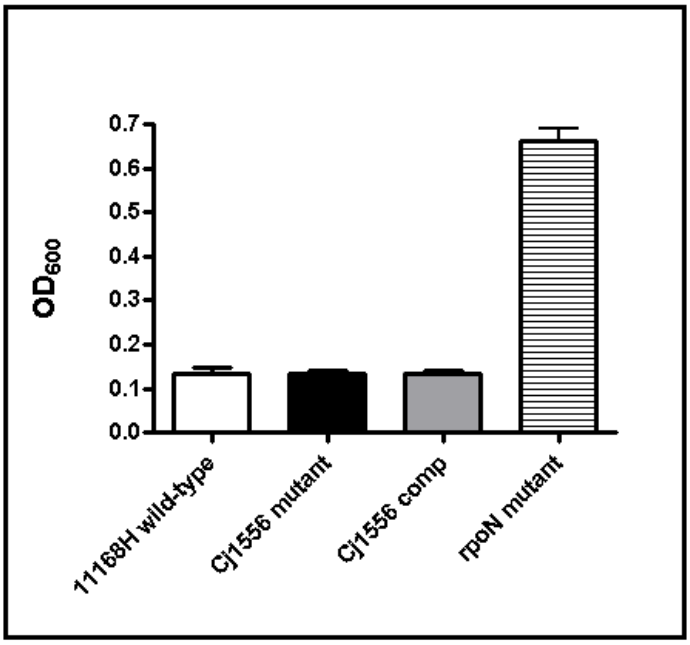



Figure 4.19. Autoagglutination of the 11168H wild-type strain, *Cj1556* mutant and *Cj1556* complement (*Cj1556* comp). *C. jejuni* that were grown on blood agar plates for 24 h at 37°C under microaerobic conditions were used to inoculate 10 ml PBS to an $OD_{600}$ of 1.0. Cultures were incubated for 24 h at 37°C under microaerobic conditions. The $OD_{600}$ was measured using 1 ml taken from the middle of a 24 h incubated culture. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

### 4.2.16 The 11168H *Cj1556* mutant exhibits reduced biofilm formation

Biofilms are commonly defined as matrix-enclosed bacterial populations adherent to each other and/or to surfaces of interfaces (Costerton *et al*., 1995). Studies have shown that *C. jejuni* can form biofilms (Joshua *et al*., 2006) and that this is an important factor

in the survival of *C. jejuni* in the environment (Reuter *et al*., 2010). Recent studies have also shown biofilm formation increased under aerobic ($O_2$) stress conditions (Reuter *et al*., 2010). The microarray data identified *cprS* as being 2.0 fold up-regulated in the *Cj1556* mutant compared to the 11168H wild-type strain. A *cprS* mutant has been shown to have enhanced and accelerated biofilm formation (Svensson *et al*., 2009). Therefore an increase in CprS production in the *Cj1556* mutant was predicted to result in a decrease in biofilm formation. Analysis of the ability to form biofilms when grown for 72 h under microaerobic conditions indicated a significant reduction when comparing the *Cj1556* mutant and the 11168H wild-type strain (Figure 4.20).

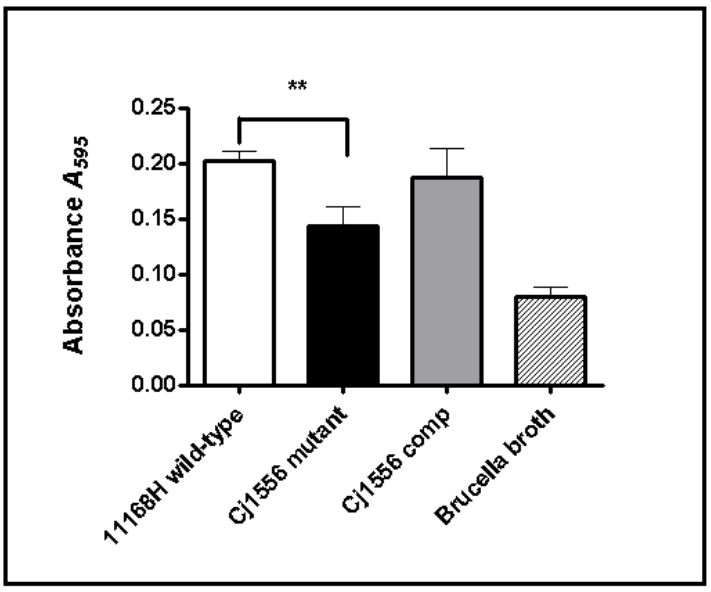



Figure 4.20. Biofilm assay on 11168H wild-type strain, *Cj1556* mutant and *Cj1556* complement (*Cj1556* comp). *C. jejuni* were grown with a starting $OD_{600}$ of 0.1, in culture for 72 h at 37°C under microaerobic conditions shaking at 75 rpm followed by crystal violet staining. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with two stars (**) indicating $p$ = 0.001- 0.01.

### 4.2.17 *Galleria mellonella* larvae exhibit increased survival after infection with the 11168H *Cj1556* mutant

*G. mellonella* larvae have recently been used as a model to study infection by *C. jejuni* and other enteric pathogens (Champion *et al*., 2009). Insect larvae such as *G. mellonella* are favorable to use as non-mammalian infection models as they can be infected at 37°C

and possess specialized phagocytic cells, termed haemocytes (Mylonakis *et al*., 2007, Bergin *et al*., 2005). Haemocytes mimic the functions of phagocytic cells in mammals and are able to degrade bacterial pathogens and also generate bactericidal compounds such as superoxide via a respiratory burst (Bergin *et al*., 2005, Lavine and Strand, 2002). Infection with the *Cj1556* mutant resulted in increased survival of *G. mellonella* larvae compared to infection with the 11168H wild-type strain (Figure 4.21). This suggests the *Cj1556* mutant is more susceptible to the host immune mechanisms resulting in increased survival of *G. mellonella*.

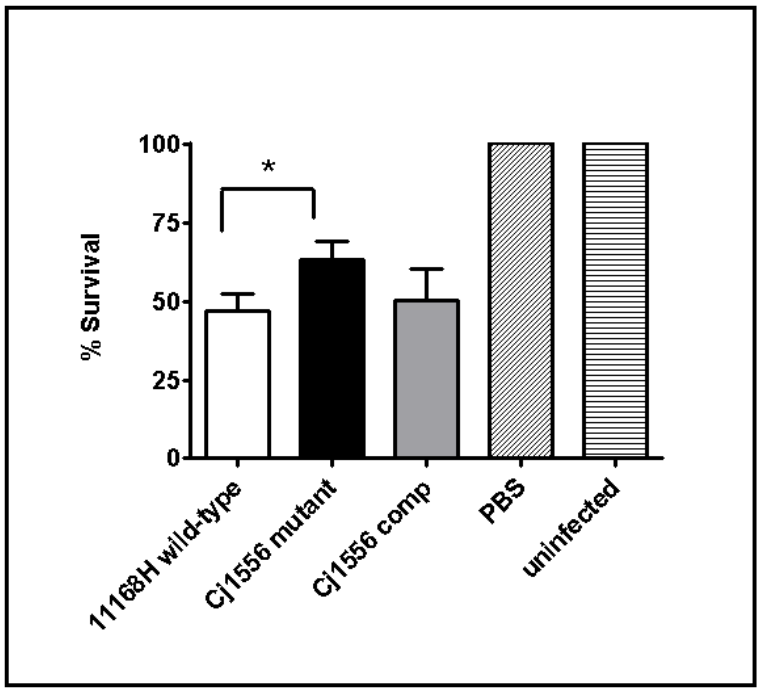



Figure 4.21. *G. mellonella* survival assay on 11168H wild-type strain, *Cj1556* mutant and *Cj1556* complement (*Cj1556* comp). *G. mellonella* larvae were injected with a 10 µl inoculum of a 24 h *C. jejuni* culture diluted to $OD_{600}$ 0.1 by micro-injection in the right foremost leg, giving an infectious dose of ≈ $10^6$ CFU. Larvae were incubated at 37°C with survival and appearance recorded at 24 h intervals. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$.

### 4.3 Discussion

*C. jejuni* survival is based upon the ability to sense and respond to the many different environments and yet the *C. jejuni* compact genome appears to have limited capacity for transcriptional control (Parkhill *et al*., 2000). Following re-annotation of the *C. jejuni* NCTC11168 genome sequence (see Chapter 3), Cj1556 was identified as a putative transcriptional regulator. Based on motif and protein interaction data, Cj1556 was hypothesized to be an important *C. jejuni* stress response regulator and therefore investigations into the ability of the 11168H *Cj1556* mutant to survive different stresses and further explore the role of Cj1556 during host-pathogen interactions were performed.

The human intestinal pathogen *C. jejuni* must survive diverse conditions in several different hosts and also in the environment. Even though *C. jejuni* is a microaerophilic organism, one of the great mysteries regarding this pathogen is the bacterium's ability to be found ubiquitously within the environment. Fundamental to the survival of *C. jejuni* is the ability to survive both aerobic ($O_2$) and oxidative stress conditions. Even though *C. jejuni* is microaerophilic, the bacterium has clearly adapted to survive when exposed to aerobic ($O_2$) stress by inhabiting conditions where atmospheric levels of oxygen are present. *C. jejuni* produces proteins which are directly involved combating both aerobic ($O_2$) and oxidative stress conditions. Important proteins identified within *C. jejuni* to counteract aerobic ($O_2$) stress are the heat response proteins HtrA and HspR; the latter encoding a regulator important for short-term aerobic tolerance (Brondsted *et al*., 2005). The *fdxA* gene encoding a ferrodoxin has been noted for being important for aerotolerance (van Vliet *et al*., 2001). In addition SpoT was found to be important for low $CO_2$ growth and aerobic survival (Gaynor *et al*., 2005). Oxidative stress is an important type of stress given that *C. jejuni* is a microaerophilic organism. ROS including superoxide, hydrogen peroxide ($H_2O_2$) and halogenated $O_2$ molecules form as a result of aerobic respiration both inside and outside the host and form a major defensive barrier against the bacterium. Often, oxygen is converted into ROS including superoxide radicals (Park, 2005). A number of proteins are synthesized by *C. jejuni* to counteract such harmful molecules and these can be grouped into i) the $O_2^-$ (superoxide) stress proteins and ii) the peroxide stress proteins (van Vliet *et al*., 1999, Farr and Kogoma, 1991). The superoxide dismutase (SOD) for example catalyses the breakdown of superoxide to $H_2O_2$ and $O_2$ and is classed as a $O_2^-$ stress proteins (Purdy and Park, 1994, Pesci *et al*., 1994). This enzyme plays a key role in the defence against oxidative stress and aerotolerance as SOD-deficient *C. jejuni* strains are less able to survive in the

air (Purdy *et al*., 1999, Stead and Park, 2000). *C. jejuni* also contains the catalase KatA which converts $H_2O_2$ to $H_2O$ and $O_2$. *C. jejuni* contains an alkyl hydroperoxide reductase encoded by *ahpC* which gives resistance to cumene hydroperoxide and aerobic ($O_2$) stress (Baillon *et al*., 1999). KatA and AhpC are classed as peroxide stress proteins. As noted above, *C. jejuni* lacks OxyR which is typically a regulator of *ahpC* and *katA* expression in response to oxidative stress in *E. coli* (Cabiscol *et al*., 2000). Most Gram-negative bacteria possess a recognition and response mechanism to oxidative stress via SoxRS and OxyR which recognise superoxide and peroxide molecules (Park, 2005). These are not present in *C. jejuni*, but a partial homologue PerR, which was noted as a functional but non-homologous version of OxyR, is present (van Vliet *et al*., 1999, Palyada *et al*., 2009). Research in *E. coli* has shown that the SoxRS system regulates the superoxide regulon (Amabile-Cuevas and Demple, 1991) and mutations in *soxR* or *soxS* fail to induce the members of their regulon (Farr and Kogoma, 1991). OxyR regulates the OxyR regulon, which is part of the peroxide stress regulon. The OxyR protein is a transcription factor that senses oxidative stress through disulfide bond formation and under conditions of oxidative stress, transcription of its regulon is induced following a change in OxyR conformation due to this disulfide bond formation (Storz *et al*., 1990, Toledano *et al*., 1994, Zheng *et al*., 1998). OxyR is considered a global regulator and *H. influenzae oxyR* mutants are unable to respond to oxidative stress (Maciver and Hansen, 1996).

Bioinformatics analysis of Cj1556 identified the presence of a helix-turn-helix motif and this suggested a role for this protein as a transcriptional activator. This motif is part of the MarR family of transcriptional regulators, which includes proteins that control virulence factor production, bacterial responses to antibiotics and oxidative stresses (Wilkinson and Grove, 2004). Using the search results from the *Campylobacter* Protein Interaction Database, a putative link between Cj1556 and Ctb was identified. Ctb is a truncated haemoglobin encoded by *Cj0465c* and is part of the *C. jejuni* nitrosative stress response regulon with Cgb (encoded by *Cj1586*), which are both under the control of NssR (encoded by *Cj0466*), a NO-sensitive regulator (Monk *et al*., 2008). No differences in bacterial survival were observed during nitrosative stress assays with the *Cj1556* mutant compared to the 11168H wild-type strain. However Ctb has also been shown to be involved in oxygen delivery (Wainwright *et al*., 2005, Wainwright *et al*., 2006). Oxidative stress assays showed that the *Cj1556* mutant has increased sensitivity compared to the 11168H wild-type strain and that the wild-type level of sensitivity to

oxidative stress was fully restored with complementation of the *Cj1556* mutation. In fact, the *Cj1556* complement demonstrated even greater resistance to $H_2O_2$ than the 11168H wild-type strain, possibly due to the strength of the promoter as the complementation vector utilises the constitutive chloramphenicol cassette promoter to express the *Cj1556* CDS and not the native *Cj1556* promoter. In addition, microarray data suggests that expression of *Cj1556* is negatively autoregulated. Thus, the wild-type *Cj1556* CDS would not be expected to be constitutively expressed as is the case in the *Cj1556* complement. This would be a possible explanation as to why the *Cj1556* complement demonstrated even greater resistance to $H_2O_2$ than the 11168H wild-type strain, due to overexpression of *Cj1556*.

The ability of the *Cj1556* mutant to interact with (adhere and invade) and invade Caco-2 cells was investigated over a 24 h time period. No significant difference in the ability of the *Cj1556* mutant to interact with or invade Caco-2 cells was observed at either 3 or 6 h post infections. However significant difference in both interaction and invasion was observed at 24 h post-infection. This suggests the *Cj1556* mutant does not have any defect in the ability to adhere to or invade Caco-2 cells, but may have a reduced ability to survive contact with host cells over longer periods of time. To further investigate long term survival of the *Cj1556* mutant, intracellular survival assays were performed. These assays indicated that the *Cj1556* mutant has a reduced ability to survive within Caco-2 cells compared to the 11168H wild-type strain. The difference in the level of survival between the *Cj1556* mutant and the 11168H wild-type strain in the intracellular survival assay at 24 h post-infection was approximately 0.5 log (Figure 4.12), very similar to the difference between the numbers of invasive bacteria for the *Cj1556* mutant compared to the 11168H wild-type strain 24 h post-infection (Figure 4.11B). However, the difference in the number of interacting bacteria between the *Cj1556* mutant and the 11168H wild-type strain 24 h post-infection was approximately 1.5 log (Figure 4.11A). This suggests that in addition to a reduced ability for intracellular survival, the *Cj1556* mutant was also more susceptible to extracellular stresses when compared to the 11168H wild-type strain. The *Cj1556* mutant also exhibited increased sensitivity to $H_2O_2$ *in vitro*, so it is reasonable to suggest ROS released by Caco-2 cells during these experiments will have an effect on *C. jejuni* survival. Standard co-culture assays will not only result in exposure of *C. jejuni* to ROS released by Caco-2 cells, but also to aerobic ($O_2$) stress as the assays are performed in a $CO_2$ incubator. The approximate atmospheric $O_2$ and $CO_2$ levels are around 21% and 0.04% respectively. During co-culture experiments, the level of $CO_2$

will be around 5% and so the $O_2$ level will be around 16-18%. Based on the relative levels of survival between the interaction, invasion and intracellular survival assays, the greater level of sensitivity exhibited by the *Cj1556* mutant during the interaction assay may be in part due to increased exposure of extracellular *C. jejuni* to aerobic ($O_2$) stress.

As such aerobic survival assays were performed to replicate the conditions during the interaction, invasion and intracellular survival assays by incubating *C. jejuni* in tissue culture media but in the absence of Caco-2 cells. A reduction in survival was observed for the *Cj1556* mutant compared to the 11168H wild-type strain under aerobic conditions, but not under microaerobic conditions. *C. jejuni* typically loses viability within IECs over 24 h with no evidence of intracellular replication (Konkel *et al*., 1992a). Recent evidence suggests that *C. jejuni* reside in membrane bound compartments termed *C. jejuni* containing vacuole (CCV), avoiding fusion with lysosomes (Watson and Galán, 2008). *C. jejuni* engulfed by macrophages must resist a combination of bacteriocidal stresses such as ROS. There are contradictory reports regarding the ability of *C. jejuni* to survive within macrophages, depending on the macrophage cell type and *C. jejuni* strain used (Day *et al*., 2000, Wassenaar *et al*., 1997). In this study, the *Cj1556* mutant was eliminated quicker when compared to the 11168H wild-type within the mouse macrophage J774A.1 cell line. Taken together, this data indicates that Cj1556 plays a multi-factorial role in bacterial survival during adhesion to and invasion of human IECs.

Following on from the interaction, invasion and survival studies, host innate immune responses was assessed by co-culturing *C. jejuni* strains with IECs and assessing IL-6 and IL-8 induction. Minimal IL-6 and IL-8 was detected when the 11168H wild-type strain and *Cj1556* mutant were co-cultured with Caco-2 cells (data not shown). This concurs with research demonstrating the poor chemokine/cytokine induction ability of Caco-2 cells when co-cultured with *C. jejuni* (MacCallum *et al*., 2006). However using a T84 cell line, significant levels of induction of both IL-6 and IL-8 by the 11168H wild-type strain and *Cj1556* mutant were observed. Previous reports have demonstrated ELISA experiments using T84 cells to detect levels of IL-8 and IL-6 over 1000 pg/ml when co-cultured with *C. jejuni* (MacCallum *et al*., 2006). In this study, these levels were not achieved most likely due to a number of factors, i) the use of 24-well plates rather than 12-well plates, ii) using an MOI of approximately 20:1 rather than 100:1, iii) variations in *C. jejuni* strains, iv) variation in T84 cell line e.g. passage number. In

addition, results from ELISA experiments using Caco-2 IECs produced IL-8 and IL-6 levels detected at approximately 0-20 pg/ml when co-cultured with *C. jejuni* strains (data not shown). Thus, the trends observed in this study using T84 cells would most likely be consistent if the increased level of chemokine/cytokine induction was achieved by modifications to the experimental protocol.

In this study, there was no significant difference in the level of IL-8 induction by the *Cj1556* mutant compared to the 11168H wild-type strain, however a significant reduction in the level of IL-6 induction by the *Cj1556* mutant compared to the 11168H wild-type strain was observed. IL-8 acts as a chemoattractant allowing the recruitment of lymphocytes and neutrophils (Hobbie *et al*., 1997, Philpott *et al*., 2000), whereas IL-6 is believed to be important for epithelial cell integrity (Friis *et al*., 2009). It is possible that less IL-6 was induced when T84 cells were co-cultured with *Cj1556* mutant compared to the 11168H wild-type strain possibly due to the decreased survival characteristic of the *Cj1556* mutant strain. Based on data from this study, co-culturing the *Cj1556* mutant strain for 24 h in a 37°C $CO_2$ incubator would result in decreased survival of the *Cj1556* mutant based on the increased sensitivity of this strain compared to the 11168H wild-type strain. This may be a possible reason for the decreased IL-6 production. This result also suggests that IL-8 may be important for an extracellular response as both *Cj1556* mutant and the 11168H wild-type strain elicited similar levels of IL-8 from T84 IECs. However, IL-6 may be more important for an intracellular response as the *Cj1556* mutant was shown to invade less compared to the 11168H wild-type strain and so elicited less IL-6 from T84 IECs.

Transcriptional analysis of the *Cj1556* mutant compared to the 11168H wild-type strain identified *Cj1556* as the most up-regulated CDS (10.4 fold). This result was somewhat of a surprise as mutated strains traditionally display a reduced transcriptional level compared to the wild-type strain. Analysis of the *Cj1556* nucleotide sequence upstream of the $Km^R$ cassette in the *Cj1556* mutant confirmed that this was the sequence printed on the oligonucleotide spotted array, leading to the hypothesis that expression of *Cj1556* is controlled by a negative autoregulation feedback mechanism. In the 11168H wild-type strain, basal levels of Cj1556 would block off further expression of *Cj1556* by inhibiting the binding of RNA polymerase to the *Cj1556* promoter site. However in the absence of Cj1556 in the *Cj1556* mutant, expression of *Cj1556* can continue. This kind of negative

autoregulation is a feature of the MarR family of transcriptional regulators (Wilkinson and Grove, 2004).

The role of Cj1556 as a putative transcriptional regulator would suggest that when expressed, Cj1556 also binds to other areas of the genome and controls expression of other proteins. The microarray data indicated down-regulation of *katA* (5.13), *perR* (5.05) and *hspR* (2.07) in the *Cj1556* mutant (Table 4.2). These genes are also involved in oxidative/aerobic ($O_2$) stress response in *C. jejuni*. It is possible Cj1556 interacts with these genes and controls their expression. To analyse this further, the promoter regions upstream of Cj1556 were searched and identified in *C. jejuni katA*, *perR* and *hspR* (Appendix 10). The *C. jejuni katA* encodes a catalase which converts $H_2O_2$ to $H_2O$ and $O_2$. As noted before, *C. jejuni* lacks OxyR which is typically a regulator of AhpC and KatA in response to oxidative stress in *E. coli* (Cabiscol *et al*., 2000). A *Cj1556* mutant may have directly or indirectly affected the expression of *katA*. Further studies would be useful for investigating the global binding of Cj1556 using methods such as ChIP-Seq or ChIP-ChIP (Ren *et al*., 2000, Bulyk *et al*., 1999, Uzzau *et al*., 2001).

As described earlier, a number of proteins are synthesized by *C. jejuni* to counteract aerobic ($O_2$) and oxidative stress inducing molecules and these can be grouped into, i) the $O_2^-$ (superoxide) stress proteins and, ii) the peroxide stress proteins (van Vliet *et al*., 1999, Farr and Kogoma, 1991). *C. jejuni* does not possess the typical recognition and response mechanism to oxidative stress via SoxRS and OxyR which recognise superoxide and peroxide molecules (Park, 2005). However, a partial homologue PerR, which was noted as a functional but non-homologous version of OxyR, is present (van Vliet *et al*., 1999, Palyada *et al*., 2009). Mutation of *Cj1556* may have directly or indirectly affected the expression of *perR*. Reduced expression of *katA* and *perR* would provide one explanation for the increased sensitivity of the *Cj1556* mutant to oxidative and aerobic ($O_2$) stress observed in this study, however further experiments are required to confirm this predicted function.

The ability of *C. jejuni* to form biofilms goes some way to explain how a bacterium with such fastidious growth requirements remains ubiquitous in the environment (Buswell *et al*., 1998, Joshua *et al*., 2006). *C. jejuni* can form three distinct forms of biofilm: cell-cell aggregates, pellicles at the air-liquid interface and glass-attached flocs (Joshua *et al*., 2006). Our understanding of the specific mechanisms underlying biofilm formation in *C.*

*jejuni* are still limited (Svensson *et al*., 2008). *C. jejuni* lacks the classical 2CRSs involved in biofilm formation that are present in other bacteria such as GacSA in *Pseudomonas aeruginosa* (Parkins *et al*., 2001). Genes involved in biofilm formation have been linked to responses to oxidative and aerobic ($O_2$) stress and *C. jejuni* biofilm formation is increased under aerobic conditions (Reuter *et al*., 2010). A *C. jejuni spoT* mutant has been found to overproduce a novel calcofluor white reactive exoploysaccharide and demonstrate enhanced biofilm formation (McLennan *et al*., 2008). Interestingly a *C. jejuni cprS* mutant has been shown to display growth defects, enhanced and accelerated biofilm formation and also to exhibit decreased oxidative stress tolerance (Svensson *et al*., 2009). The microarray data identified *cprS* as being 2.0 fold up-regulated in the *Cj1556* mutant compared to the 11168H wild-type strain. Therefore an increase in CprS production in the *Cj1556* mutant was predicted to result in a decrease in biofilm formation. Analysis of the 11168H wild-type and *Cj1556* mutant strains indicated a significant reduction in relative biofilm formation by the *Cj1556* mutant (Fig. 6). Recent studies have also shown increased biofilm formation under aerobic ($O_2$) stress conditions (Reuter *et al*., 2010). This suggests a possible role for Cj1556 in biofilm formation either directly or indirectly.

*C. jejuni* proteins associated with heat stress responses such as HspR have also been linked to oxidative and aerobic ($O_2$) stress (Andersen *et al*., 2005). The *Cj1556* mutant showed a greater level of sensitivity to 60°C stress compared to the 11168H wild-type strain. Heat stress above 55°C has been noted to accelerate the spiral-to-coccoid transition and result in cell death (Nguyen *et al*., 2006). Previous studies have identified numerous *C. jejuni* genes involved in heat shock response and HtrA and HspR also have been shown to have roles in aerobic survival, host cell adherence and invasion (Brondsted *et al*., 2005). Transcriptional analysis identified *hspR* as being approximately 2.0 fold down-regulated in the *Cj1556* mutant compared to the 11168H wild-type strain. In addition, analysis of *C. jejuni hspR* promoter sequences identified the presence of these promoter sequences upstream of *Cj1556* (Appendix 10). It is interesting to note that the *Cj1556* mutant has increased sensitivity to heat stress and this may be due to Cj1556 interacting with HspR. Even though no protein-protein interactions between the two were identified in the *Campylobacter* Protein Interaction Database, there may still exist a connection between the heat shock response and aerobic tolerance (Andersen *et al*., 2005, Brondsted *et al*., 2005).

The *G. mellonella* insect model was used to investigate pathogenicity of the *Cj1556* mutant. The *G. mellonella* insect model is suitable for identification of *C. jejuni* virulence determinants and the larvae possess specialised phagocytic cells, termed haemocytes (Bergin *et al*., 2005). The insect immune system is subdivided into humoral and cellular defence responses. Humoral defences include the production of antimicrobial peptides (Meister *et al*., 2000), reactive intermediates of oxygen or nitrogen (Bogdan *et al*., 2000) and the complex enzymatic cascades that regulate coagulation or melanisation of haemolymph (Muta and Iwanaga, 1996). Cellular defence refers to haemocyte-mediated immune responses like phagocytosis, nodulation and encapsulation (Strand and Pech, 1995, Schmidt *et al*., 2001). Haemocytes perform many of the functions of phagocytic cells in mammals, are capable of ingesting bacterial pathogens and generating bactericidal compounds such as superoxide via a respiratory burst (Bergin *et al*., 2005, Champion *et al*., 2009). Studies have shown that after infection of *G. mellonella* with *Yersinia pseudotuberculosis*, the bacteria accumulates in haemocytes, thus suggesting that *G. mellonella* may be useful for the identification of other genes associated with intracellular survival (Champion *et al*., 2009). Infection with the *Cj1556* mutant resulted in increased survival of *G. mellonella* larvae compared to survival after infection with the 11168H wild-type strain at 24 h post-infection. This suggests the *Cj1556* mutant is more susceptible to the host immune mechanisms resulting in reduced bacterial survival within *G. mellonella*. At least six types of haemocytes have been identified in insects such as *G. mellonella* with plasmatocytes and granulocytes the most abundant (Boman and Hultmark, 1987). Production of ROS has also been detected in haemocytes with evidence of both oxygen radicals and $H_2O_2$ both found in plasmatocytes of *G. mellonella* (Slepneva *et al*., 1999). This data links the increased sensitivity of the *Cj1556* mutant to $H_2O_2$ stress *in vitro* with an attenuation of virulence *in vivo* using the *G. mellonella* model of infection.

The digestive secretion bile consists of around 50% bile salts, such as cholates and deoxycholates. Bile salts exhibit potent antibacterial properties, acting as detergents to disrupt cell membranes and as DNA-damaging agents (Begley *et al*., 2005). Although bacteria inhabiting the gastrointestinal tract are able to resist the antimicrobial effects of bile, a number of studies have also shown that bile increases the virulence potential of enteric pathogens (Begley *et al*., 2005). The bile salt sDOC has been shown to increase the virulence of *C. jejuni*, enhancing bacterial ability to invade epithelial cells (Malik-Kale *et al*., 2008). Growing *C. jejuni* in the presence of a physiologically relevant

concentration of sDOC (0.1% w/v) changes the invasion kinetics such that maximal invasion of INT 407 cells occur in under 30 minutes compared to 3 h for *C. jejuni* grown in the absence of sDOC (Malik-Kale *et al*., 2008). Microarray analysis has shown that a number of *C. jejuni* virulence factors are up-regulated in the presence of 0.1% (w/v) sDOC, including *ciaB, cmeABC, dccR* and *tlyA* (Malik-Kale *et al*., 2008). Interestingly, *Cj1556* was also up-regulated in the presence of sDOC, with transcription increasing 2.8 fold (Malik-Kale *et al*., 2008). The transcriptional response of *E. coli* O157:H7 to bile treatment has also been investigated using microarrays and has identified bile-induced changes in transcription for genes encoding proteins affecting membrane structure and permeability, bile resistance, adhesion and virulence potential (Hamner *et al*., 2010). Most interestingly this data indicates that bile induces expression of the *marRAB* operon, by binding to the repressor protein MarR and thus preventing binding of MarR to the *marRAB* promoter site (Hamner *et al*., 2010). Cj1556 is a member of the MarR family of transcriptional regulators and further studies will be required to confirm whether bile can bind to the Cj1556 protein and thus prevent binding to the *Cj1556* promoter site, resulting in the up-regulation of Cj1556 in the presence of bile observed previously (Malik-Kale *et al*., 2008).

In addition to Cj1556, the *C. jejuni* NCTC11168 genome contained one other CDS with the MarR family motif; Cj1546. Cj1546 was also re-annotated as a putative transcriptional regulator with 43.6% identity and 58.4% similarity to Cj1556. (Figure 4.1). Protein Blast search results for Cj1556 identified this CDS to be present in relatively few strains/species of *Campylobacter* (Table 4.1). Homologues with relatively high identity and expect scores were present in *C. jejuni* strains CF93-6, 84-25 and 81-176, along with the species *C. fetus* 82-40. Protein Blast search results for Cj1546 identified this CDS in all *C. jejuni* strains with a significant level of identity and similarity (data not shown). Even though no standout aerobic ($O_2$) and/or oxidative stress protein-protein interactions were identified for Cj1546 using the *Campylobacter* Protein Interaction database, one hypothesis as to the function of these MarR motif containing proteins is that they both perform similar roles in relation to aerobic ($O_2$) and oxidative stress. It is tempting to speculate that whilst all *C. jejuni* strains express Cj1546, strains such as *C. jejuni* NCTC11168 and 81-176 that also express Cj1556 may have a greater chance of survival within the human host for example due to greater resistance to oxidative stresses. Further experiments will be required to confirm this hypothesis. During this study, it was not possible to construct a 11168H *Cj1546* mutant. These initial

attempts at constructing a *Cj1546* mutant may have proved unsuccessful due to the small size of the CDS (345 nucleotides) or because mutation of *Cj1546* is a lethal mutation. Adopting a similar strategy used in the construction of a *Cj1556* mutant by using larger pUC library clones from the original genome sequencing project harbouring the gene of interest may prove more efficient.

### 4.4 Conclusion

In summary, a cornerstone of *C. jejuni* survival is based upon the ability to sense and respond to the different environments the bacterium encounters within and outside hosts. This data indicates that the Cj1556 protein is involved in the regulation of both the *C. jejuni* oxidative and aerobic ($O_2$) stress responses and also plays a role in bacterial survival *in vivo*. As such the Cj1556 putative transcriptional regulator has been designated as CosR (*C*ampylobacter *o*xidative *s*tress *R*egulator).

# Chapter 5: *Cj0248* encodes a putative phosphohydrolase which has a role in both motility and virulence

## 5.1 Introduction

Re-annotation of the *C. jejuni* NCTC11168 genome sequence resulted in the identification of 15 CDSs for further study (see Section 3.4). The selection of these 15 CDSs was based on a number of different criteria such as newly identified motifs along with literature searches. The function of each CDS was unknown, however putative functions were linked to roles in virulence, signal transduction or regulation of gene expression. Following the successful construction of eight defined isogenic *C. jejuni* 11168H mutants, a number of preliminary phenotypic assays such as growth kinetics, motility, interaction and invasion assays were performed. Two mutants were chosen for further characterisation and this chapter details the characterisation of the 11168H *Cj0248* mutant.

### 5.1.1 Aims

- Construct a defined isogenic *C. jejuni* 11168H *Cj0248* mutant
- Construct a 11168H *Cj0248* complement
- Characterise the 11168H *Cj0248* mutant

## 5.2 Constructions and characterisation of the *C. jejuni* 11168H *Cj0248* mutant

### 5.2.1 Literature search results

Using *in vitro* transposition of *C. jejuni* chromosomal DNA followed by natural transformation of the transposed DNA, large random transposon mutant libraries were prepared in wild-type strain 81-176 allowing up to 16,000 mutants to be generated (Hendrixson *et al*., 2001). Subsequent screening identified 28 mutants with motility defects (Hendrixson *et al*., 2001). One such CDS was *Cj0248*. To further investigate caecal colonisation by *C. jejuni*, signature-tagged transposon mutagenesis was performed, generating 1550 *C. jejuni* 81-176 mutants of which 29 mutants (representing 22 different genes) were found to be required for wild-type levels of caecal colonisation (Hendrixson and DiRita, 2004). As such, a *C. jejuni* 81-176 *Cj0248* mutant was shown to have reduced motility (Hendrixson *et al*., 2001) and a 100- to 10,000- fold reduction in

colonisation using a chick caecal colonisation model (Hendrixson and DiRita, 2004). In addition the crystal structure of Cj0248 has been elucidated (Xu *et al*., 2006). Cj0248 contains a HD motif (PF08668), indicating a function as a possible phosphohydrolase or a signal transduction protein. Enzymes known to possess phosphohydrolase activity may be involved in nucleic acid metabolism, signal transduction and possibly other functions in bacteria (Aravind and Koonin, 1998). Cj0248 was predicted to contain a HD motif indicating a function as a possible phosphohydrolase, however structural analysis indicated that Cj0248 does not function as a metal-dependent phosphohydrolase (Xu *et al*., 2006). Thus, Cj0248 would appear to have a role in signal transduction linked to motility and colonisation. No specific function has as yet been designated. Importantly, all previous work was performed on the *C. jejuni* 81-176 strain.

### 5.2.2 Bioinformatic analysis of Cj0248

BlastP search results of the Cj0248 protein sequence identified the presence of this CDS in a number of *C. jejuni* strains and *Campylobacter* species albeit with significantly lower similarity scores (Table 5.1). In addition, homologues are also present in other genus from the epsilon proteobacterium class such as *Helicobacter*, *Wolinella* and *Sulfurospirillum*. Homologues of this CDS are annotated as "hypothetical protein" or a function based on the HD motif. No homologue has been characterised as yet. Comparison of the protein sequences of *C. jejuni* NCTC11168 and 81-176 strains identified only five amino acids that were different between the two strains. Three of these amino acid differences were conserved and one was a semi-conserved substitution. Only one of the five different amino acids was a non-conserved substitution (Figure 5.1). This suggests no functional differences between the two strains. No links were identified between the Cj0248 protein and any other protein in the *Campylobacter* Protein Interaction Database (Parrish *et al*., 2007).

Table 5.1. BlastP search results for Cj0248 protein sequence.

| Species / Strain | Identity score | Expect score |
|---|---|---|
| *Campylobacter jejuni* subsp. *jejuni* NCTC11168 | 285/285 | 2e-163 |
| *Campylobacter jejuni* subsp. *jejuni* 84-25 | 285/285 | 2e-163 |
| *Campylobacter jejuni* subsp. *jejuni* CF93-6 | 285/285 | 2e-163 |
| *Campylobacter jejuni* subsp. *jejuni* CG8486 | 285/285 | 8e-163 |
| *Campylobacter jejuni* RM1221 | 284/285 | 4e-163 |
| *Campylobacter jejuni* subsp. *jejuni* CG8421 | 284/285 | 2e-163 |
| *Campylobacter jejuni* subsp. *jejuni* HB93-13 | 283/285 | 2e-162 |
| *Campylobacter jejuni* subsp. *jejuni* 81-176 | 280/285 | 4e-161 |
| *Campylobacter jejuni* subsp. *jejuni* 260.94 | 279/285 | 9e-161 |
| *Campylobacter jejuni* subsp. *doylei* 269.97 | 278/285 | 5e-160 |
| *Campylobacter jejuni* subsp. *jejuni* 81116 | 277/285 | 6e-160 |
| *Campylobacter coli* RM2228 | 197/286 | 2e-118 |
| *Campylobacter upsaliensis* RM3195 | 197/282 | 7e-110 |
| *Campylobacter showae* RM3277 | 74/230 | 3e-33 |
| *Campylobacter fetus* subsp. *fetus* 82-40 | 70/230 | 1e-32 |
| *Campylobacter curvus* 525.92 | 74/231 | 2e-32 |
| *Campylobacter rectus* RM3267 | 71/230 | 1e-30 |
| *Campylobacter concisus* 13826 | 71/230 | 1e-29 |

```
11168H MIGDMNELLLKSVEVLPPLPDTVSKLRKYVSEANSNIETMKVAEIISSDPLMTAKLLQLA 60
81-176 MIGDMNELLLKSVEVLPPLPDTVSKLRKYVSEANSNIETMKVAEIISSDPLMTAKLLQLA 60
       ************************************************************

11168H NSPYYGFTREITTINQVITLLGVGNIINIVMADSIRDNFKIDVSPYGLNTQNFLKTCNEE 120
81-176 NSPYYGFTREITTINQVITLLGVGNIINIVTADSIRDNFKIDVSPYGLNTQNFLKTCNEE 120
       ***************************** *****************************

11168H ATFIANWLNDEDKKLSHLLVPCAMLLRLGIVIFSNFLIQNHKDKDFLAFLNKNENLALAE 180
81-176 ATFITNWLNDEDKKLSHLLVPCAMLLRLGIVIFSNFLIQNHKDKDFLAFLNKNENLALAE 180
       ****:*******************************************************

11168H NEFLGVDHISFLGFLLHRWNFDDVLIESICFVRTPHAARE KVKKSAYALAITDHLFAPHD 240
81-176 NEFLGVDHISFLGFLLHRWNFDDVLIESICFVRTPHAARE EVKKSAYALAITDHLFAPHD 240
       ************************************:*******************

11168H GSSPFNAKAAVALLKEAKTQGINFDLNNLLSKLPNKAKENLNKED 285
81-176 GSSPFNAKAAVALLEEAKTQGINFDLNNLLSKLPSKAKENLNKED 285
        ***************:*******************.**********
```

Figure 5.1. ClustalW alignment of *C. jejuni* NCTC11168 and 81-176 protein sequences. Based on a colour legend from EBI – (http://www.ebi.ac.uk/Tools/clustalw2/help.html). '*' indicate identical matches, ':' indicates conserved substitutions, "." indicates semi-conserved substitutions and a 'space' indicates a non-conserved substitution. Underlined sequence indicates location of HD motif (PF08668).

### 5.2.3 Construction of 11168H *Cj0248* mutant

The mutagenesis strategy and techniques used to create a defined isogenic 11168H *Cj0248* mutant are described in Chapter 2. Briefly, primers were designed for PCR amplification and detection of a *Cj0248* CDS fragment (0.84 kb) using *Cj0248*-F and *Cj0248*-R (see Section 2.2.5). A *Cj0248* CDS fragment was amplified from *C. jejuni* 11168H genomic DNA and ligated with the pGEM-T Easy vector (Promega), forming the construct pGEM-*Cj0248*. This construct was transformed into XL-2 cells. Screening for positive clones was performed using PCR and sequencing. Identification of positive clones containing the *Cj0248* CDS fragment was followed by growing putative positive transformants overnight in culture and isolating plasmid DNA using a QIAprep Miniprep Kit (Qiagen). To insert a unique restriction site (*Bam*HI) within the cloned *Cj0248* CDS fragment, IPCRM was performed. Briefly, IPCR amplification of the purified plasmid was performed using *Cj0248*-IPCR-F and *Cj0248*-IPCR-R primers (see Section 2.2.5). The IPCRM product was digested with *Bgl*II and *Dpn*I, followed by ligation with the *aph-3* (aminoglycoside 3'-phosphotransferase) kanamycin resistance cassette ($Km^R$) (Trieu-Cuot *et al*., 1985) and transformation into XL-2 cells. The $Km^R$ cassette was derived from pJMK30 and does not contain a transcriptional terminator. This allows the

formation of multicistronic mRNA as the $Km^R$ can be positioned in the same orientation as the CDS of interest, hence reducing the potential for polar effects (van Vliet *et al*., 1998). Kanamycin-resistant transformants were screened by PCR using *Cj0248* specific and two $Km^R$ specific primers; $Km^R$ forward-out and $Km^R$ reverse-out (see Section 2.2.5). pUC-*Cj0248*-$Km^R$ plasmids with the $Km^R$ cassette in the same orientation as the *Cj0248* CDS were selected and electroporated into 11168H wild-type as described previously (Jones *et al*., 2004, Karlyshev *et al*., 2002). Electroporated *C. jejuni* were plated onto blood agar plates and incubated at 37°C under microaerobic conditions for 2 days. Cells were harvested and resuspended in 0.5 ml PBS. 200 µl of this suspension was spread onto blood agar plates containing kanamycin. Putative *Cj0248* mutants were restreaked and screened using PCR (Figure 5.2) and sequencing.

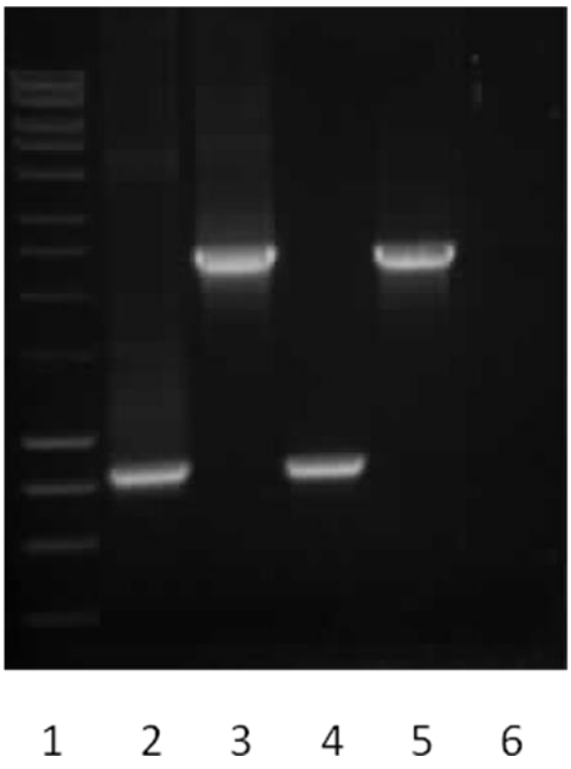



Figure 5.2. Confirmation of *Cj0248* constructs and 11168H *Cj0248* mutant using PCR with *Cj0248* specific primers. Lane 1 – 1 kb ladder / Lane 2 – pGEM-*Cj0248* plasmid DNA amplified with *Cj0248* specific primers (0.84 kb) / Lane 3 – pGEM-*Cj0248*-$Km^R$ plasmid DNA amplified with *Cj0248* specific primers (2.26 kb) / Lane 4 – *C. jejuni* 11168H wild-type strain genomic DNA amplified with *Cj0248* specific primers (0.86 kb) / Lane 5 – *C. jejuni* 11168H *Cj0248* mutant genomic DNA amplified with *Cj0248* specific primers (2.26 kb) / Lane 6 – Negative control.

After the successful construction of a *Cj0248* mutant and the selection of Cj0248 for further study, initial attempts at constructing a *Cj0248* complement proved unsuccessful. A *Cj0248* complement strain was only successfully constructed towards the end of this study (see Section 5.2.19).

### 5.2.4 The 11168H *Cj0248* mutant exhibits altered growth kinetics

To begin characterisation of the 111168H *Cj0248* mutant, growth kinetic assays were performed comparing the *Cj0248* mutant against the 11168H wild-type strain (Figure 5.3). A significant difference was observed in the growth kinetics of the wild-type strain and the *Cj0248* mutant at time points 6, 8, 10 and 12 hours. However, further incubation showed no significant difference between the *Cj0248* mutant and the 11168H wild-type strain (data not shown). As all assays were to be performed using *C. jejuni* strains grown for 24 hours, this difference during the early stages of growth was not considered as a factor that would need to be taken into account during experimental design and analysis.

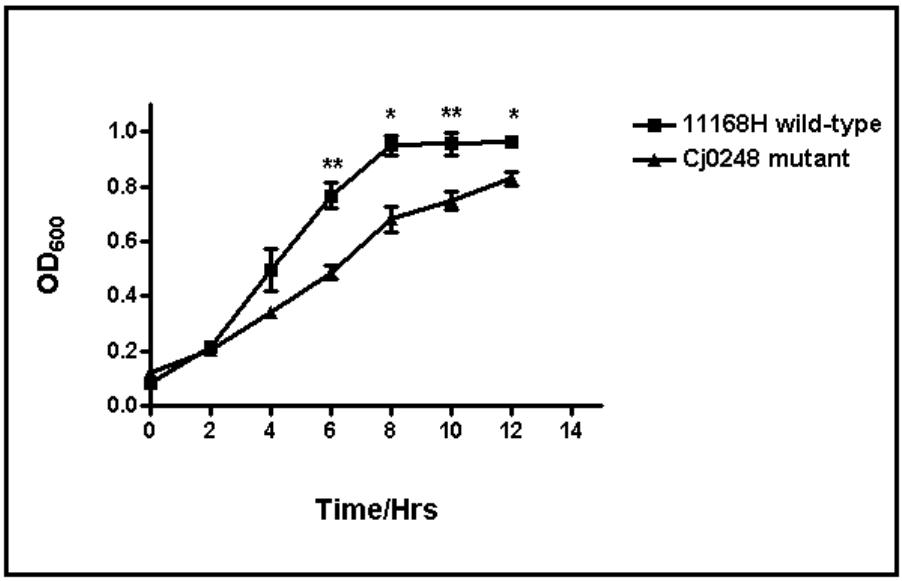



Figure 5.3. Growth kinetic assay comparing the 11168H wild-type strain to the *Cj0248* mutant. A secondary broth culture inoculation method was used where *C. jejuni* grown for 16 h was used to inoculate 15 ml preincubated Brucella broth to an $OD_{600}$ of 0.1. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were performed at selected time points. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars

represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$ and two stars (**) indicating $p = 0.001 - 0.01$.

### 5.2.5 The 11168H *Cj0248* mutant exhibits severely reduced motility

Previous published research had shown that a *C. jejuni* 81-176 *Cj0248* mutant had a severely reduced motility phenotype (Hendrixson *et al*., 2001). Consistent with this phenotype, the 11168H *Cj0248* mutant also showed a severely reduced phenotype (Figure 5.4). The severely reduced motility was statistically significant at 24, 48 and 72 h when comparing the *Cj0248* mutant to the 11168H wild-type. This data suggests Cj0248 plays a similar role in *C. jejuni* 11168H as in 81-176.

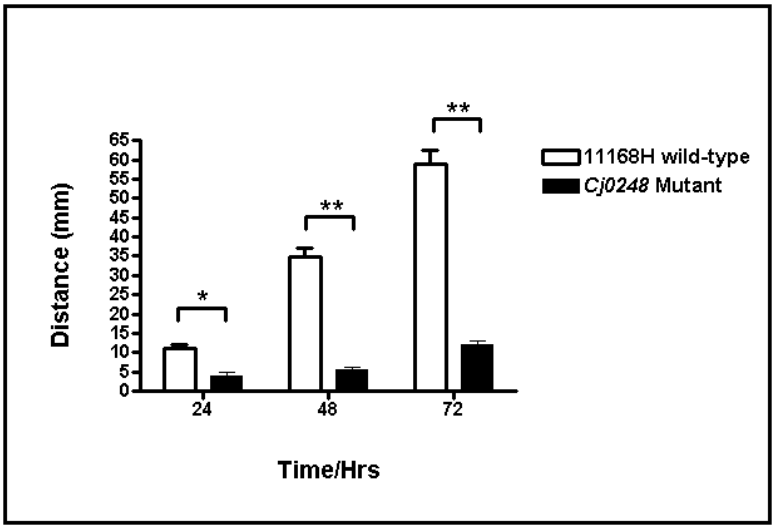



Figure 5.4. Motility assay comparing the 11168H wild-type strain and *Cj0248* mutant on semi-solid surfaces. *C. jejuni* from an overnight culture was adjusted to an $OD_{600}$ of 1.0. 5 µl of this suspension was pipetted into the centre of a motility plate using a Gilson pipette (P10). Plates were incubated at 37°C under microaerobic conditions and plate images were recorded at 24, 48 and 72 h. Motility was assessed by measuring the diameter of the halo on the plate. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$ and two stars (**) indicating $p = 0.001 - 0.01$.

### 5.2.6 The 11168H *Cj0248* mutant displays unaltered autoagglutination levels

Autoagglutination or self-association has been linked with virulence in some Gram-negative bacteria (Roggenkamp *et al*., 1995, Bieber *et al*., 1998, Brondsted *et al*., 2005). Specifically, autoagglutination has been described as a preliminary step in the formation of microcolonies (Misawa and Blaser, 2000). Microcolony formation has been identified as an important preliminary step in biofilm formation (Haddock *et al*., 2010). In *Campylobacter* species, autoagglutination has been shown to be mediated by flagella (Guerry *et al*., 2006). The autoagglutination ability of the *Cj0248* mutant was compared to that of the 11168H wild-type strain. No difference in the level of autoagglutination between the two strains was observed (Figure 5.5). A 11168H *rpoN* mutant was used as a autoagglutination negative control. Mutation of *rpoN* (which encodes the alternative sigma factor $\sigma^{54}$) renders the bacterium non-motile and without flagella (Jagannathan *et al*., 2001). This data suggests the *Cj0248* mutant has a complete or partial flagella structure and in conjunction with the motility data may suggest such flagella have impaired function.

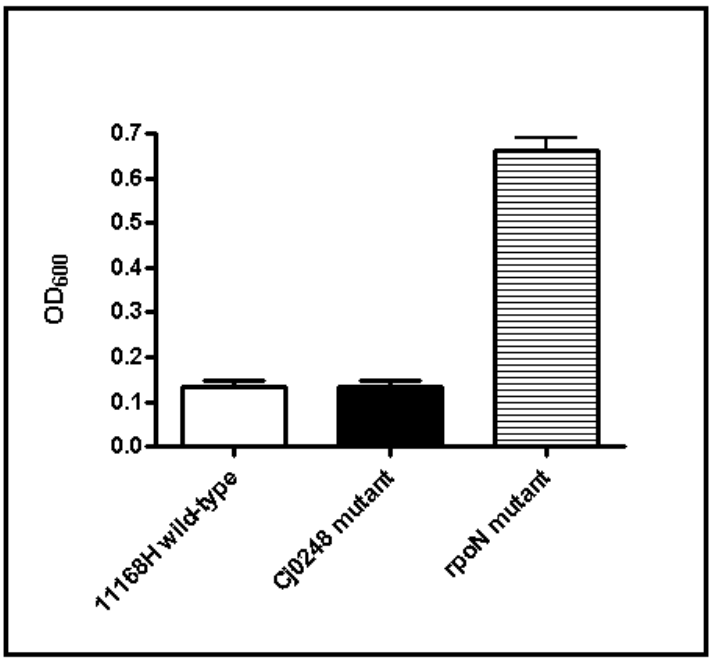



Figure 5.5. Autoagglutination ability of the 11168H wild-type strain and the *Cj0248* mutant. *C. jejuni* that were grown on blood agar plates for 24 h at 37°C under microaerobic conditions were used to inoculate 10 ml PBS to a final $OD_{600}$ of 1.0. Cultures were incubated for 24 h at 37°C under microaerobic conditions with no shaking. The $OD_{600}$ was measured using 1 ml taken from the centre of the 24 h incubated culture. A 11168H *rpoN* mutant was used a negative control which dramatically reduced autoagglutinate. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

### 5.2.7 The 11168H *Cj0248* mutant exhibits reduced biofilm formation

Biofilms are commonly defined as matrix-enclosed bacterial populations adherent to each other and/or to surfaces of interfaces (Costerton *et al*., 1995). Studies have shown that *C. jejuni* can form biofilms (Joshua *et al*., 2006) and that this is an important factor in the survival of *Campylobacter* in the environment (Reuter *et al*., 2010). Specifically, flagella have an important role in biofilm formation by allowing the survival of this microaerophilic organism in water and under other harsh environmental conditions (Guerry, 2007). To investigate this further, biofilm assays were performed where *C. jejuni* were grown for 3 days under microaerobic conditions and the level of biofilm formation was measured via crystal violet staining. Analysis of the 11168H wild-type strain and *Cj0248* mutant indicated a significant reduction in the level of biofilm

formation by the *Cj0248* mutant (Figure 5.6). Comparison between the level of biofilm formation by the *Cj0248* mutant and the *Cj1556* mutant strains also revealed differences. The level of biofilm formation for the *Cj1556* mutant was detected as approximately 0.15 $A_{595}$ whereas the level of biofilm detected for the *Cj0248* mutant data was approximately 0.10 $A_{595}$. These are relatively small differences however flagella have been described as being important for biofilm formation and this data indicates that a mutant with putatively altered flagella functionality shows a greater reduction in biofilm formation compared to a mutant with increased sensitivity to oxidative and aerobic ($O_2$) stress. The motility and autoagglutination data suggested the *Cj0248* mutant possesses complete or partial flagella, which has impaired function. This data suggests flagella with impaired function may result in decreased biofilm formation possibly as a result of less microcolony formation, which is an important step in biofilm formation.

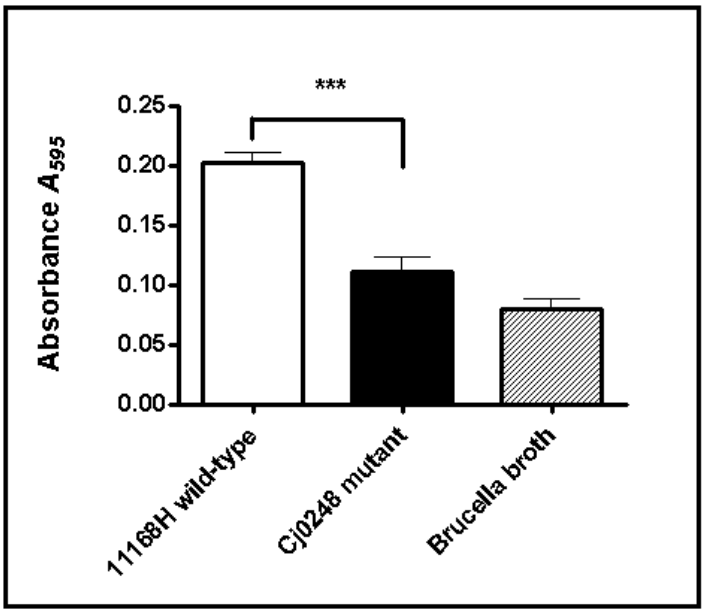
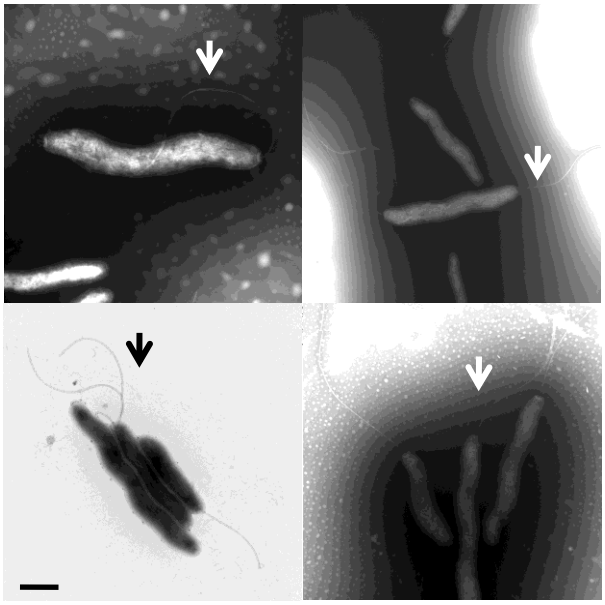



Figure 5.6. Biofilm assay on 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were inoculated with a starting $OD_{600}$ of 0.1 in culture for 72 h at 37°C under microaerobic conditions shaking at 75 rpm followed by crystal violet staining. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with three stars (***) $p < 0.001$.

### 5.2.8 Transmission electron microscopy reveals the *Cj0248* mutant possesses intact flagella

Results from the motility, autoagglutination and biofilm assays suggested *Cj0248* mutant has complete or partial flagella with impaired function. To investigate this further, TEM analysis was performed to ascertain whether the *Cj0248* mutant possessed flagella. TEM analysis was performed on *C. jejuni* cultures grown for 16 h at 37°C under microaerobic conditions and identified the presence of flagella for *Cj0248* mutant (Figure 5.7). This result would explain the autoagglutination positive phenotype of the *Cj0248* mutant (see Section 5.2.6). This result would also suggest that despite the presence of flagella, the reduced motility phenotype of the *Cj0248* mutant indicates these flagella have impaired function.

A)



B)

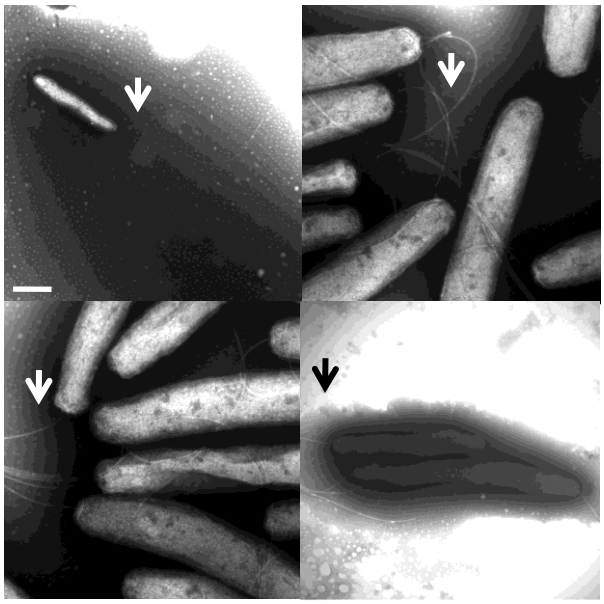


Figure 5.7. TEM analysis of the 11168H wild-type strain and *Cj0248* mutant. TEM analysis was performed on *C. jejuni* cultures grown for 16 h at 37°C under microaerobic conditions. *C. jejuni* 11168H wild-type (A) and *Cj0248* mutant (B) represented by four images from three independent experiments. Arrows indicate flagella. Magnification = 25,000x. Size bar = 500 nm. All experimental data is represented using the average of at least three biological replicates performed in triplicate.

### 5.2.9 Microarray analysis reveals changes in expression of genes involved in flagella biosynthesis

To analyse the gene expression profile of the *Cj0248* mutant compared to the 11168H wild-type strain, microarray experiments were performed using total RNA samples isolated from *C. jejuni* grown to late-log phase (16 h). Microarray experiments were performed with three biological replicates, each with one technical replicate. Statistically significant up- and down-regulated genes were selected when comparing gene expression from mutant test arrays against 11168H control arrays using ANOVA (ANalysis Of VAriance) (Bacon *et al.*, 2004, Corcionivoschi *et al.*, 2009). ANOVA was performed using a Benjamini and Hochberg False Discovery Rate as the Multiple Testing Correction within the GeneSpring software (v7.3). A total of 177 genes were differentially expressed in the *Cj0248* mutant compared to the 11168H wild-type strain based on ANOVA selection methodology. A total of 120 genes were up-regulated and 57 genes down-regulated (Appendix 11 and 12). A number of differentially expressed genes were identified that may cast further light on the phenotype of the *Cj0248* mutant (Table 5.2) and the role of Cj0248.

Table 5.2. Selection of significant genes identified from expression studies comparing the *Cj0248* mutant to the 11168H wild-type. Genes are ordered based on functional groups and not fold-change.

| Gene name | Fold change | Product function |
|---|---|---|
| *pglA* | +1.86 | GalNAc transferase |
| *flhA* | +1.81 | Flagella biosynthesis protein |
| *flgB* | -1.56 | Flagella basal-body rod protein |
| *flgC* | -4.48 | Flagella basal-body rod protein |
| *cheA* | -1.74 | Chemotaxis histidine kinase |
| *cheV* | -1.60 | Chemotaxis protein |
| *kpsD* | -1.40 | Capsule polysaccharide export system periplasmic protein |
| *kpsE* | -2.60 | Capsule polysaccharide export system inner membrane protein |
| *cbrR* | +1.40 | Two-component response regulator |
| *cmeR* | -5.16 | Transcriptional regulator |
| *rho* | -2.43 | Transcription termination factor |

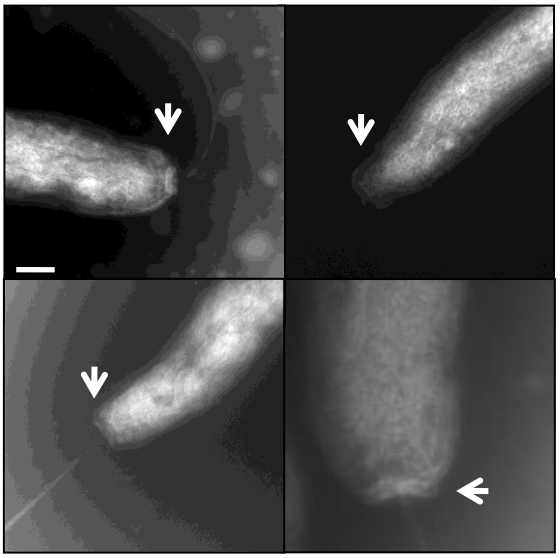
Considering the phenotype of the *Cj0248* mutant, changes in expression of genes encoding proteins with a role in flagella biosynthesis and chemotaxis were identified (see Table 5.2). Expression of *flgB* and *flgC* were identified as 1.55 and 4.48 fold down-regulated, whereas *flhA* was found to be 1.81 fold up-regulated. FlgB and FlgC are both proximal rod proteins in the flagella basal body compartment (Golden and Acheson, 2002, Konkel *et al*., 2004, Saijo-Hamano *et al*., 2004). *C. jejuni flgB* and *flgC* mutants have been shown to be defective for both motility and secretion (Konkel *et al*., 2004). FlhA is an inner membrane export apparatus protein involved in flagella protein secretion (Golden and Acheson, 2002, Miller *et al*., 1993). FlhA has been associated with a role in pathogenesis as a *C. jejuni flhA* mutant has been shown to affect motility, autoagglutination and invasion (Hendrixson and DiRita, 2003, Golden and Acheson,

2002). Expression of *cheA* and *cheV* were identified as 1.74 and 1.60 fold down-regulated. CheA and CheV are chemotaxis proteins. CheA is a histidine kinase protein part of a two-component regulatory system with the response regulator CheY, which controls the chemotaxis response to external stimuli (Chang and Miller, 2006, Hendrixson *et al*., 2001). The binding of the signal ligand is relayed by methyl-chemotaxis proteins to CheA. Analysis of the *C. jejuni* NCTC11168 genome sequence has identified MCP-like genes which encode Tlps (transducer-like proteins) forming receptors which bind to CheA and CheW (Korolik and Ketley, 2008). Diversity in the ligand binding domains of the Tlps and notable differences in the complement of response regulator domains indicate that the overall mechanism of chemotaxis single transduction in *C. jejuni* is likely to be unique (Korolik and Ketley, 2008). The *C. jejuni* NCTC11168 genome encodes 10 possible Tlp receptor proteins and two aerotaxis orthologues that potentially feed signals into the single CheA-CheW-CheY signal transduction pathway backbone (Korolik and Ketley, 2008). CheW is also part of this complex and CheV has been predicted to have a similar role to CheW as a N-terminal CheW domain fused to a C-terminal response regulator domain has been identified within CheV (Korolik and Ketley, 2008, Rosario *et al*., 1994).

### 5.2.10 Enhanced TEM analysis reveals putative flagella structural differences between 11168H wild-type strain and *Cj0248* mutant

Following on from the microarray expression studies described in Section 5.2.9, a number of genes with functions relating to flagella basal body biosynthesis were identified as differentially expressed. Further analysis was performed using TEM with a higher magnification (250,000x as opposed to 25,000x utilised in Section 5.2.8), specifically focussing on the flagella basal body region. TEM analysis was performed on *C. jejuni* cultures grown for 16 h at 37°C under microaerobic conditions. Putative structural differences in the flagella basal body region were identified when comparing the 11168H wild-type strain to the *Cj0248* mutant (Figure 5.8). The 11168H wild-type strain contains a putative defined layered flagella basal body (Figure 5.8A). The *Cj0248* mutant contains a putative altered flagella basal body region (Figure 5.8B). The locations of the proteins encoded by genes identified as differentially regulated in the microarray expression study described in Section 5.2.9 (FlhA, FlgB and FlgC) are shown in Figure 5.9.

A)



B)

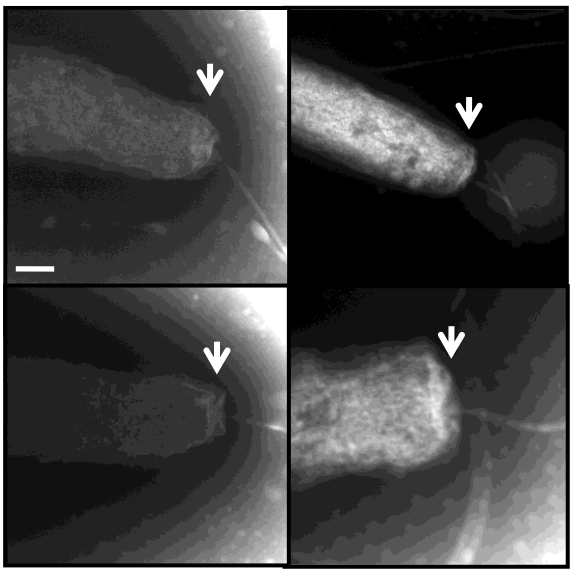


Figure 5.8. TEM analysis of 11168H wild-type strain and *Cj0248* mutant. TEM analysis was performed on *C. jejuni* culture grown for 16 h at 37°C under microaerobic conditions. 11168H wild-type strain (A) and *Cj0248* mutant (B) represented by four selected images from three independent experiments. A) 11168H wild-type strain with putative defined basal body structure, B) *Cj0248* mutant with putative altered flagella basal body region. Magnification = 250,000x. Size bar = 100 nm. All experimental data is represented using the average of at least three biological replicates performed in triplicate.

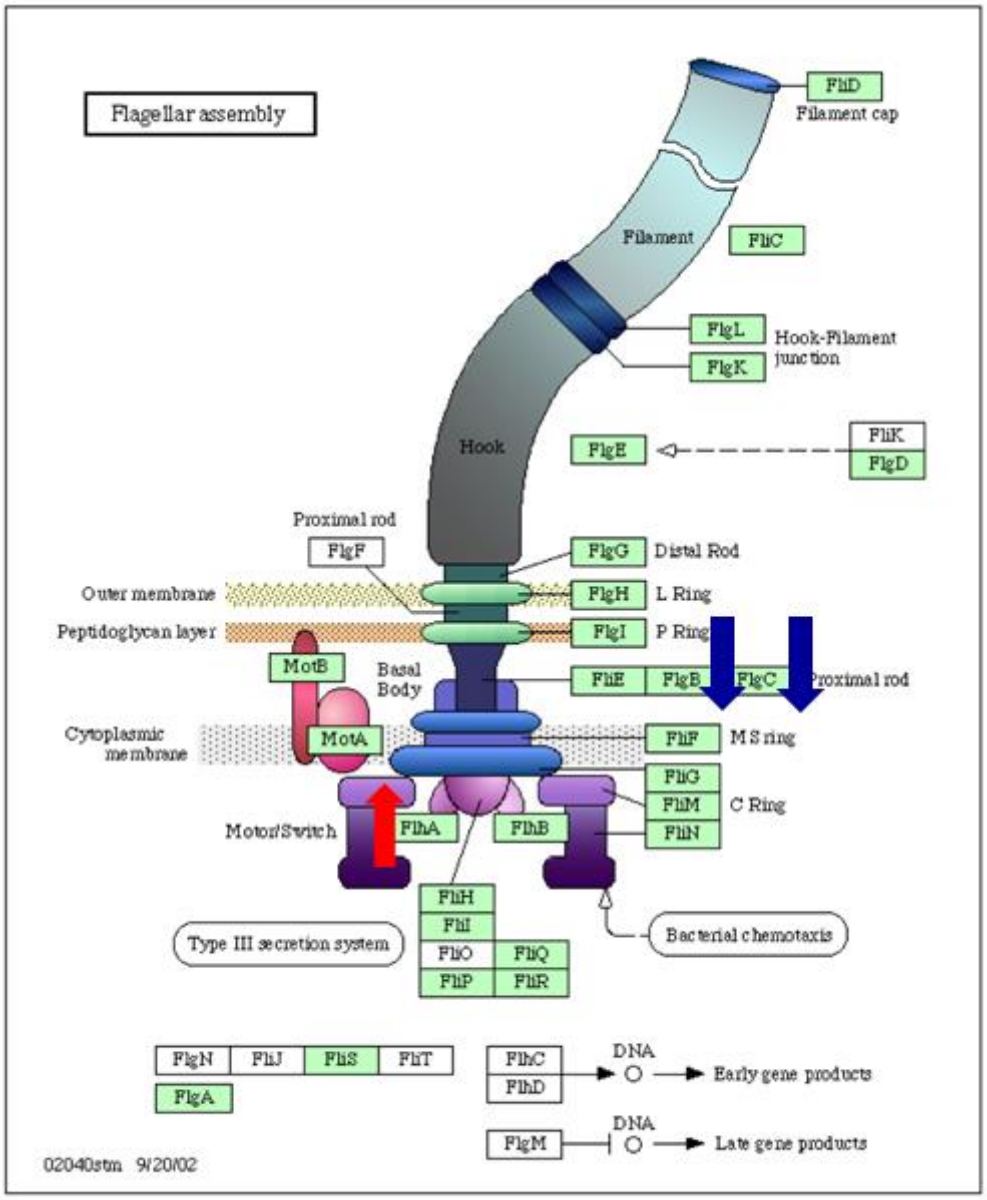



Figure 5.9. *C. jejuni* flagella structure and proteins involved in biosynthesis. Gene expression profile analysis of the *Cj0248* mutant was compared to the 11168H wild-type strain using microarray experiments from total RNA samples isolated from *C. jejuni* grown to late-log phase (16 h). Genes encoding proteins involved in flagella biosynthesis with modified expression levels in the *Cj0248* mutant are highlighted with colour arrows; *flhA* 1.81 fold up-regulated (red arrow), *flgB* and *flgC* 1.56 and 4.48 fold down-regulated respectively (blue arrows). Image obtained from GeneSpring v7.3.

### 5.2.11 The 11168H *Cj0248* mutant does not have a reduced secretion profile compared to 11168H wild-type strain

Following on from the identification of putative alterations in the flagella basal body structure and reduced motility results, further investigations were undertaken to ascertain whether the secretion profile differed between the 11168H wild-type strain and the *Cj0248* mutant. Flagella have been shown to be involved in not just motility and chemotaxis but also in the putative secretion of virulence proteins (Konkel *et al*., 2004). Thus, it was hypothesised that changes in the basal body structure may cause a possible secretion defect. The secretion of Cia proteins is dependent on a functional flagellum, indicating that this organelle has a dual function in motility and as a type III secretion system (Konkel *et al*., 2004). Research has described the identification of at least eight proteins from the culture supernatant when *C. jejuni* was grown in the presence of IECs (Konkel *et al*., 2004, Konkel *et al*., 1999b, Young *et al*., 1999). The secretion of such proteins is further induced in the presence of external compounds such as serum and/or bile salts (Guerry, 2007, Konkel *et al*., 1999b, Rivera-Amill *et al*., 2001, Malik-Kale *et al*., 2008). *C. jejuni* strains were grown in the presence of 0.1% (v/v) FCS (serum) and/or 0.1% (w/v) sDOC (bile constituent). Secretion profile analysis identified no apparent differences between the 11168H wild-type strain and *Cj0248* mutant when grown in culture for 16 h at 37°C under microaerobic conditions (Figure 5.10).

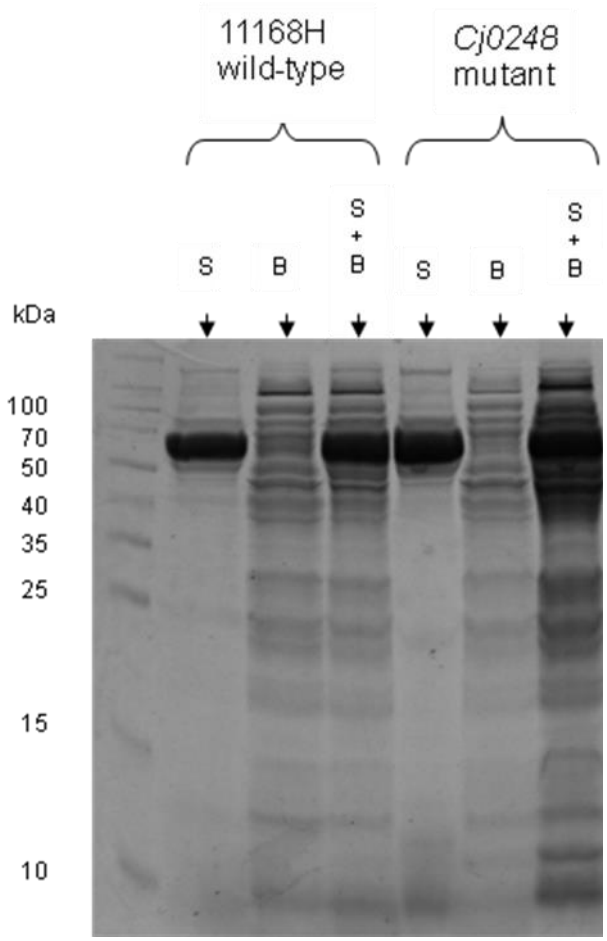



Figure 5.10. Coomassie showing secretion profile of *C. jejuni* 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were grown in culture for 16 h at 37°C under microaerobic conditions with a starting $OD_{600}$ of 0.1. Media supernatants were centrifuged, filter sterilised and concentrated. Secretion profiles were analysed by SDS-PAGE then stained with Coomassie Blue. Overnight cultures were supplemented with 0.1% (v/v) FCS serum (S) or 0.1% (w/v) sDOC (bile constituent) (B) or a combination of both. In order to semi-quantify any differences in protein secretion between strains, the starting $OD_{600}$ and growth conditions were standardized. 30 µl of final sample were analysed by SDS-PAGE. The data above is representative of experiments with at least three biological replicates, performed in triplicate.

### 5.2.12 The 11168H *Cj0248* mutant displays a reduced ability to interact (adhere and invade) with Caco-2 intestinal epithelial cells

A *C. jejuni* 81-176 *Cj0248* mutant had previously been shown to have a 100- to 10,000-fold reduction in the ability to colonise a chick caecal colonisation model when compared to the 81-176 wild-type strain (Hendrixson and DiRita, 2004). In order to

investigate this colonisation defect further, interaction assays were performed using co-culture experiments with Caco-2 IECs. The *Cj0248* mutant displayed a reduced ability to interact (adhere and invade) with Caco-2 cells after a 24 h co-culture time period compared with the 11168H wild-type strain (Figure 5.11). However no significant differences were observed when comparing the levels of interaction at 3 or 6 h. The *Cj0248* mutant interaction data is similar to the *Cj1556* mutant interaction data in that both mutants have a reduced ability to interact with Caco-2 cells compared to the 11168H wild-type strain over 24 h. However the reasons for these low interaction rates are probably different. A reason as to why the *Cj1556* mutant exhibits reduced interaction ability is the increased sensitivity to oxidative and aerobic ($O_2$) stress. In contrast, it is probable that the *Cj0248* mutant exhibits reduced interaction ability due to altered flagella functionality.

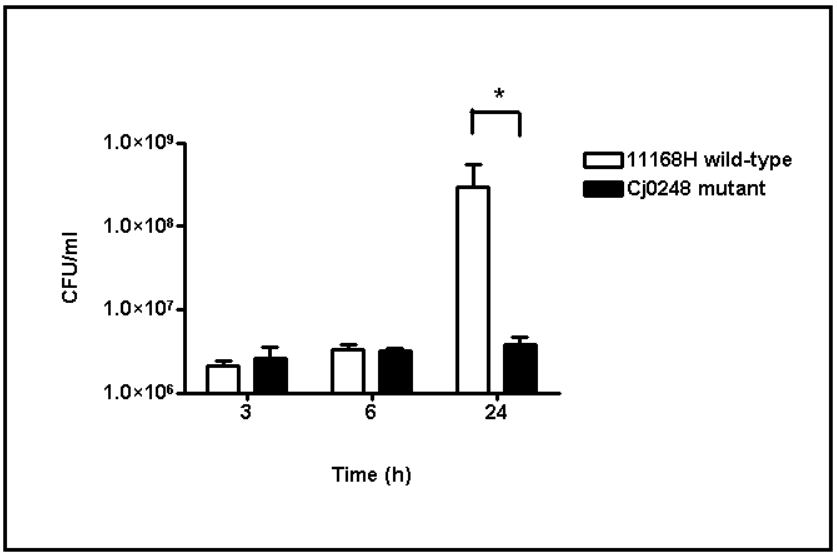



Figure 5.11. Interaction assays (adhesion and invasion) on 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were co-cultured with Caco-2 cells for 3, 6 or 24 h. Caco-2 cells were lysed and numbers of interacting bacteria assessed. The asterisk denotes a statistically significant difference (* = $p$ <0.05) in the interaction of the 11168H wild-type compared to the *Cj0248* mutant strain. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p$ = 0.01 - 0.05.

To ascertain whether the above results were due to a genuine *Cj0248* mutant phenotype and not due to increased sensitivity to Triton X-100, stress assays were performed with 0.2% (v/v) Triton X-100. No difference in the level of sensitivity to Triton X-100 was observed between 11168H wild-type strain and the *Cj0248* mutant (Figure 5.12).

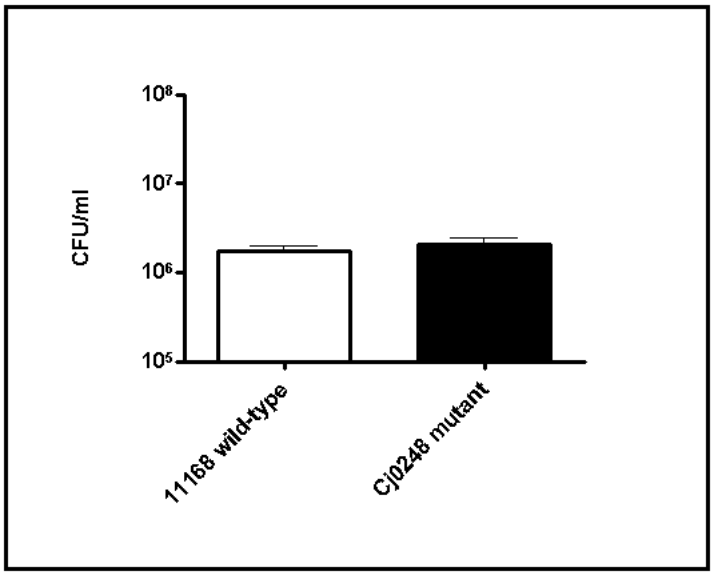



Figure 5.12. Triton X-100 0.2% (v/v) stress assay on *C. jejuni* 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were grown on blood agar plates for 24 h at 37°C under microaerobic conditions. A bacterial suspension with an $OD_{600}$ of 1.0 was prepared and incubated with 0.2% (v/v) Triton X-100 at 37°C under microaerobic conditions for 15 minutes. Serial dilutions ($10^{-1}$ - $10^{-6}$) were performed and 10 µl volumes were spotted onto blood agar plates. The plates were incubated at 37°C under microaerobic conditions for 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

### 5.2.13 The 11168H *Cj0248* mutant displays a reduced ability to invade Caco-2 intestinal epithelial cells

In addition to the interaction assays described in Section 5.2.12, the colonisation ability of the 11168H *Cj0248* mutant was further analysed by performing invasion assays using co-culture experiments with Caco-2 IECs. The *Cj0248* mutant displayed a reduced ability to invade Caco-2 cells after 3, 6 and 24 h co-culture compared with the 11168H wild-type strain (Figure 5.13). The difference in invasion was more significant at time

points 6 and 24 h compared to 3 h. Comparison of the *Cj0248* mutant invasion results to the *Cj1556* mutant invasion results showed that the *Cj0248* mutant had reduced levels of invasion at 3, 6 and 24 h, whereas the *Cj1556* mutant only showed reduced invasion at 24 h. This concurs with the hypothesis that the differences observed specifically at 24 h for interaction and invasion of the *Cj1556* mutant are more likely due to reduced intra- and extracellular survival during co-culture with Caco-2 IECs. The *Cj0248* mutant has reduced invasion compared to the 11168H wild-type at 3, 6 and 24 h. Previous data suggest the *Cj0248* mutant has impaired function. The invasion data correlates with the motility data confirming the link between the two phenotypes.

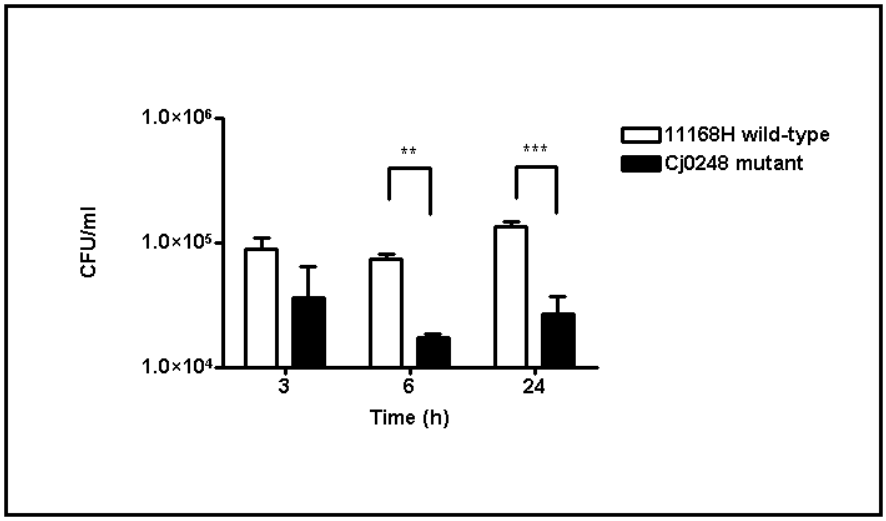


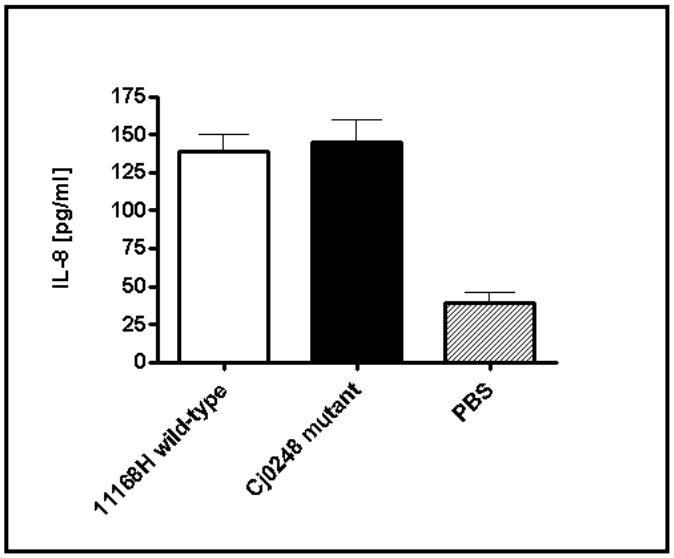

Figure 5.13. Invasion assays on 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were co-cultured with Caco-2 cells for 3, 6 or 24 h. Caco-2 cells were then incubated with gentamicin (150 μg/ml) for 2 h to kill extracellular bacteria, lysed and numbers of intracellular bacteria assessed. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$, two stars (**) indicating $p = 0.001 - 0.01$ and three stars (***) $p < 0.001$.

### 5.2.14 The 11168H *Cj0248* mutant induces a reduced IL-6 response from intestinal epithelial cells

Following on from the co-culture studies investigating the interaction and invasion properties of the *Cj0248* mutant, further co-culture investigations were performed

analysing the host innate immune response. IL-6 is an interleukin that acts as both a pro-inflammatory and anti-inflammatory cytokine and has been shown to be important for epithelial cell integrity (Friis *et al*., 2009). IL-6 is typically associated with an acute phase response (Heinrich *et al.*, 1990). Interleukin-8 is a chemokine produced by macrophages and other cell types such as epithelial cells. Both are well-characterised markers denoting a host innate immune response against pathogens (Wolff *et al*., 1998, Oppenheim *et al*., 1991). To investigate any differences between the level of IL-6 and IL-8 induced by the *Cj0248* mutant compared to the 11168H wild-type strain, co-culture experiments with two different cell lines were performed. Only minimal secretion of IL-6 and IL-8 was detected when the 11168H wild-type strain and *Cj0248* mutant were co-cultured with Caco-2 cells (data not shown). However using a T84 cell line, significant levels of induction of both IL-6 and IL-8 by the 11168H wild-type strain and *Cj0248* mutant were observed (Figure 5.14). There was no significant difference in the level of IL-8 induction by the *Cj0248* mutant compared to the 11168H wild-type strain, however a significant reduction in the level of IL-6 induction by the *Cj0248* mutant compared to the 11168H wild-type strain was observed (Figure 5.14). Comparison between the level of IL-8 induction by the *Cj0248* mutant and the *Cj1556* mutant revealed no difference when co-cultured with T84 cells for 24 h. However both mutants induced a reduced IL-6 secretion compared to the 11168H wild-type strain when co-cultured with T84 cells for 24 h. The *Cj1556* mutant IL-6 induction data was recorded at approximately 70 pg/ml, whereas the *Cj0248* mutant IL-6 induction data was recorded at approximately 45 pg/ml. There may be some biological significance to these results as IL-6 has been shown to be important for epithelial cell integrity (Friis *et al*., 2009). This data suggests, i) that bacterial ability to invade and survive within IECs is more important in IL-6 induction than IL-8 induction, ii) a *Cj0248* mutant with reduced invasion leads to less IL-6 secretion from host cells when compared to a *Cj1556* mutant with reduced survival properties.

A)




B)

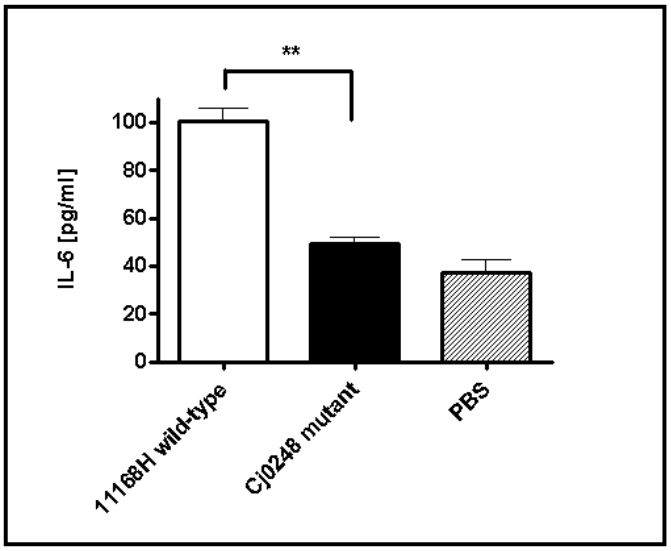



Figure 5.14. Analysis of host innate immune response during 11168H wild-type strain and *Cj0248* mutant infection. *C. jejuni* were co-cultured with T84 IECs for 24 h and the levels of IL-8 and IL-6 secreted were quantified using either a human IL-8 ELISA (A) or IL-6 ELISA (B). All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$, two stars (**) indicating $p = 0.001 - 0.01$ and three stars (***) $p < 0.001$.

### 5.2.15 *Galleria mellonella* exhibit decreased survival rates when injected with 11168H *Cj0248* mutant

*G. mellonella* larvae have recently been used as a model to study infection by *C. jejuni* and other bacteria (Champion *et al*., 2009). Insect larvae such as *G. mellonella* are favorable to use as non-mammalian infection models as they can be infected at 37°C and possess specialized phagocytic cells, termed haemocytes (Mylonakis *et al*., 2007, Bergin *et al*., 2005). Haemocytes mimic the functions of phagocytic cells in mammals and are able to degrade bacterial pathogens and also generate bactericidal compounds such as superoxide via a respiratory burst (Bergin *et al*., 2005, Lavine and Strand, 2002). When *G. mellonella* were injected with both 11168H wild-type strain and the *Cj0248* mutant, surprisingly injection with the *Cj0248* mutant resulted in decreased survival levels of *G. mellonella* larvae after 24 h infection (Figure 5.15) compared to the 11168H wild-type strain. One hypothesis to explain this result is that as *C. jejuni* is directly injected into the *G. mellonella* larvae, there is essentially no requirement for the bacterium to adhere and invade and as such the *Cj0248* mutation does not attenuate *C. jejuni* virulence in this model.

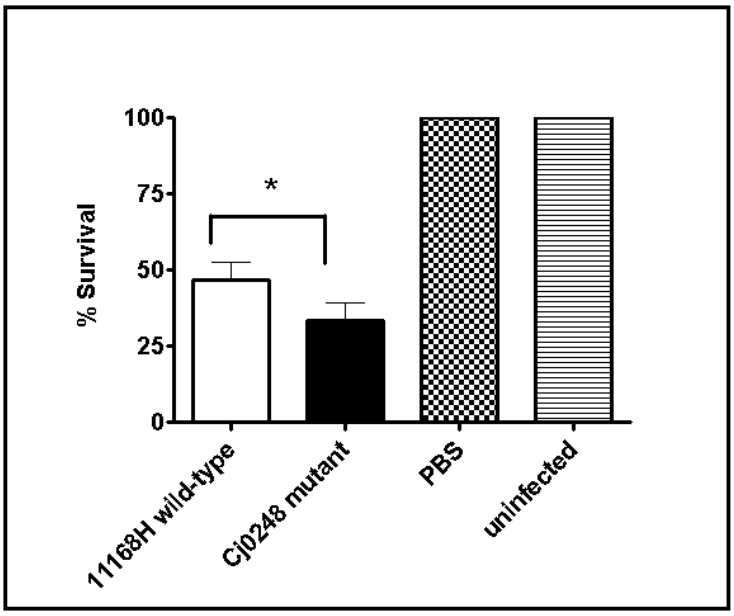



Figure 5.15. *G. mellonella* survival assay on 11168H wild-type strain and *Cj0248* mutant. *G. mellonella* larvae were injected with a 10 µl inoculum of a 24 h *C. jejuni* culture diluted to $OD_{600}$ 0.1 by micro-injection in the right foremost leg, giving an infectious dose of ≈ $10^6$ CFU. Larvae were incubated at 37°C with survival and appearance recorded at 24 h intervals. All experimental data is represented using the

average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01 - 0.05$.

### 5.2.16 The 11168H *Cj0248* mutant exhibits increased sensitivity to heat stress

In order to investigate the potential for other phenotypic changes induces by mutation of the *Cj0248* gene, a number of stress assays were performed. One such assay was heat stress. Several *C. jejuni* heat shock responses have been identified and characterised that contribute to maintenance of viability. These include the HtrA protease and the HspR regulator that have already been noted as playing a role in aerobic ($O_2$) stress, however their primary role is with heat tolerance (Brondsted *et al*., 2005). *C. jejuni* also contains homologues of well known chaperones and proteases e.g. GroESL, DnaK and HrcA (Parkhill *et al*., 2000). These are also referred to as heat shock proteins. Heat stress above 55°C has been noted to accelerate the spiral-to-coccoid transition and result in cell death (Nguyen *et al*., 2006). A range of heat stress experiments on the 11168H wild-type strain and *C. jejuni Cj0248* mutant were performed to explore any potential role of Cj0248 in heat response. No differences were observed when using 42°C for 1 h or 55°C for 15 mins respectively. However the *Cj0248* mutant did display increased sensitivity to 60°C when exposed for 5 minutes (Figure 5.16).

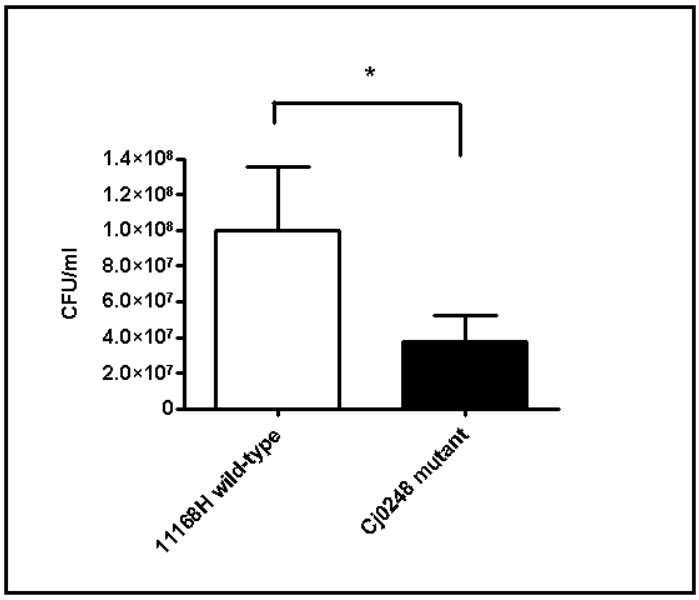



Figure 5.16. Heat stress assay on *C. jejuni* 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* strains were incubated at 60°C for 5 minutes and bacterial survival assessed. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p = 0.01$ - 0.05.

### 5.2.17 The 11168H *Cj0248* mutant exhibits similar levels of sensitivity to oxidative and nitrosative stress

Oxidative stress assays were performed using $H_2O_2$ as described in Section 2.3.4. Nitrosative stress assays were performed using acidified $NaNO_2$ as described in Section 2.3.5. No differences between the survival of the 11168H wild-type strain and *Cj0248* mutant were observed in either stress assay (data not shown).

### 5.2.18 The 11168H *Cj0248* mutant reveals unaltered growth kinetics during growth inhibition studies

Following on from the stress assays, growth inhibition assays were performed as described in Sections 2.3.6-9 to ascertain whether any modification in growth kinetics was observed due to the addition of certain compounds. Growth inhibition studies were performed using conditions to represent both oxidative and nitrosative inhibition. No differences were observed between the 11168H wild-type strain and the *Cj0248* mutant for oxidative (Figure 5.17) and nitrosative (Figure 5.18) growth inhibition. Growth inhibition studies of the *Cj0248* mutant in broth supplemented with 0.1% (w/v) sDOC

showed no difference compared to the 11168H wild-type strain (Figure 5.19). In addition, no differences were observed between the 11168H wild-type strain and *Cj0248* mutant during growth in broth supplemented with the iron chelator deferoxamine (Figure 5.20).

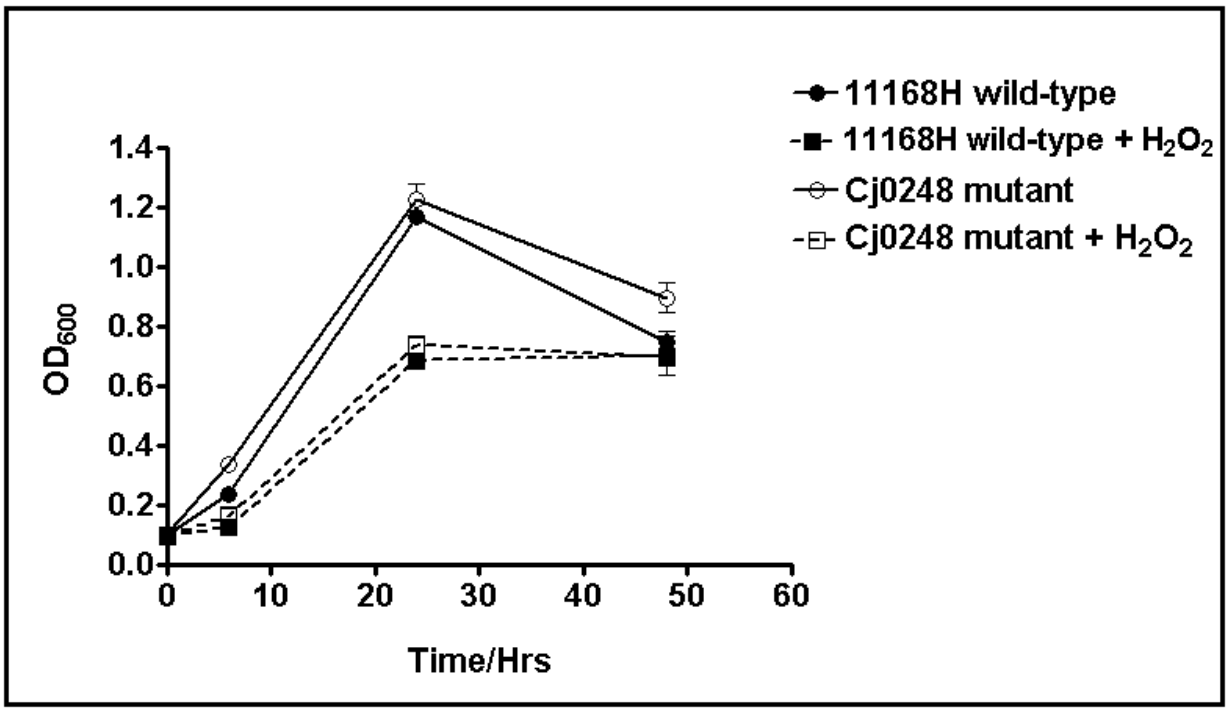



Figure 5.17. Oxidative stress growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were grown in the presence of 1 mM $H_2O_2$. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

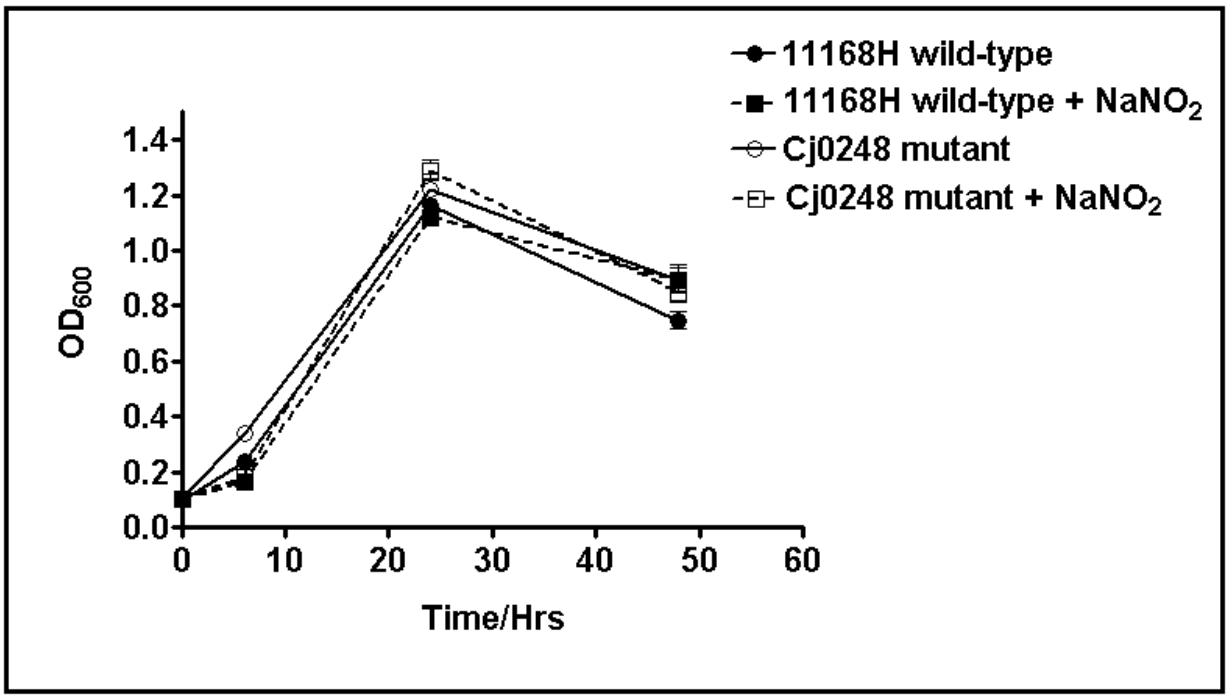



Figure 5.18. Nitrosative stress growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were grown in the presence of 0.1 mM $NaNO_2$ at pH 5. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

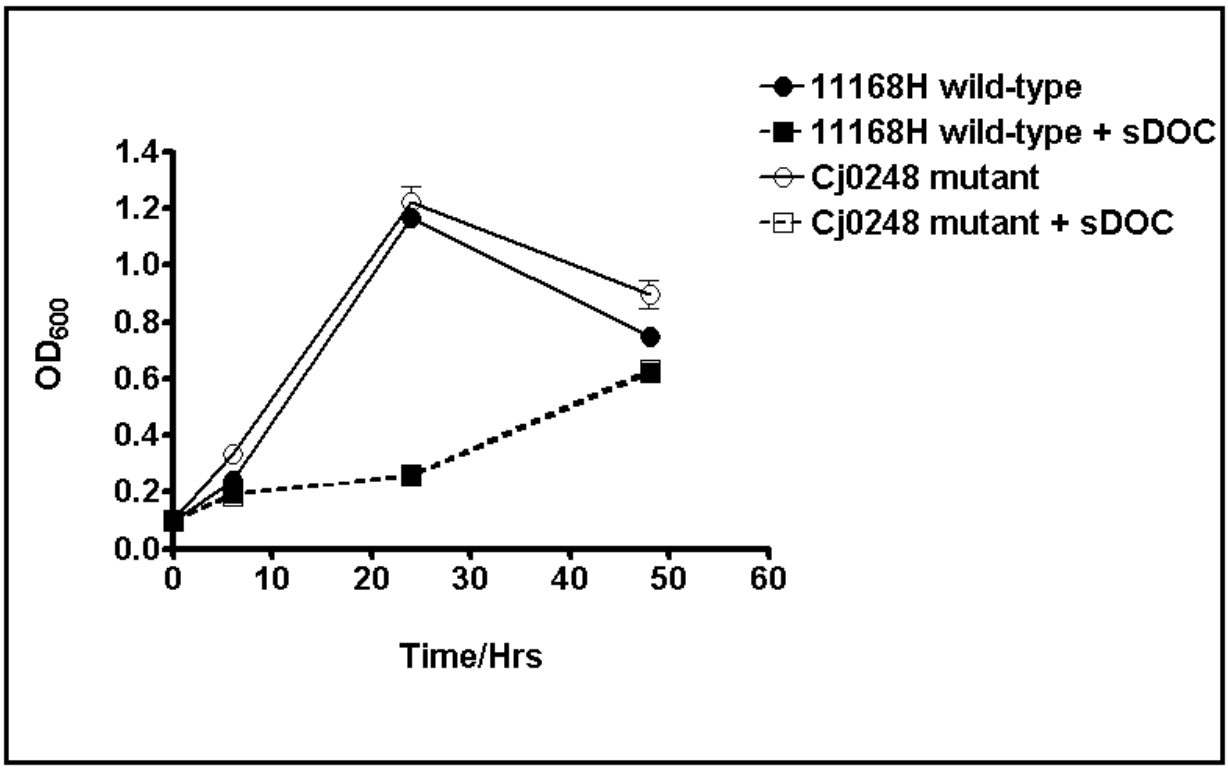



Figure 5.19. Bile stress growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were grown in the presence of 0.1% sDOC. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

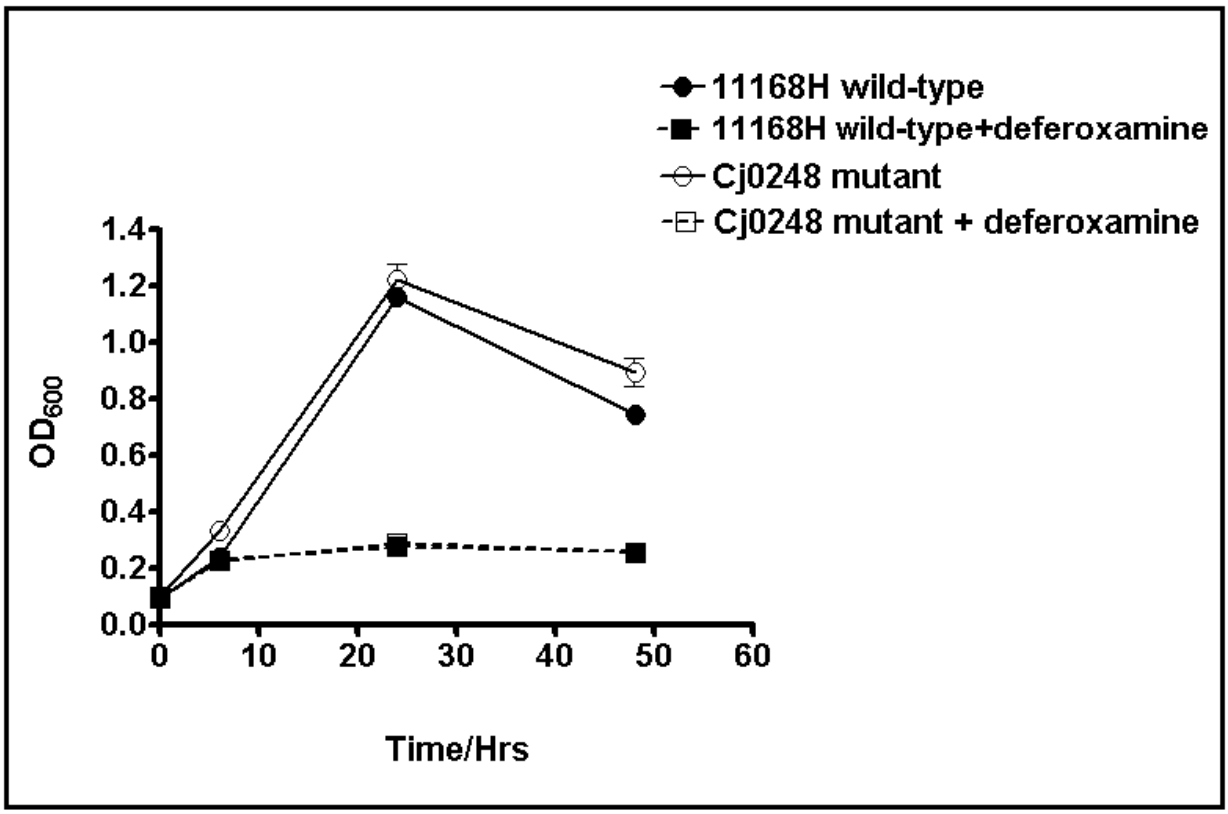



Figure 5.20. Iron limitation growth inhibition assays on *C. jejuni* 11168H wild-type strain and *Cj0248* mutant. *C. jejuni* were grown in the presence of 1 mM deferoxamine. Flasks were incubated at 37°C under microaerobic conditions, shaking at 75 rpm. $OD_{600}$ readings were recorded at 6, 24 and 48 h. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean.

### 5.2.19 Complementation of the *Cj0248* mutant

After a number of technical difficulties, complementation of the 11168H *Cj0248* mutant was successfully achieved. Complementation was performed by inserting a copy of the *Cj0248* CDS into the 11168H *Cj0248* mutant chromosome using pDENNIS, a *C. jejuni* 11168H complementation vector allowing the insertion of a functional CDS into the *Cj0223* pseudogene (Hitchen *et al*., 2010). The entire coding region for *Cj0248* was amplified by PCR using primers Comp-*Cj0248*-F and Comp-*Cj0248*-R, which introduced an *Nco*I site at the 5' end and an *Nhe*I site at the 3' end as well as including the ribosome binding site of *Cj0248* (Svensson *et al*., 2008, Wosten *et al*., 1998). Following digestion with *Nhe*I and *Nco*I, this PCR product was ligated into the pDENNIS vector. This construct was sequenced to ensure there were no mutations in the *Cj0248* nucleotide sequence. This construct was electroporated into the 11168H *Cj0248*

mutant and grown on blood agar plates. Cells were harvested and resuspended in 0.5 ml PBS. 200 µl of this suspension was spread onto blood agar plates containing kanamycin and chloramphenicol. Putative clones were restreaked and confirmed by PCR checking for the correct distance between *Cj0248* and the inserted $Km^R$ cassette (*Cj0248*-F and $Km^R$ forward-out). This confirmed the presence of the mutated CDS with the $Km^R$ cassette in the correct orientation. In addition, PCR was used to confirm the presence of an intact CDS (*Cj0248*-F and *Cj0248*-R). Sequencing was also used to confirm positive clones.

Motility assays were performed with the 11168H wild-type strain, *Cj0248* mutant and *Cj0248* complement. Partial resuscitation of the motility phenotype was observed with the *Cj0248* complement (Figure 5.21). This data gives further evidence that Cj0248 is directly involved in flagella functionality and the severely reduced motility phenotype observed from the *Cj0248* mutant is a direct consequence from inhibiting expression of Cj0248.

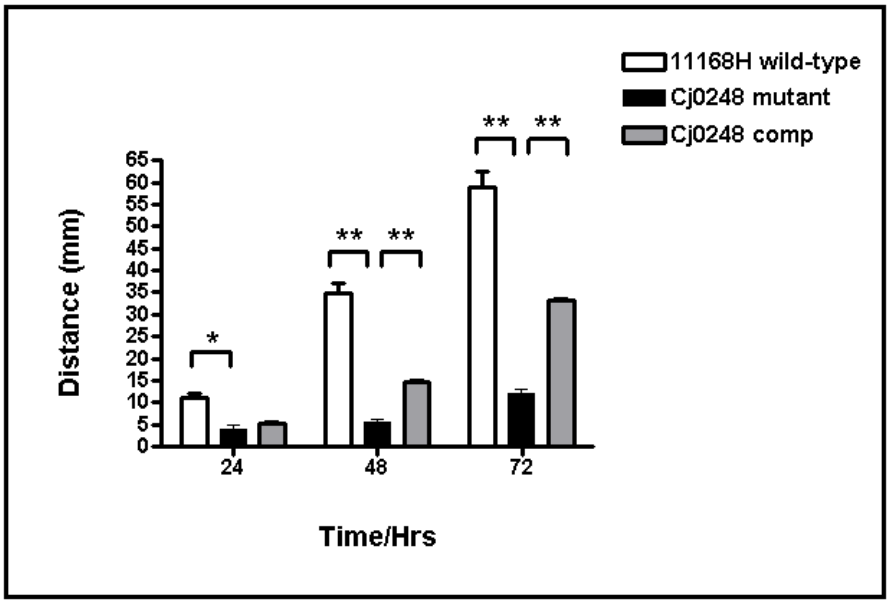



Figure 5.21. Motility assay comparing the 11168H wild-type strain, *Cj0248* mutant and *Cj0248* complement on semi-solid surfaces. *C. jejuni* from an overnight culture was adjusted to an $OD_{600}$ of 1.0. 5 µl of this suspension was pipetted into the centre of a motility plate using a Gilson pipette (P10). Plates were incubated at 37°C under microaerobic conditions and plate images were recorded at 24, 48 and 72 h. Motility was assessed by measuring the diameter of the halo on the plate. All experimental data is represented using the average of at least three biological replicates performed in triplicate. Error bars represent the standard error of mean. Variables were compared for significance using a student's *t*-test with one star (*) indicating $p$ = 0.01 - 0.05 and two stars (**) indicating $p$ = 0.001- 0.01.

### 5.3 Discussion

Motility is central to the pathogenesis of *C. jejuni* via active multi-functional flagella that allow adhesion to and invasion of IECs. A *C. jejuni* 81-176 *Cj0248* mutant had previously been shown to have a severely reduced motility phenotype (Hendrixson *et al*., 2001). Cj0248 has thus been implicated with a role in motility, though the function remained unknown. The 11168H *Cj0248* mutant displayed reduced motility when compared to the wild-type strain and this data mirrored the *C. jejuni* 81-176 data observed previously (Hendrixson *et al*., 2001). To further investigate this phenotype, autoagglutination studies were performed. In *Campylobacter* species, autoagglutination has been shown to be mediated by flagella (Guerry *et al*., 2006). Autoagglutination has been described as a preliminary step in the formation of microcolonies (Misawa and

Blaser, 2000) and microcolony formation has been identified as an important preliminary step in biofilm formation (Haddock *et al*., 2010). Autoagglutination studies were performed and no differences were observed between the 11168H wild-type strain and *Cj0248* mutant. This data suggests the *Cj0248* mutant does possess complete or partial flagella and yet the motility data suggests the flagella have impaired function. Investigation of biofilm formation showed that the *Cj0248* mutant has reduced ability to form biofilms compared to the 11168H wild-type strain. Previous research has indicated that flagella have an important role in biofilm formation conspiring to facilitate survival of this microaerophilic organism in water and under other harsh environmental conditions (Guerry, 2007). Taken together, the motility and autoagglutination data suggest the *Cj0248* mutant has complete or partial flagella, but with impaired function. The biofilm data suggests flagella with impaired function may result in decreased biofilm formation possibly as a result of less microcolony formation. It is possible Cj0248 is directly involved in biofilm formation, though the observed phenotypes are more likely due to flagella with impaired function. The presence of flagella was confirmed using TEM for both the 11168H wild-type strain and *Cj0248* mutant. This concurs with the motility, autoagglutination and biofilm data that flagella are present and the observed phenotypes are most likely due to a lack of functionality. It remains to be proved whether flagella with impaired function result in reduced microcolony formation, though it is a possibility as reduced biofilm formation was observed. An important point to note is that these biofilm assays were performed *in vitro*. Any future experiments focussing on microcolony or biofilm formation using IECs such as Caco-2 would prove negative as *C. jejuni* does not form biofilms on such tissue culture cells as they do not secrete mucin. However, using tissue culture cells such as mucin-secreting HT29 cells or *ex vivo* organ culture assays could be utilised to further investigate this phenotype.

Effective regulation of gene expression in *C. jejuni* is critical for bacterial survival. The *C. jejuni* genome contains three sigma factors: RpoD ($\sigma^{70}$), RpoN ($\sigma^{54}$) and FliA ($\sigma^{28}$) along with 34 transcription factors identified to date (Wösten *et al*., 2008). RpoD is the main sigma factor allowing binding of the RNA polymerase to the majority of *C. jejuni* promoters. RpoN has been noted to have at least 17 putative RpoN promoters in the genome (Wösten *et al*., 2008, Carrillo *et al*., 2004). The 17 identified RpoN promoters of *C. jejuni* control the transcription of 23 genes, of which 15 encode proteins that are involved in the assembly of the flagella (Wösten *et al*., 2008, Carrillo *et al*., 2004). FliA is a sigma factor that has been found to regulate the activity of at least 10 promoter sites

in *C. jejuni* that direct the transcription of 14 different genes (Wösten *et al*., 2008, Carrillo *et al*., 2004). These genes encode proteins involved in flagella apparatus biosynthesis, proteins involved in the glycosylation of the major flagellin subunits and proteins secreted through the flagella (Logan *et al*., 2002). *fliA* mutants are non-motile but still possess a flagella hook structure (Hendrixson *et al*., 2001, Jagannathan *et al*., 2001).

Flagella have been shown to be involved in not just motility and chemotaxis but also the putative secretion of virulence proteins, autoagglutination, microcolony formation and avoidance of the innate immune response (Guerry, 2007). Previous studies have demonstrated that mutations in genes involved in flagella biosynthesis are non-motile and result in reduced intestinal colonisation of animals and humans and for the invasion of IECs *in vitro* (Pavlovskis *et al*., 1991, Guerry, 2007, Nachamkin *et al*., 1993). The secretion of Cia proteins is also dependent on a functional flagellum, indicating that this organelle has a dual function in motility and as a type III secretion system (Konkel *et al*., 2004). Cia proteins have been identified from culture supernatant when *C. jejuni* was grown in the presence of IECs (Konkel *et al*., 2004, Konkel *et al*., 1999b, Young *et al*., 1999). The expression of genes encoding these proteins is further induced in the presence of other compounds such as serum and bile salts (Guerry, 2007, Konkel *et al*., 1999b, Rivera-Amill *et al*., 2001, Malik-Kale *et al*., 2008). Secretion of Cia proteins were not detected in a *flhB* mutant (defective in inner membrane export apparatus) or in *flgBCE* mutants (defective for the basal body and hook apparatus) (Larson *et al*., 2008). Secretion of Cia proteins occurred in both a *flaA* mutant and *flaB* mutant, but not a double *flaA flaB* mutant that completely lacked a flagella filament (Konkel *et al*., 2004). Secretion of Cia proteins requires a functional basal body and hook and at least one of the filament proteins (Konkel *et al*., 2004). The putative role of flagella as a secretory apparatus has proved controversial. Watson and Galan argue that there is little evidence that any component of the flagella apparatus directly secretes proteins involved in invasion when in contact with IECs (Watson and Galán, 2008). These authors argue that previous research shows that gene mutations causing a reduced motility phenotype, but not affecting the structure of the flagella apparatus, show reduced invasion into host IECs (Yao *et al*., 1994). Their argument is that these mutants, though non-motile, still harbour fully secretion positive flagella and so should be able to secrete the proposed virulence proteins.

The flagellum is a key virulence factor for *C. jejuni* enabling motility and intestinal colonisation (Hendrixson, 2006, Hendrixson and DiRita, 2004). To further investigate the reduced motility phenotype, interaction (adhesion and invasion) and invasion assays using Caco-2 IECs were performed. No significant difference was observed in the level of interaction between the 11168H wild-type and the *Cj0248* mutant after 3 and 6 h co-culture. However, a significant decrease in the level of interaction was observed after 24 h. Both the *Cj0248* mutant and the 11168H wild-type strain exhibit a similar level of autoagglutination, however the *Cj0248* mutant exhibits a reduced level of biofilm formation compared to the 11168H wild-type. It is possible that the reduction in flagella functionality in the *Cj0248* mutant results in decreased microcolony formation and as such the decrease in the level of interaction observed after 24 h when comparing the *Cj0248* mutant to the 11168H wild-type may be due to this reduction in the ability to form microcolonies. Assessing the numbers of viable 11168H wild-type and *Cj0248* mutant present after 24 h co-culture did not reveal any difference and this suggests that long term survival is not a factor (data not shown). Comparison between *Cj0248* mutant and *Cj1556* mutant reveal both have a reduction in interaction at 24 h compared to the 11168H wild-type strain. However it is probable these results are due to different reasons. The *Cj1556* mutant has been shown to have increased sensitivity to oxidative, aerobic ($O_2$) stress and reduced aerobic survival in tissue culture media compared to *Cj0248* mutant. Thus, the low interaction data for *Cj1556* mutant is more likely due to less numbers being present after 24 h, rather than a reduction in interaction ability. This is in contrast to the *Cj0248* mutant which has a reduced motility and potentially reduced microcolony formation and thus has a reduced ability to adhere and invade. It is also important to remember that this interaction data investigates the numbers of bacteria adhered and invaded.

To investigate invasion specifically, invasion assays were performed and identified a significant reduction in the level of invasion of the *Cj0248* mutant compared to the 11168H wild-type strain after 3, 6 and 24 h co-culture. The reduction in the level of invasion was more significant at time points 6 and 24 h compared to 3 h. The invasion results for the *Cj0248* mutant show different invasion dynamics when compared to the *Cj1556* mutant results. The *Cj0248* mutant shows reduced invasion at 3, 6 and 24 h, whereas the *Cj1556* mutant only exhibits reduced invasion at 24 h. This *Cj1556* mutant invasion data matches the interaction data for the *Cj1556* mutant, as significant reductions were only observed after 24 h indicating the results are more likely due to

reduced survival properties. Given that the *Cj0248* mutant has a reduced motility phenotype, it is likely the invasion phenotypes observed after 3, 6 and 24 h co-culture are linked to flagella with impaired function. Previous studies have demonstrated a 100- to 10,000- fold reduction in colonisation using a chick caecal colonisation model for the 81-176 *Cj0248* mutant compared to the 81-176 wild-type strain (Hendrixson and DiRita, 2004). This previous investigation focussed on colonisation ability, whereas these studies analysed the interaction and invasion properties over time using IECs.

Based on the reduced motility, biofilm, interaction and invasion phenotypes observed of *Cj0248* mutant compared to the 11168H wild-type strain, investigation into the host innate immune response was performed. IL-6 is an interleukin that acts as both a pro-inflammatory and anti-inflammatory cytokine and has been shown to be important for epithelial cell integrity (Friis *et al*., 2009). IL-8 acts as a chemoattractant allowing the recruitment of lymphocytes and neutrophils (Hobbie *et al*., 1997, Philpott *et al*., 2000). Both are well-characterised markers denoting a host innate immune response against pathogens (Wolff *et al*., 1998, Oppenheim *et al*., 1991). Minimal IL-6 and IL-8 was detected when the 11168H wild-type strain and *Cj0248* mutant were co-cultured with Caco-2 cells (data not shown). This concurs with research demonstrating the poor chemokine/cytokine induction ability of Caco-2 cells when co-cultured with *C. jejuni* (MacCallum *et al*., 2006). However using a T84 cell line, significant levels of induction of both IL-6 and IL-8 by the 11168H wild-type strain and *Cj0248* mutant were observed. Previous reports have demonstrated ELISA experiments using T84 cells to detect levels of IL-8 and IL-6 over 1000 pg/ml when co-cultured with *C. jejuni* (MacCallum *et al*., 2006). In this study, these levels were not achieved most likely due to a number of factors that are discussed in Section 4.3. However, the trends observed in this study using T84 cells would most likely be consistent if the increased level of chemokine/cytokine induction was achieved by modifications to the experimental protocol.

In this study, there was no significant difference in the level of IL-8 induction by the *Cj0248* mutant compared to the 11168H wild-type strain, however a significant reduction in the level of IL-6 induction by the *Cj0248* mutant compared to the 11168H wild-type strain was observed. It is possible that the reduced IL-6 induction by the *Cj0248* mutant when co-cultured with T84 cells is a result of less interaction with IECs due to flagella with impaired function. This hypothesis would be supported by the fact that reduced invasion has been observed for the *Cj0248* mutant when compared to the

11168H wild-type strain. Comparison between the level of IL-6 detected from T84 cells after co-culture with the *Cj0248* and *Cj1556* mutants revealed less induction for the *Cj0248* mutant (45 pg/ml) compared to the *Cj1556* mutant (70 pg/ml). This data suggests that, i) bacterial ability to invade and survive within IECs are important in IL-6 induction, ii) a *Cj0248* mutant with reduced motility and invasion ability induces less IL-6 secretion when co-cultured with T84 cells compared to a *Cj1556* mutant with increased sensitivity to oxidative and aerobic ($O_2$) stress.

The results of the *G. mellonella* studies identified a decreased survival of *G. mellonella* larvae when infected with *Cj0248* mutant compared to infection with the 11168H wild-type strain. The innate immune system of *G. mellonella* has a high degree of structural and functional homology with the innate immune system of mammals, with the defence against bacteria involving both cellular mechanisms such as phagocytosis, nodulization and encapsulation, along with humoral mechanisms like melanisation, haemolymph clotting and antimicrobial peptide production (Champion *et al*., 2009). A *Cj0248* mutant with putative flagella with impaired function may have been assumed to have resulted in increased survival of *G. mellonella*. However, *G. mellonella* studies identified decreased survival when injected with *Cj0248* mutant. One hypothesis to explain this result is that *C. jejuni* is directly injected into the *G. mellonella* and so there is essentially no requirement for the bacterium to adhere and invade. One method of analysing if this was true would be to feed *C. jejuni* to the larvae, though the feasibility of this has yet to be ascertained. This data would suggest there is actually no reduction in the level of virulence and survival of the *Cj0248* mutant once internalised.

Based on the reduced motility, interaction and invasion phenotypes of the *Cj0248* mutant compared to the 11168H wild-type strain, microarray expression analysis was performed on the *Cj0248* mutant comparing to the 11168H wild-type strain where RNA was isolated from bacterial culture grown to the late-log phase (16 h). Up- and down-regulated genes were identified from the study and *flgB* (*Cj0528c*) and *flgC* (*Cj0527c*) were identified as being 1.55 and 4.48 fold down-regulated in the *Cj0248* mutant. Both of these down-regulated genes encode part of the rod section of the flagella, part of the basal body between the cytoplasm and outer membrane (Konkel *et al*., 2004, Saijo-Hamano *et al*., 2004). *C. jejuni flgB* or *flgC* mutants are defective for both motility and secretion (Konkel *et al*., 2004), so the decrease in expression of these genes may be significant in understanding why *Cj0248* mutant exhibits severely reduced motility and

interaction with IECs. The results in this study are in partial agreement with previous research where *flgB* or *flgC* mutants are defective for both motility and secretion. The *Cj0248* mutant does have a reduced motility phenotype compared to the 11168H wild-type strain, however based on the secretion profile study did not identify any differences between the *Cj0248* mutant and the 11168H wild-type strain. The reduced expression of *flgB* or *flgC* may or may not be sufficient to give similar phenotypic results as *flgB* or *flgC* mutants have. Also, given that secretion mechanisms rely on a functional basal body (Konkel *et al*., 2004), it is possible that mutation of *Cj0248* has an effect on this particular structure and thus not only reducing motility, but also the invasiveness of the bacterium by inhibiting protein secretion via the flagella. However, preliminary secretion profile analysis identified no apparent differences between the 11168H wild-type strain and *Cj0248* mutant when grown in culture for 16 h at 37°C under microaerobic conditions *in vitro*. Interestingly, *flhA* was 1.8 fold up-regulated. *flhA* encodes a protein involved in the biosynthesis of the export component of the flagella system that resides in the cytoplasm. FlhA is a key component of the flagella export apparatus which belongs to the FHIPEP (flagella/Hr/invasion proteins export pore) family of bacterial export proteins involved in flagella assembly and type III secretion (Carrillo *et al*., 2004, Park *et al*., 2000, Galan *et al*., 1992). FlhA has been associated with a role in pathogenesis as an *flhA* mutant strain has been shown to affect motility, autoagglutination and invasion (Hendrixson and DiRita, 2003, Golden and Acheson, 2002). Genes encoding flagella proteins have been grouped into three classes based on the order of transcription (Carrillo *et al*., 2004). Class I include genes encoding the flagella transport apparatus and have RpoD ($\sigma^{70}$) promoters (Petersen *et al*., 2003). FlhA forms part of this flagella transport apparatus. Class II genes with RpoN ($\sigma^{54}$) promoters are required for the basal body, hook and flagella filament biosynthesis. Class III genes with FliA ($\sigma^{28}$) promoters include genes required for filament biosynthesis. Mutation of *flhA* inhibits transcription of both class II and class III flagella genes and potential virulence factors regulated by FliA ($\sigma^{28}$) and RpoN ($\sigma^{54}$) (Carrillo *et al*., 2004). The functions of FlgB, FlgC and FlhA all relate to flagella and so a lack of Cj0248 may be causing the expression changes observed in this study. This may suggest a compensatory effect where the *Cj0248* mutant has a reduction in *flgB* and *flgC* expression, leading to increased expression of *flhA* to attempt to increase expression of *flgB* and *flgC*. Further investigate is required to elucidate this potential link.

Results from the microarray data led us to focus on the flagella basal body region as potentially being important in Cj0248 function. The TEM images were analysed at greater magnification and specifically looked at the flagella basal body region. Putative structural defects in the flagella basal body region were identified in the *Cj0248* mutant when compared to the 11168H wild-type strain. The TEM images were replicated using a number of different *Cj0248* mutant clones and also using different *Cj0248* mutants produced from newly electroporated constructs. The 11168H wild-type strain contains a putative defined layered flagella basal body structure however the *Cj0248* mutant contains a putative altered flagella basal body region (see figures 5.7 and 5.8). The TEM data suggests that the flagella basal body region appears to either not have formed or has been altered in the *Cj0248* mutant. Interestingly, the altered flagella basal body region appears to be varied in structure. This variation was observed on all replicate samples. There does not seem to be one specific altered structure, but a varied level of structural alteration. No links were identified between the Cj0248 protein and any of the flagella associated proteins in the *Campylobacter* Protein Interaction Database (Parrish *et al*., 2007), however *flgB* and *flgC* were identified as being down-regulated in the *Cj0248* mutant compared to the 11168H wild-type strain in this study. It is possible that the mutation of *Cj0248* has affected the expression of basal body encoding genes and thus affected flagella functionality specifically. Further studies could focus on the localisation of Cj0248 to ascertain whether the protein location is indeed around the bacterial membrane region in proximity of the flagella.

Given the putative secretion of Cia proteins is dependent on a functional flagellum (Konkel *et al*., 2004), further investigations into the secretion profile of the 11168H wild-type strain and the *Cj0248* mutant were performed. Previous studies have shown that protein secretion via a functional flagellum was not detected when *C. jejuni* was incubated in the absence of IECs (Konkel *et al*., 2004). However in this study, *in vitro* conditions were used with the addition of serum (FCS) and/or sDOC to induce increased expression of Cia proteins. No differences were observed when analysing the secretion profiles of the 11168H wild-type strain and the *Cj0248* mutant in the presence of FCS and/or sDOC *in vitro*. The secretion profile data does not suggest any difference in the ability of the *Cj0248* mutant strain to secrete proteins, however analysis of the coomassie stain by eye is insufficient to make this statement. There may be multiple proteins present within a single band, thus proteomics analysis would determine the exact protein composition. Hence, this data suggests the reduced motility and invasion phenotype of

the *Cj0248* mutant are a result of a lack of rotational functionality of the flagella rather than a lack of secretion ability. This would support the view of Watson and Galan who argue that there is little evidence that any component of the flagella apparatus is directly secreting proteins involved in invasion when in contact with IECs and that invasion is a property of being motile via a functional flagella (Watson and Galán, 2008). Further investigation into the secretion profile could be performed by silver-staining to ascertain if any proteins are present or absent from the *Cj0248* mutant secretion profile. In addition, analysis of the secretion profile from *C. jejuni* co-cultured with IECs could also be performed.

Microarray data was analysed to further investigate possible Cj0248 function. CmeR has been characterised as the regulator for the CmeABC efflux pump and *cmeR* expression was shown to be 5.16 fold down-regulated in the *Cj0248* mutant compared to the 11168H wild-type. CmeR regulates genes encoding membrane transporters, proteins involved in C4-dicarboxylate transport and utilization, enzymes for biosynthesis of CPS, and hypothetical proteins with unknown functions (Guo *et al*., 2008). Deletion of *cmeR* or mutation in the CmeR-binding site impedes the repression and results in overexpression of *cmeABC* (Lin *et al*., 2005). The CmeABC efflux pump contributes to *Campylobacter* resistance to various antimicrobial agents and bile compounds present in the intestinal environment (Lin *et al*., 2002, Lin *et al*., 2003). The exact link between Cj0248 and CmeR is currently unknown. Finally, *cbrR* – a *Campylobacter* bile resistance regulator was identified as being 1.40 fold up-regulated. A *cbrR* mutant strain is unable to survive in the presence of bile salts and other detergents and poorly colonizes experimentally inoculated chickens (Raphael *et al*., 2005). Interestingly, both CmeABC and CbrR have been associated with bile stress resistance.

In addition to the above, expression of *cheA* and *cheV* were identified as 1.74 and 1.60 fold down-regulated. Both of these proteins are involved in chemotaxis that specifically link to the flagella (Korolik and Ketley, 2008). A key element in flagella motility is chemotaxis which influences the movement of bacteria toward appropriate environmental and host niches that support ideal bacterial growth and away from components that are less beneficial for growth or harmful to the organism (Hendrixson, 2008). A two-component regulatory system, CheAY controls chemotaxis in response to environmental stimuli. The binding of the signal ligand is relayed by the MCP to CheA, a histidine kinase that forms a complex with the MCP in conjunction with CheW (Chang

and Miller, 2006, Hendrixson *et al*., 2001). CheV is a composite protein that consists of an N-terminal CheW domain fused to a C-terminal response regulator domain (Korolik and Ketley, 2008, Rosario *et al*., 1994). CheA autophosphorylates and subsequently phosphorylates CheY, the response regulator (Young *et al*., 2007). Phosphorylated CheY interacts with FliM of the flagella motor to initiate movement (Yao *et al*., 1997). It is possible Cj0248 may be affecting flagella biosynthesis indirectly by affecting chemotaxis related genes. In addition, *C. jejuni* lacks a homologue of the phosphatase CheZ (which dephosphorylates CheY), but does possess a homologue of the poorly understood protein CheV (Marchant *et al*., 2002, Young *et al*., 2007, Parkhill *et al*., 2000). CheV has an amino terminal CheW-like domain and a carboxyl terminal CheY-like domain that has been hypothesised to act as a phosphate sink for the chemotaxis signal-transduction machinery (Young *et al*., 2007, Marchant *et al*., 2002, Pittman *et al*., 2001). It is believed this may ameliorate the effect of the absence of a CheZ phosphatase on phosphate flow through this signal-transduction pathway (Young *et al*., 2007, Marchant *et al*., 2002). Given that Cj0248 contains a HD motif and the putative function as a possible phosphohydrolase or a signal transduction protein, it is possible Cj0248 has a role relating to chemotaxis or even fulfilling the missing role of CheZ? Further studies would be to investigate the enzymatic function as a phosphohydrolase and also to perform chemotaxis assays comparing the *Cj0248* mutant to the 11168H wild-type strain.

Movement of *C. jejuni* is either based on a tumbling mode induced by clockwise-rotating dissociated flagella or a smooth, straight swimming mode induced by anti-clockwise flagella rotation (Korolik and Ketley, 2008). Tumbling motion is performed to allow reorientation, whereas the straight motion is used by the bacterium to swim to a specific concentration gradient (Korolik and Ketley, 2008). Importantly, in the absence of chemoattractants, autophosphorylation of CheA is inhibited and CheY is not phosphorylated (Korolik and Ketley, 2008). The signal transduction pathway is initiated by the MCP sensory receptors described above. The interaction between CheY and FliM determines the type of movement, as studies in *E. coli* have shown phosphorylated CheY need to occupy at least 70% of the available FliM molecules in the basal body for a change of rotational direction to occur (Bren and Eisenbach, 2001). It is predicted that when CheY binds to FliM in the flagella motor complex, this leads to a clockwise rotation, resulting in tumbling of the cell (Falke *et al*., 1997, Korolik and Ketley, 2008, Bren and Eisenbach, 2001, Spohn and Scarlato, 2001). In contrast, if a chemoattractant is

bound to the MCP, this leads to suppression of CheA activity which in turn leads to less CheY binding to FliM and hence a return to anti-clockwise rotation leading to swimming (Falke *et al*., 1997, Korolik and Ketley, 2008, Bren and Eisenbach, 2001). Expression of *cheA* has been identified as being 1.74 down-regulated and so there may be a direct or indirect role for Cj0248 in terms of the type of movement or even a defective movement phenotype, though suppression of CheA has been noted to lead to swimming as opposed to tumbling of the bacterium (Colegio *et al.*, 2001, Jahreis *et al.*, 2004).

### 5.4 Conclusion

Motility is central to the pathogenesis of *C. jejuni* via active flagella that allow adhesion to and invasion of IECs. This study described the characterisation of a *C. jejuni* 11168H mutant. *Cj0248* was previously implicated in flagella functionality and pathogenesis. In this study the role of Cj0248 in *C. jejuni* pathogenesis was further investigated by performing a number of phenotypic assays such as growth kinetics, motility, interaction and invasion assays to enhance our understanding of this CDS. It is interesting to note that Cj0248 homologues are present in almost all strains and species and even related species. The *Cj0248* mutant produced a severely reduced motility and biofilm formation phenotype compared to the 11168H wild-type strain. The *Cj0248* mutant also exhibited decreased interaction and invasion at specific time points with Caco-2 cells compared to the 11168H wild-type strain. TEM analysis identified the presence of flagella, and with the aid of microarray expression data, TEM was used to identify putative altered flagella basal body structures in the *Cj0248* mutant compared to the 11168H wild-type strain. The *Cj0248* mutant induced a reduced IL-6 response from T84 cells, but not IL-8 when compared to 11168H wild-type strain. This data indicates the 11168H Cj0248 protein is involved in motility functionality within *C. jejuni* affecting the ability to adhere and invade and thus the pathogenicity of the organism. Further studies are required to identify the function of Cj0248 and to add to the understanding of *C. jejuni* pathogenesis.

# Chapter 6: Final discussion

## 6.1 Study objectives

This investigation sought to perform a re-annotation of the *C. jejuni* NCTC11168 genome sequence followed by the selection of CDSs with unknown function with putative roles in virulence, signal transduction or regulation of gene expression. Re-annotation of the NCTC11168 genome sequence resulted in the selection of 15 CDSs for further study. The selection of these 15 CDSs was based on a number of different criteria such as newly identified motifs along with literature searches. Following the successful construction of eight defined isogenic *C. jejuni* 11168H mutants, a number of preliminary phenotypic assays such as growth kinetics, motility, interaction and invasion assays were performed. The *Cj0248* and *Cj1556* mutants displayed the most interesting phenotypic results and were selected for further investigation and characterisation.

## 6.2 Re-annotation of the *C. jejuni* NCTC11168 genome sequence

The first goal of this research project was to perform a full re-annotation of the *C. jejuni* NCTC11168 genome sequence. This was performed in collaboration with the WTSI. The original annotation was completed in 2000 (Parkhill *et al*., 2000) and this genome sequence rapidly became a valuable resource leading to a renewed impetus on *Campylobacter* research. The re-annotation led to 299 (18.2%) of CDS product functions being updated, which is a significant portion of the whole genome. In addition, 1489 (90.0%) of CDSs had additional information added regarding new motifs or literature findings. The re-annotation process identified 20 additional CDSs linked to 'small molecule metabolism', six additional CDSs linked to 'broad regulatory functions', 26 additional CDSs linked to 'cell process' and 77 CDSs linked to 'miscellaneous' functions. The latter function relates to CDSs with new motifs that have not yet been linked to a specific function. Interestingly, 122 CDSs were removed from the 'unknown and other' category mainly due to the assignment of CDSs with new motifs. Arguably the greatest updates have been with the four loci within *C. jejuni* encoding the *N*- and *O*-linked glycosylation systems, CPS and LOS biosynthesis. Over the last decade, the CDSs within these loci have been characterised and this has greatly enhanced our understanding of the pathogenesis for this bacterium. In addition, novel CDSs such as regulators and the further refinement of many CDS product functions have allowed a greater understanding of the genome, mainly due to the availability of new motifs. Thus, the re-annotation has significantly added to the information available for a large majority

of CDSs within the genome. Since the original annotation, substantial updates on CDSs within loci encoding the four main glyco-based virulence determinants have been published. In addition, the databases used in the re-annotation process e.g. PFAM and PROSITE have been greatly enhanced since the original annotation. This was evident during the re-annotation process as supporting information from motif identification was a common theme in novel product function designations. The inclusion of additional search databases such as RFAM and programs such as TMHMM and SIGNALP further increased the usefulness of the re-annotation. The use of research papers for characterised genes/proteins was an additional feature of this re-annotation. Interestingly, from the 2092 literature citations added during this re-annotation, 1056 (50.5%) had been published after the year 2000. Considering there was no literature qualifier in the original annotation, this demonstrates the depth of research that has been carried out since 2000 and further supports the need to make use of this information in the re-annotation.

The re-annotation was completed manually to allow maximum utilisation of current annotation tools and incorporation of additional sources of information not available using automated methods e.g. information from scientific papers. The advantages of carrying out a manual re-annotation outweigh those associated with performing an automated process. However manual re-annotation is a labour intensive task and it is understandable why so many sequenced genomes receive an automated annotation. The resources do not exist for manual annotation of all newly sequenced genomes, let alone re-annotations. With the advent of the next-generation sequencing technology, the number of genome sequences will increase at a far greater rate than before (Margulies *et al*., 2005). Thus, it seems the best approach to annotation or re-annotation in the future will be to improve automated techniques. Priority remains with new sequenced genomes and thus manual re-annotation projects will be performed infrequently.

### 6.2.1 Future studies

The re-annotation of the *C. jejuni* NCTC11168 genome sequence has already become somewhat outdated. A publication accompanying the genome update was released in 2007 (Gundogdu *et al*., 2007) and further research has since been performed not only characterising individual CDSs, but also utilising novel genome analysis techniques such as RNA-Seq (van Vliet, 2010). What has become clear during this project is that with the number of genome sequences currently being generated, the resources simply do not exist for all these new genome sequences to be annotated manually. Manual annotation is

time consuming and costly, however automated techniques do not yet provide the level of detail and accuracy that manual annotation provides. Ideally, an automated annotation or re-annotation tool is required which performs at a level of detail and accuracy that manual annotation provides. If such a tool was available, a system could be in place where automatic annotations and re-annotations could be established at regular time points.

### 6.3 Characterisation of Cj1556

Despite stringent microaerobic growth requirements, *C. jejuni* is ubiquitous in the aerobic environment and so must possess finely tuned regulatory systems to sense and adapt to external stimuli such as oxidative and aerobic ($O_2$) stress. Re-annotation of the *C. jejuni* NCTC11168 genome sequence identified Cj1556 (originally annotated as a hypothetical protein) as a MarR family transcriptional regulator and further bioinformatic analysis indicated a role in regulating the oxidative stress response. A *C. jejuni* 11168H *Cj1556* mutant exhibited increased sensitivity to oxidative and aerobic ($O_2$) stress, decreased ability for intracellular survival in both Caco-2 human IECs and J774A.1 mouse macrophages and a reduction in virulence in the *Galleria mellonella* infection model. Microarray analysis of gene expression changes in the *Cj1556* mutant compared to the 11168H wild-type strain indicated negative autoregulation of *Cj1556* expression and down-regulation of genes associated with oxidative and aerobic ($O_2$) stress responses, such as *katA*, *perR* and *hspR*. *cprS*, which encodes a sensor kinase involved in the regulation of biofilm formation, was also up-regulated in the *Cj1556* mutant and subsequent studies showed that the *Cj1556* mutant had a reduced ability to form biofilms. This study has identified a novel *C. jejuni* transcriptional regulator Cj1556 that is involved in oxidative and aerobic ($O_2$) stress responses and is important for the survival of *C. jejuni* in the natural environment and *in vivo*. This newly identified regulator was designated CosR (Campylobacter oxidative stress Regulator).

#### 6.3.1 Future studies

The characterisation of the 11168H *Cj1556* mutant has shown Cj1556 to be involved in oxidative/aerobic ($O_2$) stress regulation. However, what has not yet been experimentally proven is the ability of Cj1556 to bind DNA as a transcriptional regulator would be expected to. Current studies are investigating not only the binding capability of Cj1556, but also the location of binding. Specifically, this study has shown the expression of Cj1556 to be putatively negatively autoregulated and so Cj1556 would be expected to

bind upstream of *Cj1556*. In addition, microarray studies identified a number of genes involved in oxidative/aerobic ($O_2$) stress with modified expression levels in the *Cj1556* mutant when compared to the 11168H wild-type strain. Cj1556 may also bind upstream of these genes. Gel-mobility shift assays can be used to identify binding of proteins to amplified regions of genomic DNA. In this instance, recombinant Cj1556 can be hybridised to amplified fragments of upstream regions from the genes of interest and run on a native DNA retardation gel to establish whether a band shift occurs. Alternatively, techniques such as ChIP-Seq or ChIP-ChIP whereby the binding capability of Cj1556 is assessed on a genome wide scale can be performed. Additional studies should also be performed on the putative link between Cj1556 and Cj1546. One hypothesis is that Cj1546, found in all *C. jejuni* strains, is involved in the primary line of defence against oxidative/aerobic ($O_2$) stress and that Cj1556, present in only a few *C. jejuni* strains, enhances the ability of these strains to survive these stresses. To test this hypothesis, a *Cj1546* mutant should be constructed in both *C. jejuni* 11168H and 81-176, as these strains possess both Cj1546 and Cj1556 genes. This would identify whether a compensatory role exists between Cj1556 and Cj1546 in the response to aerobic ($O_2$) and/or oxidative stress. In addition, a *Cj1546* mutant should be constructed in a strain lacking Cj1556 such as *C. jejuni* 81116. This would identify whether Cj1546 has a role in aerobic ($O_2$) and/or oxidative stress response. A double *Cj1546*/*Cj1556* mutant should also be constructed in either *C. jejuni* 11168H or 81-176. This would identify whether there is any increased sensitivity to aerobic ($O_2$) and/or oxidative stress. Finally, a functional copy of Cj1556 should be inserted into the *C. jejuni* 81116 genome. This would identify whether there is indeed any increased resistance to aerobic ($O_2$) and/or oxidative stress through the addition of Cj1556 to a Cj1546 background. These mutants can be studied using the assays described in this study to test this hypothesis. MarR family motifs are also linked to a number of additional roles such as controlling virulence factor production and bacterial responses to antibiotics, so a number of different assays such as chemotaxis studies or an array of antibiotic stress assays can be performed to test these further. In addition, recent data has indicated that bile induces expression of the *E. coli marRAB* operon, by binding to the repressor protein MarR and thus preventing binding of MarR to the *marRAB* promoter site (Hamner *et al*., 2010). Further studies will be required to confirm whether bile can bind to the Cj1556 protein and thus prevent binding to the *Cj1556* promoter site, resulting in the up-regulation of Cj1556 in the presence of bile observed previously (Malik-Kale *et al*., 2008).

It is possible that Cj1556 could be used as a target candidate for the strategic control of *C. jejuni* colonisation in the poultry industry. Blocking the function of Cj1556 would result in increased susceptibility to oxidative and aerobic ($O_2$) stresses, making the possibility of an interesting control strategy if combined with spraying broiler houses with hydrogen peroxide.

### 6.4 Characterisation of Cj0248

The ability of *C. jejuni* to adhere and invade to human IECs is critical in the pathogenesis of this ubiquitous bacterium. Central to the survival of *C. jejuni* is the ability to become motile via active flagella that allow adhesion to and invasion of IECs. The flagellum has also been implicated as secretory machinery allowing proteins to be secreted that induce invasion into IECs. Cj0248 has previously been implicated in flagella functionality and pathogenesis. In this study we further investigated the phenotype of a 11168H *Cj0248* mutant by performing an array of assays designed to enhance our understanding of this CDS. We confirmed previous findings that mutation of *Cj0248* gives a severely reduced motility phenotype using a different *C. jejuni* wild-type strain. The 11168H *Cj0248* mutant also exhibited reduced interaction and invasion when co-cultured with Caco-2 cells compared to the 11168H wild-type strain. However, TEM analysis identified the presence of flagella. Using microarray expression data, TEM was used to identify an altered flagella basal body region in the *Cj0248* mutant compared to the 11168H wild-type strain. Subsequent assays demonstrated the *Cj0248* mutant had a reduced ability to form biofilms compared to the 11168H wild-type strain and also that the *Cj0248* mutant induced a reduced IL-6 response from T84 cells, but not IL-8 when compared to 11168H wild-type strain. Secretion profile analysis identified no differences in the protein profile of the *Cj0248* mutant compared to the 11168H wild-type strain. The function of Cj0248 remains unknown.

#### 6.4.1 Future studies

Further investigation is required to ascertain the function of Cj0248. The phenotypes observed in this study must be confirmed by performing assays on the recently constructed complemented mutant strain. So far this complemented mutant strain has successfully demonstrated the partial resuscitation of the motility phenotype. In addition, the functionality of Cj0248 as a putative phosphohydrolase needs to be confirmed enzymatically. Analysis of the secretion profile of the 11168H *Cj0248* mutant should be expanded using silver staining and after co-culture with IECs. It is possible that

differences in the secretion profile between *Cj0248* mutant and the 11168H wild-type strain may be observed when co-culturing *C. jejuni* in the presence of conditioned media for different time periods. To ascertain the exact composition of the protein profile, proteomics analysis should be performed on the bands from the coomassie or silver stained gel. The location of Cj0248 within *C. jejuni* should also be studied by using monoclonal or polyclonal antibodies specific for Cj0248 on different bacterial cell fractions. This would identify whether Cj0248 is indeed located around the bacterial membrane region in proximity of the flagella. Enhanced TEM results identified a putative altered flagella basal body region when comparing *Cj0248* mutant to the 11168H wild-type strain. Based on this data, further analysis specifically of the flagella apparatus would be required to ascertain if Cj0248 directly interferes with basal body formation. A starting point would be to perform enhanced TEM on the complemented mutant strain to analyse whether the flagella basal body is structurally similar to the 11168H wild-type strain. Further experiments should be performed on the 11168H wild-type strain assessing the role of Cj0248 using monoclonal or polyclonal antibodies against this protein. Using microarray data, a number of putative links were made between Cj0248 function and chemotaxis. In particular, chemotaxis is heavily linked to flagella functionality so there may be a direct role for Cj0248 relating to chemotaxis. Investigation of the role of Cj0248 could be performed using monoclonal or polyclonal antibodies against Cj0248 on the 11168H wild-type strain in conjuction with assays using chemotactic attractants (e.g. mucin) as a possible inducer for the increased expression of Cj0248.

## Appendices

### Appendix 1 – Products used in this study

dATP – Promega, Southampton, UK
dGTP – Promega, Southampton, UK
dCTP – Promega, Southampton, UK
dTTP – Promega, Southampton, UK
$MgCl_2$ - Promega, Southampton, UK
10X Buffer
*Taq* DNA polymerase M166F (2,500 units) Promega, Southampton, UK

| | |
|---|---|
| ABI Prism Terminator Ready Reaction Mix | - Applied Biosystems, Warrington, UK |
| Agarose | - Sigma-Aldrich, Poole, UK |
| Ampicillin | - Sigma-Aldrich, Poole, UK |
| Bacto agar | - BD, Oxford, UK |
| Bovine Serum Albumin (BSA) | - Sigma-Aldrich, Poole, UK |
| Brucella broth | - Fluka, Gillingham, UK |
| Campylobacter Selective Supplement (Skirrow) | - Oxoid, Basingstoke, UK |
| Columbia Agar | - Fluka, Gillingham, UK |
| Dulbecco's Modified Eagle's Medium (DMEM) | - Sigma-Aldrich, Poole, UK |
| Ethidium Bromide | - Promega, Southampton, UK |
| Foetal Bovine Serum | - Sigma-Aldrich, Poole, UK |
| GenElute Gel Extraction Kit | - Sigma-Aldrich, Poole, UK |
| Glycerol | - VWR, Lutterworth, UK |
| Horse blood in Alsevers | - TCS Biosciences, Botolph Claydon, UK |
| 1 Kb Ladder | - Invitrogen, Paisley, UK |
| Kanamycin | - Gibco/Invitrogen, Paisley, UK |
| Luria Bertani Agar (LB agar) | - Merck, Hoddesdon, UK |
| Luria Bertani Broth (LB agar) | - Difco, Northampton, UK |
| Lysozyme | - Sigma-Aldrich, Poole, UK |
| 2 Mercaptoethanol | - Stratagene, Amsterdam, Holland |
| MinElute PCR purification Kit | - Qiagen, Crawley, UK |

| | |
|---|---|
| Mueller-Hinton Broth (MH) | - Oxoid, Basingstoke, UK |
| Non-essential amino acids | - Sigma-Aldrich, Poole, UK |
| Oligonuclotide primers | - Sigma-Aldrich, Poole, UK |
| Penicillin-Streptomycin solution | - Sigma-Aldrich, Poole, UK |
| Phosphate Buffered Saline | - Sigma-Aldrich, Poole, UK |
| pGEM-T Easy vector | - Promega, Southampton, UK |
| Restriction endonucleases | - Promega, Southampton, UK,<br>- New England Biolabs, Hitchin, UK<br>- Roche, Lewes, UK |
| SOC Broth | - Fluka, Gillingham, UK |
| SCS110 Competent Cells | - Stratagene, Amsterdam, Holland |
| T4 DNA Ligase | - Promega, Southampton, UK |
| T4 DNA Ligase buffer | - Promega, Southampton, UK |
| Triton X-100 | - Sigma-Aldrich, Poole, UK |
| 0.25 % Trypsin-EDTA | - Sigma-Aldrich, Poole, UK |
| QIAquick PCR purification kit | - Qiagen, Crawley, UK |
| QIAquick gel extraction kit | - Qiagen, Crawley, UK |
| QIAprep Spin Miniprep Kit | - Qiagen, Crawley, UK |
| Puregene DNA Purification System | - Gentra Systems Inc. - Flowgene, Lichfield, UK |
| XL-2 Blue MRF' Competent cells | - Stratagene, Amsterdam, Holland |

**Appendix 2 – Media used in this study**

*Columbia blood agar plates (to make 8-10 plates)**

Dissolve 12.6 g Columbia agar in 279 ml Milli-RO water. Agar was heated in a microwave to a molten level and allowed to cool to approximately 50°C. 7% Horse blood in Alsevers (21 ml), *Campylobacter* Selective Supplement (Skirrow) and any antibiotics were added when ready to pour.

*LB agar (to make 8-10 plates)**

Dissolve 9.25 g LB agar in 250 ml Milli-RO water. Agar was heated in a microwave to a molten level and allowed to cool to approximately 50°C. Antibiotics were added when ready to pour.

*LB agar/Ampicillin/IPTG/X-Gal*

| | |
|---|---|
| LB agar | - 250 ml |
| Ampicillin (100 mg/ml) | - 250 µl |
| Isopropyl β-D-1-thiogalactopyranoside - IPTG (1 pmol/µl) | - 40 µl |
| 2% (w/v) X-GAL** | - 500 µl |

** 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

*Motility agar**

Dissolve 1.0 g Bacto agar in 250 ml Brucella broth to obtain 0.4% semi-solid media.

*Overnight Cultures (10 ml)**

- *E. coli*

Dissolve 6.25 g LB broth in 250 ml Milli-RO water. For overnight cultures, use 10 ml LB broth with appropriate antibiotics.

- *C. jejuni*

Dissolve 7.0 g Brucella broth in 250 ml Milli-RO water. For overnight cultures, use 10 ml Brucella broth with appropriate antibiotics.

*Glycerol stock solution**

- *E. coli*

Glycerol stocks should be made from overnight cultures (described above). 1.5 ml 100% (v/v) glycerol was added into the 10 ml overnight culture. Mix and then prepare 500 µl aliquots into 1.5 ml microcentrifuge tubes. Snap freeze and store at -80°C.

- *C. jejuni*

Glycerol stocks should be made from overnight plates (described above). Prepare the *C. jejuni* glycerol solution in a 10 ml Universal as:-

| | |
|---|---|
| MH broth | - 8 ml |
| 100% (v/v) FCS | - 1 ml |
| 100% (v/v) Glycerol | - 1 ml |

*50X TAE (to make 1 litre)**

| | |
|---|---|
| Tris Base | - 242 g |
| Acetic acid | - 57.1 ml |
| EDTA (0.5 M adjusted to pH 8.0) | - 100 ml |
| Milli-RO water | - up to 1 litre |

*1X TAE (to make 20 ml)**

| | |
|---|---|
| EDTA 0.5 M pH8.0 | - 40 µl |
| Tris Base 1 M pH8.0 | - 200 µl |
| Milli-RO water | - 19.7 ml |

*Agarose gel (300 ml)*

Dissolve 3.15 g agarose in 450 ml Milli-RO water. Agar was heated in a microwave to a molten level and allowed to cool to approximately 50°C. Add 0.5 µg/µl Ethidium Bromide (22.5 µl) into 450 ml when ready to pour. Mix bottle thoroughly before pouring into gel tray.

*Gel Loading Buffer (20 ml)*

| | |
|---|---|
| 100% (v/v) Glycerol | - 10 ml |
| 20% (v/v) SDS | - 500 µl |
| 5% (w/v) Bromothenol blue | - 100 µl |
| EDTA (0.5 M) | - 100 µl |
| Sterilised Milli-Q water | - up to 20 ml |

*dNTP 1.25 mM Stock (to make 1 ml total volume)*

| | |
|---|---|
| dATP (100mM) | - 12.5 µl |
| dTTP (100mM) | - 12.5 µl |
| dGTP (100mM) | - 12.5 µl |
| dCTP (100mM) | - 12.5 µl |
| Sterilised Milli-Q water | - 950 µl |

*PCR pre-mix (to make 1 ml total volume)*

| | |
|---|---|
| 10X Buffer | - 100 µl |
| dNTPs (1.25mM) | - 168 µl |
| $MgCl_2$ (25 mM) | - 60 µl |
| *Taq* polymerase (5 U/µl) | - 6 µl |
| Sterilised Milli-Q water | - up to 1000 µl |

*EBF buffer (100 ml)*

| | |
|---|---|
| 100% (v/v) Glycerol | - 15 ml |
| 10% (w/v) Sucrose | - 10 ml |
| Sterilised Milli-Q water | - up to 100 ml |

*SOC media components*

Tryptone (20 g/l)
Yeast extract (5 g/l)
Sodium Chloride (10 nM)
Potassium Chloride (2.5 mM)
Magnesium Chloride (10 mM)
Magnesium Sulphate (10 mM)
Glucose (20mM)
Final pH 7.0

*Coomassie staining solution 0.125% (w/v)*

| | |
|---|---|
| Coomassie | - 0.25 g |
| 100% (v/v) Methanol | - 150 ml |
| 100% (v/v) Acetic acid | - 50 ml |
| Sterilised Milli-Q water | - up to 500 ml |

*De-stain solution (500 ml)*

| | |
|---|---|
| 100% (v/v) Methanol | - 150 ml |
| 100% (v/v) Acetic acid | - 50 ml |
| Sterilised Milli-Q water | - up to 500 ml |

*10X running buffer**

| | |
|---|---|
| Tris base | - 30 g |
| D-glycine | - 144 g |
| Sterilised MilliQ water | - up to 1 litre |
| Adjust pH to 8.3 | |

*1X running buffer*

| | |
|---|---|
| 10X running buffer | - 100 ml |
| 10% (v/v) SDS | - 10 ml |
| Sterilised MilliQ water | - up to 1000 ml |

*1X Transfer buffer**

| | |
|---|---|
| Tris base | - 3.03 g |
| Glycine | - 14.4 g |
| 100% (v/v) Methanol | - 200 ml |
| Sterilised MilliQ water | - up to 1 litre |

*Equilibration Buffer*

| | |
|---|---|
| Sodium phosphate (1 M) | - 50 ml |
| Sodium chloride (1 M) | - 300 ml |
| Sterilised MilliQ water | - up to 1 litre |
| Adjust to pH 8.0 | |

*Wash Buffer*

| | |
|---|---|
| Sodium phosphate (1 M) | - 50 ml |
| Sodium chloride (1 M) | - 300 ml |
| Imidazole (1 M) | - 5 ml |
| Sterilised MilliQ water | - up to 1 litre |
| Adjust to pH 8.0 | |

*Elution Buffer*

| | |
|---|---|
| Sodium phosphate (1 M) | - 50 ml |
| Sodium chloride (1 M) | - 300 ml |
| Imidazole (1 M) | - 250 ml |
| Sterilised MilliQ water | - up to 1 litre |
| Adjust to pH 8.0 | |

*300 mM sodium chloride*

| | |
|---|---|
| Sodium chloride | - 17.5 g |
| Sterilised MilliQ water | - up to 1 litre |

*30 mM sodium phosphate*

| | |
|---|---|
| Sodium chloride | - 4.89 g |
| Sterilised MilliQ water | - up to 1 litre |

*5 mM imidazole*

| | |
|---|---|
| Sodium chloride | - 0.34 g |
| Sterilised MilliQ water | - up to 1 litre |

*250 mM imidazole*

| | |
|---|---|
| Sodium chloride | - 1 g |
| Sterilised MilliQ water | - up to 1 litre |

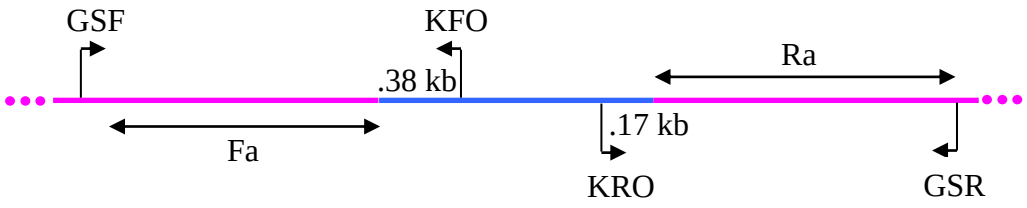
*Cell lysis solution (to make 1 ml)*

| | |
|---|---|
| EDTA (0.5 M) | - 20 µl |
| Tris Base (1 M pH8.0 ) | - 10 µl |
| 10% (v/v) Triton X-100 | - 2 µl |
| Lysozyme (20 mg/ml) | - 50 µl |
| Sterilised Milli-Q water | - to 1000 µl |

* Denotes solutions/buffers that were sterilised by autoclaving prior to use.

**Appendix 3 – Different orientations of the $Km^R$ cassette**

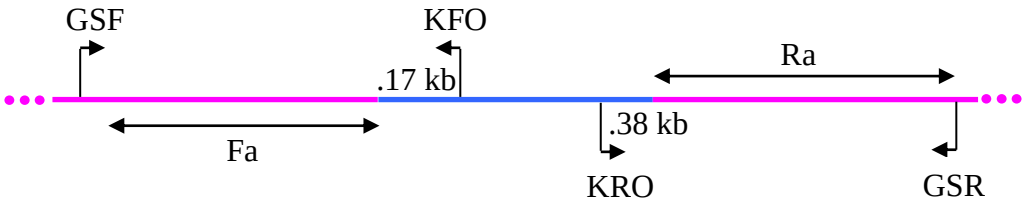
a) Correct Orientation




GSF + KFO = Fa + .38 kb

GSR + KRO = Ra + .17 kb

b) Reverse Orientation




GSF + KRO = Fa + .17 kb

GSR + KFO = Ra + .38 kb

Oligonucleotide Primer

Gene Insert

$Km^R$ cassette

GSF = Gene Specific Forward

GSR = Gene Specific Reverse

KFO = $Km^R$ forward-out

KRO = $Km^R$ reverse-out

FA = Forward gene size

RA = Reverse gene size

**Appendix 4 – Km$^R$ cassette calculations**

| ISF = Insert specific forward primer | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ISR = Insert specific reverse primer | | | | | | | | |
| KFO = Km$^R$ forward out primer | | | | | | | | |
| KRO = Km$^R$ reverse out primer | | | | | | | | |
| FA = Forward CDS size | | | | | | | | |
| RA = Reverse CDS size | | | | | | | | |
| | | | | | | | | |
| Correct orientation | | Reverse orientation | | | | | | |
| | | | | | | | | |
| GSF + KFO = Fa + .381 kb | | GSF + KFO = Fa + .173 kb | | | | | | |
| GSR + KRO = Ra + .173 kb | | GSR + KRO = Ra + .381 kb | | | | | | |
| | | | | | | | | |
| Kanamycin size = 1.43 kb | | | | | | | | |
| ( ) indicate calculation using primers | | | | | | | | |
| **CDS** | **CDS Size (kb)** | **CDS+Kan** | **Fa (kb)** | **ISF + KFO (kb)** | **ISF + KRO (kb)** | **Ra (kb)** | **ISR + KRO (kb)** | **ISR + KFO (kb)** |
| *Cj1556* | 1.494 (0.278) | 2.92 (1.708) | 0.917 | 1.298 | 1.030 | 0.571 | 0.744 | 0.952 |
| *Cj0248* | 0.858 (0.838) | 2.284 (2.264) | 0.429 | 0.810 | 0.602 | 0.429 | 0.602 | 0.810 |

## Appendix 5 – Re-annotation manuscript

BMC Genomics

BioMed Central

Research article

Open Access

# Re-annotation and re-analysis of the *Campylobacter jejuni* NCTC11168 genome sequence

Ozan Gundogdu[1], Stephen D Bentley[2], Matt T Holden[2], Julian Parkhill[2], Nick Dorrell[1] and Brendan W Wren*[1]

Address: [1]Pathogen Molecular Department, London School of Hygiene & Tropical Medicine, Keppel Street, UK and [2]Pathogen Sequencing Unit, Sanger Institute, UK

Email: Ozan Gundogdu - Ozan.gundogdu@lshtm.ac.uk; Stephen D Bentley - sdb@sanger.ac.uk; Matt T Holden - mh3@sanger.ac.uk; Julian Parkhill - parkhill@sanger.ac.uk; Nick Dorrell - Nick.dorrell@lshtm.ac.uk; Brendan W Wren* - Brendan.wren@lshtm.ac.uk

* Corresponding author

Published: 12 June 2007

*BMC Genomics* 2007, 8:162 doi:10.1186/1471-2164-8-162

Received: 16 January 2007
Accepted: 12 June 2007

This article is available from: http://www.biomedcentral.com/1471-2164/8/162



## Abstract

**Background:** *Campylobacter jejuni* is the leading bacterial cause of human gastroenteritis in the developed world. To improve our understanding of this important human pathogen, the *C. jejuni* NCTC11168 genome was sequenced and published in 2000. The original annotation was a milestone in *Campylobacter* research, but is outdated. We now describe the complete re-annotation and re-analysis of the *C. jejuni* NCTC11168 genome using current database information, novel tools and annotation techniques not used during the original annotation.

**Results:** Re-annotation was carried out using sequence database searches such as FASTA, along with programs such as TMHMM for additional support. The re-annotation also utilises sequence data from additional *Campylobacter* strains and species not available during the original annotation. Re-annotation was accompanied by a full literature search that was incorporated into the updated EMBL file [EMBL: AL111168]. The *C. jejuni* NCTC11168 re-annotation reduced the total number of coding sequences from 1654 to 1643, of which 90.0% have additional information regarding the identification of new motifs and/or relevant literature. Re-annotation has led to 18.2% of coding sequence product functions being revised.

**Conclusions:** Major updates were made to genes involved in the biosynthesis of important surface structures such as lipooligosaccharide, capsule and both *O*- and *N*-linked glycosylation. This re-annotation will be a key resource for *Campylobacter* research and will also provide a prototype for the re-annotation and re-interpretation of other bacterial genomes.

## Background

*Campylobacter jejuni* is the leading bacterial cause of human gastroenteritis in the developed world [1]. *C. jejuni* infection has also been associated with post-infection sequelae including septicaemia and neuropathies such as Guillain-Barré Syndrome (GBS) [2]. Infection has largely been linked with the consumption of contaminated poultry or meat products. Given the socioeconomic importance of this pathogen, it is surprising that the ecology, the epidemiology and, in particular, the pathogenesis are still so poorly understood [3]. The lack of information on this problematic pathogen was one of the main driving

forces for the original *C. jejuni* NCTC11168 genome project published in 2000 [4], and equally is why a re-annotation and re-analysis of the genome is required.

Since the publication of the *C. jejuni* NCTC11168 genome sequence in 2000, there has been a spectacular increase in research on this important human pathogen. One result of this has been significant revisions of the genetic loci that code for important surface structures on *C. jejuni* strains. The surface polysaccharide region has since been identified as a capsule locus (*Cj1413c – Cj1448c*) [5-7]. The flagellar modification locus has been identified as an *O*-linked glycosylation pathway (*Cj1293 – Cj1342*) [8-11]. Progress has also been made in our understanding of the lipooligosaccharide (LOS) locus. In addition, the *N*-linked glycosylation pathway has been identified in *C. jejuni* (*Cj1119 – Cj1130*) [9,12-14]. This *N*-linked general glycosylation system was initially thought to only be present in eukaryotes. To date, up to 30 proteins modified with the same heptasaccharide glycan structure have been identified. Research over the last 7 years on *C. jejuni*, coupled with the publication of a further 2 *C. jejuni* genome sequences [15,16] and another 3 *Campylobacter* species [15], has heightened the need for re-analysis of the original NCTC11168 genome sequence.

Re-annotation is defined as the process of annotating a previously annotated genome [17]. Examples of re-annotated genomes are unfortunately rare compared to the number of sequenced genomes [18,19]. Clearly the ever-increasing number of new genome sequences requires prioritisation from annotators. Automated methods can save time and resources, but will not incorporate the maximum information available from expert curators, leading to incomplete or even false designations. By contrast, manual annotation is costly and time consuming. However, manual re-annotation of genomes can significantly reduce the perpetuation of errors and thus reduce the time spent on flawed research. Outdated annotations can lead to significant gaps in our knowledge. Hence, there is a need for a research community-wide review and regular update of genome interpretations. Here we have shown the importance of genome re-annotation in terms of maintaining and increasing the usefulness of this resource, a number of years after the original genome sequencing project was completed.

In this study, we describe the re-annotation and re-analysis of the *C. jejuni* NCTC11168 genome. Manual re-annotation of all coding sequences (CDSs) was carried out using current annotation techniques. Literature searches, updates to genome structure and additional unique genome searches were carried out to produce the most comprehensive annotation of any *Campylobacter* genome to date. The re-annotation of the *C. jejuni* NCTC11168 genome also represents a useful model for the re-evaluation of other bacterial genomes.

## Results & Discussion

### *Gene number adjustment*

A complete re-annotation of the *C. jejuni* NCTC11168 genome was performed resulting in the reduction of the total number of CDSs from 1654 to 1643. This reduction was due to the merging of adjacent CDSs or the removal of CDSs. Three CDSs originally designated as pseudogenes were removed as a result of merging with adjacent pseudogenes. CDSs designated as pseudogenes were also updated to reflect the complete amino acid sequence for the encoded protein regardless of expression. Phase-variable CDSs that contained an intersecting homopolymeric region between adjacent CDSs on separate frames were merged. This allowed the complete amino acid sequence for appropriate genes to be obtained regardless of phase. Re-interpretation of phase-variable CDSs resulted in removal of seven CDSs. CDS (*Cj1520*) was removed because of the recently discovered CRISPR structural moieties [20] (See Structural modifications section in Results & Discussion). In total, 11 CDSs were removed from the re-annotated sequence (Table 1). The accurate identification of all CDSs within the genome has implications for downstream applications, such as mutagenesis, microarray design and proteome analysis.

### *Functional annotation update*

A systematic re-annotation of all CDSs was performed. For the purpose of this re-annotation, all CDSs with additional information have had an 'updated' note qualifier attached. This qualifier contains consistent free-hand descriptions on recently identified motifs, relevant similarity search results and any characterisation work carried out within *Campylobacter* species/strains or any orthologs in similar microorganisms. Additionally, the 'updated' note qualifier also contains reasoning for including 'putative' or not within the product function. Putative designations infer an accepted product function without definitive evidence. For each CDS, a full literature search was performed. In total, 64.5% of CDSs have had one or more literature qualifier added. Interestingly, from all the literature added (2092), 50.5% have been published after the year 2000. Considering there was no literature qualifier in the original annotation, this data illustrates the depth of research that has been carried out since 2000 and further supports the need to make use of this information in a re-annotation. Detailed statistics on genome modifications are given in Table 2. 18.2% of CDSs have had their product functions updated. 60.5% of CDSs with new product function have been designated with a different functional classification. Additional file 1 gives the outline of functional classification used in this annotation. This description was adopted from the Sanger Institute. A

**Table 1: CDSs removed or merged from *C. jejuni* NCTC11168 re-annotation.**

| Gene Number | Type/Description |
|---|---|
| Cj0290c/Cj0291c/Cj0292c | Pseudogenes |
| Cj0968/Cj0969 | Pseudogenes |
| Cj0031/Cj0032 | Phase-variable |
| Cj0170/Cj0171 | Phase-variable |
| Cj0628/Cj0629 | Phase-variable |
| Cj1144c/Cj1145c | Phase-variable |
| Cj1325/Cj1326 | Phase-variable |
| Cj1335/Cj1336 | Phase-variable |
| Cj1677/Cj1678 | Phase-variable |
| Cj1520 | CRISPR region identified |

different functional classification may still be within the same field as the previous function, or may be in a completely new area. 97.8% of these CDSs with a new product function and a different functional classification were given a completely new type of functional classification. Additional file's 2, 3 and 4 give in-depth data on the change and distribution of CDSs within these functional categories. Importantly, the number of CDSs in the 'Unknown and other' category has been reduced by 122. Also, the number of CDSs in the 'Miscellaneous' category has risen by 77. This is attributed to the fact that a number of CDSs have new information relating to a product function from uncharacterised motifs and thus the CDSs were not placed into a specific category as yet.

Since the original annotation, significant new information has been derived on the genetic loci encoding the four main carbohydrate surface structures. The C. jejuni N-linked glycosylation pathway (not described in the original annotation), has been fully characterised [9,12-14]. This re-annotation includes the nomenclature for the pglA-K (protein glycosylation) genes and has updated all product functions for genes Cj1119c – Cj1130c. The LOS locus (Cj1131c – Cj1152c) described in the original annotation was updated to include recent product functions and gene names including neuA1, B1, C1 and hldDE [21-24]. The O-linked glycosylation loci (Cj1293 – Cj1342) involved in flagellar glycosylation, has been updated to include neu, pse and maf genes [8-11]. Finally, the capsule locus (Cj1413c – Cj1448c), has now been updated to include kps and hdd genes [5-7].

Additional genome-wide updates were also carried out, of which a large proportion entailed adding specificity to existing product function. For example, the identification of a new PFAM or PROSITE motif has allowed the product function to become further specified e.g. putative transport protein modified to putative MFS (Major Facilitator System) transport protein. A complete list of changes throughout the *C. jejuni* NCTC11168 genome is provided in Additional File 5.

***Pseudogene & phase-variable modification***

Pseudogene identification is a challenging process where discrepancies exist between pseudogene assignment techniques [25]. Identifiers include detection of Open Reading Frames (ORF) belonging to a single CDS on multiple frames, the presence of one or more stop codon within a CDS, and extra information from the biology of the microorganism. More recently, comparative genomics has been used as a technique for pseudogene assignment [26]. The number of pseudogenes identified in the original

**Table 2: Genome wide statistics from *C. jejuni* NCTC11168 re-annotation.**

| | |
|---|---|
| CDSs with new motifs identified | +44.9% |
| CDSs with new literature qualifier | +64.5% |
| CDSs with new updated note qualifier | +90.0% |
| CDSs with hypothetical designations in product function | -7.5% |
| CDSs with conserved hypothetical in product function | +3.7% |
| CDSs with putative designation in product function | +5.9% |
| CDSs with new gene qualifier | +6.3% |
| CDSs with new product function | +18.2% |
| CDSs with new product function and with a new functional classification | +60.5% |
| CDS with new product function and a new functional classification with a different type of function | +97.8% |

annotation of the *C. jejuni* NCTC11168 genome was 20. We carried out a re-analysis on all pseudogenes in the NCTC11168 genome. The majority of revisions we carried out incorporated multiple features created from different coordinates on more than one frame. This process is often complicated with support needed from FASTA and TBLASTX search results. Completion of this re-analysis resulted in modification of 19 out of 20 pseudogenes (Table 3). The final pseudogene number was 19 due to the merging of two adjacent CDSs designated as pseudogenes (*Cj0968*/*Cj0969*).

An example of the difficulty and complexity associated with pseudogene designation is observed when viewing the CDSs *Cj0522*, *Cj0523* and *Cj0524* within *C. jejuni* NCTC11168. These three CDSs are represented as one whole CDS on a single frame within *C. jejuni* RM1221 (*Cje0628*). The three CDSs are large enough to be represented as individual CDSs and in *C. jejuni* NCTC11168 have been represented on more than one frame. The question can be asked as to whether these CDSs (which are intact in *C. jejuni* RM1221), represent a pseudogene in *C. jejuni* NCTC11168. Given the fact that in *C. jejuni* RM1221 these three CDSs do actually code for a product (Na/Pi-cotransporter, putative), it is more likely that they represent a pseudogene in *C. jejuni* NCTC11168. In this re-annotation, our intention was to carry out a full mark up of existing pseudogenes, however, the potential for a pseudogene has been noted.

The frequency and importance of pseudogene formation in microorganisms has attained added significance in recent years with the emergence of genome reduction theories and enhanced virulence through pathoadaptive mutations [27,28]. Recent studies have suggested that ever increasing non-functional genes are being identified within microorganisms and in particular are more common in genomes of recently evolved pathogens, than in their benign or free-living relatives [25]. The number and type of predicted pseudogenes within *C. jejuni* NCTC11168 and *C. jejuni* RM1221 are compared in Additional File 6. Observing CDS location rather than CDS function was carried out for this comparison. This was to ensure variation in product function naming does not exclude identical pseudogenes, which are represented on the same row. Currently, the *C. jejuni* 81–176 genome has not been fully annotated so could not be used in this comparison. This is also the case for *C. coli* RM2228, *C. lari* RM2100 and *C. upsaliensis* 3195 which only have an estimation of pseudogene numbers based on a subset of genes [15]. In *C. jejuni* NCTC11168, 63% (12/19) of the pseudogenes are shared with *C. jejuni* RM1221. In contrast to 19 pseudogenes in *C. jejuni* NCTC11168, *C. jejuni* RM1221 contains 47 pseudogenes. Assuming these are genuine pseudogenes this would imply *C. jejuni* NCTC11168 (1980 human isolate, UK) and *C. jejuni* RM1221 (2000 chicken isolate, US), share a core set of ancestral pseudogenes. Even with the variation of isolation dates, source and geographical location, there is substantial conservation of pseudogene type. It is speculative to suggest when and how the additional pseudogenes in *C. jejuni* RM1221 arose, or when and how the *C. jejuni* NCTC11168 genome lost CDSs as pseudogenes since divergence occurred.

**Table 3: Pseudogenes in *C. jejuni* NCTC11168 with modification.**

| Gene Number | Product | Modification |
|---|---|---|
| Cj0046 | pseudogene (putative sodium:sulfate transmembrane transport protein) | Features introduced within CDS |
| Cj0072c | pseudogene (putative iron-binding protein) | Features introduced within CDS |
| Cj0223 | pseudogene (putative IgA protease family protein) | Features introduced within CDS |
| Cj0292c | pseudogene (putative glycerol-3-phosphate transporter) | Merging of multiple CDS |
| Cj0444 | pseudogene (putative TonB-denpendent outer membrane receptor) | Features introduced within CDS |
| Cj0501 | pseudogene (ammonium transporter) | Features introduced within CDS |
| Cj0565 | pseudogene (conserved hypothetical protein) | Features introduced within CDS |
| Cj0654c | pseudogene (putative transmembrane transport protein) | Features introduced within CDS |
| Cj0676 | pseudogene (potassium-transporting ATPase A chain) | Features introduced within CDS |
| Cj0678 | pseudogene (potassium-transporting ATPase C chain) | Features introduced within CDS |
| Cj0742 | pseudogene (putative outer membrane protein) | Features introduced within CDS |
| Cj0752 | pseudogene (IS element transposase) | Features introduced within CDS |
| Cj0866 | pseudogene (arylsulfatase) | Features introduced within CDS |
| Cj0969 | pseudogene (putative periplasmic protein) | Merging of multiple CDS |
| Cj1064 | pseudogene (nitroreductase) | Features introduced within CDS |
| Cj1389 | pseudogene (putative C4-dicarboxylate anaerobic carrier) | Features introduced within CDS |
| Cj1395 | pseudogene (putative MmgE/PrpD family protein) | Unmodified |
| Cj1470c | pesudogene (type II protein secretion system F protein) | Features introduced within CDS |
| Cj1528 | pseudogene (putative C4-dicarboxylate anaerobic carrier) | Features introduced within CDS |

The significance of pseudogenes in early genome annotations were frequently ignored, as these were considered as sequencing artefacts. However, given the recent realisation of the importance of pseudogenes in pathoadaptive mutations, a greater significance is placed on their identification [26,29]. An example of this is the re-analysis of *Escherichia coli* K-12, which has predicted an additional 161 from the original single pseudogene identified [28]. The same study also indicated pseudogenes are continually generated, with existing pseudogenes being eliminated over a period of time [28]. Pseudogenes can accumulate in the genomes of some bacterial species, especially those undergoing processes like niche adaptation, host specialisation or weak selection strength [30]. Analysis of further *Campylobacter* strains and species along with additional epsilon proteobacteria species will aid our understanding on this emerging area of interest. Also, greater understanding of pseudogene dynamics and in particular innovative pseudogene identification techniques will yield more information about the actual number and purpose of these entities within microorganisms.

Phase-variable CDSs containing hypervariable regions were also analysed. The initial annotation gave a number of hypervariable sequences found within the *C. jejuni* genomic shotgun sequence [4]. These hypervariable sequences are scattered throughout the genome, however, there is a large cluster within both the *O*-linked glycosylation and capsule loci. Further research on these loci have illustrated the impact of phase-variation on microorganism pathogenicity [11,31,32]. Table 4 shows phase-variable CDSs that have been modified.

### *Structural modifications*

As well as CDS updates, novel features were also added to the re-annotation. For example, the incorporation of the recently identified Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) regions within *Campylobacter* [20,33,34]. CRISPR regions are thought to be mobile elements. In conjunction with this, the three CDSs upstream of the CRISPR repeats were identified as CRISPR associated proteins and this concurs with existing CRISPR structures. To date, there has only been one identified CRISPR region within *C. jejuni* and this has now been incorporated within the genome. As a result, one CDS (*Cj1520*) has been removed. This CDS was previously annotated as having five repeat regions. Thus, the genome now contains a CRISPR repeat region in place of the removed CDS.

Additional genome searches included RFAM database search to discover any non-coding RNAs. This search identified two new non-coding RNA structures. RFAM RF00169, a bacterial signal recognition particle (SRP) RNA, was identified upstream of *Cj0046*. The SRP is a universally conserved ribonucleoprotein involved in the co-translational targeting of proteins to membranes [35,36]. Also, RFAM RF00059 a thiamin pyrophosphate (TPP) riboswitch (THI element) was identified upstream of *Cj0453* (thiamin biosynthesis protein ThiC). The RFAM motif is a conserved structure (THI element), involved in thiamin-regulation [37].

The final step of the re-annotation process was the incorporation of Gene Ontology (GO) annotation. GO annotation attempts to link three structured, controlled vocabularies (ontologies) that describe gene products in terms of their associated biological processes, cellular components and molecular functions in a species-independent manner [38]. This re-annotation aims to incorporate GO annotation using two automated methods. Upon submission to EBI, the EMBL file will carry a GOA link that lists the GO annotation identified automatically by EBI. A second version of GO was generated by performing a reciprocal FASTA search using RM1221. This was created and submitted to GeneDB. GO annotation is a valuable feature in current annotation techniques that can expedite systems biology approaches to genome analysis.

## Conclusions

In summary, the re-annotation and re-analysis of the *C. jejuni* NCTC11168 genome sequence has led to substantial updates across the entire genome, incorporating a vast amount of research information performed since the original annotation in 2000 and also integrating data from

**Table 4: Merged phase-variable CDSs in *C. jejuni* NCTC11168 re-annotation.**

| Gene Number | Product | Effect |
|---|---|---|
| Cj0031/Cj0032 | putative type IIS restriction/modification enzyme | Fusion/separation |
| Cj0170/Cj0171 | hypothetical protein Cj0170 | Fusion/separation |
| Cj0628/Cj0629 | putative lipoprotein | Fusion/separation |
| Cj1144/Cj1145 | hypothetical protein Cj1144c | Fusion/separation |
| Cj1325/Cj1326 | putative methyltransferase | Fusion/separation |
| Cj1335/Cj1336 | motility accessory factor (function unknown) | Fusion/separation |
| Cj1677/Cj1678 | putative lipoprotein | Fusion/separation |

additional *Campylobacter* species and strains. Major updates include noteworthy modifications to the 4 main surface structure loci in the genome, 18.2% of genome product functions being updated and 90.0% of all CDSs now having additional information. The inclusion of literature searches and a GO annotation alongside genome wide structural modifications has resulted in *C. jejuni* NCTC11168 being the most comprehensively annotated *Campylobacter* genome to date.

## Methods

### *Sequence searches*

Manual re-annotation of all previously annotated *C. jejuni* NCTC11168 CDSs [4] was carried out based on results from BLASTP [39] and FASTA [40] sequence comparisons using non-redundant databases. Re-annotation was based, wherever possible, on characterised proteins or genes [4]. Additional functional data was provided by using the PFAM [41] and PROSITE [42] motif databases. New searches carried out in this re-annotation included running RFAM [43] database and also the programs TMHMM [44] and SIGNALP [45].

### *Literature & additional searches*

This re-annotation included a complete literature search of all CDS numbers and gene names using PubMed [46], HighWire Press [47], Scirus [48] and Google Scholar [49]. Artemis software release 8 [50] was used during re-annotation. The re-annotated sequence was submitted to the EMBL public database and also to GeneDB [51] and CampyDB [52]. The EMBL file included an 'original' and 'updated' note qualifier, a 'product' qualifier and each CDS represented with a unique 'locus_tag' qualifier. Appropriate 'gene' qualifiers were also present. The GeneDB submission included all the above and extra qualifiers 'colour' and 'literature'. This re-annotation also included for the first time a Gene Ontology (GO) annotation of the NCTC11168 genome sequence. This was created automatically on submission to EMBL and can be accessed via the GOA link. A separate GO annotation was created within GeneDB by carrying out a reciprocal FASTA comparison with *C. jejuni* RM1221 and adopting the GO annotation of orthologous CDSs.

### *Re-designation of pseudogenes*

Advances in genome annotation techniques that were unavailable during the original annotation have led to updated interpretation of pseudogenes and phase-variable CDSs. Using guidance from TBLASTX search results, we carried out a full re-analysis of all pseudogenes. CDSs designated as pseudogenes have been updated to reflect the complete amino acid sequence for the encoded protein regardless of expression. This has caused differences from the amino acid sequence of the previous annotation. Some pseudogene modifications entailed merging two or more adjacent, in frame CDSs (previously annotated as separate pseudogene CDSs), to create a single pseudogene containing internal stop codons. In other cases, pseudogene features were created with multiple coordinates representing one or more frameshift in the CDS – these had previously only detailed the start and stop coordinates so did not reflect the true position of the non-mutated CDS. In both cases the assignment of coordinates was based on matches to homologues determined through FASTA searches.

### *Re-designation of CDSs with an intersecting homopolymeric tract*

CDSs containing an intersecting homopolymeric tract were merged to reflect the complete amino acid sequence for appropriate genes regardless of phase. This is analogous to the scenario described above for frameshifted pseudogenes. This modification was carried out for two CDSs with an intersecting homopolymeric tract. The joining of such CDSs was not undertaken in the original annotation.

## Authors' contributions

OG carried out the re-annotation process and drafted the manuscript. SDB assisted with the re-annotation process. MTH assisted with running additional programs used in the re-annotation. JP, ND and BWW participated in the conception and supervised the design of the study. All authors submitted comments on drafts and read and approved the final manuscript.

## Additional material

### Additional File 1

C. jejuni *functional classification (created at Sanger Institute).*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-8-162-S1.doc]

### Additional File 2

*Distribution of functional classification before/after re-annotation.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-8-162-S2.doc]

### Additional File 3

*Changes to functional classification categories before and after the re-annotation.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-8-162-S3.doc]

### Additional File 4

*Changes to CDS functions and functional classifications.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-8-162-S4.xls]

## Additional File 5

*CDSs modified in* C. jejuni *NCTC11168 re-annotation.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-8-162-S5.doc]

## Additional File 6

*Pseudogene comparison between* C. jejuni *NCTC11168 and* C. jejuni *RM1221.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-8-162-S6.doc]

## Acknowledgements

Funding was provided by the BBSRC and the Wellcome Trust. We would like to thank C. Parry for technical support. We would like to thank M. Sebaihia and NR. Thomson for stimulating conversations. We would also like to thank TJ. Carver for computational support.

## References

1. Andersen MT, Brondsted L, Pearson BM, Mulholland F, Parker M, Pin C, Wells JM, Ingmer H: **Diverse roles for HspR in Campylobacter jejuni revealed by the proteome, transcriptome and phenotypic characterization of an hspR mutant.** *Microbiology* 2005, **151(Pt 3)**:905-915.
2. Nachamkin I, Allos BM, Ho T: **Campylobacter species and Guillain-Barre syndrome.** *Clin Microbiol Rev* 1998, **11(3)**:555-567.
3. Dorrell N, Mangan JA, Laing KG, Hinds J, Linton D, Al-Ghusein H, Barrell BG, Parkhill J, Stoker NG, Karlyshev AV, Butcher PD, Wren BW: **Whole genome comparison of Campylobacter jejuni human isolates using a low-cost microarray reveals extensive genetic diversity.** *Genome Res* 2001, **11(10)**:1706-1715.
4. Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream MA, Rutherford KM, van Vliet AH, Whitehead S, Barrell BG: **The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.** *Nature* 2000, **403(6770)**:665-668.
5. Karlyshev AV, Linton D, Gregson NA, Lastovica AJ, Wren BW: **Genetic and biochemical evidence of a Campylobacter jejuni capsular polysaccharide that accounts for Penner serotype specificity.** *Mol Microbiol* 2000, **35(3)**:529-541.
6. Karlyshev AV, Champion OL, Churcher C, Brisson JR, Jarrell HC, Gilbert M, Brochu D, St Michael F, Li J, Wakarchuk WW, Goodhead I, Sanders M, Stevens K, White B, Parkhill J, Wren BW, Szymanski CM: **Analysis of Campylobacter jejuni capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses.** *Mol Microbiol* 2005, **55(1)**:90-103.
7. Karlyshev AV, McCrossan MV, Wren BW: **Demonstration of polysaccharide capsule in Campylobacter jejuni using electron microscopy.** *Infect Immun* 2001, **69(9)**:5921-5924.
8. Thibault P, Logan SM, Kelly JF, Brisson JR, Ewing CP, Trust TJ, Guerry P: **Identification of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin.** *J Biol Chem* 2001, **276(37)**:34862-34870.
9. Szymanski CM, Logan SM, Linton D, Wren BW: **Campylobacter--a tale of two protein glycosylation systems.** *Trends Microbiol* 2003, **11(5)**:233-238.
10. Liu F, Tanner ME: **PseG of Pseudaminic Acid Biosynthesis: A UDP-SUGAR HYDROLASE AS A MASKED GLYCOSYLTRANSFERASE.** *J Biol Chem* 2006, **281(30)**:20902-20909.
11. Karlyshev AV, Linton D, Gregson NA, Wren BW: **A novel paralogous gene family involved in phase-variable flagella-mediated motility in Campylobacter jejuni.** *Microbiology* 2002, **148(Pt 2)**:473-480.
12. Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M, Wren BW: **Functional analysis of the Campylobacter jejuni N-linked protein glycosylation pathway.** *Mol Microbiol* 2005, **55(6)**:1695-1703.
13. Glover KJ, Weerapana E, Imperiali B: **In vitro assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation.** *Proc Natl Acad Sci U S A* 2005, **102(40)**:14255-14259.
14. Kelly J, Jarrell H, Millar L, Tessier L, Fiori LM, Lau PC, Allan B, Szymanski CM: **Biosynthesis of the N-linked glycan in Campylobacter jejuni and addition onto protein through block transfer.** *J Bacteriol* 2006, **188(7)**:2427-2434.
15. Fouts DE, Mongodin EF, Mandrell RE, Miller WG, Rasko DA, Ravel J, Brinkac LM, DeBoy RT, Parker CT, Daugherty SC, Dodson RJ, Durkin AS, Madupu R, Sullivan SA, Shetty JU, Ayodeji MA, Shvartsbeyn A, Schatz MC, Badger JH, Fraser CM, Nelson KE: **Major structural differences and novel potential virulence mechanisms from the genomes of multiple campylobacter species.** *PLoS Biol* 2005, **3(1)**:e15.
16. Hofreuter D, Tsai J, Watson RO, Novik V, Altman B, Benitez M, Clark C, Perbost C, Jarvie T, Du L, Galan JE: **Unique features of a highly pathogenic Campylobacter jejuni strain.** *Infect Immun* 2006, **74(8)**:4694-4707.
17. Ouzounis CA, Karp PD: **The past, present and future of genome-wide re-annotation.** *Genome Biol* 2002, **3(2)**:COMMENT2001.
18. Camus JC, Pryor MJ, Medigue C, Cole ST: **Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv.** *Microbiology* 2002, **148(Pt 10)**:2967-2973.
19. Dandekar T, Huynen M, Regula JT, Ueberle B, Zimmermann CU, Andrade MA, Doerks T, Sanchez-Pulido L, Snel B, Suyama M, Yuan YP, Herrmann R, Bork P: **Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames.** *Nucleic Acids Res* 2000, **28(17)**:3278-3288.
20. Schouls LM, Reulen S, Duim B, Wagenaar JA, Willems RJ, Dingle KE, Colles FM, Van Embden JD: **Comparative genotyping of Campylobacter jejuni by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination.** *J Clin Microbiol* 2003, **41(1)**:15-26.
21. Guerry P, Ewing CP, Hickey TE, Prendergast MM, Moran AP: **Sialylation of lipooligosaccharide cores affects immunogenicity and serum resistance of Campylobacter jejuni.** *Infect Immun* 2000, **68(12)**:6656-6662.
22. Gilbert M, Karwaski MF, Bernatchez S, Young NM, Taboada E, Michniewicz J, Cunningham AM, Wakarchuk WW: **The genetic bases for the variation in the lipo-oligosaccharide of the mucosal pathogen, Campylobacter jejuni. Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide.** *J Biol Chem* 2002, **277(1)**:327-337.
23. Linton D, Karlyshev AV, Hitchen PG, Morris HR, Dell A, Gregson NA, Wren BW: **Multiple N-acetyl neuraminic acid synthetase (neuB) genes in Campylobacter jejuni: identification and characterization of the gene involved in sialylation of lipo-oligosaccharide.** *Mol Microbiol* 2000, **35(5)**:1120-1134.
24. Valvano MA, Messner P, Kosma P: **Novel pathways for biosynthesis of nucleotide-activated glycero-manno-heptose precursors of bacterial glycoproteins and cell surface polysaccharides.** *Microbiology* 2002, **148(Pt 7)**:1979-1989.
25. Lerat E, Ochman H: **Recognizing the pseudogenes in bacterial genomes.** *Nucleic Acids Res* 2005, **33(10)**:3125-3132.
26. Lerat E, Ochman H: **Psi-Phi: exploring the outer limits of bacterial pseudogenes.** *Genome Res* 2004, **14(11)**:2273-2278.
27. Harrison PM, Gerstein M: **Studying genomes through the aeons: protein families, pseudogenes and proteome evolution.** *J Mol Biol* 2002, **318(5)**:1155-1174.
28. Ochman H, Davalos LM: **The nature and dynamics of bacterial genomes.** *Science* 2006, **311(5768)**:1730-1733.
29. Homma K, Fukuchi S, Kawabata T, Ota M, Nishikawa K: **A systematic investigation identifies a significant number of probable pseudogenes in the *Escherichia coli* genome.** *Gene* 2002, **294(1-2)**:25-33.
30. Mira A, Pushker R: **The silencing of pseudogenes.** *Mol Biol Evol* 2005, **22(11)**:2135-2138.

31. Linton D, Karlyshev AV, Wren BW: **Deciphering Campylobacter jejuni cell surface interactions from the genome sequence.** *Curr Opin Microbiol* 2001, 4(1):35-40.
32. Szymanski CM, Michael FS, Jarrell HC, Li J, Gilbert M, Larocque S, Vinogradov E, Brisson JR: **Detection of conserved N-linked glycans and phase-variable lipooligosaccharides and capsules from campylobacter cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy.** *J Biol Chem* 2003, 278(27):24509-24520.
33. Godde JS, Bickerton A: **The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes.** *J Mol Evol* 2006, 62(6):718-729.
34. Jansen R, van Embden JD, Gaastra W, Schouls LM: **Identification of a novel family of sequence repeats among prokaryotes.** *Omics* 2002, 6(1):23-33.
35. Rosenblad MA, Gorodkin J, Knudsen B, Zwieb C, Samuelsson T: **SRPDB: Signal Recognition Particle Database.** *Nucleic Acids Res* 2003, 31(1):363-364.
36. Regalia M, Rosenblad MA, Samuelsson T: **Prediction of signal recognition particle RNA genes.** *Nucleic Acids Res* 2002, 30(15):3368-3377.
37. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS: **Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms.** *J Biol Chem* 2002, 277(50):48949-48959.
38. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, 25(1):25-29.
39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, 215(3):403-410.
40. Pearson WR, Lipman DJ: **Improved tools for biological sequence comparison.** *Proc Natl Acad Sci U S A* 1988, 85(8):2444-2448.
41. Sonnhammer EL, Eddy SR, Durbin R: **Pfam: a comprehensive database of protein domain families based on seed alignments.** *Proteins* 1997, 28(3):405-420.
42. Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJ, Hofmann K, Bairoch A: **The PROSITE database, its status in 2002.** *Nucleic Acids Res* 2002, 30(1):235-238.
43. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR: **Rfam: an RNA family database.** *Nucleic Acids Res* 2003, 31(1):439-441.
44. Sonnhammer EL, von Heijne G, Krogh A: **A hidden Markov model for predicting transmembrane helices in protein sequences.** *Proc Int Conf Intell Syst Mol Biol* 1998, 6:175-182.
45. Nielsen H, Engelbrecht J, Brunak S, von Heijne G: **Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.** *Protein Eng* 1997, 10(1):1-6.
46. Entrez Pubmed NIH USA: . [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi].
47. HighWire Press SU: . [http://highwire.stanford.edu/].
48. Scirus - for scientific information only E: . [http://www.scirus.com/srsapp/].
49. Scholar G: . [http://scholar.google.com/].
50. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B: **Artemis: sequence visualization and annotation.** *Bioinformatics* 2000, 16(10):944-945.
51. Hertz-Fowler C, Peacock CS, Wood V, Aslett M, Kerhornou A, Mooney P, Tivey A, Berriman M, Hall N, Rutherford K, Parkhill J, Ivens AC, Rajandream MA, Barrell B: **GeneDB: a resource for prokaryotic and eukaryotic organisms.** *Nucleic Acids Res* 2004, 32(Database issue):D339-43.
52. Chaudhuri RR, Khan AM, Pallen MJ: **coliBASE: an online database for Escherichia coli, Shigella and Salmonella comparative genomics.** *Nucleic Acids Res* 2004, 32(Database issue):D296-9.

## Appendix 6 – Cj1556 – Protein Interaction Database Results

/note="Updated (2006) note: Pfam domain PF01638 Transcriptional regulator identified within CDS. Product modified to more specific family member based on motif match. No specific characterisation has been carried out yet, so putative kept within product function"
/product="putative transcriptional regulator
/locus_tag="Cj1556"

"Cj0082","Cj1556","cydB","Cj1556","1","C.jejuni YTH"
"Cj0465c","Cj1556","Cj0465c","Cj1556","1","C.jejuni YTH"
"Cj0389","Cj1556","serS","Cj1556","1","C.jejuni YTH"
"Cj0560","Cj1556","Cj0560","Cj1556","1","C.jejuni YTH"
"Cj0777","Cj1556","Cj0777","Cj1556","1","C.jejuni YTH"
"Cj0701","Cj1556","Cj0701","Cj1556","1","C.jejuni YTH"
"Cj0665c","Cj1556","argG","Cj1556","1","C.jejuni YTH"
"Cj1059c","Cj1556","gatA","Cj1556","1","C.jejuni YTH"
"Cj1318","Cj1556","Cj1318","Cj1556","1","C.jejuni YTH"
"Cj1336","Cj1556","Cj1336","Cj1556","1","C.jejuni YTH"
"Cj1434c","Cj1556","Cj1434c","Cj1556","1","C.jejuni YTH"
"Cj1285c","Cj1556","Cj1285c","Cj1556","1","C.jejuni YTH"
"Cj1489c","Cj1556","ccoO","Cj1556","1","C.jejuni YTH"
"Cj1422c","Cj1556","Cj1422c","Cj1556","1","C.jejuni YTH"
"Cj1556","Cj1108","Cj1556","clpA","1","Predictions from E.coli"
"Cj1556","Cj0759","Cj1556","dnaK","1","Predictions from E.coli"
"Cj1556","Cj0509c","Cj1556","clpB","1","Predictions from E.coli"
"Cj1556","Cj0476","Cj1556","rplJ","1","Predictions from E.coli"
"Cj1556","Cj0284c","Cj1556","cheA","1","Predictions from E.coli"

/note="Updated (2006) note: Pfam domain PF02322 Cytochrome oxidase subunit II identified within CDS. Also, nine probable transmembrane helices predicted by TMHMM2.0. Further support given to product function. Characterised within Escherichia coli with marginal identity score. Putative not added to product function"
/product="cytochrome bd oxidase subunit II"
/locus_tag="Cj0082"
/gene="cydB"

/note="Updated (2006) note: Characterised within Escherichia coli with acceptable identity score. Putative not added to product function"
/product="seryl-tRNA synthetase"
/locus_tag="Cj0389"
/gene="serS"

/note="Updated (2006) note: Pfam domain PF01152 Bacterial-like globin identified within CDS. This family of heme binding proteins are found mainly in bacteria. Characterised within Campylobacter jejuni. Ctb protein is shown to be involved in moderating oxygen flux within Campylobacter jejuni. NssR (Nitrosative stress sensing Regulator - Cj0466) controls the expression of a nitrosative stress-responsive regulon in Campylobacter jejuni which includes ctb (Cj0465c) and cgb (Cj1586). Putative not added to product function"
/product="group III truncated haemoglobin"
/locus_tag="Cj0465c"
/gene="ctb"

/note="Updated (2006) note: Pfam domain PF01554 MatE identified within CDS. Prosite domain PS00589 PTS_HPR_SER, Phosphotransferase system, HPr serine phosphorylation site also identified. Twelve probable transmembrane helices predicted by TMHMM2.0. Characterised members of the Multi Antimicrobial Extrusion (MATE) family function as drug/sodium antiporters. These proteins mediate resistance to a wide range of cationic dyes, fluroquinolones, aminoglycosides and other structurally diverse antibodies and drugs. MATE proteins are found in bacteria, archaea and eukaryotes. These proteins are predicted to have 12 alpha-helical transmembrane regions (which this one does). Product modified to more specific family member due to motif match. No specific characterisation with acceptable identity scores has been carried out yet. Putative kept within product function. Literature search identified paper giving further clues to product function"
/product="putative MATE family transport protein"
/locus_tag="Cj0560"
/literature="PMID:16048946"

/note="Updated (2006) note: Some characterisation work within Escherichia coli and Rattus norvegicus (Rat). Appropriate motifs present. Putative not added to product function"
/product="argininosuccinate synthase"
/locus_tag="Cj0665c"
/gene="argG"

/note="Updated (2006) note: Similar to many with marginal identity scores. Putative kept within product function"
/product="putative protease"
/locus_tag="Cj0701"

/note="Updated (2006) note: Characterisation work carried out within more than one species with marginal identity score. Putative kept within product function"
/product="putative ATP-dependent DNA helicase"
/locus_tag="Cj0777"

/locus_tag="Cj1059c"
/gene="gatA"
/literature="Pfam domain PF01425 Amidase identified within CDS. Further support given to product function. Characterised in Bacillus subtilis with acceptable identity score, so putative not added to product function"
/product="Glu-tRNAGln amidotransferase subunit A"

/note="Updated (2006) note: Pfam domains PF02621 Uncharacterized ACR, COG1427 identified within CDS. Conserved added to product funtion"
/product="conserved hypothetical protein Cj1285c"
/locus_tag="Cj1285c"

/note="Updated (2006) note: Pfam domain PF01973 Protein of unknown function DUF115 identified within CDS. Characterisation work has been carried out within Campylobacter jejuni (PMID:11895937). Product function modified to more specific family member. Identifed as part of Campylobacter jejuni O-linked glycosylation locus (Cj1293 - Cj1342)"
/product="motility accessory factor (function unknown)"
/locus_tag="Cj1318"
/gene="maf1"

/note="Updated (2006) note: Pfam domain PF01973 Protein of unknown function DUF115 identified within CDS. Some characterisation work has been carried out within Campylobacter jejuni. Product funciton updated to more specific family member. Identifed as part of Campylobacter jejuni O-linked glycosylation locus (Cj1293 - Cj1342). Coding sequences have been merged to reflect the complete amino acid sequence for this gene regardless of phase. Previous annotation gave Cj1335 and Cj1336 as seperate CDS. Merging of these CDSs has lead to loss of the downstream CDS"
/product="motility accessory factor (function unknown)"
/locus_tag="Cj1335"
/gene="maf4"

/note="Updated (2006) note: No specific characterisation has been carried out yet, so putative kept within product function. Literature search identified papers giving further clues to product funciton. Identifed as part of Campylobacter jejuni capsule locus (Cj1413c - Cj1448c)"
/product="putative sugar transferase"
/locus_tag="Cj1434c"

/note="Updated (2006) note: Prosite PS00221 MIP family signature identified within CDS. No specific characterisation has been carried out yet. Putative kept within product function. Identified within the Campylobacter jejuni Capsule locus (Cj1413c - Cj1448c)"
/product="putative sugar transferase"
/locus_tag="Cj1422c"

/note="Updated (2006) note: PF02433 Cytochrome C oxidase, mono-heme subunit/FixO subunit identified within CDS. Further support given to product function. Characterised in Helicobacter pylori with acceptable identity score. Putative not added to product function"
/product="cb-type cytochrome C oxidase subunit II"
/locus_tag="Cj1489c"
/gene="ccoO"

## Appendix 7 – Cj1546 – Protein Interaction Database Results

/note="Updated (2006) note: Pfam domain PF01638 Transcriptional regulator identified within CDS. Product modified to more specific family member based on motif match. Putative kept within product function"
/product="putative transcriptional regulator"
/locus_tag="Cj1546"

"Cj1546","Cj1166c","Cj1546","Cj1166c","1","C.jejuni YTH"
"Cj1546","Cj1193c","Cj1546","Cj1193c","1","C.jejuni YTH"
"Cj1546","Cj0865","Cj1546","dsbB","1","C.jejuni YTH"
"Cj1546","Cj0797c","Cj1546","Cj0797c","1","C.jejuni YTH"
"Cj1546","Cj0620","Cj1546","Cj0620","1","C.jejuni YTH"
"Cj1546","Cj0161c","Cj1546","moaA","1","C.jejuni YTH"
"Cj1546","Cj1108","Cj1546","clpA","1","Predictions from E.coli"
"Cj1546","Cj0759","Cj1546","dnaK","1","Predictions from E.coli"
"Cj1546","Cj0509c","Cj1546","clpB","1","Predictions from E.coli"
"Cj1546","Cj0476","Cj1546","rplJ","1","Predictions from E.coli"
"Cj1546","Cj0284c","Cj1546","cheA","1","Predictions from E.coli"

/note="Updated (2006) note: Pfam domain PF06738 Protein of unknown function (DUF1212) identified within CDS. Five probable transmembrane helices predicted by TMHMM2.0. Further support given to product function"
/product="putative integral membrane protein"
/locus_tag="Cj1166c"

/note="Original (2000) note: Cj1193c, probable periplasmic protein, len: 268 aa; no Hp match. Contains probable N-terminal signal sequence"
/product="putative periplasmic protein"
/locus_tag="Cj1193c"

/note="Updated (2006) note: Four probable transmembrane helices predicted by TMHMM2.0. This CDS has been characterised as DsbB protein (PMID:15632440). It has been demonstrated that Cj0865 is indeed a disulfide oxidoreductase. This has been shown via in silico and in vivo work in E. coli and also complementation and enzymatic tests in C. jejuni. There is an alignment of different DsbB proteins (including Cj0865) with prediction of 5 transmembranehelises (not 4 as stated in annotation); (P24-A39; F60-I74; S87-L101; F194-A214; F231-G246). There are some differences in amino acid sequences between Cj0865 and its orthologue from 81-176 (CJJ81176_0881). It is stated that its translation starts from the second Met (although the predicted N-terminal amino acid sequences of both proteins are identical), there is also an insertion of 3 amino acids at the COOH end of protein. Thus, not added to protein function"
/product="putative disulphide oxidoreductase"
/locus_tag="Cj0865"
/gene="dsbB"

/note="Original (2000) note: Cj0797c, unknown, len: 71 aa; weak similarity to TR:O68849 (EMBL:AF055586) ORF10 in Vibrio cholerae integron InVch (80 aa), fasta scores; opt: 95 z-score: 151.4 E(): 0.39, 39.2% identity in 51 aa overlap. No Hp match"
/product="hypothetical protein Cj0797c"
/locus_tag="Cj0797c"

/note="Updated (2006) note: Pfam domain PF01863 Protein of unknown function DUF45 identified within CDS. This motif has no known function. Members are found in some archaebacteria, as well as Helicobacter pylori (Epsilon Proteobacteria). The proteins are 190-240 amino acids long, with the C terminus being the most conserved region, containing three conserved histidines. This motif is similar to that found in Zinc proteases, suggesting a possible role as a protease. Conserved added to product function"
/product="conserved hypothetical protein Cj0620"
/locus_tag="Cj0620"

/note="Updated (2006) note: Pfam domains PF06463 Molybdenum Cofactor Synthesis C and PF04055 Radical SAM superfamily were identified within CDS. Further support given to product funtion. Characterisation work carried out within Escherichia coli with marginal identity score. Putative kept within product function"
/product="putative molybdenum cofactor biosynthesis protein A"
/locus_tag="Cj0161c"
/gene="moaA"

/note="Updated (2006) note: Pfam domains PF02861 Clp amino terminal domain and PF00004 ATPase family associated with various cellular activities (AAA) were identified within CDS. Further support given to product function. Characterised within Escherichia coli with acceptable identity score. Putative not added to product function"
/product="ATP-dependent Clp protease ATP-binding subunit"
/locus_tag="Cj1108"
/gene="clpA"

/note="Updated (2006) note: Characterised within Campylobacter jejuni and others to DnaK protein. Putative not added to product function"
/product="heat shock protein DnaK"
/locus_tag="Cj0759"

/note="Updated (2006) note: Characterised within Campylobacter jejuni. Putative not added to product function"
/product="ATP-dependent Clp protease ATP-binding subunit"
/locus_tag="Cj0509c"
/gene="clpB"

/note="Updated (2006) note: Prosite domain PS01109 RIBOSOMAL_L10, Eubacterial ribosomal protein L10 identified within CDS. Further support given to product function. Characterised within Escherichia coli with acceptable identity score. Putative not added to product function"
/product="50S ribosomal protein L10"
/locus_tag="Cj0476"
/gene="rplJ"

/note="Updated (2006) note: Characterised within Escherichia coli with marginal identity scores. Putative not added to product function"
/product="chemotaxis histidine kinase"
/locus_tag="Cj0284c"
/gene="cheA"

**Appendix 8 – Genes up-regulated in the *Cj1556* mutant when compared to 11168H wild-type strain using RNA from late-log phase of growth**

| Gene ID | Fold Change | Product Function |
|---|---|---|
| *Cj1556* | 10.4 | putative transcriptional regulator - Cj1556 (cosR) |
| *Cj0244* | 5.392 | 50s ribosomal protein L35 |
| *Cj0374* | 5.061 | hypothetical protein Cj0374 |
| *Cj1326* | 3.536 | putative methyltransferase Cj1326 - homopolymeric tract |
| *Cj0767c* | 3.493 | phosphopantetheine adenylyltransferase |
| *Cj1611* | 3.381 | 30S ribosomal protein S20 |
| *Cj0496* | 3.366 | hypothetical protein Cj0496 |
| *Cj1555c* | 3.058 | hypothetical protein Cj1555c |
| *Cj0487* | 2.872 | putative amidohydrolase |
| *Cj0628* | 2.73 | putative lipoprotein - homopolymeric tract |
| *Cj0491* | 2.578 | 30S ribosomal protein S12 |
| *Cj0854c* | 2.125 | putative periplasmic protein |
| *Cj1226c* | 2.019 | putative two-component sensor (histidine kinase) |
| *Cj0598* | 1.966 | putative membrane protein |
| *Cj1145c* | 1.935 | hypothetical protein Cj1144c – homopolymeric tract |
| *Cj0919c* | 1.887 | putative ABC-type amino-acid transporter permease protein |
| *Cj1423c* | 1.867 | putative D-glycero-D-manno-heptose 1-phosphate guanosyltransferase |
| *Cj1666c* | 1.852 | putative periplasmic protein |
| *Cj0310c* | 1.85 | putative efflux protein |
| *Cj1542* | 1.767 | putative allophanate hydrolase subunit 1 |
| *Cj0694* | 1.756 | putative periplasmic protein |
| *Cj0301c* | 1.684 | putative molybdenum transport system permease protein |
| *Cj0944c* | 1.674 | putative periplasmic protein |
| *Cj1566c* | 1.649 | NADH dehydrogenase I chain N |
| *Cj0475* | 1.629 | 50S ribosomal protein L1 |
| *Cj0357c* | 1.624 | putative integral membrane protein |
| *Cj0263* | 1.582 | zinc transporter |
| *Cj1163c* | 1.58 | putative cation transport protein |
| *Cj0057* | 1.575 | putative periplasmic protein |
| *Cj0534* | 1.566 | succinyl-coA synthetase alpha chain |
| *Cj0935c* | 1.564 | putative sodium:amino-acid symporter family protein |
| *Cj1327* | 1.563 | N-acetylneuraminic acid synthetase |
| *Cj1438c* | 1.559 | putative sugar transferase |
| *Cj0099* | 1.558 | putative biotin--[acetyl-CoA-carboxylase] synthetase |
| *Cj0338c* | 1.554 | DNA polymerase I |
| *Cj1497c* | 1.552 | hypothetical protein Cj1497c |
| *Cj1410c* | 1.551 | putative membrane protein |
| *Cj0041* | 1.539 | putative flagella hook-length control protein |
| *Cj1655c* | 1.533 | Na(+)/H(+) antiporters |
| *Cj0856* | 1.529 | signal peptidase I |
| *Cj0706* | 1.528 | hypothetical protein Cj0706 |
| *C0j272* | 1.527 | hypothetical protein Cj0272 |

| | | |
|---|---|---|
| *Cj1529c* | 1.511 | phosphoribosylformylglycinamidine cyclo-ligase |
| *Cj0551* | 1.508 | elongation factor P |
| *Cj0195* | 1.508 | flagellum-specific ATP synthase |
| *Cj0433c* | 1.507 | phospho-N-acetylmuramoyl-pentapeptide- transferase |
| *Cj1436c* | 1.48 | Aminotransferase |
| *Cj1718c* | 1.461 | 3-isopropylmalate dehydrogenase |
| *Cj0928* | 1.459 | putative integral membrane protein (dedA family) |
| *Cj1241* | 1.443 | putative MFS (Major Facilitator Superfamily) transporter protein |
| *Cj0343c* | 1.439 | putative integral membrane protein |
| *Cj0112* | 1.43 | putative TolB precursor protein |
| *Cj0689* | 1.427 | acetate kinase |
| *C0j606* | 1.415 | putative secretion protein HlyD |
| *Cj0693c* | 1.411 | S-adenosyl-methyltransferase |
| *Cj0853c* | 1.403 | glutamate-1-semialdehyde 2,1-aminomutase |
| *Cj0056c* | 1.4 | hypothetical protein Cj0056c |
| *Cj0474* | 1.397 | 50S ribosomal protein L11 |
| *Cj0372* | 1.373 | putative glutathionylspermidine synthase |
| *Cj0345* | 1.359 | putative anthranilate synthase component I |
| *Cj1426c* | 1.35 | putative methyltransferase family protein |
| *Cj0574* | 1.348 | acetolactate synthase large subunit |
| *Cj0335* | 1.347 | flagella biosynthetic protein |
| *Cj0899c* | 1.313 | 4-methyl-5(beta-hydroxyethyl)-thiazole monophosphate synthesis protein |
| *Cj1568c* | 1.294 | NADH dehydrogenase I chain L |
| *Cj0320* | 1.273 | putative flagella assembly protein |
| *Cj1695c* | 1.273 | 50S ribosomal protein L5 |
| *Cj0541* | 1.269 | polyprenyl synthetase |
| *Cj1290c* | 1.249 | biotin carboxylase |
| *Cj0434* | 1.232 | 2,3-bisphosphoglycerate-independent phosphoglycerate mutase |
| *Cj0860* | 1.225 | probable integral membrane protein |
| *Cj0069* | 1.173 | hypothetical protein Cj0069 |
| *Cj1625c* | 1.109 | amino acid transporter |

**Appendix 9 – Genes down-regulated in the *Cj1556* mutant when compared to 11168H wild-type strain using RNA from late-log phase of growth**

| Gene ID | Fold Change | Product Function |
|---|---|---|
| *Cj0461c* | 1.167 | putative MFS (Major Facilitator Superfamily) transport protein |
| *Cj1690c* | 1.287 | 30S ribosomal protein S5 |
| *Cj1273c* | 1.362 | putative DNA-directed RNA polymerase omega chain |
| *Cj1311* | 1.364 | putative acylneuraminate cytidylyltransferase |
| *Cj0264c* | 1.374 | molybdopterin-containing oxidoreductase |
| *Cj0424* | 1.385 | putative acidic periplasmic protein |
| *Cj1254* | 1.468 | hypothetical protein Cj1254 |
| *Cj0653c* | 1.541 | putative aminopeptidase |
| *Cj0529c* | 1.546 | putative aminodeoxychorismate lyase family protein |
| *Cj1053c* | 1.715 | putative integral membrane protein |
| *Cj1677* | 1.869 | putative lipoprotein (Homopolymeric tract) |
| *Cj0830* | 1.89 | putative integral membrane protein |
| *Cj0265c* | 2.053 | putative cytochrome C-type haem-binding periplasmic protein |
| *Cj0661c* | 2.268 | GTP-binding protein Era homologue |
| *Cj0623* | 2.347 | hydrogenase isoenzymes formation protein |
| *Cj0175c* | 2.387 | putative iron-uptake ABC transport system periplasmic iron-binding protein |
| *Cj0564* | 2.545 | putative integral membrane protein |
| *Cj1086c* | 4.695 | hypothetical protein Cj1086c |

## Appendix 10 – Promoter search results

300 nucleotides upstream and 50 nucleotides of *Cj1556*

tttagagctattgcgtacaaaacatactggctctaaatttctttttaaagcttctttgctgatttctttgcctactttaccgcttgcacataa
aattgctattttcataatatctccttttaagtatttttacataaaaaatgcaatctagaaattataaatatatttactactaaaaatcaagta
gatattattttataatatagtatctattttaggagtattatgttattttagtattttaatgagataaatttttttaaaaaatatgttataatttaaa
aaataaggatttataatgaaaaaatatcattctctttgtcctattgaaaccacgcttaatttgat

```
> test sequence
 Length of sequence - 350
 Threshold for promoters - 0.20
 Number of predicted promoters - 1
 Promoter Pos:    309 LDF- 9.08
 -10 box at pos.    294 atttataat Score    81
 -35 box at pos.    273 ttataa    Score    24

 Oligonucleotides from known TF binding sites:

 For promoter at    309:
        lrp:  ATTTTTTT at position     255 Score -  11
      lexA:  TTTTTTTA at position     256 Score -  16
    rpoD17:  TGTTATAA at position     271 Score -  12
      argR:  TTATAATT at position     273 Score -  14
       fis:  AAAAATAA at position     284 Score -   9
    rpoD17:  TTTATAAT at position     295 Score -   9
```

300 nucleotides upstream and 50 nucleotides of *ctb* (*Cj0465c*)

agcaggcatttctgctataaaatttggaggattaaaataatgcaaagtaatttctttatcacttgctgtacttttataaatgcgtatttttc
ccttgagtaaaataaaaaaatttttagcctcttcgccttcataaaataaaatactatttttttgaaatttttttaattttcctacactagataa
aagttctaaataatctttcattatataatatttccattcttaacttatgttaaatttaatttatcttatttttgctatattaacgccataaaattaa
catttaagaaaggcttatatgaaatttgaaacaattaatcaagaaagcatagcaaaactcatggaaat

```
> test sequence
 Length of sequence - 351
 Threshold for promoters - 0.20
 Number of predicted promoters - 1
 Promoter Pos:    155 LDF- 11.68
 -10 box at pos.    136 aaataaaat Score    56
 -35 box at pos.    117 tagcct    Score    19

 Oligonucleotides from known TF binding sites:

 For promoter at    155:
     rpoD17:  GAGTAAAA at position      97 Score -   8
        ihf:  AAATAAAA at position     102 Score -  13
      arcA:  AATAAAAA at position     103 Score -  12
       phoB:  TCATAAAA at position     131 Score -  11
        ihf:  AAATAAAA at position     136 Score -  13
        ihf:  AATAAAAT at position     137 Score -  10
        lrp:  TATTTTTT at position     147 Score -  11
        lrp:  ATTTTTTT at position     148 Score -  11
        fnr:  TTTTTTGA at position     150 Score -   9
        lrp:  ATTTTTTT at position     159 Score -  11
       lexA:  TTTTTTTA at position     160 Score -  16
       nagC:  TTTAATTT at position     164 Score -   7
     rpoD18:  TTAATTTT at position     165 Score -   9
```

300 nucleotides upstream and 50 nucleotides of *katA* (*Cj1385*)

aaggaatcggcttgcgattaactctattttcattacgtgcatcccagtgttctatttttttagcacctaaatcatgtaaaacagctctaat
aggagtgatttcatctgcaccaataaccaaaactgacataaattctcctttgtgatttttaaaatatattttgataattattattataatata
aaaaaactaaatttattttgaaatttattatcaataaaatgcaattattcagttaattttaatttttaatagatataatttactaattaataaaa
tttattatttaggagaaaacaatgaaaaaattgactaacgattttggaaacattatagctgataaccaaaat

```
> test sequence
 Length of sequence - 351
 Threshold for promoters - 0.20
 Number of predicted promoters - 1
 Promoter Pos:    270 LDF - 7.57
 -10 box at pos.    255 agatataat Score    77
 -35 box at pos.    237 ttaatt    Score    30
```

```
Oligonucleotides from known TF binding sites:

 For promoter at    270:
        fnr:  ATCAATAA at position     214 Score -   5
        ihf:  AATAAAAT at position     217 Score -  10
     rpoD18:  TTAATTTT at position     237 Score -   9
       nagC:  TTTAATTT at position     242 Score -   7
     rpoD18:  TTAATTTT at position     243 Score -   9
       argR:  AATTAATA at position     269 Score -  11
        ihf:  AATAAAAT at position     273 Score -  10
        lrp:  ATTTATTA at position     279 Score -   6
```

300 nucleotides upstream and 50 nucleotides of *perR* (*Cj0322*)

gtaaaatttggtttatttttagtgaaaatgttaaaattggcggtatagaaagtttaattaataatttttacaaaaatatgatttgcatgtaa
aagttgttagctttgaatatgaagacaaatttattgaacatggaaaaacaagtgaggtggaaaaaaatctagaaaaagatgtcaat
agtttgttgacgaaagttttaaaattttatcattaaataataaaaaatattatttaatataaacttcttttaatgaaaattttattattatttttac
ttaaaaaagcaaggaataaataaaatggaattactacaaatgcttaaaaaacatgagttaaaagctactccgcaa

```
> test sequence
 Length of sequence - 351
 Threshold for promoters - 0.20
 Number of predicted promoters - 1
 Promoter Pos:    309 LDF - 7.95
 -10 box at pos.    294 aaataaaat Score    56
 -35 box at pos.    272 tttact    Score    42

 Oligonucleotides from known TF binding sites:

 For promoter at    309:
       deoR:  AATTTTAT at position     257 Score -   8
     rpoD17:  TTTTACTT at position     271 Score -   8
     rpoD19:  TACTTAAA at position     274 Score -   8
     rpoD17:  AATAAATA at position     291 Score -  11
       lexA:  ATAAATAA at position     292 Score -  14
       tyrR:  TAAATAAA at position     293 Score -  10
        ihf:  AAATAAAA at position     294 Score -  13
        ihf:  AATAAAAT at position     295 Score -  10
```

300 nucleotides upstream and 50 nucleotides of *hspR* (*Cj1230*)

agagaagatgatgatttatatcaaaaggttgatatttctttaaaaactgcacttttggtggaaaaataaatataaaaacactcaaaga
aggcaaggaagaagcaataattaacataacgcctaattcaaaaaataaccaaaaaattcgtttaaaaggctatggtgtgcaaaat
agaaaaagtgatatttatggggatatgtatttgattttaaatgttgttttgccaaatttagataccttgatgaaaaatttatagaactctt
aaaagaaaaattaccttaaagaaaggaataaaatatggaacagcattatgatgaacctgtatatttaatcagtgttgtggctaaa

```
> test sequence
 Length of sequence - 351
 Threshold for promoters - 0.20
 Number of predicted promoters - 1
 Promoter Pos:     75 LDF - 6.69
 -10 box at pos.     60 tggaaaaat Score    51
 -35 box at pos.     39 tttaaa    Score    41

 Oligonucleotides from known TF binding sites:

 For promoter at     75:
       carP:  CACTTTTT at position      50 Score -   8
        fis:  AAAAATAA at position      63 Score -   9
      rpoD17:  AATAAATA at position      66 Score -  11
        fnr:  ATAAATAT at position      67 Score -   9
       lexA:  ATATAAAA at position      71 Score -  11
       lexA:  TAAAAACA at position      74 Score -   7
```

**Appendix 11 – Genes up-regulated in the *Cj0248* mutant when compared to 11168H wild-type strain using RNA from late-log phase of growth**

| Gene ID | Fold Change | Product Function |
|---|---|---|
| *Cj0795c* | 2.557 | UDP-N-acetylmuramoyl-tripeptide D-alanyl-D-alanine ligase |
| *Cj0848c* | 2.427 | hypothetical protein Cj0848c |
| *Cj0523* | 2.41 | putative membrane protein |
| *Cj1131c* | 2.244 | UDP-GlcNAc/Glc 4-epimerase |
| *Cj0660c* | 2.237 | putative transmembrane protein |
| *Cj0485* | 2.054 | putative oxidoreductase |
| *Cj0998c* | 2.041 | putative periplasmic protein |
| *Cj0816* | 1.911 | hypothetical protein Cj0816 |
| *Cj1718c* | 1.885 | 3-isopropylmalate dehydrogenase |
| *Cj0836* | 1.866 | methylated-DNA--protein-cysteine methyltransferase |
| *Cj1588c* | 1.864 | putative MFS (Major Facilitator Superfamily) transport protein |
| *Cj1125c* | 1.86 | GalNAc transferase |
| *Cj1168c* | 1.848 | putative integral membrane protein (dedA homologue) |
| *Cj0882c* | 1.812 | flagella biosynthesis protein |
| *Cj1668c* | 1.81 | putative periplasmic protein |
| *Cj0823* | 1.781 | hypothetical protein Cj0823 |
| *Cj0500* | 1.778 | putative rhodanese-like domain protein |
| *Cj0280* | 1.769 | hypothetical protein Cj0280 |
| *Cj1215* | 1.76 | putative peptidase M23 family protein |
| *Cj1016c* | 1.754 | branched-chain amino-acid ABC transport system permease protein |
| *Cj1210* | 1.734 | putative integral membrane protein |
| *Cj0174c* | 1.729 | putative iron-uptake ABC transport system permease protein |
| *Cj1340c* | 1.728 | hypothetical protein Cj1340c (1318 family) |
| *Cj1566c* | 1.69 | NADH dehydrogenase I chain N |
| *Cj1551c* | 1.689 | putative type I restriction enzyme S protein |
| *Cj1284* | 1.681 | putative K+ uptake protein |
| *Cj1615* | 1.673 | putative haemin uptake system permease protein |
| *Cj0535* | 1.662 | OORD subunit of 2-oxoglutarate:acceptor oxidoreductase |
| *Cj0850c* | 1.654 | putative MFS (Major Facilitator Superfamily) transport protein |
| *Cj1596* | 1.654 | 50S ribosomal protein L17 |
| *Cj1618c* | 1.654 | putative radical SAM domain protein |
| *Cj1040c* | 1.652 | putative MFS (Major Facilitator Superfamily) transport protein |
| *Cj1582c* | 1.642 | putative peptide ABC-transport system permease protein |
| *Cj1163c* | 1.625 | putative cation transport protein |
| *Cj1089c* | 1.618 | hypothetical protein Cj1089c |
| *Cj0661c* | 1.617 | GTP-binding protein ERA homologue |
| *Cj1098* | 1.614 | aspartate carbamoyltransferase |
| *Cj1555c* | 1.611 | hypothetical protein Cj1555c |
| *Cj0693c* | 1.611 | S-adenosyl-methyltransferase |

| | | |
|---|---|---|
| *Cj1091c* | 1.611 | leucyl-tRNA synthetase |
| *Cj1039* | 1.606 | putative undecaprenyldiphospho-muramoylpentapeptide b-N-acetylglucosaminyltransferase |
| *Cj0590* | 1.604 | putative SAM-dependent methyltransferase |
| *Cj0821* | 1.598 | UDP-N-acetylglucosamine pyrophosphorylase |
| *Cj0345* | 1.594 | putative anthranilate synthase component I |
| *Cj1653c* | 1.59 | probable lipoprotein |
| *Cj0690c* | 1.583 | possible restriction /modification enzyme |
| *Cj0179* | 1.576 | biopolymer transport protein |
| *Cj1253* | 1.567 | polyribonucleotide nucleotidyltransferase |
| *Cj0984* | 1.562 | hypothetical protein Cj0984 |
| *Cj0308c* | 1.554 | putative dethiobiotin synthetase |
| *Cj0524* | 1.551 | hypothetical protein Cj0524 |
| *Cj1547* | 1.55 | homologue of BLC protein |
| *Cj0742* | 1.541 | pseudogene (putative outer membrane protein) |
| *Cj0305c* | 1.539 | hypothetical protein Cj0305c |
| *Cj1066* | 1.539 | Nitroreductase |
| *Cj1442c* | 1.538 | putative sugar transferase |
| *Cj0091* | 1.534 | putative lipoprotein |
| *Cj1092c* | 1.529 | protein-export membrane protein |
| *Cj0437* | 1.525 | succinate dehydrogenase flavoprotein subunit |
| *Cj1661* | 1.524 | possible ABC transport system permease protein |
| *Cj0790* | 1.51 | formyltetrahydrofolate deformylase |
| *Cj1614* | 1.501 | haemin uptake system outer membrane receptor |
| *Cj0433c* | 1.493 | phospho-N-acetylmuramoyl-pentapeptide- transferase |
| *Cj1252* | 1.488 | putative periplasmic protein |
| *Cj1228c* | 1.486 | serine protease (protease DO) |
| *Cj1601* | 1.486 | phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerise |
| *Cj0646* | 1.482 | putative lipoprotein |
| *Cj1655c* | 1.481 | Na(+)/H(+) antiporters |
| *Cj1622* | 1.479 | riboflavin-specific deaminase/reductase |
| *Cj0947c* | 1.477 | putative carbon-nitrogen hydrolase |
| *Cj1055c* | 1.477 | putative sulfatase family protein |
| *Cj1612* | 1.474 | peptide chain release factor 1 |
| *Cj0540* | 1.467 | putative exporting protein |
| *Cj0861c* | 1.465 | para-aminobenzoate synthase glutamine amidotransferase component II |
| *Cj0822* | 1.46 | phosphopantothenoylcysteine decarboxylase |
| *Cj0860* | 1.451 | probable integral membrane protein |
| *Cj0259* | 1.437 | Dihydroorotase |
| *Cj0928* | 1.434 | putative integral membrane protein (dedA family) |
| *Cj0139* | 1.433 | putative endonuclease |
| *Cj1568c* | 1.432 | NADH dehydrogenase I chain L |
| *Cj0650* | 1.432 | putative ATP /GTP binding protein |
| *Cj1200* | 1.432 | putative NLPA family lipoprotein |

| | | |
|---|---|---|
| *Cj0571* | 1.427 | putative transcriptional regulator |
| *Cj1093c* | 1.427 | protein-export membrane protein |
| *Cj1217c* | 1.421 | hypothetical protein Cj1217c |
| *Cj1453c* | 1.42 | putative tRNA(Ile)-lysidine synthase |
| *Cj1647* | 1.418 | putative ABC transport system ATP-binding protein |
| *Cj0866* | 1.415 | pseudogene (arylsulfatase) |
| *Cj1019c* | 1.407 | branched-chain amino-acid ABC transport system periplasmic binding protein |
| *Cj0643* | 1.405 | two-component response regulator |
| *Cj1102* | 1.405 | tRNA pseudouridine synthase B |
| *Cj0294* | 1.4 | putative MoeB/ThiF family protein |
| *Cj1137c* | 1.4 | putative glycosyltransferase |
| *Cj0935c* | 1.397 | putative sodium:amino-acid symporter family protein |
| *Cj0727* | 1.397 | putative periplasmic solute-binding protein |
| *Cj0068* | 1.384 | Protease |
| *Cj1666c* | 1.383 | putative periplasmic protein |
| *Cj0541* | 1.376 | polyprenyl synthetase |
| *Cj0338c* | 1.371 | DNA polymerase I |
| *Cj1319* | 1.369 | putative nucleotide sugar dehydratase |
| *Cj0075c* | 1.354 | putative oxidoreductase iron-sulfur subunit |
| *Cj1447c* | 1.347 | capsule polysaccharide export ATP-binding protein |
| *Cj0491* | 1.346 | 30S ribosomal protein S12 |
| *Cj1529c* | 1.341 | phosphoribosylformylglycinamidine cyclo-ligase |
| *Cj0495* | 1.338 | putative methyltransferase domain protein |
| *Cj0499* | 1.331 | putative histidine triad (HIT) family protein |
| *Cj0079c* | 1.321 | cytolethal distending toxin A |
| *Cj1689c* | 1.312 | 50S ribosomal protein L15 |
| *Cj1504c* | 1.308 | putative selenide,water dikinase |
| *Cj0057* | 1.302 | putative periplasmic protein |
| *Cj1193c* | 1.297 | putative periplasmic protein |
| *Cj1422c* | 1.288 | possible sugar transferase |
| *Cj1320* | 1.278 | putative aminotransferase (degT family) |
| *Cj0515* | 1.276 | putative periplasmic protein |
| *Cj0606* | 1.264 | putative secretion protein HlyD |
| *Cj1185c* | 1.251 | putative ubiquinol-cytochrome C reductase cytochrome B subunit |
| *Cj0093* | 1.251 | putative periplasmic protein |
| *Cj1418c* | 1.166 | putative transferase |
| *Cj0695* | 1.163 | cell division protein ftsA |
| *Cj1498c* | 1.163 | adenylosuccinate synthetase |

**Appendix 12 – Genes down-regulated in the *Cj0248* mutant when compared to 11168H wild-type strain using RNA from late-log phase of growth**

| Gene ID | Fold Change | Product Function |
|---|---|---|
| *Cj0396c* | 1.238 | putative lipoprotein |
| *Cj0363c* | 1.25 | putative oxygen-independent coproporphyrinogen III oxidase |
| *Cj1401c* | 1.302 | putative triosephosphate isomerase |
| *Cj0842* | 1.321 | putative lipoprotein |
| *Cj1175c* | 1.35 | arginyl-tRNA synthetase |
| *Cj0689* | 1.399 | acetate kinase |
| *Cj1277c* | 1.41 | putative ABC transporter ATP-binding protein |
| *Cj1444c* | 1.412 | capsule polysaccharide export system periplasmic protein |
| *Cj1297* | 1.418 | hypotehtical protein Cj1297 |
| *Cj0952c* | 1.449 | putative HAMP containing membrane protein |
| *Cj1433c* | 1.456 | hypothetical protein Cj1433c |
| *Cj1288c* | 1.464 | glutamyl-tRNA synthetase |
| *Cj1235* | 1.473 | putative peptidase M23 family protein |
| *Cj0237* | 1.477 | carbonic anyhydrase |
| *Cj0467* | 1.477 | amino-acid ABC transporter integral membrane protein |
| *Cj0992c* | 1.493 | oxygen-independent coproporphyrinogen III oxidase |
| *Cj0314* | 1.506 | diaminopimelate decarboxylase |
| *Cj0066c* | 1.511 | 3-dehydroquinate dehydratase |
| *Cj0281c* | 1.517 | putative transaldolase |
| *Cj0895c* | 1.527 | 3-phosphoshikimate 1-carboxyvinyltransferase |
| *Cj0527c* | 1.555 | flagella basal-body rod protein |
| *Cj0315* | 1.592 | putative HAD-superfamily hydrolase, subfamily IIA |
| *Cj0285c* | 1.61 | chemotaxis protein |
| *Cj0327* | 1.639 | putative endoribonuclease L-PSP family protein |
| *Cj0261c* | 1.653 | putative SAM-dependent methyltransferase |
| *Cj1451* | 1.706 | dUTPase |
| *Cj0840c* | 1.733 | fructose-1,6-bisphosphatase |
| *Cj0284c* | 1.742 | chemotaxis histidine kinase |
| *Cj1315c* | 1.767 | imidazole glycerol phosphate synthase subunit |
| *Cj0286c* | 1.799 | hypothetical protein Cj0286c |
| *Cj0462* | 1.838 | putative radical SAM domain protein |
| *Cj0529c* | 1.953 | putative aminodeoxychorismate lyase family protein |
| *Cj1514c* | 1.957 | hypothetical protein Cj1514c |
| *Cj0854c* | 1.976 | putative periplasmic protein |
| *Cj0056c* | 2.004 | hypothetical protein Cj0056c |
| *Cj0922c* | 2.024 | ABC-type amino-acid transporter ATP-binding protein |
| *Cj0539* | 2.07 | hypothetical protein Cj0539 |
| *Cj0575* | 2.141 | acetolactate synthase small subunit |
| *Cj1359* | 2.183 | polyphosphate kinase |
| *Cj0758* | 2.364 | heat shock protein grpE |

| | | |
|---|---|---|
| *Cj1156* | 2.433 | transcription termination factor |
| *Cj0874c* | 2.451 | putative cytochrome C |
| *Cj0757* | 2.632 | putative heat shock regulator |
| *Cj1445c* | 2.674 | capsule polysaccharide export system inner membrane protein |
| *Cj0102* | 2.74 | ATP synthase F0 sector B' subunit |
| *Cj0233c* | 2.762 | putative orotate phosphoribosyltransferase |
| *Cj0859c* | 2.924 | hypothetical protein Cj0859c |
| *Cj0977* | 3.106 | hypothetical protein Cj0977 |
| *Cj0981c* | 3.185 | putative MFS (Major Facilitator Superfamily) transport protein |
| *Cj1656c* | 3.497 | hypothetical protein Cj1656c |
| *Cj1639* | 3.559 | nifU protein homologue |
| *Cj0994c* | 4.255 | ornithine carbamoyltransferase |
| *Cj0528c* | 4.484 | flagella basal-body rod protein |
| *Cj1465* | 4.854 | hypothetical protein Cj1465 |
| *Cj0368c* | 5.155 | transcriptional regulator CmeR |
| *Cj0449c* | 5.556 | hypothetical protein Cj0449c |
| *Cj0370* | 7.353 | 30S ribosomal protein S21 |

## References


AAS, F. E., VIK, A., VEDDE, J., KOOMEY, M. & EGGE-JACOBSEN, W. 2007. *Neisseria gonorrhoeae O*-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol Microbiol,* 65**,** 607-24.

ALLOS, B. M. 1997. Association between *Campylobacter* infection and Guillain-Barre syndrome. *J Infect Dis,* 176 Suppl 2**,** S125-8.

ALLOS, B. M. 2001. *Campylobacter jejuni* Infections: update on emerging issues and trends. *Clin Infect Dis,* 32**,** 1201-6.

ALTEKRUSE, S. F., STERN, N. J., FIELDS, P. I. & SWERDLOW, D. L. 1999. *Campylobacter jejuni*--an emerging foodborne pathogen. *Emerg Infect Dis,* 5**,** 28-35.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-10.

AMABILE-CUEVAS, C. F. & DEMPLE, B. 1991. Molecular characterization of the *soxRS* genes of *Escherichia coli*: two genes control a superoxide stress regulon. *Nucleic Acids Res,* 19**,** 4479-84.

ANDERSEN, M. T., BRONDSTED, L., PEARSON, B. M., MULHOLLAND, F., PARKER, M., PIN, C., WELLS, J. M. & INGMER, H. 2005. Diverse roles for HspR in *Campylobacter jejuni* revealed by the proteome, transcriptome and phenotypic characterization of an *hspR* mutant. *Microbiology,* 151**,** 905-15.

ARAVIND, L. & KOONIN, E. V. 1998. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem Sci,* 23**,** 469-72.

ARORA, S. K., RITCHINGS, B. W., ALMIRA, E. C., LORY, S. & RAMPHAL, R. 1997. A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in *Pseudomonas aeruginosa* in a cascade manner. *J Bacteriol,* 179**,** 5574-81.

ASBURY, A. K. & CORNBLATH, D. R. 1990. Assessment of current diagnostic criteria for Guillain-Barre syndrome. *Ann Neurol,* 27 Suppl**,** S21-4.

ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet,* 25**,** 25-9.

ASPINALL, G. O., MCDONALD, A. G., PANG, H., KURJANCZYK, L. A. & PENNER, J. L. 1994. Lipopolysaccharides of *Campylobacter jejuni* serotype O:19: structures of core oligosaccharide regions from the serostrain and two bacterial isolates from patients with the Guillain-Barre syndrome. *Biochemistry,* 33**,** 241-9.

BACHTIAR, B. M., COLOE, P. J. & FRY, B. N. 2007. Knockout mutagenesis of the *kpsE* gene of *Campylobacter jejuni* 81116 and its involvement in bacterium-host interactions. *FEMS Immunology & Medical Microbiology,* 49**,** 149-154.

BACON, D. J., ALM, R. A., BURR, D. H., HU, L., KOPECKO, D. J., EWING, C. P., TRUST, T. J. & GUERRY, P. 2000. Involvement of a plasmid in virulence of *Campylobacter jejuni* 81-176. *Infect Immun,* 68**,** 4384-90.

BACON, D. J., ALM, R. A., HU, L., HICKEY, T. E., EWING, C. P., BATCHELOR, R. A., TRUST, T. J. & GUERRY, P. 2002. DNA sequence and mutational analyses of the pVir plasmid of *Campylobacter jejuni* 81-176. *Infect Immun,* 70**,** 6242-50.

BACON, D. J., SZYMANSKI, C. M., BURR, D. H., SILVER, R. P., ALM, R. A. & GUERRY, P. 2001. A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. *Mol Microbiol,* 40**,** 769-77.

BACON, J., JAMES, B. W., WERNISCH, L., WILLIAMS, A., MORLEY, K. A., HATCH, G. J., MANGAN, J. A., HINDS, J., STOKER, N. G., BUTCHER, P. D. & MARSH, P. D. 2004. The influence of reduced oxygen availability on pathogenicity and gene expression in Mycobacterium tuberculosis. *Tuberculosis (Edinb),* 84**,** 205-17.

BAILLON, M. L., VAN VLIET, A. H., KETLEY, J. M., CONSTANTINIDOU, C. & PENN, C. W. 1999. An iron-regulated alkyl hydroperoxide reductase (AhpC) confers aerotolerance and oxidative stress resistance to the microaerophilic pathogen *Campylobacter jejuni*. *J Bacteriol,* 181**,** 4798-804.

BALABAN, M., JOSLIN, S. N. & HENDRIXSON, D. R. 2009. FlhF and its GTPase activity are required for distinct processes in flagellar gene regulation and biosynthesis in *Campylobacter jejuni*. *J Bacteriol,* 191**,** 6602-11.

BEGLEY, M., GAHAN, C. G. & HILL, C. 2005. The interaction between bacteria and bile. *FEMS Microbiol Rev,* 29**,** 625-51.

BERGEN, M. A. P. V., PUTTEN, J. P. M. V., DINGLE, K. E., BLASER, M. J. & WAGENAAR, J. A. 2008. Isolation, identification, subspecies differenctiation, and typing of *Campylobacter fetus*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.213-225.

BERGIN, D., REEVES, E. P., RENWICK, J., WIENTJES, F. B. & KAVANAGH, K. 2005. Superoxide production in *Galleria mellonella* hemocytes: identification of proteins homologous to the NADPH oxidase complex of human neutrophils. *Infect Immun,* 73**,** 4161-70.

BIEBER, D., RAMER, S. W., WU, C. Y., MURRAY, W. J., TOBE, T., FERNANDEZ, R. & SCHOOLNIK, G. K. 1998. Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science,* 280**,** 2114-8.

BIELASZEWSKA, M., SINHA, B., KUCZIUS, T. & KARCH, H. 2005. Cytolethal distending toxin from Shiga toxin-producing *Escherichia coli* O157 causes irreversible G2/M arrest, inhibition of proliferation, and death of human endothelial cells. *Infect Immun,* 73**,** 552-62.

BISWAS, D., FERNANDO, U., REIMAN, C., WILLSON, P., POTTER, A. & ALLAN, B. 2006. Effect of cytolethal distending toxin of *Campylobacter jejuni* on adhesion and internalization in cultured cells and in colonization of the chicken gut. *Avian Dis,* 50**,** 586-93.

BISWAS, D., ITOH, K. & SASAKAWA, C. 2003. Role of microfilaments and microtubules in the invasion of INT-407 cells by Campylobacter jejuni. *Microbiol Immunol,* 47**,** 469-73.

BLACK, R. E., LEVINE, M. M., CLEMENTS, M. L., HUGHES, T. P. & BLASER, M. J. 1988. Experimental *Campylobacter jejuni* infection in humans. *J Infect Dis,* 157**,** 472-9.

BLASER, M. J. 1997. Epidemiologic and clinical features of *Campylobacter jejuni* infections. *J Infect Dis,* 176 Suppl 2**,** S103-5.

BLASER, M. J. & ENGBERG, J. 2008. Clinical aspects of *Campylobacter jejuni* and *Campylobacter coli* infections. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.99-121.

BLASER, M. J., HOPKINS, J. A., BERKA, R. M., VASIL, M. L. & WANG, W. L. 1983. Identification and characterization of Campylobacter jejuni outer membrane proteins. *Infect Immun,* 42**,** 276-84.

BOGDAN, C., ROLLINGHOFF, M. & DIEFENBACH, A. 2000. Reactive oxygen and reactive nitrogen intermediates in innate and specific immunity. *Curr Opin Immunol,* 12**,** 64-76.

BOLOTIN, A., QUINQUIS, B., SOROKIN, A. & EHRLICH, S. D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology,* 151**,** 2551-61.

BOMAN, H. G. & HULTMARK, D. 1987. Cell-free immunity in insects. *Annu Rev Microbiol,* 41**,** 103-26.

BORK, P. 2000. Powers and pitfalls in sequence analysis: the 70% hurdle. *Genome Res,* 10**,** 398-400.

BRAS, A. M. & KETLEY, J. M. 1999. Transcellular translocation of *Campylobacter jejuni* across human polarised epithelial monolayers. *FEMS Microbiol Lett,* 179**,** 209-15.

BREN, A. & EISENBACH, M. 2001. Changing the direction of flagellar rotation in bacteria by modulating the ratio between the rotational states of the switch protein FliM. *J Mol Biol,* 312**,** 699-709.

BRENT, M. R. 2008. Steady progress and recent breakthroughs in the accuracy of automated genome annotation. *Nat Rev Genet,* 9**,** 62-73.

BRONDSTED, L., ANDERSEN, M. T., PARKER, M., JORGENSEN, K. & INGMER, H. 2005. The HtrA protease of *Campylobacter jejuni* is required for heat and oxygen tolerance and for optimal interaction with human epithelial cells. *Appl Environ Microbiol,* 71**,** 3205-12.

BULYK, M. L., GENTALEN, E., LOCKHART, D. J. & CHURCH, G. M. 1999. Quantifying DNA-protein interactions by double-stranded DNA arrays. *Nat Biotechnol,* 17**,** 573-7.

BUSWELL, C. M., HERLIHY, Y. M., LAWRENCE, L. M., MCGUIGGAN, J. T., MARSH, P. D., KEEVIL, C. W. & LEACH, S. A. 1998. Extended survival and persistence of *Campylobacter* spp. in water and aquatic biofilms and their detection by immunofluorescent-antibody and -rRNA staining. *Appl Environ Microbiol,* 64**,** 733-41.

CABISCOL, E., TAMARIT, J. & ROS, J. 2000. Oxidative stress in bacteria and protein damage by reactive oxygen species. *Int Microbiol,* 3**,** 3-8.

CALVA, J. J., RUIZ-PALACIOS, G. M., LOPEZ-VIDAL, A. B., RAMOS, A. & BOJALIL, R. 1988. Cohort study of intestinal infection with campylobacter in Mexican children. *Lancet,* 1**,** 503-6.

CAMUS, J. C., PRYOR, M. J., MEDIGUE, C. & COLE, S. T. 2002. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology,* 148**,** 2967-73.

CANDON, H. L., ALLAN, B. J., FRALEY, C. D. & GAYNOR, E. C. 2007. Polyphosphate kinase 1 is a pathogenesis determinant in *Campylobacter jejuni*. *J Bacteriol,* 189**,** 8099-108.

CARRILLO, C. D., TABOADA, E., NASH, J. H., LANTHIER, P., KELLY, J., LAU, P. C., VERHULP, R., MYKYTCZUK, O., SY, J., FINDLAY, W. A., AMOAKO, K., GOMIS, S., WILLSON, P., AUSTIN, J. W., POTTER, A., BABIUK, L., ALLAN, B. & SZYMANSKI, C. M. 2004. Genome-wide expression analyses of Campylobacter jejuni NCTC11168 reveals coordinate regulation of motility and virulence by flhA. *J Biol Chem,* 279**,** 20327-38.

CHABAN, B., VOISIN, S., KELLY, J., LOGAN, S. M. & JARRELL, K. F. 2006. Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae N*-linked glycans: insight into *N*-linked glycosylation pathways in Archaea. *Mol Microbiol,* 61**,** 259-68.

CHAMPION, O. L., COOPER, I. A., JAMES, S. L., FORD, D., KARLYSHEV, A., WREN, B. W., DUFFIELD, M., OYSTON, P. C. & TITBALL, R. W. 2009. *Galleria mellonella* as an alternative infection model for *Yersinia pseudotuberculosis*. *Microbiology,* 155**,** 1516-22.

CHANG, C. & MILLER, J. F. 2006. Campylobacter jejuni colonization of mice with limited enteric flora. *Infect Immun,* 74**,** 5261-71.

CHILCOTT, G. S. & HUGHES, K. T. 2000. Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar typhimurium and *Escherichia coli*. *Microbiol Mol Biol Rev,* 64**,** 694-708.

CLAPHAM, D. E. 1995. Calcium signaling. *Cell,* 80**,** 259-68.

CLARK, C. G., PRICE, L., AHMED, R., WOODWARD, D. L., MELITO, P. L., RODGERS, F. G., JAMIESON, F., CIEBIN, B., LI, A. & ELLIS, A. 2003. Characterization of waterborne outbreak-associated Campylobacter jejuni, Walkerton, Ontario. *Emerg Infect Dis,* 9**,** 1232-41.

COKER, A. O., ISOKPEHI, R. D., THOMAS, B. N., AMISU, K. O. & OBI, C. L. 2002. Human campylobacteriosis in developing countries. *Emerg Infect Dis,* 8**,** 237-44.

COLEGIO, O. R., GRIFFIN, T. J. T., GRINDLEY, N. D. & GALAN, J. E. 2001. In vitro transposition system for efficient generation of random mutants of Campylobacter jejuni. *J Bacteriol,* 183**,** 2384-8.

CORCIONIVOSCHI, N., CLYNE, M., LYONS, A., ELMI, A., GUNDOGDU, O., WREN, B. W., DORRELL, N., KARLYSHEV, A. V. & BOURKE, B. 2009. *Campylobacter jejuni* alters surface capsular polysaccharide when co-cultured with epithelial cells. *Infect Immun*.

COSTERTON, J. W., LEWANDOWSKI, Z., CALDWELL, D. E., KORBER, D. R. & LAPPIN-SCOTT, H. M. 1995. Microbial biofilms. *Annu Rev Microbiol,* 49**,** 711-45.

CRUSHELL, E., HARTY, S., SHARIF, F. & BOURKE, B. 2004. Enteric *Campylobacter*: purging its secrets? *Pediatr Res,* 55**,** 3-12.

DANDEKAR, T., HUYNEN, M., REGULA, J. T., UEBERLE, B., ZIMMERMANN, C. U., ANDRADE, M. A., DOERKS, T., SANCHEZ-PULIDO, L., SNEL, B., SUYAMA, M., YUAN, Y. P., HERRMANN, R. & BORK, P. 2000. Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. *Nucleic Acids Res,* 28**,** 3278-88.

DASGUPTA, N., ARORA, S. K. & RAMPHAL, R. 2000. *fleN*, a gene that regulates flagellar number in *Pseudomonas aeruginosa*. *J Bacteriol,* 182**,** 357-64.

DASGUPTA, N., WOLFGANG, M. C., GOODMAN, A. L., ARORA, S. K., JYOT, J., LORY, S. & RAMPHAL, R. 2003. A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa*. *Mol Microbiol,* 50**,** 809-24.

DAY, W. A., JR., SAJECKI, J. L., PITTS, T. M. & JOENS, L. A. 2000. Role of catalase in *Campylobacter jejuni* intracellular survival. *Infect Immun,* 68**,** 6337-45.

DE BOLLE, X., BAYLISS, C. D., FIELD, D., VAN DE VEN, T., SAUNDERS, N. J., HOOD, D. W. & MOXON, E. R. 2000. The length of a tetranucleotide repeat tract in Haemophilus influenzae determines the phase variation rate of a gene with homology to type III DNA methyltransferases. *Mol Microbiol,* 35**,** 211-22.

DE MELO, M. A., GABBIANI, G. & PECHERE, J. C. 1989. Cellular events and intracellular survival of *Campylobacter jejuni* during infection of HEp-2 cells. *Infect Immun,* 57**,** 2214-22.

DEBRUYNE, L., GEVERS, D. & VANDAMME, P. 2008. Taxonomy of the family *Campylobacteraceae*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.3-25.

DEDISTE, A., VANDENBERG, O., VLAES, L., EBRAERT, A., DOUAT, N., BAHWERE, P. & BUTZLER, J. P. 2003. Evaluation of the ProSpecT Microplate Assay for detection of *Campylobacter*: a routine laboratory perspective. *Clin Microbiol Infect,* 9**,** 1085-90.

DUNLOP, S. P., JENKINS, D., NEAL, K. R. & SPILLER, R. C. 2003. Relative importance of enterochromaffin cell hyperplasia, anxiety, and depression in postinfectious IBS. *Gastroenterology,* 125**,** 1651-9.

EISENBACH, M. 1996. Control of bacterial chemotaxis. *Mol Microbiol,* 20**,** 903-10.

ELVERS, K. T., TURNER, S. M., WAINWRIGHT, L. M., MARSDEN, G., HINDS, J., COLE, J. A., POOLE, R. K., PENN, C. W. & PARK, S. F. 2005. NssR, a member of the Crp-Fnr superfamily from *Campylobacter jejuni*, regulates a nitrosative stress-responsive regulon that includes both a single-domain and a truncated haemoglobin. *Mol Microbiol,* 57**,** 735-50.

ELVERS, K. T., WU, G., GILBERTHORPE, N. J., POOLE, R. K. & PARK, S. F. 2004. Role of an inducible single-domain hemoglobin in mediating resistance to nitric oxide and nitrosative stress in *Campylobacter jejuni* and *Campylobacter coli*. *J Bacteriol,* 186**,** 5332-41.

ENGBERG, J., BANG, D. D., AABENHUS, R., AARESTRUP, F. M., FUSSING, V. & GERNER-SMIDT, P. 2005. *Campylobacter concisus*: an evaluation of certain phenotypic and genotypic characteristics. *Clin Microbiol Infect,* 11**,** 288-95.

ENTREZ PUBMED, N. I. H., USA. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi [Accessed].

ESCHERICH, T. 1886. Beiträge zur Kenntniss der Darmbacterien. III Über das Vorkommen van Vibrionen im Darmcanal und den Stuhlgängen der Säuglinge. *Münch. Med. Wschr. ,* 33**,** 815-817, 833-835.

EVEREST, P. H., GOOSSENS, H., BUTZLER, J. P., LLOYD, D., KNUTTON, S., KETLEY, J. M. & WILLIAMS, P. H. 1992. Differentiated Caco-2 cells as a model for enteric invasion by *Campylobacter jejuni* and *C. coli*. *J Med Microbiol,* 37**,** 319-25.

EWING, C. P., ANDREISHCHEVA, E. & GUERRY, P. 2009. Functional characterization of flagellin glycosylation in *Campylobacter jejuni* 81-176. *J Bacteriol,* 191**,** 7086-93.

FALKE, J. J., BASS, R. B., BUTLER, S. L., CHERVITZ, S. A. & DANIELSON, M. A. 1997. The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu Rev Cell Dev Biol,* 13**,** 457-512.

FALQUET, L., PAGNI, M., BUCHER, P., HULO, N., SIGRIST, C. J., HOFMANN, K. & BAIROCH, A. 2002. The PROSITE database, its status in 2002. *Nucleic Acids Res,* 30**,** 235-8.

FARR, S. B. & KOGOMA, T. 1991. Oxidative stress responses in Escherichia coli and Salmonella typhimurium. *Microbiol Rev,* 55**,** 561-85.

FINLAY, B. B. & FALKOW, S. 1990. *Salmonella* interactions with polarized human intestinal Caco-2 epithelial cells. *J Infect Dis,* 162**,** 1096-106.

FISHER, M. 1956. An unusual variant of acute idiopathic polyneuritis (syndrome of ophthalmoplegia, ataxia and areflexia). *N Engl J Med,* 255**,** 57-65.

FITZGERALD, C., WHICHARD, J. & NACHAMKIN, I. 2008. Diagnosis and antimicrobial susceptibility of *Campylobacter* species. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.227-243.

FLEISCHMANN, R. D., ADAMS, M. D., WHITE, O., CLAYTON, R. A., KIRKNESS, E. F., KERLAVAGE, A. R., BULT, C. J., TOMB, J. F., DOUGHERTY, B. A.,

MERRICK, J. M. & ET AL. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science,* 269**,** 496-512.

FLORIN, I. & ANTILLON, F. 1992. Production of enterotoxin and cytotoxin in *Campylobacter jejuni* strains isolated in Costa Rica. *J Med Microbiol,* 37**,** 22-9.

FOUTS, D. E., MONGODIN, E. F., MANDRELL, R. E., MILLER, W. G., RASKO, D. A., RAVEL, J., BRINKAC, L. M., DEBOY, R. T., PARKER, C. T., DAUGHERTY, S. C., DODSON, R. J., DURKIN, A. S., MADUPU, R., SULLIVAN, S. A., SHETTY, J. U., AYODEJI, M. A., SHVARTSBEYN, A., SCHATZ, M. C., BADGER, J. H., FRASER, C. M. & NELSON, K. E. 2005. Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. *PLoS Biol,* 3**,** e15.

FOX, J. G., ROGERS, A. B., WHARY, M. T., GE, Z., TAYLOR, N. S., XU, S., HORWITZ, B. H. & ERDMAN, S. E. 2004. Gastroenteritis in NF-kappaB-deficient mice is produced with wild-type *Camplyobacter jejuni* but not with *C. jejuni* lacking cytolethal distending toxin despite persistent colonization with both strains. *Infect Immun,* 72**,** 1116-25.

FRIEDBERG, I. 2006. Automated protein function prediction--the genomic challenge. *Brief Bioinform,* 7**,** 225-42.

FRIEDMAN, C. R., HOEKSTRA, R. M., SAMUEL, M., MARCUS, R., BENDER, J., SHIFERAW, B., REDDY, S., AHUJA, S. D., HELFRICK, D. L., HARDNETT, F., CARTER, M., ANDERSON, B. & TAUXE, R. V. 2004. Risk factors for sporadic *Campylobacter* infection in the United States: A case-control study in FoodNet sites. *Clin Infect Dis,* 38 Suppl 3**,** S285-96.

FRIEDMAN, C. R., NEIMANN, J., WEGENER, H. C. & TAUXE, R. V. 2000. Epidemioogy of *Camylobacter jejuni* infections in the United States and other industrialized nations. In: I. Nachamkin and M.J. Blaser, Editors, *Campylobacter* (edn 2). pp. 121–138.

FRIIS, L. M., KEELAN, M. & TAYLOR, D. E. 2009. Campylobacter jejuni drives MyD88-independent interleukin-6 secretion via Toll-like receptor 2. *Infect Immun,* 77**,** 1553-60.

FRISAN, T., CORTES-BRATTI, X. & THELESTAM, M. 2002. Cytolethal distending toxins and activation of DNA damage-dependent checkpoint responses. *Int J Med Microbiol,* 291**,** 495-9.

GALAN, J. E., GINOCCHIO, C. & COSTEAS, P. 1992. Molecular and functional characterization of the *Salmonella* invasion gene invA: homology of InvA to members of a new protein family. *J Bacteriol,* 174**,** 4338-49.

GARENAUX, A., JUGIAU, F., RAMA, F., DE JONGE, R., DENIS, M., FEDERIGHI, M. & RITZ, M. 2008. Survival of *Campylobacter jejuni* strains from different origins under oxidative stress conditions: effect of temperature. *Curr Microbiol,* 56**,** 293-7.

GARVIS, S. G., PUZON, G. J. & KONKEL, M. E. 1996. Molecular characterization of a *Campylobacter jejuni* 29-kilodalton periplasmic binding protein. *Infect Immun,* 64**,** 3537-43.

GAYNOR, E. C., WELLS, D. H., MACKICHAN, J. K. & FALKOW, S. 2005. The *Campylobacter jejuni* stringent response controls specific stress survival and virulence-associated phenotypes. *Mol Microbiol,* 56**,** 8-27.

GILBERT, M., GODSCHALK, P. C., KARWASKI, M. F., ANG, C. W., VAN BELKUM, A., LI, J., WAKARCHUK, W. W. & ENDTZ, H. P. 2004. Evidence for acquisition of the lipooligosaccharide biosynthesis locus in *Campylobacter jejuni* GB11, a strain isolated from a patient with Guillain-Barre syndrome, by horizontal exchange. *Infect Immun,* 72**,** 1162-5.

GILBERT, M., KARWASKI, M. F., BERNATCHEZ, S., YOUNG, N. M., TABOADA, E., MICHNIEWICZ, J., CUNNINGHAM, A. M. & WAKARCHUK, W. W. 2002. The genetic bases for the variation in the lipo-oligosaccharide of the mucosal pathogen, *Campylobacter jejuni*. Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide. *J Biol Chem,* 277**,** 327-37.

GILBERT, M., PARKER, C. T. & MORAN, A. P. 2008. *Campylobacter jejuni* lipooligosaccharides: structures and biosynthesis. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.483-504.

GLOVER, K. J., WEERAPANA, E. & IMPERIALI, B. 2005. *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc Natl Acad Sci U S A,* 102**,** 14255-9.

GODDE, J. S. & BICKERTON, A. 2006. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J Mol Evol,* 62**,** 718-29.

GOLDEN, N. J. & ACHESON, D. W. 2002. Identification of motility and autoagglutination Campylobacter jejuni mutants by random transposon mutagenesis. *Infect Immun,* 70**,** 1761-71.

GOON, S., KELLY, J. F., LOGAN, S. M., EWING, C. P. & GUERRY, P. 2003. Pseudaminic acid, the major modification on *Campylobacter* flagellin, is synthesized via the Cj1293 gene. *Mol Microbiol,* 50**,** 659-71.

GRANT, C. C., KONKEL, M. E., CIEPLAK, W., JR. & TOMPKINS, L. S. 1993. Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures. *Infect Immun,* 61**,** 1764-71.

GRIFFITHS-JONES, S., BATEMAN, A., MARSHALL, M., KHANNA, A. & EDDY, S. R. 2003. Rfam: an RNA family database. *Nucleic Acids Res,* 31**,** 439-41.

GUERRY, P. 2007. *Campylobacter* flagella: not just for motility. *Trends Microbiol,* 15**,** 456-61.

GUERRY, P., ALM, R. A., POWER, M. E., LOGAN, S. M. & TRUST, T. J. 1991. Role of two flagellin genes in *Campylobacter* motility. *J. Bacteriol.,* 173**,** 4757-4764.

GUERRY, P., EWING, C. P., HICKEY, T. E., PRENDERGAST, M. M. & MORAN, A. P. 2000. Sialylation of lipooligosaccharide cores affects immunogenicity and serum resistance of *Campylobacter jejuni*. *Infect Immun,* 68**,** 6656-62.

GUERRY, P., EWING, C. P., SCHIRM, M., LORENZO, M., KELLY, J., PATTARINI, D., MAJAM, G., THIBAULT, P. & LOGAN, S. 2006. Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Mol Microbiol,* 60**,** 299-311.

GUNDOGDU, O., BENTLEY, S. D., HOLDEN, M. T., PARKHILL, J., DORRELL, N. & WREN, B. W. 2007. Re-annotation and re-analysis of the *Campylobacter jejuni* NCTC11168 genome sequence. *BMC Genomics,* 8**,** 162.

GUO, B., WANG, Y., SHI, F., BARTON, Y. W., PLUMMER, P., REYNOLDS, D. L., NETTLETON, D., GRINNAGE-PULLEY, T., LIN, J. & ZHANG, Q. 2008. CmeR functions as a pleiotropic regulator and is required for optimal colonization of Campylobacter jejuni in vivo. *J Bacteriol,* 190**,** 1879-90.

HADDOCK, G., MULLIN, M., MACCALLUM, A., SHERRY, A., TETLEY, L., WATSON, E., DAGLEISH, M., SMITH, D. & EVEREST, P. 2010. Campylobacter jejuni 81-176 forms distinct microcolonies on in vitro infected human small intestinal tissue prior to biofilm formation. *Microbiology*.

HAMNER, S., MCINNERNEY, K., WILLIAMSON, K., FRANKLIN, M. J. & FORD, T. E. 2010. Bile salts influence the ordered expression of virulence genes in *Escherichia coli* O157:H7. *Society of General Microbiology, Abstract ED12/12*.

HARRIS, N. V., KIMBALL, T. J., BENNETT, P., JOHNSON, Y., WAKELY, D. & NOLAN, C. M. 1987. *Campylobacter jejuni* enteritis associated with raw goat's milk. *Am J Epidemiol,* 126**,** 179-86.

HARRISON, P. M. & GERSTEIN, M. 2002. Studying genomes through the aeons: protein families, pseudogenes and proteome evolution. *J Mol Biol,* 318**,** 1155-74.

HARVEY, P., BATTLE, T. & LEACH, S. 1999. Different invasion phenotypes of *Campylobacter* isolates in Caco-2 cell monolayers. *J Med Microbiol,* 48**,** 461-9.

HEINRICH, P. C., CASTELL, J. V. & ANDUS, T. 1990. Interleukin-6 and the acute phase response. *Biochem J,* 265**,** 621-36.

HENDRIXSON, D. R. 2006. A phase-variable mechanism controlling the *Campylobacter jejuni* FlgR response regulator influences commensalism. *Mol Microbiol,* 61**,** 1646-59.

HENDRIXSON, D. R. 2008. Regulation of flagellar gene expression and assembly. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.545-557.

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**,** 214-24.

HENDRIXSON, D. R. & DIRITA, V. J. 2003. Transcription of sigma54-dependent but not sigma28-dependent flagellar genes in *Campylobacter jejuni* is associated with formation of the flagellar secretory apparatus. *Mol Microbiol,* 50**,** 687-702.

HENDRIXSON, D. R. & DIRITA, V. J. 2004. Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Mol Microbiol,* 52**,** 471-84.

HICKEY, T. E., MCVEIGH, A. L., SCOTT, D. A., MICHIELUTTI, R. E., BIXBY, A., CARROLL, S. A., BOURGEOIS, A. L. & GUERRY, P. 2000. *Campylobacter jejuni* cytolethal distending toxin mediates release of interleukin-8 from intestinal epithelial cells. *Infect Immun,* 68**,** 6535-41.

HIGHWIRE PRESS, S. U. Available: http://highwire.stanford.edu/ [Accessed].

HINDIYEH, M., JENSE, S., HOHMANN, S., BENETT, H., EDWARDS, C., ALDEEN, W., CROFT, A., DALY, J., MOTTICE, S. & CARROLL, K. C. 2000. Rapid detection of *Campylobacter jejuni* in stool specimens by an enzyme immunoassay and surveillance for *Campylobacter upsaliensis* in the greater Salt Lake City area. *J Clin Microbiol,* 38**,** 3076-9.

HITCHEN, P., BRZOSTEK, J., PANICO, M., BUTLER, J. A., MORRIS, H. R., DELL, A. & LINTON, D. 2010. Modification of the *Campylobacter jejuni* flagellin glycan by the product of the Cj1295 homopolymeric tract containing gene. *Microbiology*.

HOBBIE, S., CHEN, L. M., DAVIS, R. J. & GALAN, J. E. 1997. Involvement of mitogen-activated protein kinase pathways in the nuclear responses and cytokine production induced by *Salmonella typhimurium* in cultured intestinal epithelial cells. *J Immunol,* 159**,** 5550-9.

HOFREUTER, D., TSAI, J., WATSON, R. O., NOVIK, V., ALTMAN, B., BENITEZ, M., CLARK, C., PERBOST, C., JARVIE, T., DU, L. & GALAN, J. E. 2006. Unique features of a highly pathogenic *Campylobacter jejuni* strain. *Infect Immun,* 74**,** 4694-707.

HOMMA, K., FUKUCHI, S., KAWABATA, T., OTA, M. & NISHIKAWA, K. 2002. A systematic investigation identifies a significant number of probable pseudogenes in the Escherichia coli genome. *Gene,* 294**,** 25-33.

HORVATH, P. & BARRANGOU, R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science,* 327**,** 167-70.

HU, L. & KOPECKO, D. J. 1999. *Campylobacter jejuni* 81-176 associates with microtubules and dynein during invasion of human intestinal cells. *Infect Immun*, 67**,** 4171-82.

HU, L. & KOPECKO, D. J. 2008. Cell biology of human host cell entry by *Campylobacter jejuni*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.297-313.

HU, L., MCDANIEL, J. P. & KOPECKO, D. J. 2006. Signal transduction events involved in human epithelial cell invasion by *Campylobacter jejuni* 81-176. *Microb Pathog,* 40**,** 91-100.

HU, L., TALL, B. D., CURTIS, S. K. & KOPECKO, D. J. 2008. Enhanced microscopic definition of *Campylobacter jejun*i 81-176 adherence to, invasion of, translocation across, and exocytosis from polarized human intestinal Caco-2 cells. *Infect Immun,* 76**,** 5294-304.

HUGHES, K. T., GILLEN, K. L., SEMON, M. J. & KARLINSEY, J. E. 1993. Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science,* 262**,** 1277-80.

JACOB, R. 1990. Calcium oscillations in electrically non-excitable cells. *Biochim Biophys Acta,* 1052**,** 427-38.

JACOBS, B. C., BELKUM, A. V. & ENDTZ, H. P. 2008. Guillain-Barre syndrome and *Campylobacter* Infection. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.245-261.

JAGANNATHAN, A., CONSTANTINIDOU, C. & PENN, C. W. 2001. Roles of *rpoN*, *fliA*, and *flgR* in expression of flagella in *Campylobacter jejuni*. *J Bacteriol,* 183**,** 2937-42.

JAHREIS, K., MORRISON, T. B., GARZON, A. & PARKINSON, J. S. 2004. Chemotactic signaling by an Escherichia coli CheA mutant that lacks the binding domain for phosphoacceptor partners. *J Bacteriol,* 186**,** 2664-72.

JANSEN, R., VAN EMBDEN, J. D., GAASTRA, W. & SCHOULS, L. M. 2002. Identification of a novel family of sequence repeats among prokaryotes. *Omics,* 6**,** 23-33.

JERRETT, I. V., MCORIST, S., WADDINGTON, J., BROWNING, J. W., MALECKI, J. C. & MCCAUSLAND, I. P. 1984. Diagnostic studies of the fetus, placenta and maternal blood from 265 bovine abortions. *Cornell Vet,* 74**,** 8-20.

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**,** 1225-36.

JIN, S., SONG, Y. C., EMILI, A., SHERMAN, P. M. & CHAN, V. L. 2003. JlpA of *Campylobacter jejuni* interacts with surface-exposed heat shock protein 90alpha and triggers signalling pathways leading to the activation of NF-kappaB and p38 MAP kinase in epithelial cells. *Cell Microbiol,* 5**,** 165-74.

JONES, K. 2001. The Campylobacter conundrum. *Trends Microbiol,* 9**,** 365-6.

JONES, M. A., MARSTON, K. L., WOODALL, C. A., MASKELL, D. J., LINTON, D., KARLYSHEV, A. V., DORRELL, N., WREN, B. W. & BARROW, P. A. 2004. Adaptation of *Campylobacter jejuni* NCTC11168 to high-level colonization of the avian gastrointestinal tract. *Infect Immun,* 72**,** 3769-76.

JOSHUA, G. W., GUTHRIE-IRONS, C., KARLYSHEV, A. V. & WREN, B. W. 2006. Biofilm formation in *Campylobacter jejuni*. *Microbiology,* 152**,** 387-96.

KAKUDA, T. & DIRITA, V. J. 2006. Cj1496c encodes a *Campylobacter jejuni* glycoprotein that influences invasion of human epithelial cells and colonization of the chick gastrointestinal tract. *Infect Immun,* 74**,** 4715-23.

KAMAL, N., DORRELL, N., JAGANNATHAN, A., TURNER, S. M., CONSTANTINIDOU, C., STUDHOLME, D. J., MARSDEN, G., HINDS, J.,

LAING, K. G., WREN, B. W. & PENN, C. W. 2007. Deletion of a previously uncharacterized flagellar-hook-length control gene *fliK* modulates the sigma54-dependent regulon in *Campylobacter jejuni*. *Microbiology,* 153**,** 3099-111.

KARLINSEY, J. E., TANAKA, S., BETTENWORTH, V., YAMAGUCHI, S., BOOS, W., AIZAWA, S. I. & HUGHES, K. T. 2000. Completion of the hook-basal body complex of the *Salmonella typhimurium* flagellum is coupled to FlgM secretion and *fliC* transcription. *Mol Microbiol,* 37**,** 1220-31.

KARLYSHEV, A. V., CHAMPION, O. L., CHURCHER, C., BRISSON, J. R., JARRELL, H. C., GILBERT, M., BROCHU, D., ST MICHAEL, F., LI, J., WAKARCHUK, W. W., GOODHEAD, I., SANDERS, M., STEVENS, K., WHITE, B., PARKHILL, J., WREN, B. W. & SZYMANSKI, C. M. 2005a. Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol Microbiol,* 55**,** 90-103.

KARLYSHEV, A. V., CHAMPION, O. L., JOSHUA, G. W. P. & WREN, B. W. 2005b. The polysaccharide capsule of *Campylobacter jejuni*. In *Campylobacter* Molecular and Cellular Biology, J.M. Ketley, and M.E. Konkel, eds (Horizon Bioscience), pp.249 - 258.

KARLYSHEV, A. V., EVEREST, P., LINTON, D., CAWTHRAW, S., NEWELL, D. G. & WREN, B. W. 2004. The *Campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology,* 150**,** 1957-64.

KARLYSHEV, A. V., LINTON, D., GREGSON, N. A., LASTOVICA, A. J. & WREN, B. W. 2000. Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. *Mol Microbiol,* 35**,** 529-41.

KARLYSHEV, A. V., LINTON, D., GREGSON, N. A. & WREN, B. W. 2002. A novel paralogous gene family involved in phase-variable flagella-mediated motility in *Campylobacter jejuni*. *Microbiology,* 148**,** 473-80.

KARLYSHEV, A. V., MCCROSSAN, M. V. & WREN, B. W. 2001. Demonstration of polysaccharide capsule in *Campylobacter jejuni* using electron microscopy. *Infect Immun,* 69**,** 5921-4.

KARLYSHEV, A. V., WREN, B. W. & MORAN, A. P. 2008. *Campylobacter jejuni* capsular polysaccharide. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.505-522.

KARMALI, M. A. & FLEMING, P. C. 1979. *Campylobacter* enteritis in children. *J Pediatr,* 94**,** 527-33.

KELLY, J., JARRELL, H., MILLAR, L., TESSIER, L., FIORI, L. M., LAU, P. C., ALLAN, B. & SZYMANSKI, C. M. 2006. Biosynthesis of the N-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J Bacteriol,* 188**,** 2427-34.

KETLEY, J., GUERRY, P. & PANIGRAHI, P. 1996. Pathogenic mechanisms. In *Campylobacter*, *Helicobacters,* and Related Organisms, D.G. Newell, J.M. Ketley, and R.aA. Feldman, eds (New York, Plenum Press), pp.537-544.

KETLEY, J. M. 1997. Pathogenesis of enteric infection by *Campylobacter*. *Microbiology,* 143 ( Pt 1)**,** 5-21.

KLOSE, K. E. & MEKALANOS, J. J. 1998. Distinct roles of an alternative sigma factor during both free-swimming and colonizing phases of the *Vibrio cholerae* pathogenic cycle. *Mol Microbiol,* 28**,** 501-20.

KOGA, M., GILBERT, M., LI, J., KOIKE, S., TAKAHASHI, M., FURUKAWA, K., HIRATA, K. & YUKI, N. 2005. Antecedent infections in Fisher syndrome: a common pathogenesis of molecular mimicry. *Neurology,* 64**,** 1605-11.

KONKEL, M. E., GRAY, S. A., KIM, B. J., GARVIS, S. G. & YOON, J. 1999a. Identification of the enteropathogens *Campylobacter jejuni* and *Campylobacter coli* based on the *cadF* virulence gene and its product. *J Clin Microbiol,* 37**,** 510-7.

KONKEL, M. E., HAYES, S. F., JOENS, L. A. & CIEPLAK, W., JR. 1992a. Characteristics of the internalization and intracellular survival of *Campylobacter jejuni* in human epithelial cell cultures. *Microb Pathog,* 13**,** 357-70.

KONKEL, M. E. & JOENS, L. A. 1989. Adhesion to and invasion of HEp-2 cells by *Campylobacter* spp. *Infect Immun,* 57**,** 2984-90.

KONKEL, M. E., KIM, B. J., RIVERA-AMILL, V. & GARVIS, S. G. 1999b. Bacterial secreted proteins are required for the internaliztion of *Campylobacter jejuni* into cultured mammalian cells. *Mol Microbiol,* 32**,** 691-701.

KONKEL, M. E., KLENA, J. D., RIVERA-AMILL, V., MONTEVILLE, M. R., BISWAS, D., RAPHAEL, B. & MICKELSON, J. 2004. Secretion of virulence proteins from *Campylobacter jejuni* is dependent on a functional flagellar export apparatus. *J Bacteriol,* 186**,** 3296-303.

KONKEL, M. E., MEAD, D. J., HAYES, S. F. & CIEPLAK, W., JR. 1992b. Translocation of *Campylobacter jejuni* across human polarized epithelial cell monolayer cultures. *J Infect Dis,* 166**,** 308-15.

KORLATH, J. A., OSTERHOLM, M. T., JUDY, L. A., FORFANG, J. C. & ROBINSON, R. A. 1985. A point-source outbreak of campylobacteriosis associated with consumption of raw milk. *J Infect Dis,* 152**,** 592-6.

KOROLIK, V. & KETLEY, J. 2008. Chemosensory signal transduction pathway of *Campylobacter jejuni*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.351-366.

KOVATS, R. S., EDWARDS, S. J., CHARRON, D., COWDEN, J., D'SOUZA, R. M., EBI, K. L., GAUCI, C., GERNER-SMIDT, P., HAJAT, S., HALES, S., HERNANDEZ PEZZI, G., KRIZ, B., KUTSAR, K., MCKEOWN, P., MELLOU, K., MENNE, B., O'BRIEN, S., VAN PELT, W. & SCHMID, H. 2005. Climate variability and campylobacter infection: an international study. *Int J Biometeorol,* 49**,** 207-14.

KOWARIK, M., YOUNG, N. M., NUMAO, S., SCHULZ, B. L., HUG, I., CALLEWAERT, N., MILLS, D. C., WATSON, D. C., HERNANDEZ, M., KELLY, J. F., WACKER, M. & AEBI, M. 2006. Definition of the bacterial *N*-glycosylation site consensus sequence. *Embo J,* 25**,** 1957-66.

KUROKI, S., SAIDA, T., NUKINA, M., HARUTA, T., YOSHIOKA, M., KOBAYASHI, Y. & NAKANISHI, H. 1993. *Campylobacter jejuni* strains from patients with Guillain-Barre syndrome belong mostly to Penner serogroup 19 and contain beta-N-acetylglucosamine residues. *Ann Neurol,* 33**,** 243-7.

KUTSUKAKE, K., OHYA, Y. & IINO, T. 1990. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. *J Bacteriol,* 172**,** 741-7.

LANGDON, R. H., CUCCUI, J. & WREN, B. W. 2009. N-linked glycosylation in bacteria: an unexpected application. *Future Microbiol,* 4**,** 401-12.

LARA-TEJERO, M. & GALAN, J. E. 2001. CdtA, CdtB, and CdtC form a tripartite complex that is required for cytolethal distending toxin activity. *Infect Immun,* 69**,** 4358-65.

LARSON, C. L., CHRISTENSEN, J. E., PACHECO, S. A., MINNICH, S. A. & KONKEL, M. E. 2008. *Campylobacter jejuni* secretes proteins via the flagellar type III secretion system that contribute to host cell invasion and gastroenteritis. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.315-332.

LASTOVICA, A. J. & ALLOS, B. M. 2008. Clinical significance of *Campylobacter* and reltated species other than *Campylobacter jejuni* and *Campylobacter coli*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.123-149.

LAVINE, M. D. & STRAND, M. R. 2002. Insect hemocytes and their role in immunity. *Insect Biochem Mol Biol,* 32**,** 1295-309.

LERAT, E. & OCHMAN, H. 2004. Psi-Phi: exploring the outer limits of bacterial pseudogenes. *Genome Res,* 14**,** 2273-8.

LERAT, E. & OCHMAN, H. 2005. Recognizing the pseudogenes in bacterial genomes. *Nucleic Acids Res,* 33**,** 3125-32.

LIN, A. E., KRASTEL, K., HOBB, R. I., THOMPSON, S. A., CVITKOVITCH, D. G. & GAYNOR, E. C. 2009. Atypical roles for *Campylobacter jejuni* amino acid ATP binding cassette transporter components PaqP and PaqQ in bacterial stress tolerance and pathogen-host cell dynamics. *Infect Immun,* 77**,** 4912-24.

LIN, J., AKIBA, M., SAHIN, O. & ZHANG, Q. 2005. CmeR functions as a transcriptional repressor for the multidrug efflux pump CmeABC in *Campylobacter jejuni*. *Antimicrob Agents Chemother,* 49**,** 1067-75.

LIN, J., MICHEL, L. O. & ZHANG, Q. 2002. CmeABC functions as a multidrug efflux system in *Campylobacter jejuni*. *Antimicrob Agents Chemother,* 46**,** 2124-31.

LIN, J., SAHIN, O., MICHEL, L. O. & ZHANG, Q. 2003. Critical role of multidrug efflux pump CmeABC in bile resistance and in vivo colonization of *Campylobacter jejuni*. *Infect Immun,* 71**,** 4250-9.

LINTON, D., DORRELL, N., HITCHEN, P. G., AMBER, S., KARLYSHEV, A. V., MORRIS, H. R., DELL, A., VALVANO, M. A., AEBI, M. & WREN, B. W. 2005. Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. *Mol Microbiol,* 55**,** 1695-703.

LINTON, D., KARLYSHEV, A. V., HITCHEN, P. G., MORRIS, H. R., DELL, A., GREGSON, N. A. & WREN, B. W. 2000. Multiple *N*-acetyl neuraminic acid synthetase (*neuB*) genes in *Campylobacter jejuni*: identification and characterization of the gene involved in sialylation of lipo-oligosaccharide. *Mol Microbiol,* 35**,** 1120-34.

LINTON, D., KARLYSHEV, A. V. & WREN, B. W. 2001. Deciphering *Campylobacter jejuni* cell surface interactions from the genome sequence. *Curr Opin Microbiol,* 4**,** 35-40.

LIU, F. & TANNER, M. E. 2006. PseG of Pseudaminic Acid Biosynthesis: A UDP-SUGAR HYDROLASE AS A MASKED GLYCOSYLTRANSFERASE. *J Biol Chem,* 281**,** 20902-9.

LIU, R. & OCHMAN, H. 2007. Origins of flagellar gene operons and secondary flagellar systems. *J Bacteriol,* 189**,** 7098-104.

LOGAN, S. M. 2006. Flagellar glycosylation - a new component of the motility repertoire? *Microbiology,* 152**,** 1249-62.

LOGAN, S. M., KELLY, J. F., THIBAULT, P., EWING, C. P. & GUERRY, P. 2002. Structural heterogeneity of carbohydrate modifications affects serospecificity of *Campylobacter* flagellins. *Mol Microbiol,* 46**,** 587-97.

LOGAN, S. M., SCHOENHOFEN, I. C. & GUERRY, P. 2008. *O*-linked flagellar glycosylation in *Campylobacter*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.471-481.

MACCALLUM, A., HADDOCK, G. & EVEREST, P. H. 2005a. *Campylobacter jejuni* activates mitogen-activated protein kinases in Caco-2 cell monolayers and in vitro infected primary human colonic tissue. *Microbiology-Sgm,* 151**,** 2765-2772.

MACCALLUM, A., HARDY, S. P. & EVEREST, P. H. 2005b. *Campylobacter jejuni* inhibits the absorptive transport functions of Caco-2 cells and disrupts cellular tight junctions. *Microbiology,* 151**,** 2451-8.

MACCALLUM, A. J., HARRIS, D., HADDOCK, G. & EVEREST, P. H. 2006. *Campylobacter jejuni*-infected human epithelial cell lines vary in their ability to secrete interleukin-8 compared to in vitro-infected primary human intestinal tissue. *Microbiology,* 152**,** 3661-5.

MACIVER, I. & HANSEN, E. J. 1996. Lack of expression of the global regulator OxyR in Haemophilus influenzae has a profound effect on growth phenotype. *Infect Immun,* 64**,** 4618-29.

MACKICHAN, J. K., GAYNOR, E. C., CHANG, C., CAWTHRAW, S., NEWELL, D. G., MILLER, J. F. & FALKOW, S. 2004. The *Campylobacter jejuni dccRS* two-component system is required for optimal *in vivo* colonization but is dispensable for *in vitro* growth. *Mol Microbiol,* 54**,** 1269-86.

MALIK-KALE, P., PARKER, C. T. & KONKEL, M. E. 2008. Culture of *Campylobacter jejuni* with sodium deoxycholate induces virulence gene expression. *J Bacteriol,* 190**,** 2286-97.

MARCHANT, J., WREN, B. & KETLEY, J. 2002. Exploiting genome sequence: predictions for mechanisms of *Campylobacter* chemotaxis. *Trends Microbiol,* 10**,** 155-9.

MARGULIES, M., EGHOLM, M., ALTMAN, W. E., ATTIYA, S., BADER, J. S., BEMBEN, L. A., BERKA, J., BRAVERMAN, M. S., CHEN, Y. J., CHEN, Z., DEWELL, S. B., DU, L., FIERRO, J. M., GOMES, X. V., GODWIN, B. C., HE, W., HELGESEN, S., HO, C. H., IRZYK, G. P., JANDO, S. C., ALENQUER, M. L., JARVIE, T. P., JIRAGE, K. B., KIM, J. B., KNIGHT, J. R., LANZA, J. R., LEAMON, J. H., LEFKOWITZ, S. M., LEI, M., LI, J., LOHMAN, K. L., LU, H., MAKHIJANI, V. B., MCDADE, K. E., MCKENNA, M. P., MYERS, E. W., NICKERSON, E., NOBILE, J. R., PLANT, R., PUC, B. P., RONAN, M. T., ROTH, G. T., SARKIS, G. J., SIMONS, J. F., SIMPSON, J. W., SRINIVASAN, M., TARTARO, K. R., TOMASZ, A., VOGT, K. A., VOLKMER, G. A., WANG, S. H., WANG, Y., WEINER, M. P., YU, P., BEGLEY, R. F. & ROTHBERG, J. M. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* 437**,** 376-80.

MARKS, P. W. & MAXFIELD, F. R. 1990a. Local and global changes in cytosolic free calcium in neutrophils during chemotaxis and phagocytosis. *Cell Calcium,* 11**,** 181-90.

MARKS, P. W. & MAXFIELD, F. R. 1990b. Transient increases in cytosolic free calcium appear to be required for the migration of adherent human neutrophils. *J Cell Biol,* 110**,** 43-52.

MARRAFFINI, L. A. & SONTHEIMER, E. J. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet,* 11**,** 181-190.

MARSHALL, J. K., THABANE, M., GARG, A. X., CLARK, W. F., SALVADORI, M. & COLLINS, S. M. 2006. Incidence and epidemiology of irritable bowel syndrome after a large waterborne outbreak of bacterial dysentery. *Gastroenterology,* 131**,** 445-50; quiz 660.

MCFADYEAN, J. & STOCKMAN, S. 1913. Report of the Departemental Committee Appointed by the Board of Agriculture and Fisheries to Enquire into Epizootic Abortion. His Majesty's Stationery Office, London.

MCLENNAN, M. K., RINGOIR, D. D., FRIRDICH, E., SVENSSON, S. L., WELLS, D. H., JARRELL, H., SZYMANSKI, C. M. & GAYNOR, E. C. 2008. *Campylobacter jejuni* biofilms up-regulated in the absence of the stringent

response utilize a calcofluor white-reactive polysaccharide. *J Bacteriol,* 190**,** 1097-107.

MEISTER, M., HETRU, C. & HOFFMANN, J. A. 2000. The antimicrobial host defense of *Drosophila*. *Curr Top Microbiol Immunol,* 248**,** 17-36.

MILLER, G., DUNN, G. M., SMITH-PALMER, A., OGDEN, I. D. & STRACHAN, N. J. 2004. Human campylobacteriosis in Scotland: seasonality, regional trends and bursts of infection. *Epidemiol Infect,* 132**,** 585-93.

MILLER, S., PESCI, E. C. & PICKETT, C. L. 1993. A Campylobacter jejuni homolog of the LcrD/FlbF family of proteins is necessary for flagellar biogenesis. *Infect Immun,* 61**,** 2930-6.

MILLER, W. G., BATES, A. H., HORN, S. T., BRANDL, M. T., WACHTEL, M. R. & MANDRELL, R. E. 2000. Detection on surfaces and in Caco-2 cells of *Campylobacter jejuni* cells transformed with new *gfp*, *yfp*, and *cfp* marker plasmids. *Appl Environ Microbiol,* 66**,** 5426-36.

MILLER, W. G. & MANDRELL, R. E. 2005. Prevalence of *Campylobacter* in the food and water supply: Incidence, outbreak, isolation and detection. In *Campylobacter* Molecular and Cellular Biology, J.M. Ketley, and M.E. Konkel, eds (Horizon Bioscience), pp.101-163.

MIRA, A. & PUSHKER, R. 2005. The silencing of pseudogenes. *Mol Biol Evol,* 22**,** 2135-8.

MISAWA, N. & BLASER, M. J. 2000. Detection and characterization of autoagglutination activity by Campylobacter jejuni. *Infect Immun,* 68**,** 6168-75.

MONK, C. E., PEARSON, B. M., MULHOLLAND, F., SMITH, H. K. & POOLE, R. K. 2008. Oxygen- and NssR-dependent globin expression and enhanced iron acquisition in the response of campylobacter to nitrosative stress. *J Biol Chem,* 283**,** 28413-25.

MONTEVILLE, M. R., YOON, J. E. & KONKEL, M. E. 2003. Maximal adherence and invasion of INT 407 cells by *Campylobacter jejuni* requires the CadF outer-membrane protein and microfilament reorganization. *Microbiology,* 149**,** 153-65.

MORAN, A. P. 1997. Structure and conserved characteristics of *Campylobacter jejuni* lipopolysaccharides. *J Infect Dis,* 176 Suppl 2**,** S115-21.

MORAN, A. P. & O'MALLEY, D. T. 1995. Potential role of lipo-polysaccharides of *Campylobacter jejuni* in the development of Guillain-Barre syndrome. *J. Endotoxin Res.***,** 233-235.

MORAN, A. P. & PENNER, J. L. 1999. Serotyping of *Campylobacter jejuni* based on heat-stable antigens: relevance, molecular basis and implications in pathogenesis. *J Appl Microbiol,* 86**,** 361-77.

MORAN, A. P., PRENDERGAST, M. M. & APPELMELK, B. J. 1996. Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease. *FEMS Immunol Med Microbiol,* 16**,** 105-15.

MUTA, T. & IWANAGA, S. 1996. The role of hemolymph coagulation in innate immunity. *Curr Opin Immunol,* 8**,** 41-7.

MYLONAKIS, E., CASADEVALL, A. & AUSUBEL, F. M. 2007. Exploiting amoeboid and non-vertebrate animal model systems to study the virulence of human pathogenic fungi. *PLoS Pathog,* 3**,** e101.

NACHAMKIN, I., YANG, X. H. & STERN, N. J. 1993. Role of *Campylobacter jejuni flagella* as colonization factors for three-day-old chicks: analysis with flagellar mutants. *Appl Environ Microbiol,* 59**,** 1269-73.

NAITO, M., FRIRDICH, E., FIELDS, J. A., PRYJMA, M., LI, J., CAMERON, A., GILBERT, M., THOMPSON, S. A. & GAYNOR, E. C. 2010. Effects of sequential *Campylobacter jejuni* 81-176 lipooligosaccharide core truncations on biofilm formation, stress survival, and pathogenesis. *J Bacteriol,* 192**,** 2182-92.

NGUYEN, H. T., CORRY, J. E. & MILES, C. A. 2006. Heat resistance and mechanism of heat inactivation in thermophilic campylobacters. *Appl Environ Microbiol,* 72**,** 908-13.

NIELSEN, H., ENGELBRECHT, J., BRUNAK, S. & VON HEIJNE, G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng,* 10**,** 1-6.

NIELSEN, H., HANSEN, K. K., GRADEL, K. O., KRISTENSEN, B., EJLERTSEN, T., OSTERGAARD, C. & SCHONHEYDER, H. C. 2010. Bacteraemia as a result of *Campylobacter* species: a population-based study of epidemiology and clinical risk factors. *Clin Microbiol Infect,* 16**,** 57-61.

NOTHAFT, H., AMBER, S., AEBI, M. & SZYMANSKI, C. 2008. *N*-linked protein glycosylation in *Campylobacter*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.447-470.

NOTHAFT, H., LIU, X., MCNALLY, D. J., LI, J. & SZYMANSKI, C. M. 2009. Study of free oligosaccharides derived from the bacterial *N*-glycosylation pathway. *Proc Natl Acad Sci U S A,* 106**,** 15019-24.

NOTHAFT, H., LIU, X., MCNALLY, D. J. & SZYMANSKI, C. M. 2010. *N*-linked protein glycosylation in a bacterial system. *Methods Mol Biol,* 600**,** 227-43.

NOVIK, V., HOFREUTER, D. & GALAN, J. E. 2010. Identification of *Campylobacter jejuni* genes involved in its interaction with epithelial cells. *Infect Immun*.

OCHMAN, H. & DAVALOS, L. M. 2006. The nature and dynamics of bacterial genomes. *Science,* 311**,** 1730-3.

OELSCHLAEGER, T. A., GUERRY, P. & KOPECKO, D. J. 1993. Unusual microtubule-dependent endocytosis mechanisms triggered by *Campylobacter jejuni* and *Citrobacter freundii*. *Proc Natl Acad Sci U S A,* 90**,** 6884-8.

OLDFIELD, N. J., MORAN, A. P., MILLAR, L. A., PRENDERGAST, M. M. & KETLEY, J. M. 2002. Characterization of the *Campylobacter jejuni* heptosyltransferase II gene, *waaF*, provides genetic evidence that extracellular polysaccharide is lipid A core independent. *J Bacteriol,* 184**,** 2100-7.

OLSON, C. K., ETHELBERG, S., PELT, W. V. & TAUXE, R. V. 2008. Epidemiology of *Campylobacter jejuni* Infections in Industrialized nations. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.163-189.

OPPENHEIM, J. J., ZACHARIAE, C. O., MUKAIDA, N. & MATSUSHIMA, K. 1991. Properties of the novel proinflammatory supergene "intercrine" cytokine family. *Annu Rev Immunol,* 9**,** 617-48.

OUZOUNIS, C. A. & KARP, P. D. 2002. The past, present and future of genome-wide re-annotation. *Genome Biol,* 3**,** COMMENT2001.

PALYADA, K., SUN, Y. Q., FLINT, A., BUTCHER, J., NAIKARE, H. & STINTZI, A. 2009. Characterization of the oxidative stress stimulon and PerR regulon of *Campylobacter jejuni*. *BMC Genomics,* 10**,** 481.

PALYADA, K., THREADGILL, D. & STINTZI, A. 2004. Iron acquisition and regulation in *Campylobacter jejuni*. *J Bacteriol,* 186**,** 4714-29.

PARK, S. F. 2002. The physiology of *Campylobacter* species and its relevance to their role as foodborne pathogens. *Int J Food Microbiol,* 74**,** 177-88.

PARK, S. F. 2005. *Campylobacter jejuni* stress responses during survival in the food chain and colonization. In *Campylobacter* Molecular and Cellular Biology, J.M. Ketley, and M.E. Konkel, eds (Horizon Bioscience), pp.311-330.

PARK, S. F., PURDY, D. & LEACH, S. 2000. Localized reversible frameshift mutation in the *flhA* gene confers phase variability to flagellin gene expression in *Campylobacter coli*. *J Bacteriol,* 182**,** 207-10.

PARKER, C. T., GILBERT, M., YUKI, N., ENDTZ, H. P. & MANDRELL, R. E. 2008. Characterization of lipooligosaccharide-biosynthetic loci of *Campylobacter jejuni* reveals new lipooligosaccharide classes: evidence of mosaic organizations. *J Bacteriol,* 190**,** 5681-9.

PARKHILL, J., WREN, B. W., MUNGALL, K., KETLEY, J. M., CHURCHER, C., BASHAM, D., CHILLINGWORTH, T., DAVIES, R. M., FELTWELL, T., HOLROYD, S., JAGELS, K., KARLYSHEV, A. V., MOULE, S., PALLEN, M. J., PENN, C. W., QUAIL, M. A., RAJANDREAM, M. A., RUTHERFORD, K. M., VAN VLIET, A. H., WHITEHEAD, S. & BARRELL, B. G. 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature,* 403**,** 665-8.

PARKINS, M. D., CERI, H. & STOREY, D. G. 2001. *Pseudomonas aeruginosa* GacA, a factor in multihost virulence, is also essential for biofilm formation. *Mol Microbiol,* 40**,** 1215-26.

PARRISH, J., YU, J., LIU, G., HINES, J., CHAN, J., MANGIOLA, B., ZHANG, H., PACIFICO, S., FOTOUHI, F., DIRITA, V., IDEKER, T., ANDREWS, P. & FINLEY, R. 2007. A proteome-wide protein interaction map for *Campylobacter jejuni*. *Genome Biology,* 8**,** R130.

PAVLOVSKIS, O. R., ROLLINS, D. M., HABERBERGER, R. L., JR., GREEN, A. E., HABASH, L., STROCKO, S. & WALKER, R. I. 1991. Significance of flagella in colonization resistance of rabbits immunized with *Campylobacter* spp. *Infect Immun,* 59**,** 2259-64.

PAZZAGLIA, G., BOURGEOIS, A. L., EL DIWANY, K., NOUR, N., BADRAN, N. & HABLAS, R. 1991. Campylobacter diarrhoea and an association of recent disease with asymptomatic shedding in Egyptian children. *Epidemiol Infect,* 106**,** 77-82.

PEARSON, B. M., GASKIN, D. J., SEGERS, R. P., WELLS, J. M., NUIJTEN, P. J. & VAN VLIET, A. H. 2007. The complete genome sequence of *Campylobacter jejuni* strain 81116 (NCTC11828). *J Bacteriol,* 189**,** 8402-3.

PEARSON, W. R. & LIPMAN, D. J. 1988. Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A,* 85**,** 2444-8.

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**,** 18717-25.

PEI, Z., BURUCOA, C., GRIGNON, B., BAQAR, S., HUANG, X. Z., KOPECKO, D. J., BOURGEOIS, A. L., FAUCHERE, J. L. & BLASER, M. J. 1998. Mutation in the *peb1A* locus of *Campylobacter jejuni* reduces interactions with epithelial cells and intestinal colonization of mice. *Infect Immun,* 66**,** 938-43.

PEI, Z. H., ELLISON, R. T., 3RD & BLASER, M. J. 1991. Identification, purification, and characterization of major antigenic proteins of *Campylobacter jejuni*. *J Biol Chem,* 266**,** 16363-9.

PENNER, J. L. & HENNESSY, J. N. 1980. Passive hemagglutination technique for serotyping *Campylobacter fetus* subsp. *jejuni* on the basis of soluble heat-stable antigens. *J Clin Microbiol,* 12**,** 732-7.

PESCI, E. C., COTTLE, D. L. & PICKETT, C. L. 1994. Genetic, enzymatic, and pathogenic studies of the iron superoxide dismutase of *Campylobacter jejuni*. *Infect Immun,* 62**,** 2687-94.

PETERSEN, L., LARSEN, T. S., USSERY, D. W., ON, S. L. & KROGH, A. 2003. RpoD promoters in *Campylobacter jejuni* exhibit a strong periodic signal instead of a -35 box. *J Mol Biol,* 326**,** 1361-72.

PETERSON, M. C. 1994. Rheumatic manifestations of *Campylobacter jejuni* and *C. fetus* infections in adults. *Scand J Rheumatol,* 23**,** 167-70.

PHILPOTT, D. J., YAMAOKA, S., ISRAEL, A. & SANSONETTI, P. J. 2000. Invasive *Shigella flexneri* activates NF-kappa B through a lipopolysaccharide-dependent innate intracellular response and leads to IL-8 expression in epithelial cells. *J Immunol,* 165**,** 903-14.

PHONGSISAY, V., PERERA, V. N. & FRY, B. N. 2006. Exchange of lipooligosaccharide synthesis genes creates potential Guillain-Barre syndrome-inducible strains of *Campylobacter jejuni*. *Infect Immun,* 74**,** 1368-72.

PHONGSISAY, V., PERERA, V. N. & FRY, B. N. 2007. Expression of the *htrB* gene is essential for responsiveness of *Salmonella typhimurium* and *Campylobacter jejuni* to harsh environments. *Microbiology,* 153**,** 254-62.

PITTMAN, M. S., GOODWIN, M. & KELLY, D. J. 2001. Chemotaxis in the human gastric pathogen *Helicobacter pylori*: different roles for CheW and the three CheV paralogues, and evidence for CheV2 phosphorylation. *Microbiology,* 147**,** 2493-504.

POLY, F., READ, T., TRIBBLE, D. R., BAQAR, S., LORENZO, M. & GUERRY, P. 2007. Genome sequence of a clinical isolate of *Campylobacter jejuni* from Thailand. *Infect Immun,* 75**,** 3425-33.

POPE, J. E., KRIZOVA, A., GARG, A. X., THIESSEN-PHILBROOK, H. & OUIMET, J. M. 2007. Campylobacter reactive arthritis: a systematic review. *Semin Arthritis Rheum,* 37**,** 48-55.

POPTSOVA, M. S. & GOGARTEN, J. P. 2010. Using comparative genome analysis to identify problems in annotated microbial genomes. *Microbiology,* 156**,** 1909-17.

PRENDERGAST, M. M. & MORAN, A. P. 2000. Lipopolysaccharides in the development of the Guillain-Barre syndrome and Miller Fisher syndrome forms of acute inflammatory peripheral neuropathies. *J Endotoxin Res,* 6**,** 341-59.

PROUTY, M. G., CORREA, N. E. & KLOSE, K. E. 2001. The novel sigma54- and sigma28-dependent flagellar gene transcription hierarchy of *Vibrio cholerae*. *Mol Microbiol,* 39**,** 1595-609.

PURDY, D., BUSWELL, C. M., HODGSON, A. E., MCALPINE, K., HENDERSON, I. & LEACH, S. A. 2000. Characterisation of cytolethal distending toxin (CDT) mutants of *Campylobacter jejuni*. *J Med Microbiol,* 49**,** 473-9.

PURDY, D., CAWTHRAW, S., DICKINSON, J. H., NEWELL, D. G. & PARK, S. F. 1999. Generation of a superoxide dismutase (SOD)-deficient mutant of *Campylobacter coli*: evidence for the significance of SOD in Campylobacter survival and colonization. *Appl Environ Microbiol,* 65**,** 2540-6.

PURDY, D. & PARK, S. F. 1994. Cloning, nucleotide sequence and characterization of a gene encoding superoxide dismutase from *Campylobacter jejuni* and *Campylobacter coli*. *Microbiology,* 140 ( Pt 5)**,** 1203-8.

RAPHAEL, B. H., PEREIRA, S., FLOM, G. A., ZHANG, Q., KETLEY, J. M. & KONKEL, M. E. 2005. The *Campylobacter jejuni* response regulator, CbrR, modulates sodium deoxycholate resistance and chicken colonization. *J Bacteriol,* 187**,** 3662-70.

REGALIA, M., ROSENBLAD, M. A. & SAMUELSSON, T. 2002. Prediction of signal recognition particle RNA genes. *Nucleic Acids Res,* 30**,** 3368-77.

REN, B., ROBERT, F., WYRICK, J. J., APARICIO, O., JENNINGS, E. G., SIMON, I., ZEITLINGER, J., SCHREIBER, J., HANNETT, N., KANIN, E., VOLKERT, T. L., WILSON, C. J., BELL, S. P. & YOUNG, R. A. 2000. Genome-wide location and function of DNA binding proteins. *Science,* 290**,** 2306-9.

REUTER, M., MALLETT, A., PEARSON, B. M. & VAN VLIET, A. H. 2010. Biofilm formation by *Campylobacter jejuni* is increased under aerobic conditions. *Appl Environ Microbiol,* 76**,** 2122-8.

RICHARD, A., OBERHELMAN & DAVID, N., TAYLOR 2000. Campylobacter infections in developing countries. In *Campylobacter* Second Edition, I. Nachmkin and M.J. Blaser, eds (ASM Press), pp.297-313.

RICHARDSON, N. J., KOORNHOF, H. J. & BOKKENHEUSER, V. D. 1981. Long-term infections with *Campylobacter fetus* subsp. *jejuni*. *J Clin Microbiol,* 13**,** 846-9.

RIETSCHEL, E. T., KIRIKAE, T., SCHADE, F. U., ULMER, A. J., HOLST, O., BRADE, H., SCHMIDT, G., MAMAT, U., GRIMMECKE, H. D., KUSUMOTO, S. & ET AL. 1993. The chemical structure of bacterial endotoxin in relation to bioactivity. *Immunobiology,* 187**,** 169-90.

RIVERA-AMILL, V., KIM, B. J., SESHU, J. & KONKEL, M. E. 2001. Secretion of the virulence-associated *Campylobacter* invasion antigens from *Campylobacter jejuni* requires a stimulatory signal. *J Infect Dis,* 183**,** 1607-16.

RODIONOV, D. A., VITRESCHAK, A. G., MIRONOV, A. A. & GELFAND, M. S. 2002. Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *J Biol Chem,* 277**,** 48949-59.

ROGGENKAMP, A., NEUBERGER, H. R., FLUGEL, A., SCHMOLL, T. & HEESEMANN, J. 1995. Substitution of two histidine residues in YadA protein of *Yersinia enterocolitica* abrogates collagen binding, cell adherence and mouse virulence. *Mol Microbiol,* 16**,** 1207-19.

ROPPER, A., WIJDICKS, E. & TRUAX, B. 1991. Guillain-Barre Syndrome. F.A Davis Company, Philadelphia.

ROSARIO, M. M., FREDRICK, K. L., ORDAL, G. W. & HELMANN, J. D. 1994. Chemotaxis in Bacillus subtilis requires either of two functionally redundant CheW homologs. *J Bacteriol,* 176**,** 2736-9.

ROSENBLAD, M. A., GORODKIN, J., KNUDSEN, B., ZWIEB, C. & SAMUELSSON, T. 2003. SRPDB: Signal Recognition Particle Database. *Nucleic Acids Res,* 31**,** 363-4.

RUSSELL, R. G. & BLAKE, D. C., JR. 1994. Cell association and invasion of Caco-2 cells by *Campylobacter jejuni*. *Infect Immun,* 62**,** 3773-9.

RUTHERFORD, K., PARKHILL, J., CROOK, J., HORSNELL, T., RICE, P., RAJANDREAM, M. A. & BARRELL, B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics,* 16**,** 944-5.

SAIJO-HAMANO, Y., UCHIDA, N., NAMBA, K. & OOSAWA, K. 2004. In vitro characterization of FlgB, FlgC, FlgF, FlgG, and FliE, flagellar basal body proteins of Salmonella. *J Mol Biol,* 339**,** 423-35.

SALAUN, L., LINZ, B., SUERBAUM, S. & SAUNDERS, N. J. 2004. The diversity within an expanded and redefined repertoire of phase-variable genes in *Helicobacter pylori*. *Microbiology,* 150**,** 817-30.

SAMBROOK, J. & RUSSELL, D. W. 2001. *Molecular Cloning A laboratory Manual,*.

SANDER, C. & SCHNEIDER, R. 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins,* 9**,** 56-68.

SARASTE, M., SIBBALD, P. R. & WITTINGHOFER, A. 1990. The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci,* 15**,** 430-4.

SAUERWEIN, R. W., BISSELING, J. & HORREVORTS, A. M. 1993. Septic abortion associated with Campylobacter fetus subspecies fetus infection: case report and review of the literature. *Infection,* 21**,** 331-3.

SAUNDERS, N. J., PEDEN, J. F., HOOD, D. W. & MOXON, E. R. 1998. Simple sequence repeats in the Helicobacter pylori genome. *Mol Microbiol,* 27**,** 1091-8.

SCHAAD, U. B. 1982. Reactive arthritis associated with Campylobacter enteritis. *Pediatr Infect Dis,* 1**,** 328-32.

SCHMIDT, O., THEOPOLD, U. & STRAND, M. 2001. Innate immunity and its evasion and suppression by hymenopteran endoparasitoids. *Bioessays,* 23**,** 344-51.
SCHOENHOFEN, I. C., MCNALLY, D. J., BRISSON, J. R. & LOGAN, S. M. 2006a. Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: synthesis from UDP-*N*-acetylglucosamine by a single enzymatic reaction. *Glycobiology,* 16**,** 8C-14C.
SCHOENHOFEN, I. C., MCNALLY, D. J., VINOGRADOV, E., WHITFIELD, D., YOUNG, N. M., DICK, S., WAKARCHUK, W. W., BRISSON, J. R. & LOGAN, S. M. 2006b. Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*: enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J Biol Chem,* 281**,** 723-32.
SCHOLAR, G. Available: http://scholar.google.com/ [Accessed].
SCHONBERGER, L. B., HURWITZ, E. S., KATONA, P., HOLMAN, R. C. & BREGMAN, D. J. 1981. Guillain-Barre syndrome: its epidemiology and associations with influenza vaccination. *Ann Neurol,* 9 Suppl**,** 31-8.
SCHOULS, L. M., REULEN, S., DUIM, B., WAGENAAR, J. A., WILLEMS, R. J., DINGLE, K. E., COLLES, F. M. & VAN EMBDEN, J. D. 2003. Comparative genotyping of Campylobacter jejuni by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. *J Clin Microbiol,* 41**,** 15-26.
SCIRUS - FOR SCIENTIFIC INFORMATION ONLY, E. Available: http://www.scirus.com/srsapp/ [Accessed].
SCOTT, N. E., PARKER, B. L., CONNOLLY, A. M., PAULECH, J., EDWARDS, A. V., CROSSETT, B., FALCONER, L., KOLARICH, D., DJORDJEVIC, S. P., HOJRUP, P., PACKER, N. H., LARSEN, M. R. & CORDWELL, S. J. 2010. Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, HCD and ETD-MS applied to the *N*-linked glycoproteome of *Campylobacter jejuni*. *Mol Cell Proteomics*.
SEBALD, M. & VERON, M. 1963. [Base DNA Content and Classification of Vibrios.]. *Ann Inst Pasteur (Paris),* 105**,** 897-910.
SELLARS, M. J., HALL, S. J. & KELLY, D. J. 2002. Growth of *Campylobacter jejuni* supported by respiration of fumarate, nitrate, nitrite, trimethylamine-*N*-oxide, or dimethyl sulfoxide requires oxygen. *J Bacteriol,* 184**,** 4187-96.
SHENOY, A. R. & VISWESWARIAH, S. S. 2003. Site-directed mutagenesis using a single mutagenic oligonucleotide and *Dpn*I digestion of template DNA. *Anal Biochem,* 319**,** 335-6.
SIKIC POGACAR, M., RUBESA MIHALJEVIC, R., KLANCNIK, A., BRUMINI, G., ABRAM, M. & SMOLE MOZINA, S. 2009. Survival of stress exposed *Campylobacter jejuni* in the murine macrophage J774 cell line. *Int J Food Microbiol,* 129**,** 68-73.
SKIRROW, M. B. 1977. Campylobacter enteritis: a "new" disease. *Br Med J,* 2**,** 9-11.
SKIRROW, M. B. 1987. A demographic survey of campylobacter, salmonella and shigella infections in England. A Public Health Laboratory Service Survey. *Epidemiol Infect,* 99**,** 647-57.
SKIRROW, M. B. 1991. Epidemiology of *Campylobacter* enteritis. *Int J Food Microbiol,* 12**,** 9-16.
SKIRROW, M. B. & BLASER, M. J. 2002. *Campylobacter jejuni*. In *Infections of the Gastrointestinal Tract*, M.J. Blaser, P.D. Smith, J.I. Ravdin, H.B. Greenberg and R.L. Guerrant, eds (Lippincott Williams and Wilkins, Baltimore), pp.719-739.
SLEPNEVA, I. A., GLUPOV, V. V., SERGEEVA, S. V. & KHRAMTSOV, V. V. 1999. EPR detection of reactive oxygen species in hemolymph of *Galleria mellonella*

and *Dendrolimus superans sibiricus* (Lepidoptera) larvae. *Biochem Biophys Res Commun,* 264**,** 212-5.

SMITH, J. L. & BAYLES, D. O. 2006. The contribution of cytolethal distending toxin to bacterial pathogenesis. *Crit Rev Microbiol,* 32**,** 227-48.

SMITH, T. F. & ZHANG, X. 1997. The challenges of genome sequence annotation or "the devil is in the details". *Nat Biotechnol,* 15**,** 1222-3.

SONNHAMMER, E. L., EDDY, S. R., BIRNEY, E., BATEMAN, A. & DURBIN, R. 1998a. Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res,* 26**,** 320-2.

SONNHAMMER, E. L., EDDY, S. R. & DURBIN, R. 1997. Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins,* 28**,** 405-20.

SONNHAMMER, E. L., VON HEIJNE, G. & KROGH, A. 1998b. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol,* 6**,** 175-82.

SORVILLO, F. J., LIEB, L. E. & WATERMAN, S. H. 1991. Incidence of campylobacteriosis among patients with AIDS in Los Angeles County. *J Acquir Immune Defic Syndr,* 4**,** 598-602.

SPOHN, G. & SCARLATO, V. 2001. Motility, chemotaxis and flagellar. In *Helicobacter pylori*: physiology and genetics. Edited by H.L.T Mobley, G.L. Mendz and S.L. Hazall. (ASM Press). pp239-248. .

ST MICHAEL, F., SZYMANSKI, C. M., LI, J., CHAN, K. H., KHIEU, N. H., LAROCQUE, S., WAKARCHUK, W. W., BRISSON, J. R. & MONTEIRO, M. A. 2002. The structures of the lipooligosaccharide and capsule polysaccharide of *Campylobacter jejuni* genome sequenced strain NCTC 11168. *Eur J Biochem,* 269**,** 5119-36.

STEAD, D. & PARK, S. F. 2000. Roles of Fe superoxide dismutase and catalase in resistance of *Campylobacter coli* to freeze-thaw stress. *Appl Environ Microbiol,* 66**,** 3110-2.

STERN, N. J. 2008. *Salmonella* species and *Campylobacter jejuni* cecal colonization model in broilers. *Poult Sci,* 87**,** 2399-403.

STORZ, G., TARTAGLIA, L. A. & AMES, B. N. 1990. Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science,* 248**,** 189-94.

STOTHARD, P. & WISHART, D. S. 2006. Automated bacterial genome analysis and annotation. *Curr Opin Microbiol,* 9**,** 505-10.

STRAND, M. R. & PECH, L. L. 1995. Immunological basis for compatibility in parasitoid-host relationships. *Annu Rev Entomol,* 40**,** 31-56.

SVENSSON, S. L., DAVIS, L. M., MACKICHAN, J. K., ALLAN, B. J., PAJANIAPPAN, M., THOMPSON, S. A. & GAYNOR, E. C. 2009. The CprS sensor kinase of the zoonotic pathogen *Campylobacter jejuni* influences biofilm formation and is required for optimal chick colonization. *Mol Microbiol,* 71**,** 253-72.

SVENSSON, S. L., FRIRDICH, E. & GAYNOR, E. C. 2008. Survival strategies of *Campylobacter jejuni*: stress responses, the viable but nonculturable state, and biofilms. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.571-590.

SZYMANSKI, C. M., BURR, D. H. & GUERRY, P. 2002. *Campylobacter* protein glycosylation affects host cell interactions. *Infect Immun,* 70**,** 2242-4.

SZYMANSKI, C. M., LOGAN, S. M., LINTON, D. & WREN, B. W. 2003a. *Campylobacter*--a tale of two protein glycosylation systems. *Trends Microbiol,* 11**,** 233-8.

SZYMANSKI, C. M., MICHAEL, F. S., JARRELL, H. C., LI, J., GILBERT, M., LAROCQUE, S., VINOGRADOV, E. & BRISSON, J. R. 2003b. Detection of conserved *N*-linked glycans and phase-variable lipooligosaccharides and capsules from *campylobacter* cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy. *J Biol Chem,* 278**,** 24509-20.

SZYMANSKI, C. M. & WREN, B. W. 2005. Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol,* 3**,** 225-37.

SZYMANSKI, C. M., YAO, R., EWING, C. P., TRUST, T. J. & GUERRY, P. 1999. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol Microbiol,* 32**,** 1022-30.

TAUXE, R. V., HARGRETT-BEAN, N., PATTON, C. M. & WACHSMUTH, I. K. 1988. Campylobacter isolates in the United States, 1982-1986. *MMWR CDC Surveill Summ,* 37**,** 1-13.

TAYLOR, D. E. & TRACZ, D. M. 2005. Mechanisms of Antimicrobial Resistance in *Campylobacter*. In *Campylobacter* Molecular and Cellular Biology, J.M. Ketley, and M.E. Konkel, eds (Horizon Bioscience), pp.193-204.

TAYLOR, D. N., ECHEVERRIA, P., PITARANGSI, C., SERIWATANA, J., BODHIDATTA, L. & BLASER, M. J. 1988. Influence of strain characteristics and immunity on the epidemiology of *Campylobacter* infections in Thailand. *J Clin Microbiol,* 26**,** 863-8.

THIBAULT, P., LOGAN, S. M., KELLY, J. F., BRISSON, J. R., EWING, C. P., TRUST, T. J. & GUERRY, P. 2001. Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J Biol Chem,* 276**,** 34862-70.

THORNLEY, J. P., JENKINS, D., NEAL, K., WRIGHT, T., BROUGH, J. & SPILLER, R. C. 2001. Relationship of *Campylobacter* toxigenicity in vitro to the development of postinfectious irritable bowel syndrome. *J Infect Dis,* 184**,** 606-9.

TOLCIN, R., LASALVIA, M. M., KIRKLEY, B. A., VETTER, E. A., COCKERILL, F. R., 3RD & PROCOP, G. W. 2000. Evaluation of the Alexon-trend ProSpecT *Campylobacter* microplate assay. *J Clin Microbiol,* 38**,** 3853-5.

TOLEDANO, M. B., KULLIK, I., TRINH, F., BAIRD, P. T., SCHNEIDER, T. D. & STORZ, G. 1994. Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell,* 78**,** 897-909.

TONG, Z., ZHOU, D., SONG, Y., ZHANG, L., PEI, D., HAN, Y., PANG, X., LI, M., CUI, B., WANG, J., GUO, Z., QI, Z., JIN, L., ZHAI, J., DU, Z., WANG, X., YU, J., HUANG, P., YANG, H. & YANG, R. 2005. Pseudogene accumulation might promote the adaptive microevolution of Yersinia pestis. *J Med Microbiol,* 54**,** 259-68.

TRACZ, D. M., KEELAN, M., AHMED-BENTLEY, J., GIBREEL, A., KOWALEWSKA-GROCHOWSKA, K. & TAYLOR, D. E. 2005. pVir and bloody diarrhea in Campylobacter jejuni enteritis. *Emerg Infect Dis,* 11**,** 838-43.

TRIEU-CUOT, P., GERBAUD, G., LAMBERT, T. & COURVALIN, P. 1985. *In vivo* transfer of genetic information between gram-positive and gram-negative bacteria. *Embo J,* 4**,** 3583-7.

UZZAU, S., FIGUEROA-BOSSI, N., RUBINO, S. & BOSSI, L. 2001. Epitope tagging of chromosomal genes in Salmonella. *Proc Natl Acad Sci U S A,* 98**,** 15264-9.

VALVANO, M. A., MESSNER, P. & KOSMA, P. 2002. Novel pathways for biosynthesis of nucleotide-activated glycero-manno-heptose precursors of bacterial glycoproteins and cell surface polysaccharides. *Microbiology,* 148**,** 1979-89.

VAN ALPHEN, L. B., BLEUMINK-PLUYM, N. M., ROCHAT, K. D., VAN BALKOM, B. W., WOSTEN, M. M. & VAN PUTTEN, J. P. 2008. Active migration into the subcellular space precedes *Campylobacter jejuni* invasion of epithelial cells. *Cell Microbiol,* 10**,** 53-66.

VAN VLIET, A. H. 2010. Next generation sequencing of microbial transcriptomes: challenges and opportunities. *FEMS Microbiol Lett,* 302**,** 1-7.

VAN VLIET, A. H., BAILLON, M. A., PENN, C. W. & KETLEY, J. M. 2001. The iron-induced ferredoxin FdxA of *Campylobacter jejuni* is involved in aerotolerance. *FEMS Microbiol Lett,* 196**,** 189-93.

VAN VLIET, A. H., BAILLON, M. L., PENN, C. W. & KETLEY, J. M. 1999. *Campylobacter jejuni* contains two fur homologs: characterization of iron-responsive regulation of peroxide stress defense genes by the PerR repressor. *J Bacteriol,* 181**,** 6371-6.

VAN VLIET, A. H., WOOLDRIDGE, K. G. & KETLEY, J. M. 1998. Iron-responsive gene regulation in a *Campylobacter jejuni fur* mutant. *J Bacteriol,* 180**,** 5291-8.

VÉRON, M. & CHATELAIN, R. 1973. Taxonomic study of the genus *Campylobacter* Sebald and Véron and designation of the neotype strain for the type species *Campylobacter fetus* (Smith and Taylor) Sebald and Véron. *Int. J. Syst. Bacteriol. ,* 23**,** 122-134.

VIJAYAKUMAR, S., MERKX-JACQUES, A., RATNAYAKE, D. B., GRYSKI, I., OBHI, R. K., HOULE, S., DOZOIS, C. M. & CREUZENET, C. 2006. Cj1121c, a novel UDP-4-keto-6-deoxy-GlcNAc C-4 aminotransferase essential for protein glycosylation and virulence in *Campylobacter jejuni*. *J Biol Chem,* 281**,** 27733-43.

WAINWRIGHT, L. M., ELVERS, K. T., PARK, S. F. & POOLE, R. K. 2005. A truncated haemoglobin implicated in oxygen metabolism by the microaerophilic food-borne pathogen *Campylobacter jejuni*. *Microbiology,* 151**,** 4079-91.

WAINWRIGHT, L. M., WANG, Y., PARK, S. F., YEH, S. R. & POOLE, R. K. 2006. Purification and spectroscopic characterization of Ctb, a group III truncated hemoglobin implicated in oxygen metabolism in the food-borne pathogen *Campylobacter jejuni*. *Biochemistry,* 45**,** 6003-11.

WALKER, R. I. 2005. Considerations for development of whole cell bacterial vaccines to prevent diarrheal diseases in children in developing countries. *Vaccine,* 23**,** 3369-85.

WANG, W. L. & BLASER, M. J. 1986. Detection of pathogenic *Campylobacter* species in blood culture systems. *J Clin Microbiol,* 23**,** 709-14.

WASSENAAR, T. M. & BLASER, M. J. 1999. Pathophysiology of *Campylobacter jejuni* infections of humans. *Microbes Infect,* 1**,** 1023-33.

WASSENAAR, T. M., ENGELSKIRCHEN, M., PARK, S. & LASTOVICA, A. 1997. Differential uptake and killing potential of *Campylobacter jejuni* by human peripheral monocytes/macrophages. *Med Microbiol Immunol (Berl),* 186**,** 139-44.

WATSON, R. O. & GALAN, J. E. 2008. *Campylobacter jejuni* survives within epithelial cells by avoiding delivery to lysosomes. *PLoS Pathog,* 4**,** e14.

WATSON, R. O. & GALÁN, J. E. 2008. Interaction of *Campylobacter jejuni* with host cells. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.289-296.

WHEELER, J. G., SETHI, D., COWDEN, J. M., WALL, P. G., RODRIGUES, L. C., TOMPKINS, D. S., HUDSON, M. J. & RODERICK, P. J. 1999. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive. *Bmj,* 318**,** 1046-50.

WHITEHOUSE, C. A., BALBO, P. B., PESCI, E. C., COTTLE, D. L., MIRABITO, P. M. & PICKETT, C. L. 1998. *Campylobacter jejuni* cytolethal distending toxin causes a G2-phase cell cycle block. *Infect Immun,* 66**,** 1934-40.

WHITFIELD, C. 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem,* 75**,** 39-68.

WHITFIELD, C. & ROBERTS, I. S. 1999. Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol Microbiol,* 31**,** 1307-19.

WILKINSON, S. P. & GROVE, A. 2004. HucR, a novel uric acid-responsive member of the MarR family of transcriptional regulators from *Deinococcus radiodurans*. *J Biol Chem,* 279**,** 51442-50.

WISTROM, J. & NORRBY, S. R. 1995. Fluoroquinolones and bacterial enteritis, when and for whom? *J Antimicrob Chemother,* 36**,** 23-39.

WOLFF, B., BURNS, A. R., MIDDLETON, J. & ROT, A. 1998. Endothelial cell "memory" of inflammatory stimulation: human venular endothelial cells store interleukin 8 in Weibel-Palade bodies. *J Exp Med,* 188**,** 1757-62.

WOOLCOCK, P. R., MOORE, J. D., MCFARLAND, M. D. & PANIGRAHY, B. 1996. Isolation of paramyxovirus serotype 7 from ostriches (*Struthio camelus*). *Avian Dis,* 40**,** 945-9.

WOOLDRIDGE, K. G., WILLIAMS, P. H. & KETLEY, J. M. 1996. Host signal transduction and endocytosis of *Campylobacter jejuni*. *Microb Pathog,* 21**,** 299-305.

WOSTEN, M. M., BOEVE, M., GAASTRA, W. & VAN DER ZEIJST, B. A. 1998. Cloning and characterization of the gene encoding the primary sigma-factor of *Campylobacter jejuni*. *FEMS Microbiol Lett,* 162**,** 97-103.

WOSTEN, M. M., WAGENAAR, J. A. & VAN PUTTEN, J. P. 2004. The FlgS/FlgR two-component signal transduction system regulates the *fla* regulon in *Campylobacter jejuni*. *J Biol Chem,* 279**,** 16214-22.

WÖSTEN, M. M. S. N., MOURIK, A. V. & PUTTEN, J. P. M. V. 2008. Regulation of Genes in *Campylobacter jejuni*. In *Campylobacter* Third Edition, I. Nachmkin, C.M. Szymanski, and M.J. Blaser, eds (ASM Press), pp.611-624.

XU, Q., SCHWARZENBACHER, R., MCMULLAN, D., ABDUBEK, P., AGARWALLA, S., AMBING, E., AXELROD, H., BIORAC, T., CANAVES, J. M., CHIU, H. J., DEACON, A. M., DIDONATO, M., ELSLIGER, M. A., GODZIK, A., GRITTINI, C., GRZECHNIK, S. K., HALE, J., HAMPTON, E., HAN, G. W., HAUGEN, J., HORNSBY, M., JAROSZEWSKI, L., KLOCK, H. E., KOESEMA, E., KREUSCH, A., KUHN, P., LESLEY, S. A., MILLER, M. D., MOY, K., NIGOGHOSSIAN, E., PAULSEN, J., QUIJANO, K., REYES, R., RIFE, C., SPRAGGON, G., STEVENS, R. C., VAN DEN BEDEM, H., VELASQUEZ, J., WHITE, A., WOLF, G., HODGSON, K. O., WOOLEY, J. & WILSON, I. A. 2006. Crystal structure of virulence factor CJ0248 from *Campylobacter jejuni* at 2.25 A resolution reveals a new fold. *Proteins,* 62**,** 292-6.

YAO, R., BURR, D. H., DOIG, P., TRUST, T. J., NIU, H. & GUERRY, P. 1994. Isolation of motile and non-motile insertional mutants of *Campylobacter jejuni:* the role of motility in adherence and invasion of eukaryotic cells. *Mol Microbiol,* 14**,** 883-93.

YAO, R., BURR, D. H. & GUERRY, P. 1997. CheY-mediated modulation of *Campylobacter jejuni* virulence. *Mol Microbiol,* 23**,** 1021-31.

YOUNG, G. M., SCHMIEL, D. H. & MILLER, V. L. 1999. A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proc Natl Acad Sci U S A,* 96**,** 6456-61.

YOUNG, K. T., DAVIS, L. M. & DIRITA, V. J. 2007. *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat Rev Microbiol,* 5**,** 665-79.

YOUNG, N. M., BRISSON, J.-R., KELLY, J., WATSON, D. C., TESSIER, L., LANTHIER, P. H., JARRELL, H. C., CADOTTE, N., ST. MICHAEL, F., ABERG, E. & SZYMANSKI, C. M. 2002. Structure of the *N*-Linked Glycan Present on Multiple Glycoproteins in the Gram-negative Bacterium, *Campylobacter jejuni*. *J. Biol. Chem.,* 277**,** 42530-42539.

YOUNG, V. B. & MANSFIELD, L. S. 2005. *Campylobacter* Infection - Clinical Context. In *Campylobacter* Molecular and Cellular Biology, J.M. Ketley, and M.E. Konkel, eds (Horizon Bioscience), pp.1-12.

YUKI, N., TAKI, T., INAGAKI, F., KASAMA, T., TAKAHASHI, M., SAITO, K., HANDA, S. & MIYATAKE, T. 1993. A bacterium lipopolysaccharide that elicits Guillain-Barre syndrome has a GM1 ganglioside-like structure. *J Exp Med,* 178**,** 1771-5.

ZHENG, M., ASLUND, F. & STORZ, G. 1998. Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science,* 279**,** 1718-21.