1. Introduction

1.1 Background

Campylobacter jejuni is the most common cause of bacterial diarrhoeal disease worldwide. Little is known about the ability of this organism to cause disease. A wide range of phenotypic and genotypic diversity has been reported for this species along with a range of disease outcomes.

1.1.1 Classification

C. jejuni belongs to the delta-epsilon group of proteobacteria within the family *Campylobactereaceae,* which also includes the genera *Helicobacter* and *Arcobacter*. The delta-epsilon group is significantly divergent from the gamma subgroup which contains many human enteropathogens such as the salmonellae, *Escherichia coli* and *Shigella* [1]. The genus *Campylobacter* now includes about twenty species and subspecies, eight of which are known to cause human gastrointestinal disease [2]. *Campylobacter* was previously thought to be purely a pathogen of animals until the 1970s when it was discovered that *Campylobacter* caused diarrhoea in man [3].

Campylobacter are Gram-negative with a low G+C content chromosome (30%). They are non-spore forming, spiral rod shaped bacteria 0.2-0.8 μm wide and 0.5-5 μm long. Cells are typically motile and move in a corkscrew-like motion propelled by a single polar flagellum. They require a microaerobic environment for growth and are thermophilic with an optimum growth temperature of 42-43°C [1].

1.1.2 Physiology and metabolism

1.1.2.1 Growth

Many aspects of the physiology and metabolism of these organisms remain poorly understood due in part to difficulties in cultivating members of the *Campylobacter* genus. The majority of *Campylobacter* strains need to be cultured in a microaerobic environment consisting of 5-10% (v/v) oxygen and 5-10% (v/v) carbon dioxide in complex growth media with additional supplements [4]. *C. jejuni* is susceptible to a wide variety of antimicrobial treatments and food processing methods such as drying, freezing and salting. *C. jejuni* is also sensitive to osmotic stress and osmotic pressure, oxygen concentrations above 5% and has not been reported to grow at temperatures below 30°C [2].

After several days of *in vitro* culture, cells have been noted to change from spiral to coccoid forms accompanied by a loss of culturability. It remains controversial as to whether this represents a survival mechanism or is a degenerative form but it has been reported that the change to coccoid forms does not require *de novo* protein synthesis. This suggests that the change to coccoid forms is not actively controlled and therefore may represent cell injury [5]. *C. jejuni* is not thought to mount a stationary-phase response to limited nutrient availability, or the build up of toxic waste products, which is characterized in many other bacteria by increased resistance to environmental stress. Resuscitation of aged cultures has been demonstrated but is more likely to represent growth in numbers of residual viable cells rather than reversal of a viable but not culturable state [6]. It is possible that a subpopulation within stationary phase cultures of *Campylobacter* may be better able to cope with injury [7].

1.1.2.2 Transport and iron uptake

The genome sequences of *C. jejuni* show that it has a limited capacity for biosynthesis and therefore many transport systems exist for the acquisition of essential amino acids, other nutrients and ions from the external environment [4;8;9].

Iron is essential for all organisms as it is a cofactor of many enzymes e.g. peroxidases and cytochromes; in addition it is used in electron transport and redox reactions [10]. Iron-uptake systems are often considered to be virulence factors as iron availability is limited in mammalian host tissues. In response to low iron availability bacteria may produce low-molecular weight iron chelators called siderophores [10]. It has been suggested that certain strains of *C. jejuni* may produce siderophores; however, no siderophores have been characterized from *C. jejuni* [11]. *C. jejuni* may be able to scavenge siderophores produced by other bacteria in the intestinal tract [12] as several systems for the uptake of iron complexed to siderophores have been discovered. These uptake systems include the *ceuBCDE* operon for the uptake of enterochelin and *cfrA* which has been proposed as a ferric enterobactin receptor [13] and which is only present in some strains [14]. Also cj0178 may be the receptor of an as yet unidentified iron source [10]. In addition a haemin/haemoglobin uptake system exists encoded by *chuABCD* with a *chuA* mutant being unable to use haemoglobin or haemin as an iron source [13]. Haem compounds may be released by the host at the site of inflammation [12] and therefore be accessible to *C. jejuni* once it leaves the intestinal tract.

Excess iron can be toxic to cells causing oxidative stress therefore iron uptake is tightly regulated. In *C. jejuni* there are two iron-response regulators, the ferric uptake regulator Fur and PerR, which regulates peroxide stress defence proteins AhpC and KatA [4], underlining the link between iron uptake and oxidative stress resistance. *C. jejuni* also possesses the ability to store iron for iron-limited environments as ferritin (*cft*) [15].

1.1.2.3 Carbon metabolism

C. jejuni has no phosphofructokinase, a glycolysis enzyme, and therefore cannot significantly metabolize externally supplied sugars. *C. jejuni* has been proposed to obtain carbon and energy requirements from tricarboxylic acid (TCA) cycle intermediates, some organic acids and amino acids; in particular *C. jejuni* has been shown to grow in culture using serine, glutamate, aspartate, asparagine, glutamine and proline as the sole carbon source [16]. A recent study has shown that *Campylobacter* strains fall into three distinct metabolic groups: 91% of *C. jejuni* strains tested were able to oxidize α–ketoglutarate, succinate, fumarate and aspartic acid; 7% of *C. jejuni* strains were unable to metabolize α-ketoglutatarate and 2% of *C. jejuni* strains were unable to oxidize succinate, fumarate and aspartic acid [17]. This highlights the fact that there is metabolic diversity between different strains of *C. jejuni*. Proteases might be important to the nutrition of the organism under carbon-limiting conditions by breaking down proteins into constituent parts that can be fed into TCA cycle [16].

1.1.2.4 Electron transport

C. fetus has been shown to grow anaerobically by respiring formate and fumarate in a similar way to *Wolinella succinogenes* along with some other members of the *Campylobacter* genus. *C. jejuni* in contrast has been reported not to be able to grow anaerobically [18] even though the genome sequence of strain NCTC 11168 revealed the presence of genes for fumarate reductase and other genes known to be involved in anaerobic electron transport pathways from other bacteria [4]. Electron acceptors other than oxygen may be important for growth in the avian gut and also the mammalian gut where oxygen is limited. The respiratory chain in *C. jejuni* appears to be highly branched and complex with many cytochromes [4] suggesting an ability to adapt to different environmental conditions.

1.1.3 *Campylobacter* infection

1.1.3.1 Epidemiology

About 90% of human *Campylobacter* isolates in England and Wales are *C. jejuni* with most of the remainder being *C. coli* [19]. Reports of campylobacteriosis are not normally distinguished at the species level so the peak of 58,059 cases in 1998 were reported to the Communicable Disease Surveillance Centre (CDSC) simply as *Campylobacter* [20]. More recent figures would tend to suggest a decrease in cases over the past few years, with the Health Protection Agency (HPA) receiving 47,597 laboratory reports of *Campylobacter* in faecal isolates during 2002, 7,317 less than the previous year [21]; this trend is also apparent in the USA [22].

In the UK there is a seasonal variation in incidence, with a peak in late spring, a lesser peak in autumn and a winter low. Regional variation also occurs with a greater incidence in rural rather than urban populations [16]. The incidence of campylobacteriosis is highest in males under 1 year old with a second peak occurring in adults aged 25-34. Incidence is higher in males than females for all age groups [20].

Even though cases of acute gastroenteritis caused by *Campylobacter* now outnumber those caused by *Salmonella,* outbreaks of campylobacteriosis are rarer than outbreaks of salmonellosis; only 12 general outbreaks of *Campylobacter*, affecting 239 people were reported to the Communicable Disease Surveillance Centre (CDSC) between 1995 and 1996, compared to 233 outbreaks of salmonellosis involving 4,946 people [20]. Part of the reason for this could be that *Campylobacter* does not multiply in foods; however only a low dose is required to cause disease: 50-100 cells if not lower, depending on the infecting strain [2].

C. jejuni is found naturally in the gastrointestinal tract of birds (particularly poultry), cattle and domestic pets, where it rarely causes disease. Transmission to humans has been reported from a variety of sources including raw or undercooked meat, especially poultry

[2;20]. The Food Standards Agency (FSA) has quoted that an average of 50% of retail chickens in the United Kingdom are contaminated by *Campylobacter* [23] although reports of contamination vary dramatically with location, sampling season and different producers of raw retail chickens [24;25].

Other sources of infection are unpasteurised milk, bird-pecked milk on doorsteps and untreated water: *Campylobacter* may be shed into surface water by birds and can survive for many weeks at low temperatures. However most infections remain unexplained by recognised risk factors [20].

1.1.3.2 Disease outcomes

C. jejuni can cause a spectrum of disease ranging from asymptomatic colonisation to severe inflammatory diarrhoea and has also been associated with bacteraemia, endocarditis, meningitis, urinary tract infection and other extraintestinal diseases including Guillain-Barré syndrome (GBS) and Miller Fisher syndrome (MFS) [26]. Post infective complications are rare with about 1% developing reactive arthritis and about 0.1% developing GBS [16]. GBS is an autoimmune-mediated disorder of the peripheral nervous system resulting in paralysis [27] and MFS is considered a variant of GBS that causes paralysis of ocular muscles, ataxia and a loss of tendon reflex [28]. GBS has been associated with *C. jejuni* serotypes 0:41[29] and 0:19[30]; with the latter serotype (0:19) the risk of developing GBS may be as high as 1 in 150 [29]. Amplified fragment length polymorphism (AFLP) analysis has shown that strains associated with GBS or MFS do not belong to a distinct genetic group [31] suggesting that host factors are a major determining factor in the onset of these disorders.

In its uncomplicated form campylobacteriosis is characterised by fever, abdominal cramping and diarrhoea (with or without faecal leukocytes). The incubation period is normally 1-7 days after which profuse diarrhoea frequently lasts 2-3 days accompanied by acute abdominal pains [26]. The average duration of illness calculated from nine outbreaks

was 4.6 days although one third of patients from these outbreaks were ill for more than seven days [32]. The disease is usually self-limiting and in the majority of cases people recover without the aid of antibiotics; the relapse rate being 5-10% [26]. The proportion of patients admitted to hospital varies but is generally cited as between 5-10% and fatalities are rare, usually only occurring in the elderly or the immunocompromised [32]. In circumstances where antibiotic treatment is necessary, for example in prolonged or systemic infection, erythromycin is used as the drug of choice but fluoroquinolones and tetracycline may also be used [26;32].

There is a marked discrepancy between disease outcomes in developed and developing countries. In developed countries campylobacteriosis can be quite severe with bloody diarrhoea a common feature of infection whereas in developing countries diarrhoea is more likely to be watery. In developing countries the disease predominantly affects young children possibly relating to developed immunity resulting in subsequent asymptomatic infection [33].

It has been reported that travellers abroad are more likely to develop disease symptoms similar to those they would develop if they contracted the disease in their own country. This suggests that host susceptibility is an important factor and that differences in disease outcome are unlikely to be solely due to strain differences in geographically separate areas. Indeed in human volunteer studies the same strain can cause different severity of illness in different people [16].

Immunocompromised people are more at risk of developing disease, with one report suggesting that in Acquired Immunodeficiency Syndrome (AIDS) patients in Los Angeles during the period 1983-1987 the incidence of disease was 39 times higher than in the general population. This may however be an overestimate as AIDS patients are considered to be more likely to report to health services on development of symptoms [34]. It has also been

suggested that immunocompromised people often develop a more severe form of disease [26]. There is also a strong association between reactive arthritis, a rare complication, and people who have the human lymphocyte antigen HLA-B27 [34]. These facts highlight the role of the immune system in severity of disease.

The reasons for variable host response are not clear but may depend on a combination of the virulence of the infecting strain, the challenge dose, and the susceptibility of the patient [32].

1.1.4 Pathogenesis

1.1.4.1 Models of infection

Although *C. jejuni* causes a large number of infections each year the bacterium is nutritionally fastidious and supposedly extremely susceptible to environmental stresses (section 1.1.2.1). This apparent paradox highlights the fact that *C. jejuni* remains a poorly understood pathogen.

Part of the reason for the poor understanding of this pathogen is due to a lack of suitable animal models to assess virulence [35]. *Campylobacter* has the ability to colonize the intestinal tract of many animals including humans, pigs, cattle and birds but in most hosts *Campylobacter* behaves as a commensal gut organism. Only primates and possibly ferrets show disease outcomes similar to those in humans with other animal models e.g. chickens and mice being used for colonization studies. So far factors which may explain why only a restricted number of species succumb to disease have been elusive [16].

Human volunteer challenges have been used in the past to study pathogenicity. In these reports large oral inocula were needed to cause illness [36] whereas low doses have been implicated in waterborne outbreaks [29] so even this may be a poor reflection of how *C. jejuni* causes infection naturally.

Despite the above limitations to experimental determination of *C. jejuni* pathogenicity certain factors have been shown to be important for disease progression as described below.

1.1.4.2 Motility

The primary stage of infection or colonization involves *Campylobacter* moving towards cell surfaces. The distinctive type of corkscrew-like movement displayed by *Campylobacter* is thought to be an adaption allowing penetration of the mucus overlying the intestinal epithelium [12]. This movement is mediated by flagella. There are two genes encoding flagellin in *C. jejuni: flaA* and *flaB*, and these genes are arranged in tandem on the chromosome and show a high degree of sequence identity to each other. The flagella are constructed from multimers of flagellin; FlaA flagellin protein is the major component with a small amount of FlaB flagellin protein. These flagellin proteins are attached by a hook protein to a basal structure which is embedded in the membrane and, along with the stator units (MotA and MotB plus FliMNG), acts as a motor for rotation [29].

The flagella are post-translationally modified by phosphorylation and glycosylation [37] and they exhibit phase and antigenic variation [15]. The *flaA* and *flaB* genes are independently transcribed by different types of promoter, σ^{28}- and σ^{54}-dependent respectively. The expression of *flaB* seems to be environmentally regulated by temperature and pH [16]. It has also been demonstrated that the flagella are able to secrete proteins into the extracellular milieu [38].

Motility is directed by chemotaxis which allows the organism to locate and move towards the mucus layer of the gut [16]. Mucin, L-serine and L-fucose all act as chemoattractants for *C. jejuni* and bile acids act as chemorepellants [16]. Ten proteins containing methyl-accepting chemotaxis domains have been identified within the genome sequence of *C. jejuni* strain NCTC 11168 [8;12].

1.1.4.3 Adherence

Several adhesins have been described in *Campylobacter* including flagellin, lipopolysaccharide (LPS) and a number of membrane proteins [16]. In the case of flagella it is difficult to separate adhesion from motility functions. Binding to host cells is proposed to be mediated by proteins synthesized constitutively as heat killed bacteria are still able to bind [39]. Different strains have been noted to vary in their ability to adhere to epithelial cells *in vitro* [40].

A mutant of PEB1, a homologue of cluster 3 binding proteins of bacterial ABC transporters, showed a 50-100 fold decrease in adherence and 15-fold decrease in invasion of epithelial cells in culture [41]. An outer membrane protein CadF (Campylobacter adhesion to fibronectin) is required for adherence and may stimulate invasion upon binding to fibronectin [42]. A lipoprotein, JlpA, has also been shown to be an adhesin [43].

CheY, which affects the rate of flagellar motor switching, has been shown to have an effect on both adherence and invasion. *C. jejuni* mutants containing two copies of *cheY* are non-adherent and non-invasive and *cheY*$^-$ strains are hyperadherent and hyperinvasive *in vitro* [44].

1.1.4.4 Cellular Invasion

Strains of *C. jejuni* differ in their ability to invade human cell lines *in vitro* [16;40]. Once *C. jejuni* has attached to a gut epithelial surface a subpopulation goes on to invade the epithelial cells; this invasion has been correlated with inflammatory disease [12]. At least two mechanisms of invasion have been proposed. The first mechanism involves actin reorganization and accumulation within the mammalian cell, beneath the site of bacterial attachment, followed by microfilament-mediated uptake. The second mechanism is microfilament independent and instead utilises a microtubule mediated uptake system

involving coated pit formation. In addition both clathrin-coated pits and clathrin-independent caveolae have been implicated in endocytosis of the pathogen [45]. This data suggests that there may be several different mechanisms at work.

C. jejuni invasion is dependent on both *de novo* synthesized bacterial proteins and host cell signal transduction [12]: at least 8 bacterial proteins are produced and secreted upon co-cultivation of *C. jejuni* and the human embryonic intestinal cell line, INT-407. One of these proteins is CiaB which has been identified as necessary for internalization into cells [46].

In addition to intracellular invasion, translocation across cell monolayers has been observed and has been shown to be inhibited by chloramphenicol, suggesting that *de novo* bacterial protein synthesis is required. Some strains are able to translocate across confluent cell monolayers despite being classed as non-invasive and indeed electron microscopy studies have shown bacteria passing through and between host cells suggesting that both paracellular and transcellular routes are used [15].

1.1.4.5 Intracellular survival

After invasion of the intestinal epithelial cells, *C. jejuni* appear to be largely confined within endosomal vacuoles [45]. Bacterial numbers in INT407 cells have been shown to decrease after phagosome-lysosome fusion after which most bacteria have adopted a coccoid morphology and there is little evidence of intracellular multiplication [45]. *Campylobacter* have also been observed free in the cytoplasm of cells *in vitro* and *in vivo* [45].

Campylobacter may translocate across cells allowing the bacteria to reach the bloodstream and deeper tissues [29]. *Campylobacter* are sensitive to complement-mediated lysis and as such are thought to be rapidly killed upon traversing the epithelium [29]. The host inflammatory response leads to polymorphonuclear leucocytes and monocytes infiltrating intestinal epithelium [15]. If *Campylobacter* do survive early immune responses

and circulate in the bloodstream they will eventually be taken up by macrophages where the bacteria may survive for up to seven days [29]. Strain differences in serum resistance [47] and phagocyte-mediated killing have been reported [16]. Mechanisms of intracellular survival are unknown although serum resistance has been linked to sialylation of lipooligosaccharide [48].

1.1.4.6 Toxins

As the levels of invading *Campylobacter* are thought to make up less than 1% of applied bacteria on a monolayer of cells in culture, the action of toxins has been proposed as a pathogenesis factor [49], indeed certain aspects of *Campylobacter* disease would be consistent with the action of toxins.

There are two major classes of proteinaceous toxins which would be relevant to the observed pathogenicity of *Campylobacter*; enterotoxins and cytotoxins. Enterotoxins bind to cellular receptors, enter the cell and elevate intracellular cAMP levels causing excess secretion of fluid resulting in watery diarrhoea. Cytotoxins kill target cells by inhibition of cellular protein synthesis or inhibition of actin filament formation [2] which may be consistent with diarrhoea containing blood and inflammatory cells [16].

At least six different types of toxin have been proposed to be encoded by *Campylobacter* strains including cholera-like toxin and various cytotoxins [29]. Reports of enterotoxin production vary widely between isolates with some isolates producing enterotoxin and some not. No correlation was found between enterotoxin production and prevalent Lior or Penner serotypes (see section 1.1.5.1 for explanation of serotype schemes). It has been suggested that enterotoxin production results in watery type diarrhoea as opposed to inflammatory bloody diarrhoea. However this could be determined by host factors (section 1.1.3.2) [49].

Several different cytotoxins have been proposed to be encoded by *Campylobacter* including cytolethal distending toxin (CDT). The early effects of CDT are similar to those of an enterotoxin but after several days cells show distension and death [49]. CDT appears to be able to cause some diarrhoeal symptoms in the rat-ileal-loop assay [50] but the mechanisms of action of this toxin are unknown. CDT is encoded by *cdtABC* which have been found in all *C. jejuni* strains studied so far [29;51]. The CDT locus is found in other *Campylobacter* species although it shows high sequence divergence between species [50]. The amount of toxin produced by isolates varies even though all strains seem to possess the genes that encode the toxin [49].

Haemolytic toxins have also been described, with 92% of *C. jejuni* strains tested showing haemolysis on blood agar. Hepatotoxin and shiga-like toxin have also been proposed but studies on these toxins are contradictory [49].

With the exception of CDT, genes encoding proposed toxins have not been isolated or identified from sequencing projects [29]. The importance of these proposed toxins in disease remains unclear.

1.1.4.7 Surface polysaccharide structures

All strains produce lipooligosaccharide (LOS); a lipid A molecule joined to core oligosaccharide. LOS with a structure mimicking human ganglioside GM1, thought to be produced to evade the immune system, has been postulated to be an important factor in the development of GBS [52]. In addition it was thought that about one third of strains also produced a high molecular-weight lipopolysaccharide (LPS) which contained an O-chain consisting of repeating oligosaccharide. This has since been shown to be a capsular polysaccharide. Interestingly, capsule has been shown to be present even in strains previously not thought to produce LPS [53]. The capsule has also been proved to be the basis of the Penner typing scheme (section 1.1.5.1), and it has been proposed to aid surface

spreading, contribute to serum resistance, phagocytic killing and cell toxicity [15]. Capsular polysaccharide is poorly immunogenic which may aid in resistance to host-specific immune response [35] and may protect the cells from desiccation when in the environment.

1.1.4.8 Environmental survival

Anything which aids growth or survival in either the host or general environment could be considered a virulence factor if it aids transmission. Oxygen stress defences are used to deal with toxic oxygen metabolites produced during normal metabolism, during transmission or when in contact with host immune defences [12]. Superoxide stress defence is mediated by the superoxide dismutase (SOD) SodB whilst peroxide stress defence is mediated by catalase and alkyl hydroperoxide reductase (AhpC) [12].

The majority of *Campylobacters* are phenotypically catalase-positive (KatA). Catalase reduces oxygen stress by detoxifying hydrogen peroxide to oxygen and water [10]. Prior exposure to oxidative stress has been shown to increase rates of invasion, and catalase has been shown to contribute to intramacrophage survival. However it does not play a role in intraepithelial cell survival [54].

SodB is a superoxide dismutase which catalyzes the breakdown of superoxides into hydrogen peroxide and oxygen; mutants in *sodB* are attenuated in intracellular survival and colonization [10]. AhpC alkyl hydroperoxide reductase converts reactive hydroperoxides to the corresponding alcohols [10].

The ability of *Campylobacter* to persist outside the host environment may be an important factor in transmission and therefore a determinant of which strains may infect individuals as *C. jejuni* is thought to be susceptible to a wide range of environmental stresses and does not grow at ambient temperatures. *Campylobacter* lose culturability at different rates when transferred to water [16]. *Campylobacter* are found in natural water sources throughout the year but appear to survive better when the water is cold [34]. Some strains

are able to survive for up to four weeks at 4-10°C: this persistence was increased when bacteria were introduced with biofilms of indigenous water flora [16]. Factors determining strain variation in persistence are currently poorly understood but different isolates may vary in the virulence determinants they carry [16].

1.1.4.9 Plasmids

Recently the role of plasmids in *Campylobacter* virulence has been studied [55;56]. However, not all highly invasive strains have plasmids; the prevalence of plasmids in *Campylobacter* has been estimated by several sources as 19-53% [55]. In a survey of 688 isolates from diverse sources 32% were found to harbour plasmid DNA of size ranging 2-162 Kb with 16% harbouring multiple plasmids. No plasmid type was common to all *Campylobacter* [57].

Although plasmids may be involved in virulence there are a large number of clinical isolates that possess no plasmids so chromosomal determinants must also be of importance in virulence. Plasmids have also been implicated in the spread of antibiotic resistance. Tetracycline resistance is largely associated with plasmids [57] but other resistances may be chromosomally mediated [58]. Two *C. jejuni* plasmids both containing type IV secretion systems have been identified in strain 81-176; pVir and pTet [55]. These plasmids will be discussed further in chapter 3.

1.1.4.9 Chicken colonization

The colonization of chickens by *Campylobacter* is thought to be an important consideration in disease causation by the organism as consumption of poultry products has been implicated in transmission of the bacteria to humans, although the source of most infections remains unidentified [20]. Levels of chicken colonization are high with up to 10^9 *Campylobacter*

being recovered from a single chicken [59]. Different strains of *Campylobacter* are known to differ in their ability to colonize chickens [60].

1.1.4.10 Pathogenesis Summary

Phenotypic differences have been observed for traits implicated in virulence such as adherence, invasive properties, toxin production, serum resistance, chicken colonization potential, aerotolerance and temperature tolerance [29]. These phenotypic differences may correlate with differences in clinical outcome of disease, survival of the bacterium in the environment and transmission of the bacterium between hosts. Phenotypic difference may reflect underlying genotypic diversity.

1.1.5 Subtyping and diversity

Subtyping methods have been developed for *Campylobacter* and are frequently used in surveillance and epidemiological studies. In order to trace sources of infection, discrimination between different strains is necessary.

1.1.5.1 Serotyping

Two well established subtyping techniques are serotyping schemes. The Lior scheme is based on heat-labile antigens and a bacterial agglutination method [61] which now recognizes over 100 serotypes of *C. jejuni* [29]. The Penner scheme is based on heat-stable antigens using a passive hemagglutination technique [62]. The Penner scheme has recently been shown to be based on capsular polysaccharide [53] and currently identifies more than 60 serotypes of *C. jejuni* [29]. Phage typing has been used to give finer discrimination, and there are currently 336 serotype-phage type combinations. However there are a number of strains which remain untypable using traditional methods (19% of human isolates) [19]. Serotyping has been used to monitor *Campylobacter* on contaminated foodstuffs. A survey

of poultry food products in Denmark found that 85% of *Campylobacter* isolates were *C. jejuni* with certain Penner serotypes being more common than others [63].

Serotyping techniques are labour intensive and as such are largely limited to reference laboratories [34]. Molecular techniques already in use for other bacteria have been adapted for *Campylobacter* in order to provide more accessible typing procedures [64].

1.1.5.2 Molecular subtyping

Flagellin typing (*fla* typing) utilises restriction fragment length polymorphism (RFLP) in the PCR-amplified flagellin locus [29]. Due to variations in procedure, results from different laboratories cannot be directly compared which limits the usefulness of this assay for tracking infections [64]. In addition, due to recombination in the flagellin locus, the long term applicability of this assay for subtyping has been called into question as it does not accurately represent the entire genome [65]. However, *fla* typing has been used to demonstrate not only diversity between environmental isolates of *C. jejuni* but also to link certain *fla* types from environmental isolates to those from cases of human campylobacteriosis in the same geographical area [66].

Pulsed-field gel electrophoresis (PFGE) also shows diversity between isolates of *C. jejuni* [29;66]. The method is based on digestion of the bacterial chromosome by restriction enzymes that cleave the DNA infrequently. A major problem with this is that differences in electrophoretic conditions can lead to apparent differences in the profiles obtained even for the same DNA preparation which may make the comparison of different PFGE patterns unreliable [64].

Ribotyping involves gel electrophoresis of digested genomic DNA followed by Southern blot hybridization with a probe specific for rRNA genes. There is limited discriminatory power for this typing method due to there being only 3 rRNA gene copies

present in the genomes of *Campylobacter* species which means isolates can not be reliably identified at the subspecies level [64].

Randomly amplified polymorphic DNA (RAPD) analysis uses short non-specific primers to arbitrarily amplify DNA products under low-stringency PCR conditions. Band patterns consist of both weak and strong amplicons which can complicate interpretation. In addition up to 14% of strains examined may be untypeable due to DNase activity and there is poor reproducibility [64]. RAPD analysis has been used to distinguish between invasive and non-invasive isolates based on band differences; a distinct RAPD profile was found in 63% of invasive strains but was also found in 16% of non-invasive strains [67]. RAPD analysis has also shown genetic diversity between *C. jejuni* isolates from human faeces, seawater and poultry products [68].

Amplified fragment length polymorphism (AFLP) analysis involves the complete digestion of chromosomal DNA with two restriction enzymes, one with a 4 bp recognition site and the other with a 6 bp recognition site, followed by PCR amplification based on restriction sites [64]. The bacterial subtypes recognized by one technique often do not correlate with the subtypes determined by other typing techniques [29]. However AFLP analysis and multilocus sequence typing (MLST) have been shown to give similar genetic groupings when performed on the same isolates [69]. In one study AFLP analysis identified more than 100 different profiles amongst human, chicken and cattle isolates of *C. jejuni* and showed that isolates from human and cattle were more likely to show similar banding patterns than those from chickens [69].

A major limiting factor in some subtyping schemes is the reproducibility and comparison of results between laboratories as well as the fact that a number of strains are untypable using certain techniques. In order to provide a standardised test that is easy to perform and compare between laboratories, an MLST scheme for *Campylobacter* has been

set up based on seven housekeeping loci [70]. The results of initial typing using 194 strains indicate that *C. jejuni* is genetically diverse with a weakly clonal population structure. Using this technique 155 sequence types (STs) were observed with 26% being unique. Some STs were consistent with Penner serotype however some displayed a high level of diversity within serotypes [70].

1.1.5.3 Subtyping Summary

These subtyping techniques demonstrate a wide range of phenotypic and genotypic diversity between *Campylobacter* isolates in different environmental and clinical settings, although it is difficult to compare between techniques and even between laboratories. It is difficult, based on these techniques, to know the full extent of genotypic diversity within the species *C. jejuni.*

1.2 Genomic studies

In some bacteria an increase in observed pathogenicity has been attributed to the uptake and incorporation of virulence genes which in some cases cluster in regions known as pathogenicity islands [71], for example the *cagA* pathogenicity island of *Helicobacter pylori* [72]. The *cag* pathogenicity island is a 40 Kb island flanked by direct repeats; this island encodes 31 CDSs including the CagA cytotoxin and a type IV secretion system [73]. Pathogenicity islands are characterized by different G+C content to the core chromosome G+C content, instability, integration at specific loci e.g. tRNA genes, presence of mobility elements and the presence of direct repeats [71]. Pathogenic *Yersina* can be divided into low-pathogenicity strains, which induce mild intestinal infection in humans, and high-pathogenicity strains, which induce severe systemic infection in humans [74]. Several genes responsible for the high-pathogenicity phenotype are clustered on a genomic island termed the high-pathogenicity island [75]. This high-pathogenicity island present in *Yersinia* spp. is particularly unstable and can be lost at frequencies of up to 10^{-5} per generation [76]. Virulence genes may also be located on transmissible genetic elements such as transposons [71], plasmids [77;78] or bacteriophages [79;80]. However, in *C. jejuni* the degree to which genetic differences contribute to variations in disease outcome and epidemiological characteristics is as yet unclear [29]. In the current genomics age large scale methods have been adopted to explore bacterial strain diversity.

1.2.1 Genome sequencing

1.2.1.1 *C. jejuni* strain NCTC 11168

C. jejuni strain NCTC 11168 was isolated from a case of human campylobacteriosis in 1977 and is a commonly used laboratory strain. In 2000 the sequence of *C. jejuni* strain NCTC 11168 was published [8]. The genome of strain NCTC 11168 is 1641481 bp with 94.3% predicted to code for proteins. Out of 1654 predicted Coding Sequences (CDSs) approximately 22% of *C. jejuni* genes had no matches to previously identified genes with known function. Only 55.4% of *C. jejuni* CDSs had orthologues in the closely related bacterium *Helicobacter pylori*. The majority of predicted CDSs did not appear to be organized into operons or clusters. Exceptions to this include the lipooligosaccharide (LOS) and capsular polysaccharide biosynthesis clusters. Interestingly these polysaccharide biosynthesis clusters have a lower G+C content than the rest of the chromosome. The sequence data was largely unable to elucidate novel candidate genes for the production of toxins, adhesins, invasins and other classical virulence determinants with the exception of components of sialylation pathways and the cytolethal distending toxin genes (*cdtABC*). There were also a lack of bacteriophage, inserted sequence (IS) elements and obvious pathogenicity islands [8].

One striking discovery from the genome sequence was the identification of 24 regions of sequence polymorphism which mainly consisted of poly G/C tracts which alter in length due to slipped-strand mispairing. Slipped-strand mispairing involves denaturation and displacement of the strands of DNA in a duplex followed by mispairing of complementary bases within a short repeat, e.g. a homopolymeric tract. When slipped-strand mispairing is followed by replication or repair this can lead to the insertion or deletion of one or more bases within the homopolymeric tract [81]. If the number of bases inserted or deleted is not a multiple of three this in turn can alter the expression of specific proteins by shifting their

translational reading frame. Tract length variation resulting in translational frameshifting has been shown in other bacteria to be responsible for phase variation whereby bacteria randomly vary surface properties or antigenicity [82]. The phase-variable genes in *C. jejuni* predominantly cluster in the lipo-oligosaccharide, capsule and flagellar biosynthesis regions indicating that *C. jejuni* is also using this mechanism to alter surface properties.

1.2.1.2 *C. jejuni* strain RM1221

More recently the genome sequence of *C. jejuni* strain RM1221, which was isolated from the skin of a retail chicken, has been published [9] and at 1777831 bp was larger than that of strain NCTC 11168. Strain RM1221 was predicted to encode 1884 proteins which was 230 more than in strain NCTC 11168. The LOS and capsular polysaccharide biosynthesis loci were predicted to encode many different CDSs to that of strain NCTC 11168. The genomes of strain NCTC 11168 and strain RM1221 were shown to be syntenic but this synteny was disrupted by four genomic islands in strain RM1221. Three of the islands were phage derived with the fourth likely to be of plasmid origin [9]. Strain RM1221 also appeared to be devoid of functional IS elements.

1.2.2 Comparative genomic studies

1.2.2.1 Microarrays

Microarrays have been popular in recent years, with many bacterial species being studied, and have proved useful for comparing diversity with respect to sequenced strains. Dorrell *et al.* [51] have used a microarray approach to reveal extensive genetic diversity within *C. jejuni*: 21% of genes in strain NCTC 11168 were absent or highly divergent in one or more of the 11 *C. jejuni* strains tested. Genes for virulence determinants hypothesised to be necessary for *C. jejuni* to cause disease in humans including the cytolethal distending toxin, flagellar structural proteins, phospholipase A, the PEB antigenic surface proteins, and

proteins potentially involved in host pathogen interactions such as CiaB, CadF and CheY were conserved in all the strains tested [51]. With regard to strain diversity this approach can only show genes that are missing or significantly divergent compared to strain NCTC 11168 and not replacements or insertions that may exist within the genome of these different strains. There have also been two other comparative papers using strain NCTC 11168 microarrays and various other strains [83;84]. Pearson *et al.* [83] showed that between 2.6% and 10.2% of the NCTC 11168 CDSs were absent or divergent in the 18 *C. jejuni* strains tested. Variable CDSs were located in seven plasticity regions on the genome of strain NCTC 11168 which included the LOS and capsular polysaccharide biosynthesis loci as well as flagellin biosynthesis and post translational modification loci. Taboada *et al.* [84] compared 51 *C. jejuni* strains to NCTC 11168 showing that 20% of NCTC 11168 CDSs were divergent in at least one of the test strains. Most of these variable genes were located in 16 plasticity regions including surface polysaccharide biosynthesis and modification loci and restriction modification loci as well as regions containing hypothetical CDSs.

C. jejuni strain ATCC 43431 has been analysed using a shotgun microarray to identify genes unique to this strain in comparison to strain NCTC 11168. This method identified 130 complete and incomplete CDSs [85]. Many LOS and capsule associated genes were discovered along with restriction modification genes, integrases and hypothetical genes. However, these CDSs were only fragments and were not expanded to give genomic context.

1.2.2.2 Subtractive hybridization

Subtractive hybridization has been used for some time as a method for identifying genes expressed in one cell type but not another by hybridizing cDNA to RNA [86;87]. The method has since been adapted for identifying DNA differences between bacterial strains [88;89]. Subtractive hybridization was evaluated as a method of comparing genomes using

the two sequenced strains of *H. pylori* where it was shown to identify 95% of CDSs unique to one strain compared to the other [90].

Ahmed *et al.* [91] have used subtractive hybridization as a technique to identify gene fragments in strain 81116 that were not present in strain NCTC 11168. Strain 81116 is also a human campylobacteriosis isolate but has been proposed to show greater colonization potential of chickens than the strain NCTC 11168. In strain 81116, 24 fragments that were unique to this strain (less than 75% identity at the nucleotide level to NCTC 11168) were identified and used to hybridize to genomic DNA from 9 other strains: one insert was unique to 81116, one was present in all 9 tested strains and the rest showed variable distributions [91]. Gene fragments identified included those with similarity to restriction-modification enzymes, arsenic-resistance genes and cytochrome C oxidase III genes [91]. However these fragments were not characterized further to obtain entire genes and to assess their distribution across the genome. The method of subtractive hybridization has drawbacks, including limited coverage of the genome, and the production of only small fragments of novel DNA which must then be cloned after manipulation which can introduce biases.

1.2.2.3 Differential hybridization in other organisms

Others have addressed the problem of identifying novel sequences in one bacterial strain compared to another. Liang *et al.* used a differential hybridization approach to identify differences between *Pseudomonas aeruginosa* strain X24509 and the sequenced strain PA01 [92]. They developed a method of differential hybridization using arrayed libraries of cloned DNA fragments and found a genomic island (PAGI-1) that was present in 85% of pathogenic isolates. PAGI-1 was sequenced and CDSs within it predicted. Several CDSs showed sequence similarity to known genes including dehydrogenase genes, genes coding for proteins implicated in detoxification of reactive oxygen species and transcriptional regulators. The role of PAGI-1 in *Pseudomonas aeruginosa* is unknown but it may encode

genes that have a role in protecting the cell against oxidative damage and the transcriptional regulators may control the expression of chromosomal genes providing a selective advantage for strains that have acquired PAGI-1. This study showed that differential hybridization is a valid approach for identifying virulence factors in conjunction with sequence data as Liang *et al*. were able to identify the insertion site of PAGI-1 [92]. This method has several advantages: i) there are no cloning steps after manipulation, ii) a breadth of coverage can be achieved by generating large libraries, and iii) the length of sequences studied can be modified by altering the insert sizes of the clone libraries.

1.3 Aims of this thesis

C. jejuni has been demonstrated to be both genotypically and phenotypically diverse. Differences in phenotype between strains can often be due to novel genes or islands present in one strain compared to another. These genotypic differences may relate to different clinical outcomes, epidemiological characteristics or environmental persistence. In order to explore this possibility:

i) Genomic DNA arrays of *C. jejuni* strains with different phenotypic characteristics will be created.

ii) DNA present in the strains to be tested, which is not present in the sequenced strain NCTC 11168, will be identified using a differential genomic DNA hybridization approach with small-insert libraries.

iii) Strain specific DNA will be characterized by sequencing and annotation in order to identify potential virulence or survival factors.

iv) The extent and context of novel regions containing these factors will be determined relative to the chromosome of strain NCTC 11168 by sequencing larger-insert libraries.