Characterisation of Two Unusual Phylogenetic Clades of Norway Rat Associated *Campylobacter jejuni*

A thesis submitted for the degree of Doctor of Philosophy in Microbiology

Microbiology Research Group
Biomedical Sciences
School of Biological Sciences
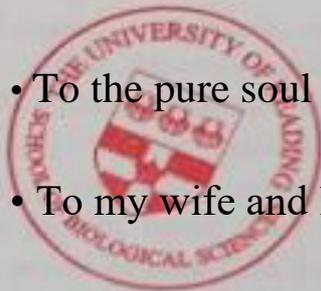University of Reading
Reading, UK

Othman A. Mohammed
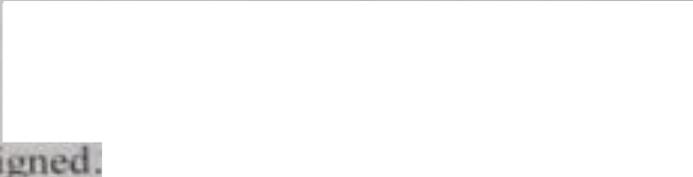
January 2018

Declaration

Dedication

I confirm that this thesis is my own work and the use of all material from other sources has been properly acknowledged.

- To the pure soul of my mother and father;

- To my wife and little daughter;

- To my dear sisters and brothers;

- To those who taught me a word.

Signed.

Date....24/04 /2018

Abstract

Campylobacter jejuni is the most common cause of bacterial gastroenteritis worldwide. Isolates of this pathogen show substantial genetic variability, which results in differences in physiological and virulence properties. Sequence types commonly associated with gastroenteritis in humans belong to multilocus sequence typing (MLST) clonal complexes 45 and 21. From a recent survey of over 300 farm-associated rat *Campylobacter* isolates, two groups of *C. jejuni* strains, Rat Group 1 (RG-1) and Entner–Doudoroff pathway positive Rat Group 2 (RG-2), were identified as distinct clades based on wgMLST (whole genome multilocus sequence typing). While the RG-2 strains are related to ED-positive CC-45 strains, the RG-1 group was not represented in the extensive database of chicken and human isolates. The main aim of this study was to characterise both RG-1 and RG-2 strains to identify features that might relate to their apparently specific niche association.

Both morphology and genetic markers confirmed that both RG-1 and RG-2 groups belong to the species *C. jejuni*. All strains in both groups grew more poorly in Mueller Hinton Broth (MHB) than generalist strains also isolated from Norway rats. Quantitation of amino acids utilised during growth demonstrated the same catabolic preference for the amino acids Ser, Asp, Glu and Pro, in both cases. Interestingly, rat mucins enhanced the poor growth of RG-1 strains, suggesting provision of some important growth factor. This together with the inability of the RG-1 strain Dg147, to colonise chickens would be consistent with a more defined host specificity of RG-1 strains, possibly with rats as a preferred host. The draft Dg147 genome was closed and gene-by-gene comparison identified 1561 core and 269 accessory genes, within this group. All RG-1 strains lacked the following three genes as defined by a cgMLST set of core *C. jejuni/coli* genes (www.pubmlst/campylobacter); *cj0145*, encoding a TAT-dependent alkaline phosphatase (PhoA), *cfbpB* (*cj0174c*) encoding a putative iron-uptake ABC transport system permease protein and *cj1721c*, encoding an uncharacterised beta-barrel outer membrane protein. The majority of the accessory genes were from two large phage genome insertions. Further genome comparisons with a set of pathogenic *C. jejuni* strains identified 35 core RG-1 genes that were absent from the human pathogen strains and potentially represent rat or niche-associated genes for future study.

This study also provides the first demonstration of the ability of *C. jejuni* strains to grow on glucose. Following closure of the genome of one RG-2 strain, Dg275, it became evident that the *glc* locus is inserted into both the *rrnA* and *rrnB* operons in this and other strains, but not in *rrnC*. Diauxic growth was seen in both minimal media and MHB supplemented with Ser and glucose. In MHB, Ser and Asp were firstly depleted in the exponential phase of growth followed by initiation of glucose utilisation and a greatly extended stationary phase of growth. This raises interesting questions regarding regulation of the *glc* locus in *C. jejuni*. There appeared to be no advantage in the ability to use glucose in colonisation of chickens. All strains recovered from one of seven tested chickens had lost the ability to utilise glucose due to a point mutation within *glk* creating a pseudogene, a result consistent with the absence of ED positive strains in chicken isolates in the database. During this study, demonstration of efficient biofilm formation and chicken colonisation by generalist strains isolated from Norway rats, highlights the potential of these pests as a reservoir for *C. jejuni* on the farmyard. Future studies should focus on the ability of RG-1 strains to colonise laboratory rats and therefore as a potential model as well as benefits of glucose utilisation to RG-2 strains.

Dedication

• To the pure soul of my mother and father;

• To my wife and little daughter;

• To my dear sisters and brothers;

• To those who taught me a word.

Acknowledgements

List of Abbreviations

AGE	Agarose Gel Electrophoresis
ANOVA	Analysis of variance
BA	Blood Agar
BIGSdb	Bacterial Isolate Genome Sequence Database
BIOLOG	Biolog phenotypic microarray
Blast	Basic local alignment search tool
CC	Clonal Complex
CDT	Cytolethal Distending Toxin
CFU	Colony Forming Unit
CPS	Capsular Polysaccharide
DW	Distilled Water
DMEMf	Dulbecco's Modified Eagle's Medium plus iron ascorbate (IA)
dNPTs	Deoxy Nucleoside Triphosphates
dpi	Days Post Infection
ED	Entner-Doudoroff
EMP	Embden-Meyerhof-Parnas
EDTA	Ethylenediaminetetraacetic acid
GGT	Glutamate by γ-Glutamyl Transpeptidase
IA	Iron Ascorbate
glc	Glucose
LOS	Lipooligosaccharide
mCCDA	Modified Charcoal Cefoperazone Deoxycholate Agar
MEM	Minimum Essential Medium
MHB	Mueller Hinton Broth
Min	Minutes
µl	Microliter
MLST	Multilocus Sequence Typing
mM	Milimolar
Mpx	Megapixels
MRD	Maximum Recovery Diluent
MW	Molecular Weight
NCBI	National Center for Biotechnology Information
nH2O	Nanopure water
ODxxx	Optical Density at wavelength xxx nm
ORF	Open Reading Frame

PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
rMLST	Ribosomal Multilocus Sequence Typing
rrn	Ribosomal RNA
pmole	Pico mole
RG-1	Rat group 1
RG-2	Rat group 2
RM-1	Round Motility 1
RM-2	Round Motility 2
RPM	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
RT	Room Temperature
SEM	Scanning Electron Microscopy
Spp.	Species
ST	Sequence Type
TAE	Tris-acetat EDTA buffer
TE	Tris-EDTA
TEM	Transmission Electron Microscopy
TSB	Tryptic Soy Broth without dextrose
wgMLST	Whole Genome Multi Locus Sequence Typing
WGS	Whole Genome Sequence

Amino acid	Three letter code	Amino acid	Three letter code
Proline	Pro	Alanine	Ala
Serine	Ser	Leucine	Leu
Aspartate	Asp	Methionine	Met
Asparagine	Asn	Tryptophan	Trp
Glutamate	Glu	Valine	Va
Glutamine	Gln	Histidine	His
Arginine	Arg	Tyrosine	Tyr
Lysine	Lys	Cysteine	Cys
Threonine	Thr	Isoleucine	Ile
Glycine	Gly	Phenylalanine	Phe

Contents

XIII

List of Figures

XV

XVI

List of Tables

List of Appendices

Chapter 1

Literature Review

1.1   The genus *Campylobacter*

Theodor Escherich was the first scientist to describe the spiral shape of a bacterium in diarrheal stools from infant and kitten samples in 1886 (Kist, 1986). The first successful vibrio shape of the bacterium in pure culture was obtained by McFaydean and Stockman in 1906 (Skirrow, 2006). They isolated the bacterium from foetal tissues of pregnant ewes. Vibrio like organisms were isolated from many animal sources including from an aborted lamb in 1913 (*Vibrio fetus*) as well as from the intestines of sheep and cattle (*Vibrio jejuni*) in 1931. The taxonomy of genus *Campylobacter* has been reviewed by Debruyne et al. (2008) and Veron and Chatelain (1973). Species *V. sputorum*, *V. coli*, *V. fetus, V. bubulus,* and *V. jejuni* were transferred to the *Campylobacter* genus by Sebald and Véron in 1963, using biochemical and serological studies to distinguish these species from known species of *Vibrio* (On, 2001). The genus *Campylobacter* was established as a new distinct genus of bacteria. Improved selective isolation involving procedures that used enriched medium containing lysed horse blood, sodium succinate, and antibiotics, helped to expand the number of species identified within the genus *Campylobacter* (Doyle and Roman, 1982; Itoh et al., 1980). In addition, some important bacteria were subsequently excluded from this genus. Romaniuk et al., (1987) used 16S rRNA sequencing to show that *C. pylorus* was in fact more closely related to *Wolinella succinogenes* (**Figure** 1.1). Goodwin et al. (1989) used different criteria, such as cellular fatty acid content, growth properties and morphology to show that *C. pylorus* was distinct from *Wollinella*. In light of this, they established the new genus, *Helicobacter*, to which they assigned *C. pylori*. The *Campylobacter* genus belongs to the family *Campylobacteriaceae,* the order *Campylobacteriales* and the class Epsilonproteobacteria (Debruyne et al., 2008).

Vandamme et al. (1991) and On (2001) revised and divided the *Campylobacter* genus into the following species and subspecies; *C. jejuni* subsp. *jejuni, C. jejuni* subsp. *doylei, C. coli, C. upsaliensis, C. lari, and C. helveticus*. They are the most common *Campylobacter* species of gastroenteritis. *C. concisus, C. sputorum, C. curvus, C. rectus, C. gracilis, C. showae,* and *C. hominis* are the species that have mostly been identified in the human oral cavity. *C. hyointestinalis* subsp. *hyointestinalis, C. hyointestinalis* subsp. *lawsonii, C. fetus* subsp. *fetus,*

2

and C. fetus subsp. *venerealis* are closely related phylogenetically. Efficient isolation rapidly established *Campylobacter* as one of the most common causes of bacterial induced gastroenteritis (Kim et al., 2016, Oyarzabal and Fernández, 2016). NCTC11168 genome, as example of a human isolate of *C. jejuni* subsp. *jejuni*, was fully sequenced and published on 2000 (Parkhill et al., 2000). It has a circular chromosome of 1,641,481 base pairs (bp) with 30.6% G+C content and encodes a predicted 1643 proteins (Gundogdu et al., 2007). This is much smaller than the genome of *Escherichia coli* O157:H7 that is 5.5 Mbp (Hayashi et al., 2001). Sequencing of additional *C. jejuni* genomes has highlighted extensive sequence variation between strains (Pendleton et al., 2013).
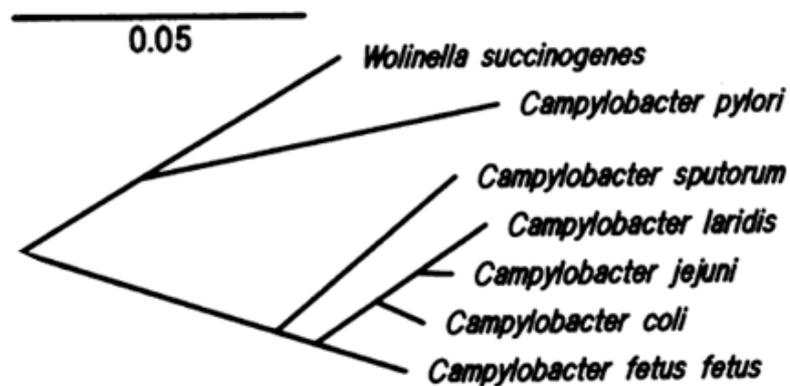
Figure 1.1: Phylogenetic relationship of some *Campylobacter* species with *Campylobacter pylori* and *Wollinella succinogenes*. The tree is based on complete 16S rRNA sequences. The scale bar shows an evolutionary distance of 0.05 nucleotide changes (Romaniuk et al., 1987).

1.2 General features of *Campylobacter jejuni*

Campylobacter spp. are Gram-negative, fastidious, microaerophilic bacteria with spiral rods with a slender morphology approximately 0.5-5μm long and 0.2-0.9μm wide and often possess one or two flagella (**Figure 1.2**). *C. jejuni* requires 5-10% O_2 as well as a CO_2 concentration of 5-10% to grow (Kelly, 2001, Ketley, 1997). *C. jejuni* can grow within the temperature range of 32 to 45°C, has an optimal growth temperature of 37 to 42°C, but can survive in the environment at temperatures lower than 32°C. These optimal temperatures are reflective of the gut temperatures of mammals and birds, a typical habitat of *Campylobacter* spp., in particular *C. jejuni*. Whilst the typical *Campylobacter* cell morphology is spiral or curve-shaped under optimal growth conditions, the bacterium undergoes a change in

morphology to a coccoid form in unfavourable conditions, such as under aerobic conditions (Jang et al., 2007), prolonged growth in a liquid medium, or under starvation conditions (**Figure 1.2**) (Griffiths, 1993, Thomas et al., 2002). Whether coccoid cells are viable but in a non-culturable physiological state (VBNC) remains an open question (Buck et al., 1983). A VBNC state would permit bacterial cells to remain in a dormant state and their forms should have specific features distinguishing them from dead cells. It has been reported that with *C. jejuni* the VBNC (coccoid) membrane is able to retain its integrity unlike deceased spiral cells that has lost the integrity of the membrane structure (He and Chen, 2010). The possibility that coccoid cells might still be capable of infection, colonise certain hosts and transform into a culturable state has also be considered (Chaisowwong et al., 2012, Ikeda and Karlyshev, 2012, Li et al., 2014).
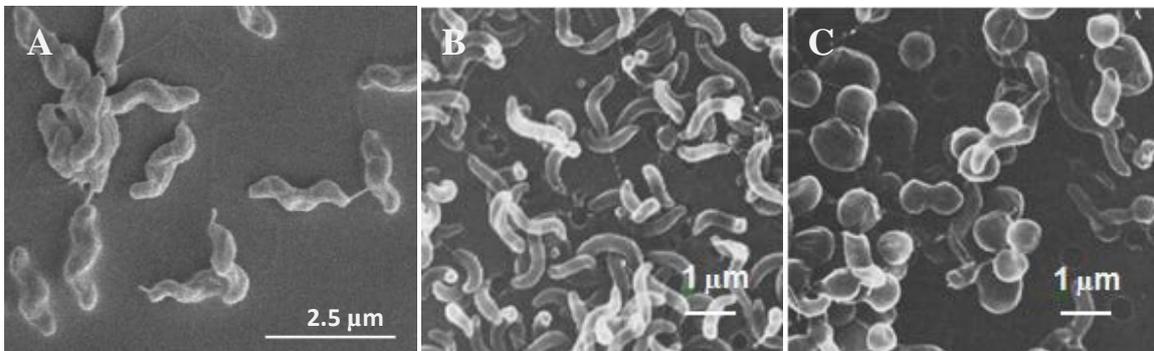
Figure 1.2: Scanning electron micrographs showing the cell morphology of *C. jejuni*. (**A**) Helical cell shape of 81-176 (Esson et al., 2017). (**B**) Helical morphology of HC *C. jejuni* strain changed to (**C**) the coccoid 'viable but non-culturable' form (Rollins and Colwell, 1986).

1.3   Epidemiology of *Campylobacter jejuni*

Campylobacter spp. are the most common cause of bacterial gastroenteritis in the UK, resulting in substantial public health and economic problems (Ketley, 1997). Campylobacteriosis is also the major cause of bacterial gastroenteritis in other industrialised and developing countries worldwide (Meerburg and Kijlstra, 2007a, Maćkiw et al., 2012, Cody et al., 2009). The main species linked to human gastroenteritis is *C. jejuni* and to a lesser extent *C. coli.* Both are commonly found in the intestinal tract of wild and farmed animals (PHE, 2017b). The number of *Campylobacter* related cases has been reported to be greater than the total number of diarrhoeal disease cases caused by *Shigella*, *Salmonella* species, and *Escherichia coli* O157:H7 put together (Blaser, 1997). Maćkiw et al. (2012) cited 190,566 confirmed cases of campylobacteriosis in 2008 in the European Union, whilst the number of laboratory reports for *Campylobacter* spp. cases in England and Wales alone was stated to be 65,032 in 2012 (PHE, 2014). Differences in published rates of infection in different countries may be attributable, in full or in part to different methods of identification and reporting of campylobacteriosis. From 2006 to 2015, in England and Wales, reported *Campylobacter* infections increased from ~46,000 in 2006 to ~65,000 in 2012, before then decreasing slightly in the following three years (**Figure 1.3**). Infection with *Campylobacter* was approximately three times higher than with *Salmonella* and *E. coli* together in this similar geographic region (PHE, 2017a). It has been estimated that the total number of human cases of *Campylobacter* infection in the UK is likely to be closer to 340,000-500,000 each year (Tam et al., 2012, Cody et al., 2009).
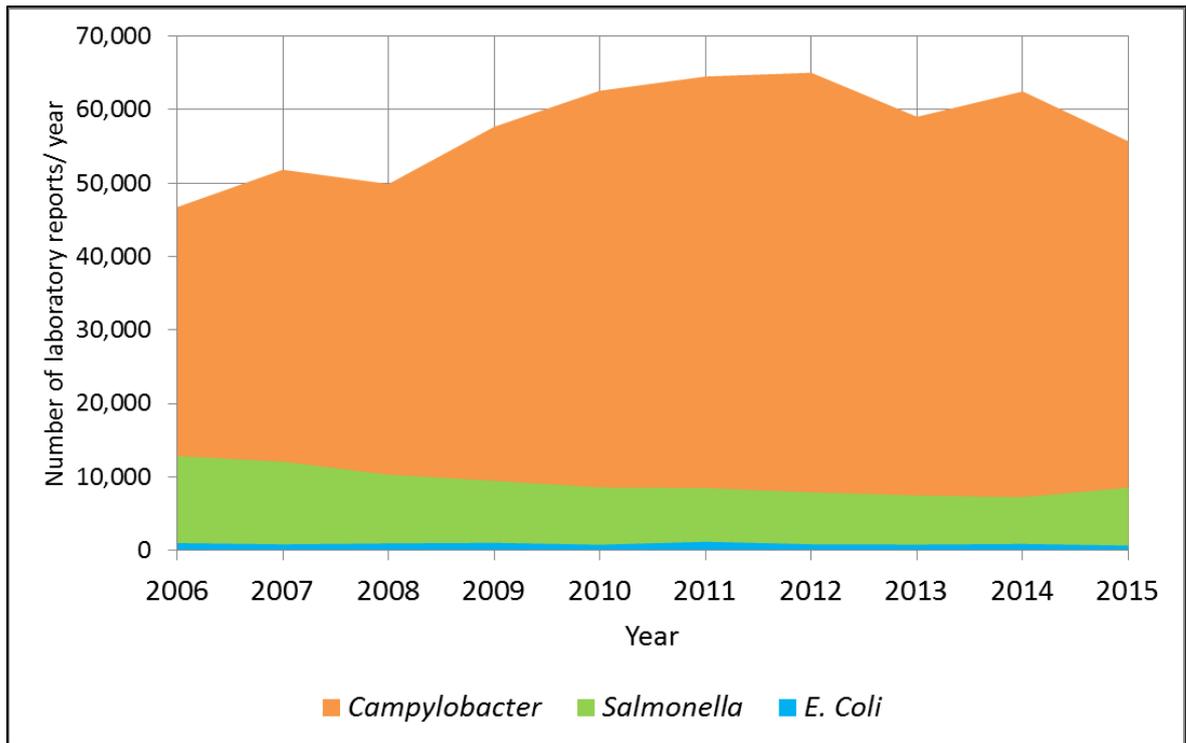
Figure 1.3: Annual laboratory reports of *Campylobacter, Salmonella* and *E. coli* in England and Wales. This data is taken from Public Health England, (PHE, 2017a).

It is well established that in developing countries, young children (less than 5 years old) are most likely to become infected with *Campylobacter*, with *Campylobacter* induced gastroenteritis being less common in adolescents and adults (Megraud et al., 1990, Pazzaglia et al., 1991). In contrast, in the industrialised world, rates of campylobacteriosis are high in young children, decrease in older children but then increase again in young adults and stay high throughout the adult population (Janssen et al., 2008). One review attributed this to acquired immunity, due to the high level of exposure in developing countries (Zilbauer et al., 2008). However, in view of the high levels of campylobacteriosis in industrialised countries, one might expect the same to occur here in the UK. One example might be that elderly people in Scotland seem to be infected with uncommon serotypes (Olson et al., 2008). This could indicate that repeated exposure has led to acquired immunity to common serotypes in older people. Alternately, it may reflect the susceptibility of older people, and resistance of healthy adults, to some strains of *C. jejuni* that are not commonly associated with gastroenteritis. Other factors have been reported to influence rates of infection, including gender and season of the year. *C. jejuni* diarrhoeal disease is slightly lower in females than males in the UK and, although it occurs throughout the year, the dominant time is during late spring and summer

(PHE, 2017a). There is evidence that human infection by different clonal complexes of *C. jejuni* varies during the year, except for the most common ST-21 clonal complex. Cases of this clonal complex are reported more or less throughout the year (Sopwith et al., 2006). Infectious cases due to ST 45, 283, and 42 clonal complexes arose throughout the summer; whereas cases arising from ST 353 and 403 clonal complexes increased during the winter (Cody et al., 2012). Reasons for the variations in seasonal cases and small differences in gender susceptibility are not fully understood, but may be in part related to lifestyle and to exposure to different reservoirs in winter, spring, and summer (Blaser, 1997, Meerburg and Kijlstra, 2007a).

Campylobacter is commonly associated with the intestinal microbiota of many domesticated animals and birds such as geese (Aydin et al., 2001) cats and dogs (Baker et al., 1999), and poultry (Ellis-Iversen et al., 2012). During slaughter and processing, meat for retail becomes contaminated (Allen et al., 2007, Rosenquist et al., 2006). This applies to all animals (poultry, cattle, swine, and sheep) that are commonly used as a food source. Pets, including birds, cats, and dogs, might also have a role in carrying the bacteria and in transmission to other organisms. Sources of human *Campylobacter* infection are mainly from contaminated food, especially undercooked poultry such as chicken and turkey (Kaspars et al., 2014), unpasteurised milk, beef, and pork (Meerburg and Kijlstra, 2007b, Maćkiw et al., 2012). It is estimated that chickens are the main reservoir of *Campylobacter* and are responsible for 80% of human campylobacteriosis (Hermans et al., 2011). Another way of acquiring the infection is from touching contaminated animals, in particular during the processing of poultry, however direct transmission of *Campylobacter* between people appears to be rare and insignificant epidemiologically (Cody et al., 2009, Meerburg and Kijlstra, 2007a, Blaser, 1997). Furthermore, wild animals and contaminated environmental water also appear to have a less significant role in transmission of *Campylobacter* directly to humans (Wilson et al., 2008, Mughini-Gras et al., 2016). Farm associated Norway rats (unpublished data MacIntyre, Prescott, Maiden, McCarthy) and other rodents can be infected with this bacterium. Therefore, Norway rats have the potential to act as a reservoir and to contribute to transfer of *Campylobacter* within the farm environment (**Figure 1.4**) (Meerburg and Kijlstra, 2007a).
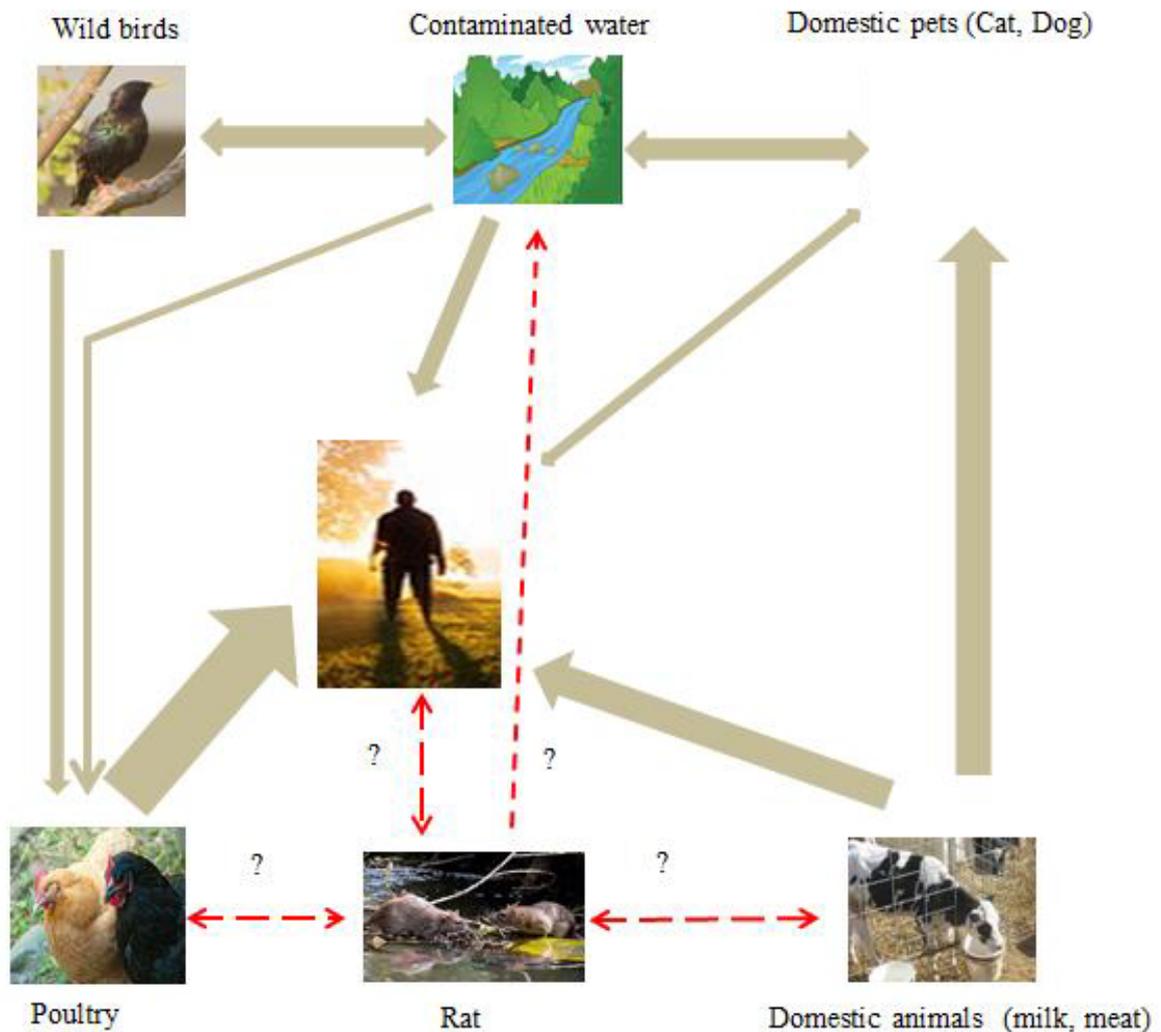
Figure 1.4: Sources and routes of *Campylobacter jejuni* transmission. There are a number of different reservoirs for infectious *C. jejuni*. The gastrointestinal tract of chickens, in particular the mucous layer, is colonised by bacteria, including *C. jejuni*, which can pass between chickens through the faecal-oral route. Drinking water may also be polluted with *C. jejuni* by faeces of animals and wild birds. Humans are directly infected by ingestion of polluted animal products such as milk and meat, especially domestic fowl or to a lesser extent from contaminated drinking water (Young et al., 2007, Dasti et al., 2010, Johnson et al., 2017). Block arrows indicate proven sources of *C. jejuni* infection, while dashed red arrows and question marks indicate potential routes of *C. jejuni* infection via rats.

To date (December 2017), data from 60,705 *C. jejuni/coli* isolates and 513 from non *jejuni/coli Campylobacter* have been submitted to the BIGSdb database. Approximately half of these isolates, 33,130 isolates (54.7%), are from the UK. The predominant species is *C. jejuni* (49,721 isolates, 82%), and the rest are from *C. coli* (10,645 isolates, 17.5%), *Campylobacter* spp. (196 isolates, 0.3%), *Campylobacter fetus* (3 isolates, <0.01%), *C. jejuni* subsp *doylei* (15 isolates, 0.02%), *C. lari* (9 isolates, 0.01%), and *C. hyointestinalis* (one isolate, <0.01%). The database also documents the source of the isolates, from which the sequence data was obtained. The majority of isolates (31,633 isolates, 52.2%) are from human sources and a significant number (15,314 isolates, 25.27%) are from chickens (**Table 1.1**). Not surprisingly, there is a correlation between the large number of human isolates entered and gastroenteritis in the database. Some sources (6.49%) are assigned as carriers, which include chickens, cattle, sheep, ducks, turkeys, and starlings. Among the 60,705 entries, 143 isolates (assigned as other animals) are from the University of Reading. These isolates were primarily isolated from farm associated Norway rats (137 isolates), plus 6 isolates are from a selection of animals from the same farms which are the focus of this study.

Table 1.1: Sources of *Campylobacter* isolates in the PubMLST database. (Last accessed December 2017) (Jolley and Maiden, 2010).

Source	Frequency of isolates	Percent of total isolates (%)
Human stool	28625	47.24
Human unspecified	2850	4.70
Human blood culture	158	0.26
Chicken	10332	17.05
Chicken offal or meat	4807	7.93
Other animals, includes Norway rats	* 143 sequenced isolates from the University of Reading	0.23
	459, ~ half assigned as carriers	0.75
Cattle	2417	3.99
Cattle faeces	282	0.47
Pig	1758	2.59
Wild bird	1413	2.33
Environmental waters	1568	2.59
Sheep	721	1.19
Sheep faeces	180	0.30
Starling	342	0.56
Turkey	380	0.63
Goose	236	0.39
Dog	269	0.44
Duck	559	0.92
Cow's milk	200	0.33
Other food	113	0.19
Farm environment	238	0.39
Broiler environment	175	0.29
Beef offal or meat	125	0.24
All others (28 values)	677	1.12
No value (unassigned)	1543	2.55

* *Campylobacter* Norway rat associated isolates

1.4   Subtyping of *Campylobacter jejuni/coli*

Typing is an essential tool in the epidemiological study of infectious diseases. Subtyping of bacterial species can be helpful for clarifying sources and modes of bacterial transmission in the food chain, describing geographical distribution, and identifying methods of control. In addition to the Penner serotyping scheme, many different sequence based methods have been tested for *C. jejuni* subtyping. These include antigen gene sequence typing (AGST), used for *flaA*, short variable region (SVR) typing, used for sequencing of the outer membrane protein *porA*. Currently multilocus sequence typing (MLST) based on sequencing seven house-keeping genes, and more recently whole genome sequencing (WGS) (Colles and Maiden, 2012) and core genome multilocus sequence type (cgMLST) (Cody et al., 2017) are used widely to type and monitor the phylogenetic relationship of bacterial strains.

1.4.1   The Penner serotyping scheme

Penner serotyping has been extensively used in epidemiological studies (Albert et al., 1992). This method is based on capsule polysaccharide (CPS) of *Campylobacter*. In this method, 47 antisera are used to differentiate *C. jejuni* serotypes. A recent study reported that 8 serotypes account for 50% of *C. jejuni* linked diarrhoeal cases worldwide and that serotypes HS4 complex, HS2 and HS1/44 are the most dominant *C. jejuni* serotypes associated with clinical disease (Pike et al., 2013). The operon encoding enzymes involved in *C. jejuni* polysaccharide capsule synthesis is located in a region of the genome that is hypervariable between strains and therefore differences in the synthesised sugars and polymers can be used to raise different antisera preparations and to differentiate strains. In addition, lipopolysaccharide can also be used to classify strains based on serotyping. Although making these antisera is complex and expensive, and has limited the use of this method (Preston and Penner, 1989, Karlyshev et al., 2000), serotyping has been a potential alternative method for identification of *C. jejuni* in large-scale surveillance experiments (Pike et al., 2013).

1.4.2   PorA protein

Outer membrane proteins (OMPs) are some of the most abundant proteins in Gram-negative bacteria. These proteins have unique structural properties most commonly based on a transmembrane amphipathic β-barrel. The major outer membrane protein (PorA) is composed of around 16 to 18 β-strands arranged in an amphipathic beta-barrel, with small turns in the base of the barrel towards the periplasmic space and long loops at the top of the barrel facing towards the outside of the cell (Cowan et al., 1992, Ferrara et al., 2016). The typical role of these proteins is to permit the passage of small hydrophilic molecules, such as sugars and amino acids, through the outer membrane (Zhang et al., 2000). It has been suggested that PorA may also have an important role in host adaptation as this protein has been implicated in antibiotic resistance and adherence to host cells (Cody et al., 2009, Ferrara et al., 2016). PorA exhibits the heat-modifiable characteristic typical of outer membrane (OM) proteins (Zhang et al., 2000). As PorA is a major surface protein, the surface loops are recognised by both bacteriophage and antibody and are highly variable. Thus, subtyping of *C. jejuni* PorA could be useful for localised epidemiological analysis either before undergoing more intensive molecular typing or combined with WGS phylogenetic analysis (Cody et al., 2009, Mohan et al., 2017, Jay-Russell et al., 2013).

1.4.3   Multi-Locus Sequence Typing (MLST) subtyping of *Campylobacter*

For MLST, fragments of 400-500bp from seven housekeeping genes, *aspA* (aspartase A), *glnA* (glutamine synthetase), *glyA* (serine hydroxymethyltransferase), *gltA* (citrate synthase), *glmM* (phosphoglucomutase, recorded as *pgm)*, *uncA* (ATP synthase subunit) and *tkt* (transketolase), are amplified and sequenced (Dingle et al., 2005, Lang et al., 2010). For each gene, sequences are compared to existing sequences in a database. Any new sequence, which is not identical to an existing sequence in the database, is defined as a new allele and allocated a new number. The allele profile of all seven different digit numbers are also identified as a sequence type (ST), again dependent on those present in the database. Any new allele profile is assigned as a novel ST and given as the next number in series (Sails et al., 2003). STs are also grouped into clonal complexes (CCs) where a CC is defined as a group of phylogenetically related STs sharing at least 4 alleles. This then represents a group of associated bacteria that may have evolved from a common ancestor (Griekspoor, P. et al.,

2013; Sheppard, S. K. et al., 2013). The availability of an online database (https://pubmlst.org/campylobacter) that documents ST number and metadata has contributed significantly to the development and clarification of the epidemiology of *Campylobacter* isolates on a local and global scale (Jolley and Maiden, 2010, Colles and Maiden, 2012). There are some clonal complexes from multiple host sources, while others appear to be host specific. Strains from ST21CC and ST45CC are frequently isolated from different hosts including humans, chickens, cattle, and sheep (Colles and Maiden, 2012). These CCs are defined as generalist strains. They presumably harbour genotypes, which have permitted them to colonise multiple hosts. In a longitudinal six-year study from 2003 to 2009 on 3300 clinical *C. jejuni* strains from Oxfordshire, it was found that the relative incidence of clonal complexes, ST 45 (1.087), ST 42 (1.227), ST 464 (1.612) and ST 52 (1.218) increased (Cody et al., 2012), while ST 658 (0.828), ST 574 (0.846) and ST 607 (0.818) decreased. This study also concluded that the clonal complexes are varying dependant on the season and resistance of strains to antibiotic. The most common clonal complex is ST 21 (22.8%) and the incidence of ST21CC was at least 2.5 times higher than other CCs. Some clonal complexes have a high association with specific hosts such as ST257CC, which is common in poultry in Europe. ST42CC and ST61CC are common in ruminants, and ST682CC and ST177CC are mainly from starlings in the UK and New Zealand (Wilson et al., 2008). The ST 137, 583, 681, 1324, and 2354 clonal complexes have also been identified in association with wild birds (French et al., 2009). The second most common *Campylobacter* species, *C. coli* ST828CC is commonly isolated from pigs with a low percentage from human, chicken and turkey sources (de Haan et al., 2010).

1.5 Whole genome sequencing and impact on comparative genomics and phylogeny

1.5.1 Bacterial genome sequencing techniques

Over the last two decades, whole bacterial genomes have been extensively sequenced. This sequencing has helped bioinformaticians and scientists to understand bacterial life, the role of bacteria in different environments, how these microorganisms evolve and interact with each other and their hosts. For *C. jejuni*, access to WGS has elucidated how this small genome encodes metabolic functions and variable products through to advanced comparative genomics and extensive phylogenetic analysis of a vast array of clinical and agricultural isolates. Over this period of 23 years, techniques have been evolved from Sanger sequencing based whole-genome shotgun sequencing to high throughout and single molecule long read sequencing (**Figure 1.5**), reviewed in Loman and Pallen (2015).

1.5.1.1   Whole-genome shotgun sequencing

Craig Venter and Hamilton Smith performed the first successful sequencing of a bacterial genome, *Haemophilus influenzae,* in 1995, using shotgun cloning of the genome and Sanger sequencing (**Figure 1.5**) (Fleischmann et al., 1995). The Sanger chain-termination or the dideoxy DNA sequencing technique was invented by Frederick Sanger in 1977 (Sanger et al., 1977). In this technique, during DNA synthesis insertion of dideoxynucleotides (ddNTPs), chemical analogues of deoxyribonucleotides (dNTPs) missing the 3′ hydroxyl group, stop synthesis (Chidgeavadze et al., 1984). This leads to DNA fragments of different lengths that are separated on polyacrylamide gels. Current methods establish the DNA sequence using different fluorescent probes for each ddNTP and provide accurate reads up to 700 to 1000bp (Heather and Chain, 2016). However, this takes 0.5- 3h and has a high cost. Sanger sequencing was used for earlier genome sequencing of short subcloned fragments. This sequencing method still provides the highest accuracy (0.001 single pass error rate %), but because of the intensive labour and time required for shotgun cloning, picking clones and

plasmid preparation, Sanger sequencing has been surpassed by high throughput (next-generation) sequencing for WGS (**Table 1.2**).
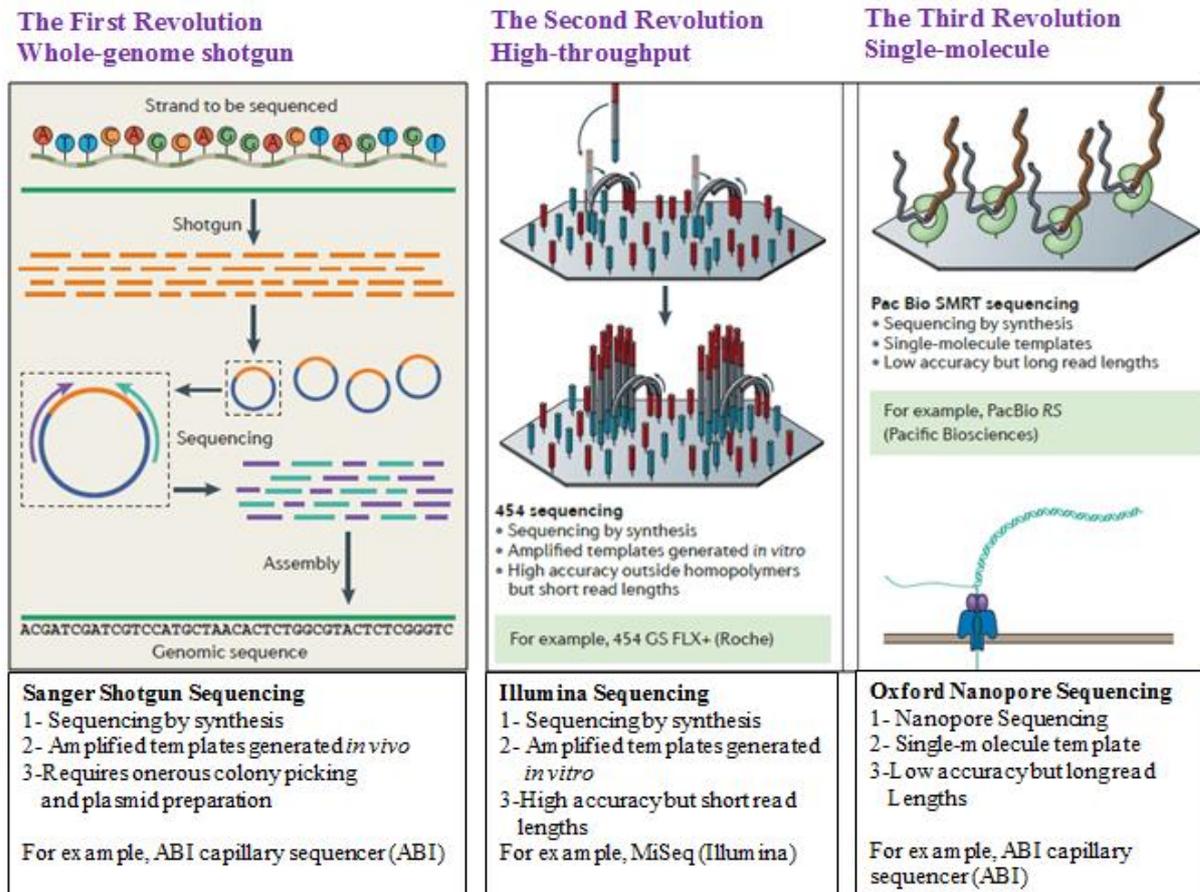
Figure 1.5: Whole genome sequencing, 1995-2018. There have been three revolutions in bacterial sequencing technologies from 1995 to 2018, from whole-genome shotgun sequencing, high-throughput sequencing to single-molecule long-read sequencing. Adapted from Loman and Pallen (2015).

1.5.1.2   High-throughput sequencing

High-throughput sequencing was the next revolution in sequencing technologies. This became available in 2007 and permitted cheap and rapid sequencing of many bacterial strains (**Figure 1.5** at above) (Reuter *et al*., 2015). This enabled the extensive analysis of strain related genomic differences and phylogeny (Holt et al., 2008). Illumina as one of the high-throughput sequencing technologies can generate many millions of highly accurate reads in a short time. The first steps in this sequencing technique are to fragment the DNA of the genome into short sections of around 200 to 600bp and then attach adaptors to their ends. The adaptor at one-end attaches to oligonucleotides on a solid surface and the other end is used to anneal to a primer and start DNA amplification. Bridge amplification occurs as the newly synthesized end bends over like a hairpin and hybridizes to the surface bound oligonucleotide. This process leads to clustering of double stranded fragments distributed over the surface of the slide. Similar to Sanger sequencing, modified fluorescent nucleotides are used to stop the reaction, but these only fluoresce on insertion. Fluorescence in each cluster is read, the fluorescent label and reversible terminator are then chemically removed, and the second round proceeds. This technique also has high accuracy (**Table 1.2**), but differs from the Sanger method by sequencing thousands of identical DNA fragments at the same time. This multiple-fold reading of the same DNA compensates for absence of proof-reading of the DNA polymerase. Illumina WGS is so much faster and cheaper than shot-gun cloning and Sanger sequencing as it does not involve cloning of individual fragments into vectors. The number of nucleotides determined in a single run is between 75-150bp. Generation of multiple overlapping sequences permits assembly of short reads into large contigs (Schurch & van Schaik, 2017; Willey et al., 2017) . However, Illumina read lengths are insufficiently large to cover repeat elements such as ribosomal RNA operons or genomic polymorphism. This is in contrast to the more advanced third-generation systems including Pacific Biosciences (PacBio) which enable long reads from a single molecule (English et al., 2012).

Table 1.2: Sequencing platforms comparison. Adapted from (Rhoads and Au, 2015).

Technique	Generation	Read length (bp)	Single pass error rate (%)	No. of reads per run	Time per run
Sanger ABI 3730 x 1	First	600-1000	0.001	96	0.5-3h
Illumina HiSeq 2500 (High Output)	Second	2 x 125	0.1	8×10^9 (paired)	7-60h
Illumina HiSeq 2500 (Rapid Run)	Second	2 x 250	0.1	1.2×10^9 (paired)	1-6days
PacBio RS II: P6-C4	Third	$1.0–1.5 \times 10^4$	13	$3.5–7.5 \times 10^4$	0.5- 4h
Oxford Nanopore MinION	Third	$2–5 \times 10^3$	38	$1.1–4.7 \times 10^4$	50h

1.5.1.3   Single-molecule, long-read sequencing

The first single-molecule, long-read sequencing technique, was released by PacBio on 2011 (Chin et al., 2013, English et al., 2012), and then Oxford Nanopore's MinION was became available in 2014 (**Figure 1.5** at above) (Quick et al., 2014). PacBio runs in a short time of 30min to 3h with relatively low throughput, but produces reads of 10 to 60 kb. The ability to produce long reads from a single molecule has permitted the system to cover repetitive regions in bacterial genomes (Schurch and van Schaik, 2017). In addition, this technique provides information that is helpful in the identification of DNA base pairs modification such as DNA methylation (Rhoads and Au, 2015). A limitation of this system is the much higher error rate of approximately 13% (**Table 1.2** at above) as well as the high cost compared to Illumina sequencing. Another alternative approach of producing long reads is nanopore sequencing as with Oxford Nanopore's MinION. In this, the DNA strand passes through a protein nanopore that is inserted in an artificial membrane. A voltage is applied and the current changes as different bases of one strand of the DNA pass through the pore. This will be processed to determine base pairs in the DNA strand (Schurch and van Schaik, 2017). Similar to PacBio, this can cover large-scale repeats in bacterial genomics (Karlsson et al., 2015), but results in less accurate reads than the strand synthesis short-read technologies (Loman and Watson, 2015).

1.5.2   Impact of bacterial sequencing on comparative genomics

With the availability of efficient Sanger sequencing and whole-genome shotgun sequencing examples of many reference bacterial genomes were sequenced. The published sequence of *C. jejuni* NCTC11168 genome explained the limited metabolic capability of this bacteria, identified a very low number of regulators and also highlighted the common occurrence of hypervariable homopolymeric repeats, for example, in genes involved in surface glycosylation (Parkhill et al., 2000). With availability of a complete annotated genome, DNA microarrays were used for comparative genomics to investigate diversity among strains. In one study, comparing 11 *C. jejuni* strains to NCTC11168 microarray, thirty genes that were divergent or absent in some strains, were identified. These genes mainly related to biosynthesis of surface bacterial structures such as capsule (Dorrell et al., 2001). In another study comparing a group of 18 strains of agricultural and clinical source to NCTC11168, 7 large hypervariable regions were identified. These contained large stretches of contiguous genes indicative of extensive horizontal gene transfer (Pearson et al., 2003).

The availability of rapid, cheap, whole genome sequencing changed the approach to comparative genomics. By July 2016, there were 72,981 enteries of WGS data for 120 different bacterial species in the National Center for Biotechnology Information (NCBI) database (Schurch and van Schaik, 2017). In addition, there are many independent databases, such as PUBMLST/ *Campylobacter,* in which there are now greater than 6,000 WGS publically available. WGS has been extensively used in the epidemiological analysis of infectious disease.

Yahara et al. (2017) analysed genetic variation among a high number of *C. jejuni/coli* sequenced isolates from poultry and clinical sources. They used 7343 MLST characterized isolates in the PubMLST/*Campylobacter* database to select 600 *C. jejuni* and *C. coli* isolates from different steps of poultry processing and clinical cases. They identified genetic differences that were overrepresented in clinical isolates of *C. jejuni* ST21CC and ST45CC strains, compared to related agricultural isolates. Clinical associated SNPs were associated with a small number of genes, for example *nuoK*, an NADH oxidoreductase, that is involved in aerobic survival and *cj1377c*, shown to have a role in formate metabolism. Interestingly, different sets of SNPs were identified in strains belonging to ST21CC and ST45CC,

suggesting different requiremens for adaptation from chicken, through slaughter and processing to humans. ST21CC genes associated with human disease included *kpsC* and *kpsD* genes that belong to capsular polysaccharides biosynthetic pathway and participate to adhesion and biofilm formation (Karlyshev et al., 2000). Human disease in ST45CC strains was related to SNP signature in *cj1373* and *cj1375* genes, which are predicted to encode efflux proteins involved in antimicrobial resistance and detoxification. For *Campylobacter* attributing source of infectious disease via bacterial genomic analysis is complicated by the high level of HGT and genome variation.

Availability of high number of bacterial genomes has been used to identify core and accessory genomes among *C. jejuni* strains. Core genes are defined as the set of genes present in all strains belonging to a taxonomic group of bacteria, while accessory genes are present in some, but not all strains and pan-genomes define the full complement of genes in the group studied (Tettelin et al., 2005).

Three studies have defined core genomes of *C. jejuni* or *C. jejuni/ coli.* The definitions of core genes are determined by the group of strains studied and parameters used. Friis et al. (2010) compared the M1 genome to 12 additional *C. jejuni* strains including NCTC11168, 81116 and 81-176. They classify two-thirds of the genome (1,295 genes) as core genes and one-third as accessory genes. Meric et al. (2014) compared the genomes of 7 published strains of *C. jejuni* subsp. *jejuni* (NCTC11168, 81116, 81-176, M1), *C. jejuni* subsp. *doylei* 269.97 and *C. coli* (76339, CVM N29710) to establish genomic diversities across these different species and subspecies. Across this group, they identified 1,035 core genes and a high number of 2,792 accessory genes across all strains following a BLAST match of ≥70% identity over ≥50% of the locus length. Most recently, Cody et al. (2017) have published a definition of core genome MLST (cgMLST) genes for clinical *C. jejuni* and *C. coli* isolates using the NCTC11168 coding sequences (1,643) as *C. jejuni* reference strain. They defined a gene as core if it was present in 95 % of strains and shared ≥70% identity over ≥50% of the locus length with the reference strain. This classified a higher number of core genes, 1,343 genes, as present in 95% of the 2,207 draft genomes from clinical isolates in Oxfordshire. The benefit of this scheme is can be used to analyse phylogenetic relationship of strains.

1.5.3   Phylogenetic analysis

To compare genomes and investigate the evolutionary relationship between different species of organisms, in particular in taxonomy, phylogenetic trees are typically used. With the large increase in avaiablilty of WGS data phylogenies are used in almost all sectors of biology. In addition to showing species relationships on the tree of life, phylogenetic analysis are used to explain the relation among paralogues of a gene family (Mäser et al., 2001), to investigate pathogen epidemiology, and to study evolution of organisms (Grenfell et al., 2004). Phylogenetic trees contain nodes that are linked by branches. Each node shows the birth of a novel lineage, while each branch shows the persistence of a lineage along time (Yang and Rannala, 2012). Phylogenetic trees represent genetic relationships among different species, whereas the nodes show evolution events.

In much more complex data, a phylogenetic tree may not fully represent complex information, for example where there has been extensive horizontal gene transfer. In this, a phylogenetic network can provide a more respresentative picture (Huson and Bryant, 2006, van Iersel et al., 2010). A phylogenetic network is used to visualize evolutionary relationships among DNA sequences, genes, genomes, or species (Huson and Scornavacca, 2011). The network is employed to reticulate the following critical events; hybridisation, recombination, horizontal gene transfer (HGT), or gene duplication, which might have occurred between isolates. This differs from phylogenetic trees by presenting many linked networks. In a phylogenetic network, taxa are shown by nodes and taxa evolutionary relationships are shown by edges, whereas in a phylogenetic tree, edges are represented as branches. Internal nodes of phylogenetic networks show ancestral species, and any of those nodes that have more than two parents relate to some changes in genetic information such as recombination (Huson and Bryant, 2006).

1.6 *Campylobacter jejuni* colonisation and pathogenesis

The most common method of *Campylobacter* transmission appears to be from a contaminated environmental source into poultry and then finally to humans. After ingestion of contaminated food or fluids, *C. jejuni* passes through the hostile acidic conditions of the stomach to colonise the lower parts of the human gastrointestinal tract. Many host factors control the ability of invading bacteria to colonise and cause disease. Such factors include bile salts, digestive enzymes, stomach acidity, mucin, peristalsis, and the immune response. Hence, the bacterium is faced with many stresses such as competition with other microbes, antimicrobial bile salts, iron limitation, oxidative stress, and the host immune response. The bacterium must overcome these unfavourable conditions to be able to survive and colonise the host. *Campylobacter* virulence determinants capable of overcoming these features include production of a polysaccharide capsule, motility, chemotaxis, stress response factors, invasion, post-translation glycosylation, microcolony formation and toxins (Ketley, 1997).

Biofilms have an important role in the survival of bacteria in unfavourable environmental conditions as well as in antibiotic resistance (Phillips and Schultz, 2012, Chmielewski and Frank, 2003). Biofilm is the attachment of bacterial cells either to themselves or to another surface, often combined with the formation of the extracellular matrix to form a raised structure. This structure helps both mono and mixed *C. jejuni* cultures to survive in suboptimal environments, such as under aerobic conditions, and increases resistance of the bacteria (Reuter et al., 2010, Ica et al., 2012). For instance, *C. jejuni* can often make biofilm structures at the surface of the watering system found in chicken house drinking equipment (Hanning et al., 2008). It has been identified that *C. jejuni* can make biofilms with other species including *Escherichia coli*, *Enterococcus. faecalis, Pseudomonas aeruginosa* and *Staphylococcus simulans* (Teh et al., 2010) and heterogeneous biofilms containing both spiral and coccoid forms have been observed with the *C. jejuni* strain 81-176 on different surfaces (Gunther and Chen, 2009). *C. jejuni* cells can make different forms of biofilm in a liquid environment (Joshua et al., 2006). The majority of *Campylobacter* biofilms described are bound microcolonies on a solid surface, monitored using a crystal violet binding assay with microcolonies forming on polystyrene microtitre plates (Reeser et al., 2007) or bound to a glass fibre filter in a liquid culture and visualised by scanning electron microscopy (Kalmokoff et al., 2006). However, pellicles also known as floating biofilms are formed by

some strains of *C. jejuni* and *C. coli* when cells aggregate to each other in a solution (Vegge et al., 2016, Joshua et al., 2006).

Furthermore, in an *ex vivo* study, it has been demonstrated that *C. jejuni* can attach to the surface of human ileal tissues, but not colon tissue, initially forming microcolonies that subsequently develop into a biofilm. This study also demonstrate the inability of non-motile cells (*flaA* or *flaB* mutants) to effectively bind to the surface and form the biofilm (Haddock et al., 2010). Proteomic analysis of 11168H showed that the flagellar motility complex has an important role in the attachment of the bacterium to solid surfaces in the process of biofilm formation (Kalmokoff et al., 2006). This indicates that the flagellum of *C. jejuni* participates in the process of biofilm formation and therefore motility and biofilm formation both appear to have important roles for flagella in *C. jejuni* induced diarrhoea in humans (Reeser et al., 2007). In accord with this, *Campylobacter* biofilm formation is influenced by motility genes and genes that contribute to chemotaxis. A list of genes require for flagella assembly, chemotaxis and motility as well as surface proteins and adhesins that might contribute to biofilm formation is shown in **Table 1.3**.

Table 1.3: Motility, potential adhesins and surface encoding genes that affect biofilm formation in *Campylobacter* (Ben et al., 2014, Haddock et al., 2010).

Flagellar assembly	**Flagellar regulator and chemotaxis**	**Flagellar glycosylation**	**Adhesins**	**Surface proteins**
flaA, flaB, fliK, flgD, flgE, fliY, fliM, flhF, fliI, fliF, fliG, fliH, flhB, motB, motA, fliN, fliO, fliE, flgC, flgB, flaG, fliD, fliS, flgH, flgG2, flgG, flgA, fliP, flhA, flgL, fliW, fliR, flgI, flgJ, flgK, pflA, fliQ	*cheA, cheV, cheY fliA, maf4, rpoN, rpoD, pEtN, cj0062c, cj0248, cj0883c, cj1026c*	*pseB, pseC, pseF, pseH, pseI, pseE, pseA, pseD, maf1, pseG, maf1*	*peb1A, flaC, peb4, ciaC, cadF*	*gne,waaC, htrB, lgtF, cg1136, cj1137, cj1138, wlaN, cstIII, neuB1, neuC1, neuA1 (cgtA), waaF, waaV, gmhA, hldE, hldD, gmhB, lgtS, cstII, galT, spoT*

Colonisation is an essential first stage in *C. jejuni* induced gastroenteritis. Spiral shape, motility and chemotaxis all help the bacterium pass through the mucus and reach the epithelial layer. The ability of a non-motile *C. jejuni* strain to invade the intestines of animals is dramatically reduced (Yao et al., 1994). Nachamkin et al. (1993) studied three-day-old chickens, inoculated with doses of 3.0 x10^4 to 6.6 x 10^8 motile wild type or flagellar mutants of *C. jejuni*. Only fully motile wild type strains colonised the ceca of the chickens. This supports the role of flagella in invasion and colonisation in different hosts. The major flagellin (FlaA) and the minor flagellin (FlaB) are the main structural components of the flagellum (Guerry, 2007). They are commonly glycosylated and glycosylation appears to be important in virulence (Guerry et al., 2006). *Campylobacter* motility is also dependent on chemotaxis; sensing and responding to the environment. Bacteria use chemotaxis to move toward any favourable nutrient and environment. Hence, this process is important in host colonisation by *C. jejuni.* Many of the metabolic substances act as chemoattractants including Ser, Cys, Glu, pyruvate, and Gln (Guccione et al., 2008, Velayudhan et al., 2004), electron donors such as formate, and electron acceptors such as nitrite and nitrate (Weingarten et al., 2008), and CheW, CheV, CheA and CheY are key chemotactic proteins (Zautner et al., 2012, Yao et al., 1997).

There is evidence that the flagellar system has a secondary role, in addition to motility, and acts as a type 3 secretion system (T3SS) for secretion of proteins during host cell invasion (Guerry, 2007). Konkel et al. (2004) provided initial evidence that the flagellar export system of *C. jejuni* mediates secretion of Cia proteins, which are involved in invasion of INT407 human intestinal cells. This was confirmed by mutations in *flgB* and *flgC* genes, *flE2* hook gene (genes required for the flagellar basal body) and the *flaA* and *flaB* filament genes.

Once *C. jejuni* has penetrated the mucous of the intestine, it may multiply, interact with epithelial cells of its host that ultimately leads to invasion of the epithelial barrier (**Figure** 1.6) (van Putten et al., 2009). A relationship between the ability of the bacteria to adhere to tissue culture cells *in vitro* and the severity of symptoms patients suffered from gastroenteritis has been reported (Fauchere et al., 1986). A number of surface exposed proteins of *C. jejuni* have been implicated as adhesins, including a putative fibronectin-binding protein, Cj1349c, putative fibronectin domain-containing lipoprotein Cj1279c, the outer membrane protein CadF, autotransporter CapA, and outer membrane PEB1 protein involved in Asp/Glu uptake.

In addition, PorA has been shown to interact with tissue culture cells (Flanagan et al., 2009, Ziprin et al., 1999, Monteville et al., 2003). Epithelial cell invasion is considered a key step in disease and gastroenteritis. The ability to invade the epithelial cells is strain dependent (Newell et al., 1985). It addition to adhesins, Cia invasive antigens and a number of other factors including polysaccharide capsule (CPS), lipooligosaccharide sialylation and FlaC secreted protein have been linked to invasion of host cells by *C. jejuni* (Dasti et al., 2010).

Bacterial toxins often have a major contribution to the disease profile in gastroenteritis resulting in watery or bloody diarrhoea. Cytolethal distending toxin (CDT) is produced by a number of different Gram-negative pathogens, including *Escherichia coli* (Toth et al., 2003), *Shigella* and *Vibrio* (Okuda et al., 1995). CDT is a key *Campylobacter* toxin and is produced by several species of *Campylobacter* including *C. jejuni, C. coli, C. fetus, C. lari,* and *C. upsaliensis* (Dasti et al., 2010). The CDT toxin is encoded by the following adjacent genes *cdtA*, *cdtB*, and *cdtC*, which encode proteins of approximately 30, 29 and 21 kDa, respectively. CdtB has DNA nuclease activity, but all three CDT subunit proteins are required for toxin delivery and activity (Lee et al., 2003, Taieb et al., 2015). Intact CDT causes destruction of host cell DNA and negatively influences the process of cell division and proliferation by blocking the G_2 cell cycle. It causes damage to the epithelial cells of the intestines, in particular intestinal crypt cells (**Figure** 1.6) (Walker et al., 1986, Whitehouse et al., 1998, Lee et al., 2003). However, a recent report (Mortensen et al., 2011) provided evidence that the CDT toxin may not be essential for development of campylobacteriosis in humans. In this case study, it was shown that both CDT-negative and CDT-positive *C. jejuni* strains caused campylobacteriosis without clinical differences between patients who were infected with either type of strain.

Surface carbohydrates, including capsule, lipooligosaccharide (LOS), and O- and N-linked glycans are important in *C. jejuni* colonisation and virulence. The outer membrane component LOS, is composed of oligosaccharide and lipid A, influences a number of aspects of virulence including epithelial invasion, adhesion and immune evasion. LOS sialylation may increase invasive ability and reduce the immune response (Louwen et al., 2008). Flagellar proteins of *C. jejuni* are modified by O-linked glycans and membrane protein N-linked glycans (Szymanski and Wren, 2005). The role of glycosylation in the ability of *C. jejuni* to cause disease has been investigated. For example, this modification might protect

bacterial surface proteins from cleavage by gut proteases (Alemka et al., 2013). Glycosylation of PorA has also been shown to enhance bacterial microcolony formation, adhesion to Caco-2 cells, and chicken colonisation (Mahdavi et al., 2014). In some strains of *C. jejuni* the carbohydrate of LOS mimics the neuronal gangliosides of peripheral nerve cells in humans. It is believed that this molecular similarity causes trouble by inducing an autoimmune reaction against cranial and peripheral nerve cells. Based on this a number of neurological diseases have been attributed to *C. jejuni* infection, in particular Guillain–Barré syndrome (GBS) (Ang, 2002). GBS causes paralysis in approximately 1 to 2 per 100,000 campylobacteriosis cases (Yu et al., 2006).

The bacterial capsule is typically composed of polysaccharide polymers extending from the cell surface and thus coating the bacterium with a hydrophilic matrix. Hence, capsule production is often critical to survival of the bacterium in the environment and in hosts by conferring immune evasion. A number of *Campylobacter* species (*C. coli*, *C. jejuni,* and *C. lari*) produce capsular polysaccharides (Guerry et al., 2012). Several different structures of *C. jejuni* CPS, varying in structural complexity (sugar composition and linkage), have been characterised (Aspinall et al., 1992, Aspinall et al., 1995, Muldoon et al., 2002), and forty-seven different serotypes of *C. jejuni* are distinguished using the Penner typing scheme, which is primarily based on the capsular polysaccharide. The KspM protein transports capsular polysaccharides. Mutation in *kspM* decreased the ability of *C. jejuni* to invade INT407 cells, decreased protection from human serum, and reduced colonisation of poultry (Bacon et al., 2001).
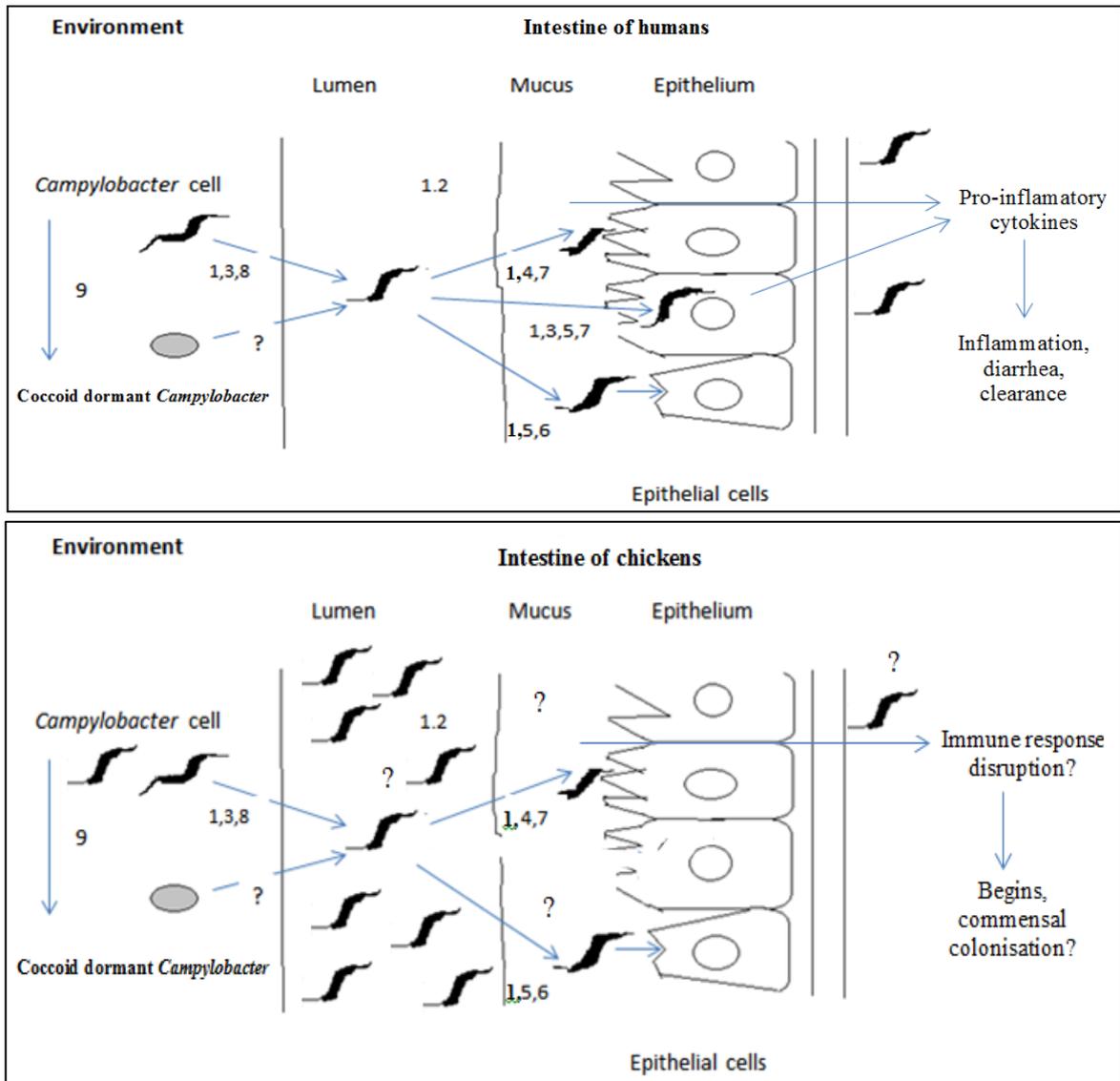
Figure 1.6: Overview of the stages of *Campylobacter* colonisation in human and chicken intestines. *C. jejuni* cells traverse the mucus layer and attach to the epithelial cells. In chickens, they primarily remain there and colonise. During disease in humans, they invade the epithelial cells resulting in a pro-inflammatory response, an increase in cytokine production, inflammation and diarrhea. In chickens, it has been proposed that disruption of the immune response may be important in successful colonisation. Numbers refer to putative virulence factors; 1-Motility, 2-Chemotaxis, 3- Oxidative stress defense, 4- Adhesion, 5- Invasion, 6- Toxin production, 7- Iron acquisition, 8- Temperature stress response, 9- Coccoid dormant stage. Adapted from van Vliet and Ketley (2001) and Young et al. (2007).

1.7 *Campylobacter jejuni* physiology

1.7.1   Oxidative stress response

C. jejuni is an oxygen sensitive bacterium, but requires low levels of oxygen used in respiration and nucleic acid synthesis (Sellars et al., 2002). It can also protect itself from oxidative stresses encountered during host invasion. A consequence of use of oxygen in respiration is production of toxic reactive oxygen species (ROS), such as hydrogen peroxide and anion superoxide (Kim et al., 2015a). Accumulation of ROS in bacterial cells can damage macromolecules such as DNA (Imlay, 2003). ROS molecules are produced from the combination of iron with oxygen (Ratledge and Dover, 2000). *C. jejuni* uses different electron acceptors and donors, such as nitrate, fumarate, and nitrite, to harmonise the stress under both microaerobic and aerobic conditions (Sellars et al., 2002). Despite possessing several oxidative stress defense genes, such as *katA* (catalase), *ahpC* (alkyl hydroperoxide reductase), and *sodB* (superoxide dismutase), *C. jejuni* cannot grow in atmospheric oxygen tension (Kim et al., 2015a).

1.7.2   Iron uptake in *Campylobacter*

Iron homeostasis is crucial to many organisms, including bacteria, to avoid toxicity that may result from high levels of iron (Ratledge and Dover, 2000). Use of comparative genomics among *C. jejuni* strains has identified several iron uptake systems (**Figure** 1.7) (Miller et al., 2009). *C. jejuni* can use different forms of iron including from haemin and haemoglobin. Making a mutation in the *chuA* iron regulated gene resulted in the failure of *C. jejuni* to grow with haem (Pickett et al., 1992). This operon consists of *chuC* (ATPase), *chuD* (periplasmic binding protein), *chuB* (permease), and *chuA*. P19 is a periplasmic protein that is part of another high-affinity iron transporter. The P19 system is encoded by a cluster of eight *cj1658* to *cj1663* genes in the NCTC11168 genome (Miller et al., 2009). It is suggested that the Ftr1-P19 system can transport iron in the absence and presence of rhodotorulic acid. Mutations in strain NCTC11168 meant that a P19 mutant was not able to use iron provided as ferri-rhodotorulic acid. Siderophores are another important iron transport system in bacteria.

These act as ferric-iron chelators (Schröder et al., 2003). *Campylobacter* spp. are not able to synthesise siderophores that transport ferric iron, but they can utilise enterochelin and ferrichrome as a source of iron (Richardson and Park, 1995). The *ceuBCDE* operon encodes a binding-protein-dependent ABC transporter system for the uptake of ferric iron across the IM. This operon plus the *cfrA* gene have role in ferri-enterochelin uptake and mutation in either leads to decreased colonisation of *C. jejuni* in the chicken intestine (Palyada et al., 2004). Mutation of *feoB* in NCTC11168 resulted in a decreased ability to uptake Fe^{2+} from the host (**Figure** 1.7) (Naikare et al., 2006).
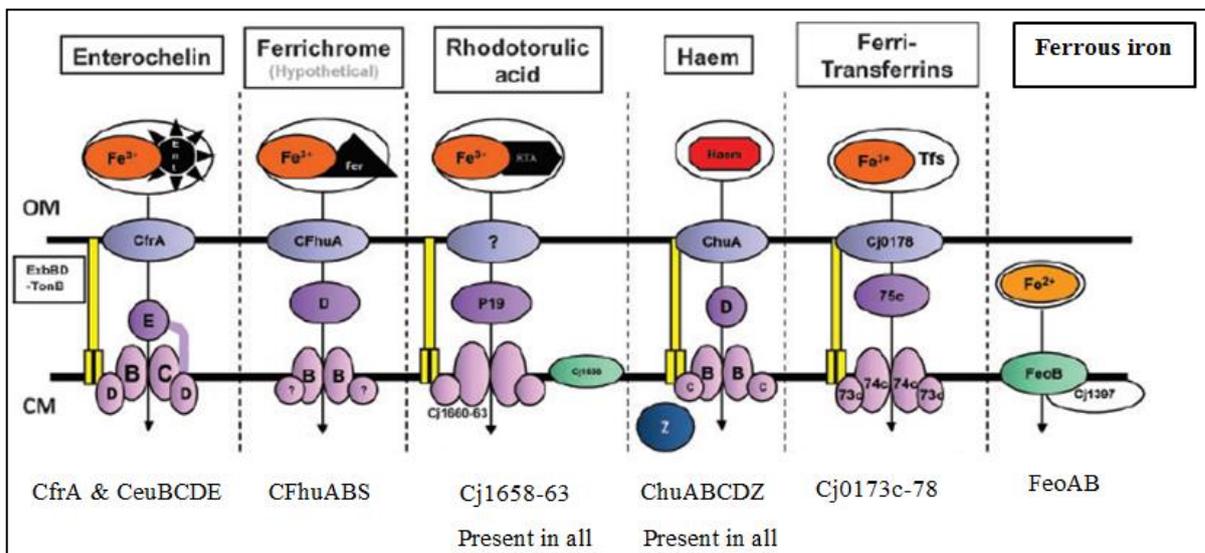
Figure 1.7: The iron-uptake systems of *C. jejuni* strains. Ferric iron transporters require outer membrane (OM) receptors. These receptors transfer the iron ferric complex to the periplasm using the proton-motive force of the energy transduction complexes, ExbBD-TonB. Within the periplasm, ferric complexes are released to the ABC transport system for transport across the cytoplasmic membrane to the cytosol. Adapted from Miller et al. (2009)

1.7.3   Amino acid and organic acid metabolism

The study of physiological features is important in clarifying the mechanisms involved in *C. jejuni* pathogenicity, and survival in different hosts, the environment and the food chain. *Campylobacter* isolates have generally been defined as being unable to ferment or oxidise common carbohydrates such as glucose (Vorwerk et al., 2015). Instead, they primarily depend on Ser, Asp and Glu and to a lesser extent Pro amino acids and TCA intermediates as carbon and energy sources (Kelly, 2001, Stahl et al., 2012). *C. jejuni* can acquire its preferred amino acids during growth in the gut. There are only a few amino acids are catabolised by *C. jejuni*. NCTC11168, 81-176, and 81116 strains have all been shown to utilise Ser, Asp, Glu, and Pro, in this order of preference (Leach et al., 1997, Guccione et al., 2008). *C. jejuni* 81-176, but NCTC11168, can also use either Asn or Gln as carbon and energy. The ability to use these amino acids is dependant on the presence of periplasmic asparaginase and γ-glutamyl transpeptidase (GGT), respectively (Hofreuter et al., 2008). Despite the small size of the *C. jejuni* genome, it encodes enzymes essential for the operation of the oxidative citric acid cycle, such as succinyl-CoA synthetase, a nicotinamide adenine dinucleotide-linked malate dehydrogenase, and fumarate reductase. This highlights the critical function of the TCA cycle as an energy source. **Figure 1.8** shows the transport systems for amino acids catabolised by C. *jejuni* and where they feed into the TCA cycle. Amino acid transporters, such as Peb1A and SdaC, have been identified as having a critical role in *C. jejuni* colonisation of the host intestinal tract (Muraoka and Zhang, 2011, Velayudhan et al., 2004, Hofreuter et al., 2008).
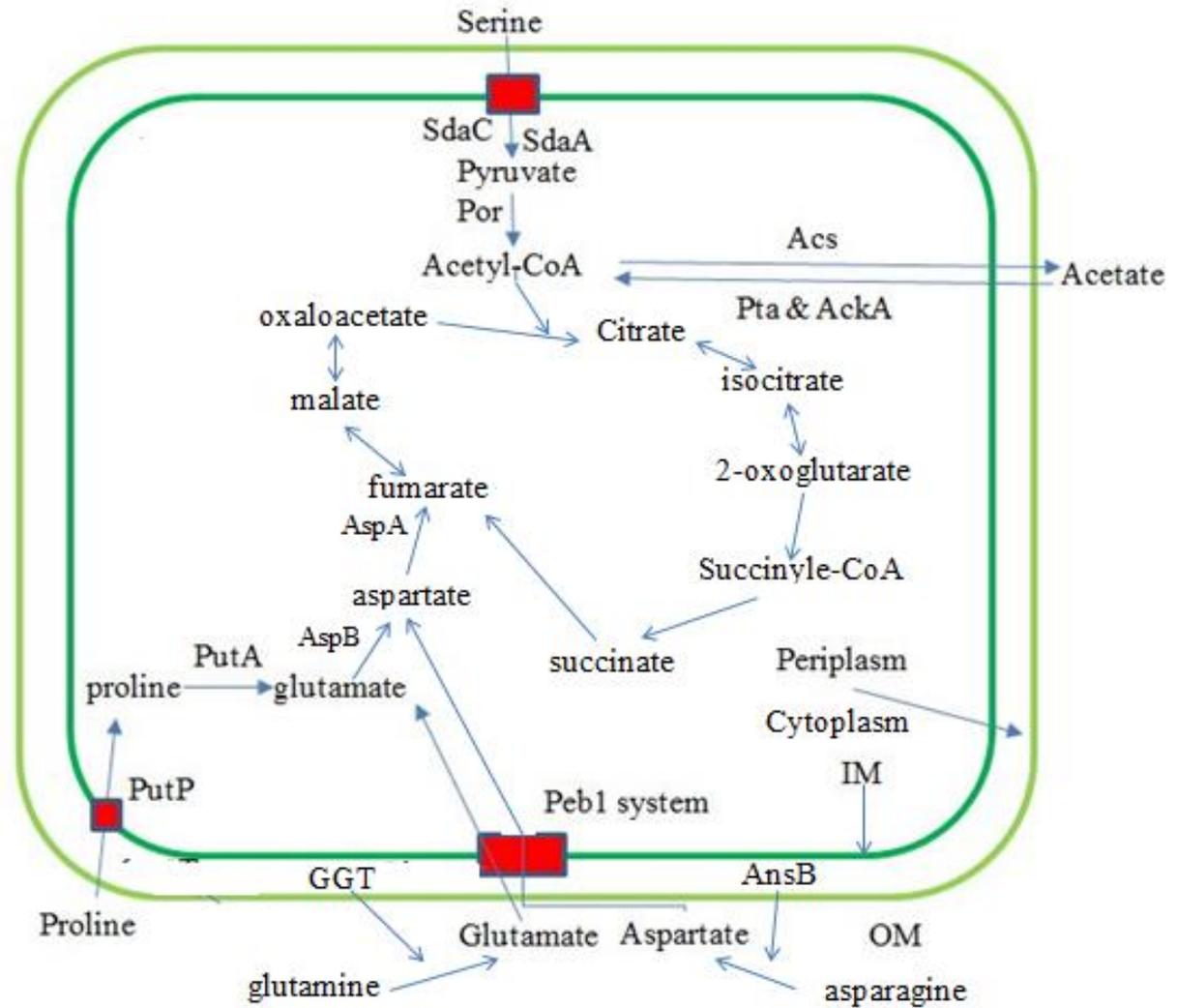
Figure 1.8: Major pathways of amino acid utilisation in *Campylobacter jejuni.* Key pathways and gene products located in the inner membrane (IM), outer membrane (OM), periplasm (between the membranes) and cytoplasm (primary place of metabolism) are shown. Red rectangles and square represent specific amino acid transport systems across the IM, transport across the OM is assumed to be by diffusion (PorA). Key enzymes/transport proteins are presented next to reactions. Amino acid utilisation: Ser - Ser transporter (SdaC), conversion of Ser to pyruvate (SdaA), pyruvate oxidoreductase (Por); Glu and Asp - Peb1 system, an ABC transporter including Peb1A the periplasmic glutamate/aspartate binding protein, aspartase (AspA) that catalyses deamination of Asp to fumarate, conversion of glutamine to glutamate by γ-glutamyl transpeptidase (GGT); Pro - Pro transporter (PutP), proline dehydrogenase (PutA) converts Pro to Glu. Acetyl-CoA synthetase, phosphotransacetylase (Pta), acetate kinase (AckA). Adapted from Guccione et al., 2008, Hofreuter, 2014, Stahl et al., 2012 and Guccione et al., 2017.

SdaC and SdaA proteins have important roles in the metabolism of Ser. SdaC transports this from the periplasm into the cell and SdaA acts as an L-serine dehydratase. Although Ser is not required for growth of *C. jejuni in vitro*, Ser utilisation is essential to the colonisation of chicks. Mutation of *sdaC* or *sdaA* led to complete failure to colonise chicks by 11168H, while the parent strain, a hypermotile variant, colonised 90-100% of three-week-old chicks within one week (Velayudhan et al., 2004).

The mechanism of uptake for both Asp and Glu is via the Peb1 system (**Figure 1.8** at above). This system contains the following gene products; Peb1A (Cj0921c), Cj0920c, Cj0919c, and PebC (Cj0922c) (Leon-Kempis et al., 2006). The Peb1A protein is an important transport binding protein for both Glu and Asp. Both substrates have a role *in vivo* as carbon sources for *C. jejuni*. Mutation of *cj0921c* in *C. jejuni* led to complete loss of transport of Glu and greatly impaired uptake of Asp. When glutamate has been transported into the cell, the enzyme AspB (Cj0762c) converts the amino acid to Asp. Asp is then converted into fumarate by AspA (Cj0087), which is fed into the citric acid cycle (Guccione et al., 2008).

It has been found that *C. jejuni* has reduced preference for Pro compared to Ser, Asp, and Glu. In addition, the uptake of Pro significantly increases during the stationary phase (Wright et al., 2009). The PutP protein (Cj1502c) transports Pro into the cell. Pro is then catabolised by PutA and converted into Glu. Glu can enter the TCA cycle after being metabolised to Asp. Asn is another less favoured amino acid. Periplasmic asparaginase (AnsB) can convert Asn to Asp and is present in many strains, if *C. jejuni* also has an Asn transporter (Hofreuter et al., 2008). Mutation of the *ansB* gene had no effect on the ability to colonise the intestines, but was shown to lead to a substantial defect in liver colonisation in infected mice. This may indicate that asparagine metabolism is important in certain tissues of the host.

C. jejuni obtains much of its energy through the uptake and utilisation of citric acid cycle compounds. Oxaloacetate, fumarate, 2-oxoglutarate and pyruvate, which can enter the TCA cycle directly, are also products of the metabolism of the above amino acids. *C. jejuni* is able to transport and utilise many of the TCA intermediates. For instance, permease KgtP (Cj1619) transports 2-oxoglutarate, and DcuA (Cj0088) and DcuB (Cj0671) transport malate, fumarate, and succinate (Guccione et al., 2008). *C. jejuni* can oxidise succinate to fumarate by fumarate reductase Frd ABC. During respiratory metabolism by *C jejuni*, the low levels of

O_2 have roles as terminal electron acceptors. Alternative electron acceptors including nitrate, nitrite, and fumarate can also be used to facilitate growth of the bacterium under oxygen-limited conditions (Pittman and Kelly, 2005). However, *C. jejuni* cannot grow under strictly anaerobic conditions (Pittman and Kelly, 2005, Kelly, 2001) most likely due to the basic requirement of oxygen during DNA synthesis (Sellars et al., 2002, Pittman and Kelly, 2005).

1.7.4   Carbohydrate utilisation by certain *Campylobacter* strains

As described above, *C. jejuni* uses key amino acids as primary energy and carbon sources as well as some organic acids and TCA intermediary compounds. It was previously believed that this microorganism could neither ferment nor oxidise different types of sugars, especially glucose. This can be explained by the fact that *C. jejuni* has an incomplete Embden-Meyerhof-Parnas (EMP) glycolysis pathway. It lacks a key enzyme, 6-phosphofructokinase, (PfkA) and therefore cannot convert fructose-6-phosphate to fructose 1, 6 phosphate. Despite this it can undergo the reverse reaction using the enzyme fructose-1, 6-bisphosphatase (*fbp*) and hence can synthesise glucose as required for polymer and polysaccharide synthesis via gluconeogenesis (Velayudhan and Kelly, 2002, Kelly, 2001, Hofreuter et al., 2008, Guccione et al., 2008, Parkhill et al., 2000, Pendleton et al., 2013). **Figure 1.9** shows EMP enzymes encoded in *C. jejuni* genome and the reverse reaction of gluconeogenesis. A gene encoding glucokinase (*glk*) and glucose specific transport system also appear to be absent from the *C. jejuni* NCTC11168 genome.

The first evidence of the ability of *C. jejuni* subsp. *jejuni* to metabolise carbohydrate was published by Muraoka and Zhang (2011) and Stahl et al. (2011), when they demonstrated that the sugar fucose could be metabolised by some *C. jejuni* strains. A recent study Dwivedi et al. (2016) reported that over 50% of *C. jejuni* strains harbour the 9kb operon (*cj0480c - cj0490*), which contains all of the genes required for uptake and metabolism of fucose. The following sequence types ST (21, 48, 354, 206, 257) CCs of *C. jejuni* were found to commonly have this operon, while ST (45, 42, 283, 464, 353) CCs were generally missing the operon. Importance of fucose utilisation *in vivo* has been demonstrated in a competitive piglet colonization assay (Stahl et al., 2011). Mutation of the *cj0486* fucose permease abolished uptake of fucose and strains with this mutation were no longer able to

competitively colonise piglets (Stahl et al., 2011). This correlates with the pig gut being rich in fucose (Lien et al., 1997). It also highlights the importance of metabolic differences in *Campylobacter* pathogenicity and importance of the bacterial metabolic ability being a potential indicator of host specificity.

The Entner–Doudoroff pathway (ED pathway) is an alternative pathway for glucose utilisation used by many bacteria and plants. This pathway was initially observed in the genome sequence of *C. jejuni* subsp. *doylei* 269.97 in 2008 (Miller, 2008). Seven genes (~9kb) encoding enzymes for catabolising glucose including a glucose permease were found to be located in one of the three copies of the rRNA operon between the isoleucyl-tRNA and 23S rRNA. More recently, a similar organisation of the ED pathway genes was also found in the genome of *C. coli* CHW470 and CHW467 within the *rrnA* operon (**Figure 1.10**). Glucose metabolism via the ED pathway is as follows: after uptake of the glucose through GlcP glucose permease, the glucose is phosphorylated by glucokinase (Glk) and metabolised by the following enzymes: Glucose-6-phosphate dehydrogenase (Zwf) converts glucose-6-phosphate to 6-phosphogluconolactonate to enable entry into the Entner-Doudoroff pathway. The Pgl enzyme then converts 6-phosphogluconolactonate to 6-phosphogluconate. The Edd enzyme converts 6-phosphogluconate to 2-keto-3-deoxy-6-phosphogluconate, which is then hydrolysed and converted into pyruvate plus glyceraldehyde 3-phosphate by the action of the Eda enzyme (**Figure 1.9**) (Conway, 1992, Vorwerk et al., 2015). Pgi1 glucose-6- phosphate isomerase (CAMP1435) is conserved in *C. jejuni* and *C. coli*. The Pgi enzyme, is encoded by the *glc* locus, was renamed to Pgi2 by Vorwerk et al. (2015). The Pgi2 protein shares only low identity, 25%, with the Pgi1 despite having the same function. They also showed that these plasticity regions were transferred from ED-positive *C. coli* strains into an ED-negative strain by natural transformation of chromosomal DNA. ED-positive *C. coli* strains have been shown to benefit from glucose utilisation in different ways. One strain was shown to benefit from glucose metabolism, while with a different strain glucose enhanced the level of capsule formation (Vegge et al., 2016). *Helicobacter pylori* also shows a preference for catabolism of amino acids and also metabolises glucose via an ED pathway encoded by the three key genes *edd, eda* and *zwf* genes (Mendz et al., 1994). An *in vivo* study on the ability of mutated *H. pylori* to colonise mice showed that glucose utilisation by this pathway was not essential for colonisation in this host (Wanken et al., 2003).
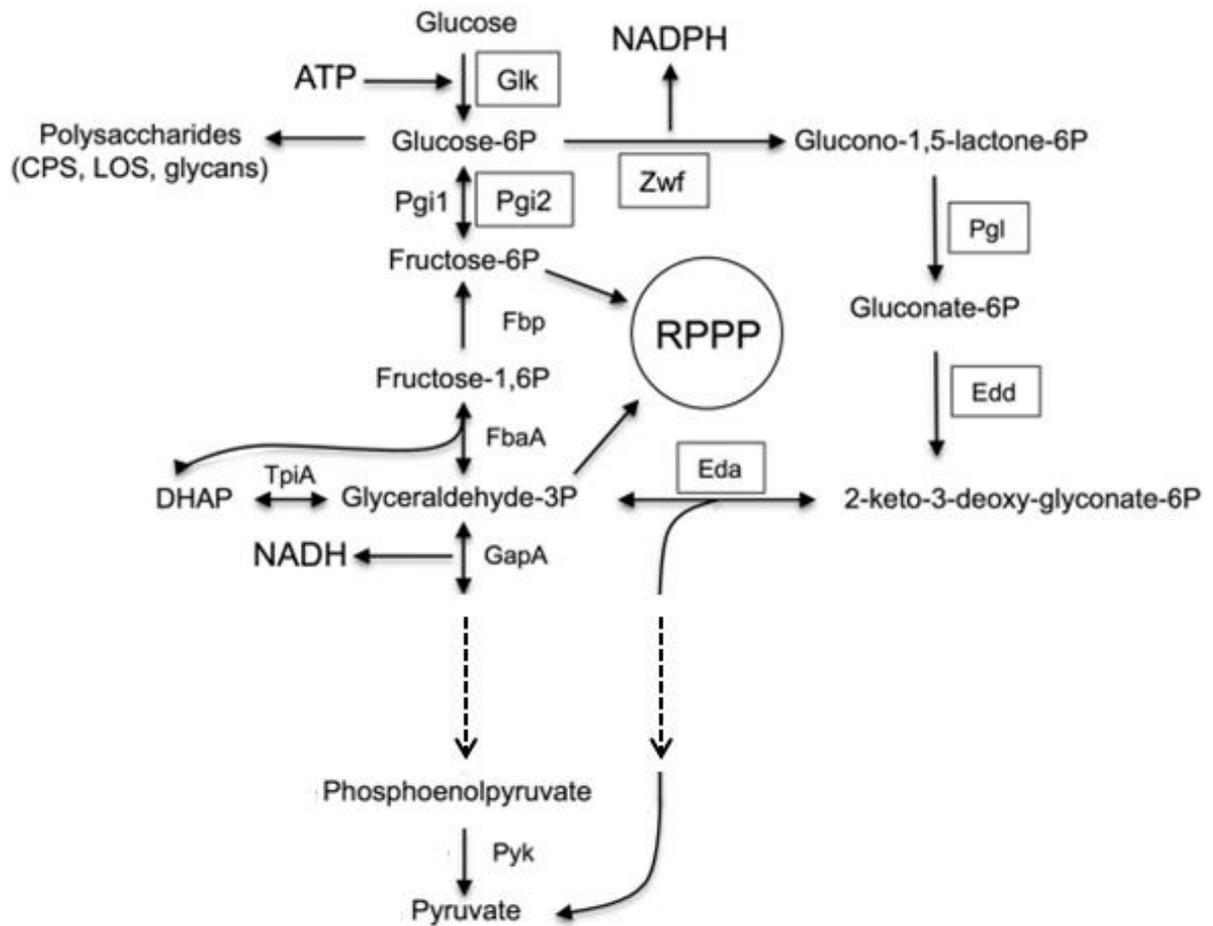
Figure 1.9: The Entner-Doudoroff pathway of glucose metabolism in *Campylobacter* isolates.

The main enzymes are encoded by *edd* (6-phosphogluconate dehydratase) and *eda* (2-keto-3-deoxy-6-phosphogluconate aldolase) genes and catalyse hexose gluconate-6P to 2-keto-3-deoxy-gluconate-6P. The other ED pathway genes are necessary for completion of the glucose metabolism, *Glk* (glucosekinase), *zwf* (glucose-6-phosphate dehydrogenase) and *pgl* (6-phosphogluconolactonase). Pgi1 and Pgi2 are both phosphoglucose isomerases and have the same function. Fbp (fructose-1,6-bisphosphatase). Enzymes, encoded on the *glc* locus, that participate in the ED pathway are labelled with boxes. RPPP is the reductive pentose phosphate pathway. Dashed arrows indicate more reactions between glycerolaldehyde-3P and phosphoenolpyruvate that are not shown. Image adapted from Vegge et al., 2016.
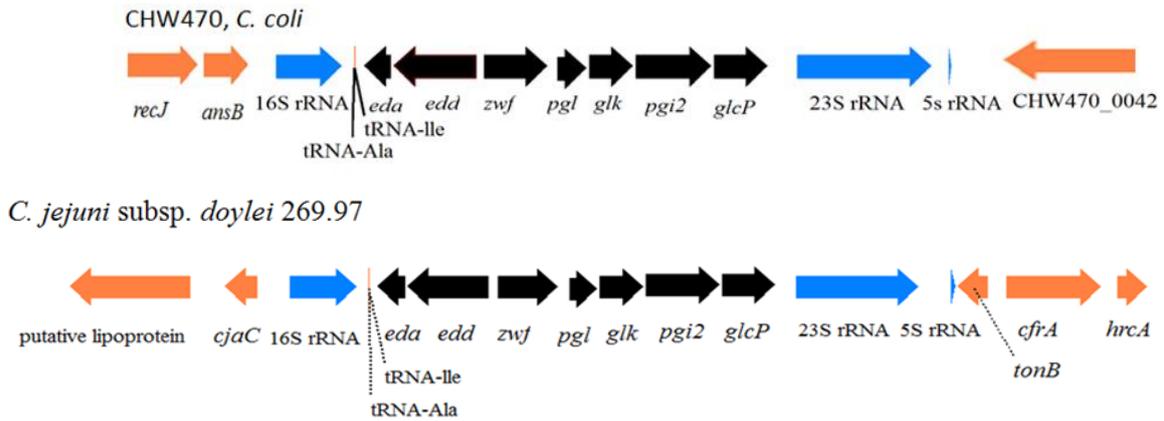
Figure 1.10: Diagram of the ED pathway encoding the *glc* locus integrated into ribosomal RNA operons. The figure shows schematic presentation of the *glc* locus in the *Campylobacter* species *C. coli* CHW470 and *C. jejuni* subsp. *doyle*i 269.97. Blue colour-coded arrows are the *rrn* operon; black arrows, the *glc* locus; orange arrows, genes flanking the *rrn* loci (Vorwerk et al., 2015).

1.8 Models of *Campylobacter jejuni* infection

The association of *C. jejuni* with animals can be considered as two broad categories: non-pathogenic colonisation and colonisation leading to disease.

1.8.1 Animal models

Most recently, infant rabbits and piglets have been successfully used as a model to study diarrhoeagenic strains of *Campylobacter* (Shang et al., 2016, de Vries et al., 2017). Newborn New Zealand White rabbits were inoculated with $1 \times 10^{2\text{-}10}$ CFU of NCTC11168 and ATCC 33560. Infectious symptoms in infant rabbits were similar to those observed in many campylobacteriosis cases, such as invasion in many cases of human campylobacteriosis resulting in diarrhoea. The bacterium consistently colonised the large intestine of the rabbits causing inflammation and cytokines induction. Furthermore, the piglet model has been successfully established as an animal model to investigate *C. jejuni* pathogenicity. The piglets were orally infected with 5×10^9 CFU of *C. jejuni* 11168, 81-176 or L115 (isolated from a paediatric diarrhoeal disease case). All strains successfully infected the gastrointestinal tract of the piglet at 3 dpi. At 5 dpi, viable cells reach approximately 10^8 CFU/g in ileal content,

10^{10} CFU/g in colonic content, and 10^{12} CFU/g in the faeces that was collected from the rectum. Critical infection symptoms, including increased temperature, watery diarrhoea reflecting a leaky epithelium and decreased bile re-absorption, were recorded. This established animal model can be used to study virulence factors of *C. jejuni*. Both of these recently established models can be used for *in vivo* studies of *C. jejuni* (Shang et al., 2016, de Vries et al., 2017). Ferrets have also been shown to be colonised by *C. jejuni* with similar symptoms to those reported in human, such as inflammation and diarrhoea (Fox et al., 1987, Nemelka et al., 2009), but use of ferrets has been limited due to high costs.

While mice have not traditionally been useful models for *C. jejuni* induced human enteritis, genetically manipulated mice have proven to be more suitable (Watson et al., 2007). There are an increasing number of congenic mouse strains available, some that have been shown to be useful as disease model for *Campylobacter* infection (Mansfield et al., 2007). Pathological lesions induced in C57BL/6 IL-10 deficient mice resembled those seen in humans. This included destruction of crypt epithelium and crypt abscesses. In this study, a high dose of 10^{6-10} CFU was necessary to infect, colonise and cause enteritis in 50-80% of mice inoculated with *C. jejuni* NCTC11168. The caecal and colonic crypts are the major site of colonisation by *C. jejuni*, while recovery of *C. jejuni* from sites other than the GI tract is more variable, suggestive of a systemic infection (Mansfield et al., 2007, Stahl et al., 2014). There are some challenges of using this model, such as inoculation of mice with the bacterium and investigation of infection by clinical examination. However, the mouse model might help to test known or presumed virulence factors of *C. jejuni*, which should in turn clarify mechanisms of disease by this bacterium.

1.8.2 *Galleria mellonella* model

Galleria mellonella larvae are the caterpillar stage of the wax moth. Although these insects are a poorly characterised model in comparison to other invertebrate models such as *Caenorhabditis elegans* and *Drosophila melanogaster*, *G. mellonella* larvae have been used quite widely as an alternative to an animal model for the study of bacterial virulence (Senior et al., 2011, Champion et al., 2010, Tsai et al., 2016). Advantages of this model include; commercial availability of cheap larvae, small size (around 1-2 cm long) and no dedicated housing, ethical consideration or animal license requirement. Significantly, unlike other invertebrate models, trials with *G. mellonella* can be performed at 37°C, a requirement for *Campylobacter* studies. Inoculation is generally via injection at the hind legs, directly into the right foreleg, permitting inoculation with precise bacterial numbers. The ability to work with large numbers of test samples provides good statistical evaluation at reasonable cost.

Galleria do not possess an acquired immune response, but have the ability to produce both a cellular and humoral innate response, with activation of phagocytic cells and release of opsonin-like proteins and small antimicrobial peptides including lysozyme. During this process, phenoloxidase is activitated and converts phenols to quinines that polymerise and accumulate as melanin around the bacteria (Kavanagh and Reeves, 2004). Although the distinct function of melanisation in host defence is not fully understood, melanisation, together with loss of motility and death over a set period, are often used as markers of virulence (Nappi and Christensen, 2005, Senior et al., 2011, Tsai et al., 2016). Virulence, recorded in a dose-response manner as LD50, varies greatly between different pathogens and does not necessarily correspond to virulence in mammals. *Pseudomonas aeruginosa* was reported to be particularly virulent with 25 CFUs resulting in 100% mortality after 24h (Hill et al., 2014). An LD50 at 48h of 2.57×10^3 CFU was reported for an enteropathogenic strain of *E. coli* (Leuko and Raivio, 2012), while sensitivity of these insects to *C. jejuni* strains is several logs lower. In one study, after infecting insects with 10^6 CFU and incubation for 24h, approximately 40% of the *Galleria* survived infection with 11168H and 81-176, while 24% survived infection with strain M1 (Champion et al., 2010). This cheap readily available model can be used to obtain preliminary information regarding virulence.

1.8.3   Chicken model of colonisation

Understanding steps in colonisation and survival of *C. jejuni* in chickens is clearly important in developing strategies to control entry of this bacterium into the food chain (Newell, 2001). In an early study, one-day-old chicks were inoculated with 5×10^8 *C. jejuni* in PBS. After 7 dpi, the bacterial cells were found to principally colonise the lower part of the gastrointestinal tract, specifically the caecum (10^{4-7} CFU/g) (Beery et al., 1988). There has also been a report of recovery of *C. jejuni* from other extraintestinal organs such as liver, spleen, and kidney, suggesting transient systemic infection (Firlieyanti et al., 2016). The following reasons have encouraged scientists to use this natural model for characterising chicken colonisation by different isolates of *C. jejuni*. Chickens are a natural and susceptible host to *Campylobacter*. Study on this model is easy compared to other animal models. The hatched chickens are usually free of *Campylobacter*. This permits the evaluation of the ability of *C. jejuni* strains to colonise the chickens, and lastly it is estimated that chickens as a main reservoir of *Campylobacter* are responsible for 80% of human campylobacteriosis (Hermans et al., 2011). The chicken model has been extensively used to understand factors involved in colonisation by *C. jejuni*, (Byrne et al., 2007, Hermans et al., 2011, Hendrixson and DiRita, 2004, Naikare et al., 2006, Bingham-Ramos and Hendrixson, 2008, Jones et al., 2004). Among genes identified are those with a role in motility (*fliA*, *flaA*, *flgK*, *flg*R, *rpoN*, *maf5*, *cj1321-cj1325/6*), capsule formation and glycosylation (*pgiH*, *kpsM*, *cj1496c*), chemotaxis (*cheY*, *cheB*, *cheR*, *tlp1*, *luxS*, *docB*, *docC*, *acfB*), invasion and adhesion (*docB*, *ciaB*, *docC*, *tlp1*, *cadF*, *pldA*, *peb1A*, *flpA*), iron regulation (*cfrA*, *ceuE*, *chuA*, *feoB*, *fur*, *znuA*, *cj0178*), and oxidative stress responses (*docA*, *cj0358*, *perR*, *sodB*, *kata*, *ahpC*).

1.9 Aims of this project

1.9.1 Background on *Campylobacter* spp. isolated from Norway rats

C. jejuni is the most common cause of bacterial gastroenteritis and is transmitted primarily via infected chicken meat. This bacterium has the ability to evolve and transmit amongst a large number of different hosts, such as humans, chickens, and wild birds. To address the possibility that farm associated Norway rats act as a vector for the transmission of *Campylobacter*, 350 *Campylobacter* strains had been isolated from Norway rats on 8 farms (C, E, H, G, F, I, J, K) in S.E. England between 2010 and 2013 (Stuart et al., 2011). The bacterial strains were isolated from the intestine of trapped rats or from fresh rodent faecal samples collected in the early morning and stored in peptone water. The samples were then transported and processed for selective isolation of *C. jejuni* complex (*C. jejuni*, *C. coli*, *C. lari*, *C. hyointestinalis*, *C. upsaliensis*) on mCCDA at 37 or 42°C. Assignment to species level was then achieved with mass spectrometry, using a Bruker Biotyper. The majority of isolates were identified as *C. jejuni*. Genomic DNA from 350 isolates was extracted and 143 samples (including 6 isolates from farmed animals) selected for Illumina Whole Genome Sequencing (WGS), genome assembly and analysis via the BIGSdb database (collaboration with Maiden and McCarthy group, Department Zoology, Oxford and the Sanger centre, Cambridge). Genomic data for all 143 isolates is now available via the PubMLST/*Campylobacter* website (http://pubmlst.org/campylobacter/) and can be identified by either their Reading Dg number or PubMLST ID number.

Isolates were classified by MLST typing and included many well-defined clonal complex groups including ST21CC, ST45CC, ST42CC ST22CC and ST828CC. However, phylogenetic comparison of both whole genome sequences (WGS) and ribosomal multilocus sequence types (rMLST) of the 137 *Campylobacter* isolates identified two novel clades of *C. jejuni* (**Figure 1.11**). Thirty of the isolates belonged to a clearly distinct clade that could not be assigned to an existing CC group and was therefore designated as 'rat group 1' or RG-1. There was no evidence of closely related strains to RG-1 strains in the BIGSdb database, which contained 30,000 isolates from chickens, humans, wild birds, or any other niche. An additional thirty-nine isolates belonging to ST45CC formed a phylogenetic clade of ED-

positive CC-45 strains, distinct from other ED negative strains belonging to ST45CC. Hence, the ED positive CC-45 strains were designated as 'Rat Group 2' or RG-2. Because both sets of strains had been isolated over 3 years from 7 or 8 different farms. It was hypothesised that the RG-1 and RG-2 strains are widespread within farm-associated rats in the S.E of England and that divergence of these groups from the common clonal complexes of *C. jejuni* may be a consequence of adaptation to a specific host or environment.
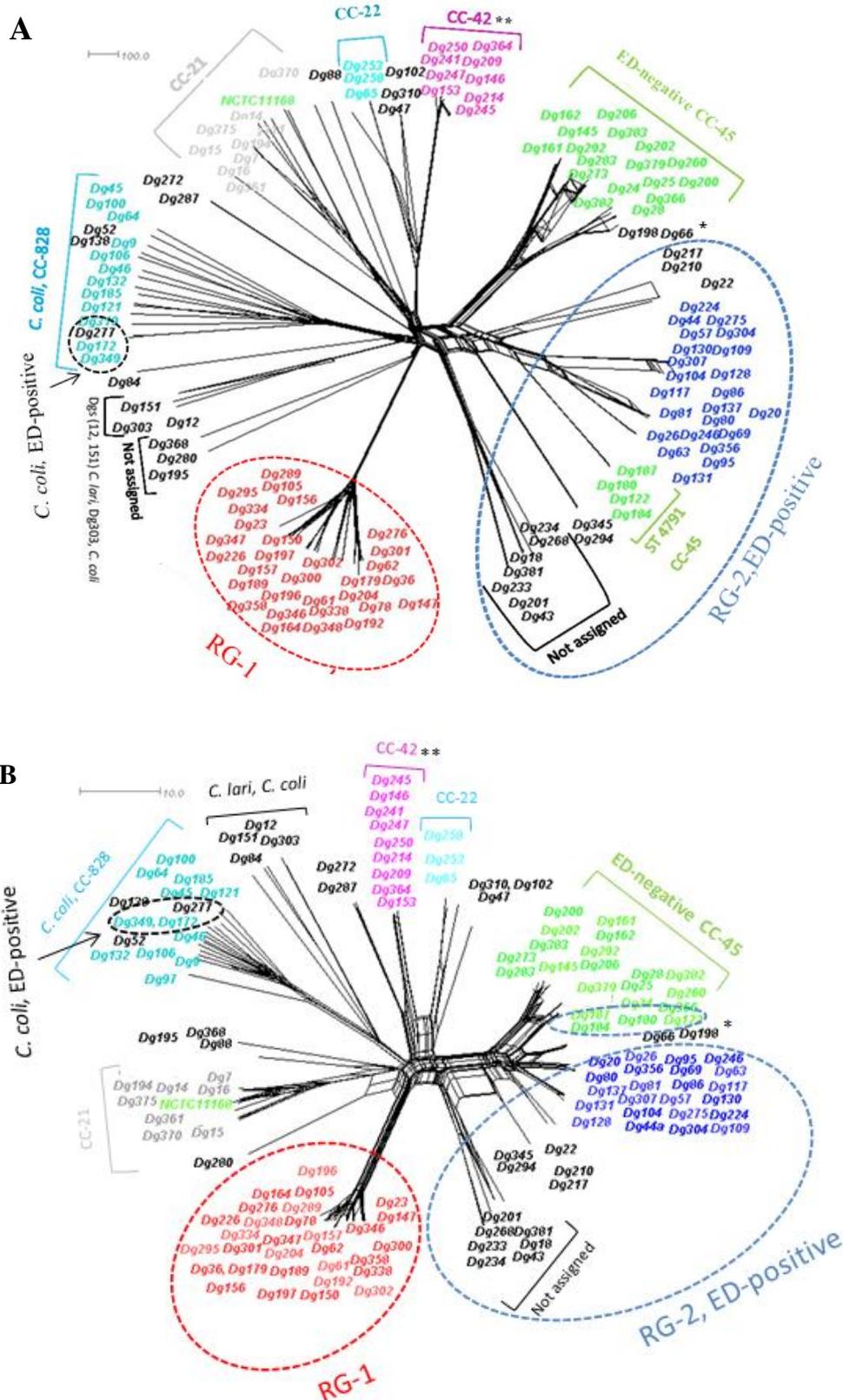
Figure 1.11: Phylogenetic network of genomes from Norway rat associated *Campylobacter* strains. (**A**) wgMLST and (**B**) rMLST of 137 *C. jejuni/ coli* strains isolated from rat faecal pellets or intestine from 8 farms in S.E. England over 3 years. RG-1, red dotted circle; RG-2, blue dotted circle, positive for *glc* locus (ED positive); black dotted circle, *C. coli* positive for

glc locus. CC-clonal complex and ST-sequence type are indicated. All data is from Norway rats except NCTC11168, (*) 81116, and (**) 81-176, well studied *C. jejuni* strains isolated from humans. Draft genomes of the farm- associated rats and genomes of the human associated strains were compared using genome comaparator via the PubMLST database. The outcome Nexus file was uploaded into SplitsTree4 to visualize the phylogenetic networks (Huson and Bryant, 2006).

1.9.2   Aims and objectives

The overall aim of this project was to characterise the RG-1 and ED-positive RG-2 strains that had been isolated from Norway rats. Glucose metabolism had never been demonstrated for *C. jejuni* and the novelty of the RG-1 strains suggested that they might be adapted to a specific niche, possibly the Norway rat. *In vitro* studies were used to investigate growth, motility and biofilm formation, with particular focus on preference for different amino acids and glucose as nutrients. Properties of RG-1 and RG-2 strains were compared to those of rat associated isolates belonging to recognised generalist clonal complexes, associated with chicken carriage and human disease. *In vivo Galleria* and chicken trials were conducted to test the hypothesis that the RG-1 and RG-2 strains might be less efficient colonisers of chickens due to adaptation to a different niche. The draft genome of one selected strain of each group was fully assembled to a closed genome, initially to localise the *glc* locus in RG-2, but also for full genome comparisons. A combination of the experimental and bioinformatics analysis has identified possible factors of interest for future study, on the significance of glucose metabolism to *C. jejuni*, and the possibility of a *C. jejuni* rat model of colonisation and virulence based on RG-1 strains.

Chapter 2

Materials and Methods

2.1 Media, buffers and chemicals

Chemicals and media were purchased through Thermo Fisher Scientific, unless otherwise indicated. Composition and preparation of the media and buffers are listed in **Appendix 2.** The following media were routinely used; 1.2% Blood Agar (BA) and Mueller-Hinton Agar (MHA) for preparation of inoculum and routine culture and 2% BA was for colony isolation and Colony Forming Unit (CFU) determination. Modified *Campylobacter* blood-free selective agar (mCCDA) was used for *C. jejuni* quantitation and recovery from animal trials, Mueller-Hinton (MH) broth was used for monitoring growth, amino acid, and glucose levels. Brain Heart Infusion (BHI) also used for monitoring growth, where indicated, and for stocking bacterial samples. *Campylobacter* motility MH and BH agar plates (0.4% agar) were used for monitoring motility. Minimum Essential Media Eagle (MEM-α) and Dulbecco's Modified Eagle Medium (DMEM) were for monitoring amino acid and glucose utilisation. Bolton broth medium was used during *Campylobacter* isolation to encourage multiplication of cells. Maximal Recovery Diluents (MRD) and Phosphate-Buffered Saline (PBS) were used for serial dilution preparation for CFU calculations.

2.2 Bacterial strains and routine culture conditions

The majority of bacteria used in this study were *C. jejuni* isolated from fecal rat pellets or intestine for which draft genome sequence data were obtained in a previous study. Data about these isolates including source, date of isolation, sequence type (ST) and/or clonal complex (CC) number is listed in **Table 2.1 A-E**. **Table 2.1 A-C** details farm-rat associated *C. jejuni* and *C. coli* strains. **Table 2.1 D** summarises subcultures of stocks, isolated following chicken trials performed in this study. **Table 2.1 E** lists published human and chicken isolates. Strains were stocked as original Dg stocks as well as first round motility stocks, where they were given the designation RM-1. RM-1 strains were used in all experiments unless otherwise indicated. Strains were routinely cultured from (-80°C) frozen stocks on 2% BA (w/w) plates containing 7% (v/v) defibrinated horse blood and incubated at 37°C for 44-48h under microaerophilic conditions using a Thermo Scientific Oxoid CampyGen Atmosphere generation system for 2.5L (CN0025A) or 3.5L (CN0035A) jars. A single colony was then

streaked on a 1.2% (w/w) BA or MHA plate and incubated at 37°C for 24-36h under microaerophilic conditions, unless otherwise indicated. These subcultured plates were used to prepare an inoculum of standardised cells (**see section 2.2.2**). Where indicated, 100µl of suspended antibiotic Bolton selective supplement (SR0183E, Thermo Scientific), containing vancomycin, cefoperazone, trimethoprim and cycloheximide, was added to 9.9ml of the liquid medium to prevent contamination. During the chicken trial, cultures were incubated in a Whitley M35 workstation (Don Whitley Scientific), at 42°C for selective isolation of *Campylobacter* in a gaseous atmosphere of 5 % O_2, 5% CO_2, 2% H_2 and 88 % N_2.

Table 2.1: Strains used during this project.

Table 2.1, A: Distinct clade, farm- associated rat, *C. jejuni* group 1 (RG-1)				
Strain Dg/RM-1 [1]	ID [2]	ST[3]	CC[3]	Date, farm site[4]
Dg23b	25941	6562	-	29/07/2011, H
Dg36b	25967	-	-	22/08/2011, H
Dg61	24189	6561	-	25/08/2011, K
Dg62a	24190	6561	-	25/08/2011, K
Dg78a	25939	5129	-	11/11/2011, E
Dg105	26050	-	-	05/12/2011, C
Dg147*	26008	6562	-	21/05/2012, G
Dg150	26046	6562	-	21/05/2012, G
Dg156	26053	6562	-	21/05/2012, G
Dg157	26060	-	-	22/05/2012, C
Dg164	26032	-	-	22/05/2012, C
Dg179	25984	-	-	22/05/2012, X
Dg189	24194	6562	-	22/05/2012, X
Dg192	24196	6564	-	28/05/2012, E
Dg196	24191	6561	-	30/05/2012, I
Dg197	25992	6562	-	30/05/2012, I
Dg204	26010	6561	-	30/05/2012, I
Dg226	26042	6562	-	26/04/2012, X
Dg276	26065	-	-	03/07/2012, X
Dg289	26014	-	-	19/07/2012, C
Dg295a	26030	-	-	19/07/2012, C

Dg300	25968	6562	-	19/07/2012, X
Dg301b	25972	-	-	19/07/2012, X
Dg302	25976	6562	-	19/07/2012, X
Dg334	25946	7279	-	25/07/2012, G
Dg338	26069	6562	-	25/07/2012, G
Dg346	26017	7278	-	26/04/2012, C
Dg347	26023	-	-	26/04/2012, C
Dg348	26025	-	-	26/04/2012, C
Dg358	26001	6564	-	30/04/2012, E

Table 2.1, B: ED-positive, farm- associated rat *C. jejuni* group 2 (RG-2)

Isolate	ID	ST	CC	Date, farm site
Dg95	26020	7259	45	05/12/2011, G
Dg275*	26058	7259	45	03/07/2012, X
Dg26	25959	7259	45	29/07/2011, H
Dg69	25962	5130	-	11/11/2011, E
Dg356	24192	5130	-	30/04/2014, E
Dg80	25987	5130	-	11/11/2011, E
Dg57	25947	7259	45	22/08/2011, C
Dg246	26063	7259	45	27/06/2012, G
Dg130	26015	7259	45	05/12/2011, X
Dg104	26043	7259	45	05/12/2011, C
Dg63	25975	7258	-	25/08/2011, L
Dg224	25989	7259	45	26/04/2012, X
Dg304	25991	7259	45	19/07/2012, X
Dg307	25990	7259	45	19/07/2012, X
Dg81	25995	5130	-	11/11/2011, E
Dg20	25937	7259	45	14/07/2011, G
Dg26	25959	7259	45	29/07/2011, H
Dg44	25950	7259	45	22/08/2011, X
Dg117	26064	7259	45	05/12/2011, X
Dg210	26027	7276	-	25/04/2012, G
Dg217	25945	7276	-	25/04/2012, G

Dg22	25974	3471	-	29/07/2011, H
Dg122	25999	4791	45	05/12/2011, X
Dg180	26034	4791	45	22/05/2012, X
Dg184	26040	4791	45	22/05/2012, X
Dg187	26054	4791	45	22/05/2012, X
Dg345	26009	7256	45	26/04/2012, C
Dg294	26022	7256	45	19/07/2012, C
Dg233	25997	-	-	26/06/2012, E
Dg234	26049	-	-	26/06/2012, E
Dg268	25948	-	-	03/07/2012, C
Dg43	25978	-	-	22/08/2011, X
Dg381	26002	-	-	03/08/2012, I
Dg18	25977	-	-	14/07/2011, G
Dg201	25993	-	-	30/05/2012, I
Dg131	25988	7259	45	05/12/2011, X
Dg86	26004	5130	-	11/11/2011, E
Dg128	26007	7259	45	05/12/2011, X
Dg109	26019	7259	45	05/12/2011, C
Dg137	26031	7259	45	05/12/2011, G

Table 2.1, C: Other farm-associated rat *C. jejuni* and *C. coli* isolates

Isolate	ID	ST	CC	Date, farm site
Dg200	25985	45	45	30/05/2012, I
Dg145	24193	45	45	21/05/2012, G
Dg292	25964	45	45	17/07/2012, C
Dg206	26018	45	45	25/04/2012, G
Dg162	26024	45	45	22/05/2012,C
Dg202	26026	45	45	30/05/2012, I
Dg273	26038	45	45	03/07/2012, X
Dg383	26062	45	45	03/08/2012, J
Dg24	25944	583	45	29/07/2011, H
Dg25	25955	583	45	09/07/2011, H
Dg28	25963	583	45	09/07/2011, H
Dg260	26051	583	45	03/07/2012, C

Dg366	26013	583	45	30/04/2012, J
Dg379	26037	137	45	02/05/2012, I
Dg382	26055	137	45	03/08/2012, I
Dg283	25960	1326	45	03/07/2012, X
Dg161	26016	45	45	22/05/2012,C
Dg7a	25965	21	21	23/06/2011, E
Dg14b	25951	21	21	23/06/2011, E
Dg16 b	25973	21	21	23/06/2011, E
Dg194	26061	21	21	28/05/2012, E
Dg361	25994	21	21	30/04/2012, E
Dg370	26041	50	21	30/04/2012, J
Dg153	24195	42	42	21/05/2012, G
Dg209	25981	42	42	25/04/2012, G
Dg245	26056	42	42	27/06/2012, G
Dg247	25957	42	42	27/06/2012, G
Dg172 (*C. coli*)	26067	1016	828	22/05/2012,C
Dg349 (*C.coli*)	25986	1016	828	26/04/2012, C

Table 2.1, D: *C. jejuni* strains recovered from chicken trial

Ch95-92**, Ch95-93, Ch95-94**, Ch95-95**, Ch95-97
Ch275-99**, Ch275-102, Ch275-103**, Ch275-104, Ch275-105
Ch194-106, Ch194-107, Ch194-108, Ch194-109, Ch194-110, Ch194-111, Ch194-112
Ch370-116
Ch200-120**, Ch200-121, Ch200-122**, Ch200-123, Ch200-124, Ch200-125, Ch200-126
Ch153-127, Ch153-128, Ch153-129, Ch153-130, Ch153-131, Ch153-132, Ch153-133
HPC5-134, HPC5-135, HPC5-136, HPC5-137, HPC5-138, HPC5-139, HPC5-140

Table 2.1, E: Reference strains

Isolate	ID	ST	CC	Reference
NCTC11168	48	43	21	From human stool, 1977, UK (Parkhill et al., 2000)
11168H	48	43	21	Variant of the reference isolate. 11168, provided by Andrey Karlyshev (Karlyshev et al., 2002)
81116 (NCTC11828)	1220	267	283	From human stool, 1981, UK (Pearson et al., 2007)
HPC5		356	-	Provided by Ian Connerton (Loc Carrillo et al., 2005)

[1]All strains are *C. jejuni* unless indicated *C. coli*. Dg designates cultures stocked from BA, in 2012/13, during genome preparation for the WGS data deposited in BIGSdb database (Jolley and Maiden, 2010). Strains were restocked as RM-1 following a single round of subculture on motility MHA plates (as described in **section 2.2.4**). RM stocks were used in all subsequent studies. Ch prefix (**Table 2.1, D**) designates strains recovered from chicken trials, -second number indicates chicken number from which strains were recovered.

[2]ID, identity number used in PUBMLST (www.pubmlst/campylobacter) to access the WGS data deposited in BIGSdb database (Jolley and Maiden, 2010)

[3] ST and CC, sequence type and clonal complex as defined by MLST. -, indicates no number assigned due to variability of *pgm* (*glmM*) locus for RG-1 strains and *gltA* locus for RG-2 strains

* genome closed in this study; ** Ch strains recovered from chicken trial and sent for WGS

2.2.1   Stocking of *Campylobacter* strains

Strain numbers are as defined during the initial study of isolation and sequencing of the bank of Reading University *Campylobacter* rat isolates (unpublished data). In this earlier study, during preparation of genomic DNA for WGS, a single colony had been streaked on two BA plates; 'a' was used primarily for genomic isolation and plate 'b' for stocking at -80°C in BHI (15% glycerol, v/v). Hence, strains were identified by the original D number (original purified stocks) or by a Dg number (the original D number + 'g' to denote it had been restocked at the time of DNA extraction). Dg stocks were used as a source of bacterial culture in this study. The first round motility (RM-1) stocks were prepared as described under motility test. All stocks were prepared in a class 2 Biological Safety Cabinet to avoid contamination and stored at -80°C.

2.2.2   Preparation of standardised bacterial cell suspension

A standardised cell suspension prepared from growth on a BA plate was used as inoculum for standard growth studies, motility tests, Biofilm assay, Galleria studies and chicken trials and from growth on an MHA plate for Biolog studies and testing specific growth supplements. To prepare an inoculum, strains were sub-cultured on 2% BA (w/w), incubated at 37°C for 42-48h, followed by subculture of a single colony onto routine BA or MHA plates and further incubation for 24-36h. Using a category 2 biological safety cabinet, one loopful (~5µl) of growth was taken and evenly suspended in 1ml of fresh MHB or minimal medium, as required, in a screw-capped Bijou tube. Suspended cells were centrifuged at 8000rpm for 2

minutes at RT (Mini spin, Eppendorf). The supernatant was removed and the pelleted, washed cells gently re-suspended in 1ml of fresh media. The OD_{600} was monitored using 600µl of suspended cells in a 1cm path length cuvette and a WPA CO8000 cell density meter. The remaining sample was used to standardise samples to 0.02 OD_{600} by dilution with appropriate media. A loopful of standardised culture was routinely plated on a BA plate and incubated for 48h aerobically at 37°C to confirm absence of contamination. **Figure 2.1** confirms a linear correlation of 2-fold serial dilution of *C. jejuni* from OD_{600}1.05 down to 0.02 in PBS. An OD_{600} of 0.2 from BA corresponded to an average of ~7.4 x 10^8 CFU of *C.jejuni*/ml. **See Appendix 1** for calculated CFU correlation for individual strains.
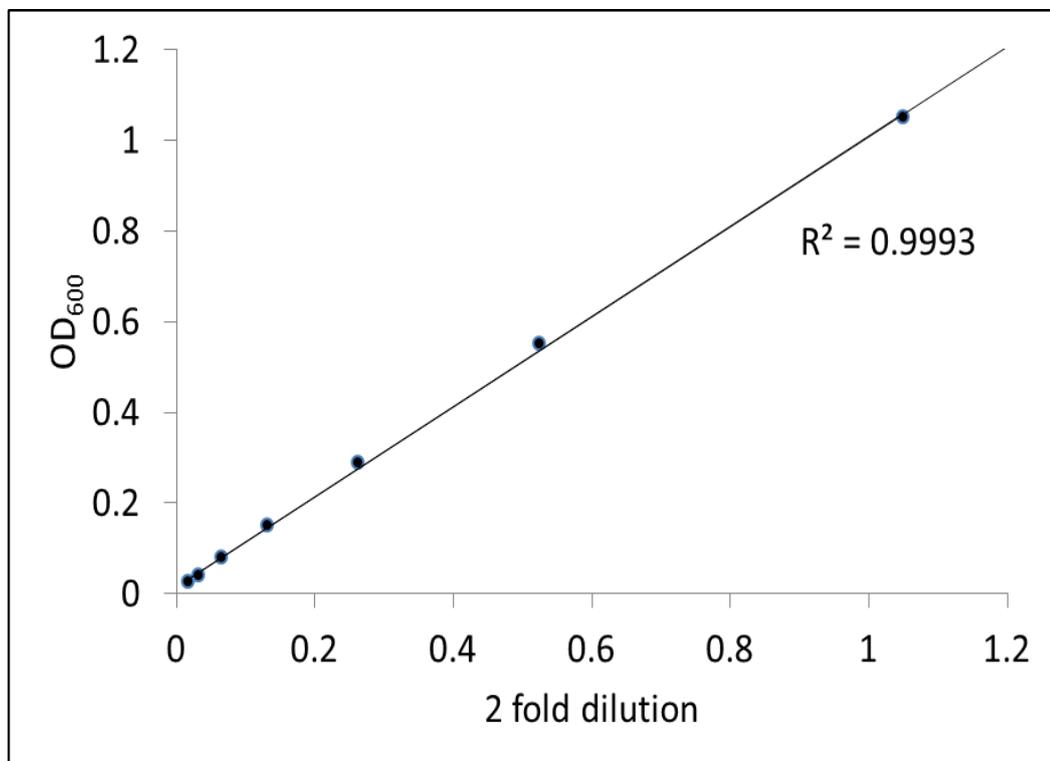
Figure 2.1: Correlation between serial dilution of *C. jejuni* Dg153 and optical density. A two-fold serial dilution of cells in PBS from 1.05 to 0.02 OD_{600} was prepared and read using a WPA CO8000 Cell Density Meter.

2.2.3 Viable cell count (CFU) and doubling time

The Miles and Misra technique (Miles et al., 1938) was followed to enumerate viable *Campylobacter* cells. For growth studies, 10-fold serial dilutions from 10^{-1} to 10^{-7} were

prepared in a 96-well plate (Fisher Thermo Scientific, 10687551) by adding 20μl sample to 180μl of Phosphate-Buffered Saline (PBS) or Maximum Recovery Diluent (MRD). Triplicate serial dilutions were prepared for each sample. Plates, 2%-BA or 2%-CCDA, were predried for ~30min at 37°C with the lid slightly opened. Samples (10μl) from 10^{-4} to 10^{-7} dilutions were then spotted, 4-6 times, onto the predried plates, using a multichannel pipettor with VWG (613-2324) or Starlab (S1113-1706) tips. Spots were allowed to dry and plates were incubated at 37°C for 42-48h. A magnifying lens was used to count CFU, ideally between 2-30 colonies /drop. For growth analysis, doubling (generation) time was calculated. Cell numbers were taken from exponential time points. Doubling time was calculated using the following formula $= \frac{duration \; x \log{(2)}}{\log{(final\;conc)} - \log{(initial\;conc)}}$ (Roth V, 2006 http://www.doubling-time.com/compute.php).

2.2.4 *Campylobacter* motility test

A motility test was performed as previously described with slight modification (Guerry et al., 1991). A yellow sterile pipette tip was used to stab 5μl of standardised 0.2 OD_{600} bacterial suspensions into the center of a motility plate (containing 0.4% agar, w/w). The low density of the agar facilitates bacterial swimming within the motility agar plate resulting in an extensive visible area of growth around the inoculation point. Following incubation of the culture plates at 37°C for 72h, the radius of the ring was measured in millimeter (mm). To isolate bacteria from this plate, 2-3 stabs were taken from the motility area of the 0.4% MHA plate using a sterile 3ml plastic dropper. The stab of motility agar plates, which contained motile cells, was suspended in BHI, then directly sub-cultured on BA plates and incubated for 30-42h at 37°C. The cells were then suspended in a 2ml micro-tube containing 1ml of 15% glycerol and stocked as the 'first round motility' stock culture (RM-1). For the second round of MHA culture, the procedure was repeated, the RM-1 stock was streaked on BA plate, and a repeat motility test performed. Bacteria from this second round were stocked as 'second round motility' (RM-2). No antibiotic was used in any of the motility tests hence samples were confirmed as contamination free by inoculating 1-2 loopfuls of the motility area on a blood agar plate and checking for aerobic growth for 48h at 37°C.

2.2.5 Summary of Analytical Profile Index (API) protocol

Bacterial cells were harvested from 38h incubated BA plates and resuspended in API 0.85% medium (3ml) to turbidity roughly equivalent to 6 McFarland. This suspension was used immediately. Instructions of the manufacturer were used (BioMeriux SA 080501-en-2012.03). The cupules of the enzymatic test half of an API-CAMPY strip were inoculated with approximately 80-100µl of the cell suspension and incubated for 24h at 37°C under aerobic conditions. The strip was then read after adding reagents and waiting for 5min. Results were compared to a colour chart.

2.2.6 Microtiter plate growth assay

Growth of *Campylobacter* isolates in 96-well plates was used to monitor the initial growth of different strains in rich media, optimised growth in minimal media, and to compare utilisation of different amino acids and glucose as a carbon source. Sterile 96-well polystyrene plates (Corning, 10687551 or Greiner, plate 655161 with lid 656161) were used with 180µl of fresh medium and 20µl of inoculum, prepared in the same medium. Each strain was plated in triplicate. Test wells were inoculated with 20µl standardised cell suspension equating to an OD of 0.02 (~1.5 x 10^6 CFU/20µl) and immediately incubated at 37^oC in 3.5L jars with a gas pack, for the desired time. For any incubation period longer than 48h, the gas generating pack was replaced. Inoculum and incubated cultures were confirmed to be free of facultative aerobic contaminants. In some experiments, where indicated, Bolton's Broth Selective antibiotic was added to ensure absence of any contamination. When Bolton broth supplement antibiotic was added to the growth, culture there was no significant effect on bacterial cell growth (**supplementary Figure 2.1**).

In initial experiments, prior to reading cell density, cells were gently re-suspended by rocking the plate on a Biometra WT12 rocker for 2min but this was insufficient to resuspend cells for a number of strains. Therefore, to more accurately monitor growth for all strains, the bacterial cells were resuspended by pipetting up and down 4 times with a 200µl BIOHIT multichannel pipette. This was sufficient to read plates for incubations up to 24h, but evaporation affected readings for incubation times longer than this. Therefore, 100µl of resuspended cell culture was transferred into a new 96-well plate with 100µl of fresh medium. Bacterial density

(OD_{600}) was monitored in a Spectra MAX 340 plate reader. An OD_{600} of 0.2 using the WPA CO8000 cell density meter or the Thermo Spectronic Helios Alpha spectrophotometer was equivalent to a reading of 0.078 at OD_{600} using the Spectra MAX 340pc spectrophotometer with 200µl of culture in a 96-well plate.

2.2.7   Nutrient dependent growth of *Campylobacter*

The microtiter plate growth assay was followed to study amino acid utilisation by both RG-1 and RG-2 strains and glucose utilisation by RG-2 strains. Fresh stocks of the amino acids (400mM) were made in the established MEM-α medium as described by Guccione et al. (2008). The stock was filter sterilised using Fisher Scientific (15206869) filter steriliser. Each well contained 0.002 OD_{600} of bacterial cells and 20mM of the carbon source.

2.2.8   Growth in flasks with shaking

Monitoring growth under shaking condition was done for RG-2 strains when they were grown in MHB supplemented with 10% Foetal Bovine serum (FBS) (v/v) and 1% Bolton Broth supplement antibiotic, with or without 20mM glucose. This method was also used to monitor growth in established DMEMf medium (**Appendix 2**) with added serine (20mM) and with or without glucose (20mM). Ten mililiter of the medium was added into a 25ml conical flask, incubated for up to 72h at 37°C with shaking at 150 rpm in a 2.5L jar, using 2.5L atmosphere generation system CampyGen bags (Thermo Scientific, CN0025A). Aliquots were taken at the required time of incubation and gas packs were replaced. The bacterial cells were removed from the samples by centrifuging at 13,000 rpm for 5min. The supernatant was then stored at -20°C for subsequent amino acid and glucose analysis. Before monitoring the carbon sources, the samples were thawed out and filter sterilised using 0.22µm Millex- GV filter tip (SLGV004SL).

2.3   Quantification of amino acids and glucose

2.3.1   Amino acids quantification

RG-1 strains were grown in MEM-FBS medium (**Appendix 2**) supplemented with either serine or proline and incubated under microtiter plate growth assay conditions (**see section 2.2.6**). RG-2 strains were grown in MHB plus 10% FBS (v/v), with or without glucose, or DMEMf medium with serine, with or without glucose (and without FBS) and incubated in shaking flasks (**see section 2.2.8**). Samples of culture superntant taken during growth were stored at -20°C. The stored samples (200 - 600μl) were thawed out, filter sterilised (Merck Millipore, code SLGV004SL) and diluted in distilled water (DW) to within the range of the standard mixture (200μM) of the EZ-Faast kit (KG0-7166). Solid phase extraction was the first step of the analysis. This step was done by passing the liquid culture medium through a sorbent packed tip for removal of interfering compounds and binding of the amino acids to the tip. In the second step, the amino acids were removed from the sorbent packed tip and derivatised by using 4 reagents, step by step. The derivatised organic compound was lastly analysed by gas chromatography-mass spectrometry (GC-MS), (Agilent Technologies, 6890N, Network GC system). GC separated constituents of the organic solution into pulses and was followed by ion separation by MS.

2.3.2   Glucose quantification

Following growth, free glucose in MHB or DMEMf culture media, was monitored by either an enzymatic glucose oxidase assay (Assay range 1-200μM) or high-pressure anion chromatography. For these studies, freshly filtered 400mM D-glucose (Sigma-Aldrich, code, G8270) and Ser (400mM) were prepared in the appropriate media. The strains were grown in DMEMf + 20mM Ser ± 20mM glucose or in MHB +10% FBS ± 20mM glucose. Both media were finally incubated in flasks with shaking condition.

2.3.2.1   Glucose Assay Kit (Merk, product code, CBA086)

The Merk glucose assay kit was routinely used for colourimetric measurement of glucose remaining in culture media following bacterial growth. The amount of glucose was proportional to color density from 1-10 nmole/well (**Figure 2.2**). The kit was stable at 4°C for 24 – 48h once opened, but could not be refrozen. Therefore, the first trial was used to get an idea of how much glucose was left at each time point during growth. Based on this, the other samples were then diluted in DW to within the range of the kit (1- 200 μM). The reagent mix was initially prepared from 46μl glucose assay buffer, 2μl glucose probe and 2μl glucose enzyme. A 50μl of this reagent mix was added to sample (final volume 50μl adjusted with glucose assay buffer) in a 96-well Greiner plate. Each sample was tested in duplicate or triplicate, as shown. The plate was incubated for 30min at 37°C, protected from light, and absorbance at 570nm was measured with a Spectra MAX 340pc spectrophotometer. This kit was also used to measure free glucose concentrations in the standard medium. There was no glucose in DMEMf or TSB, ~3mM in each of MHB and FBS, and ~3.7mM in BHI.
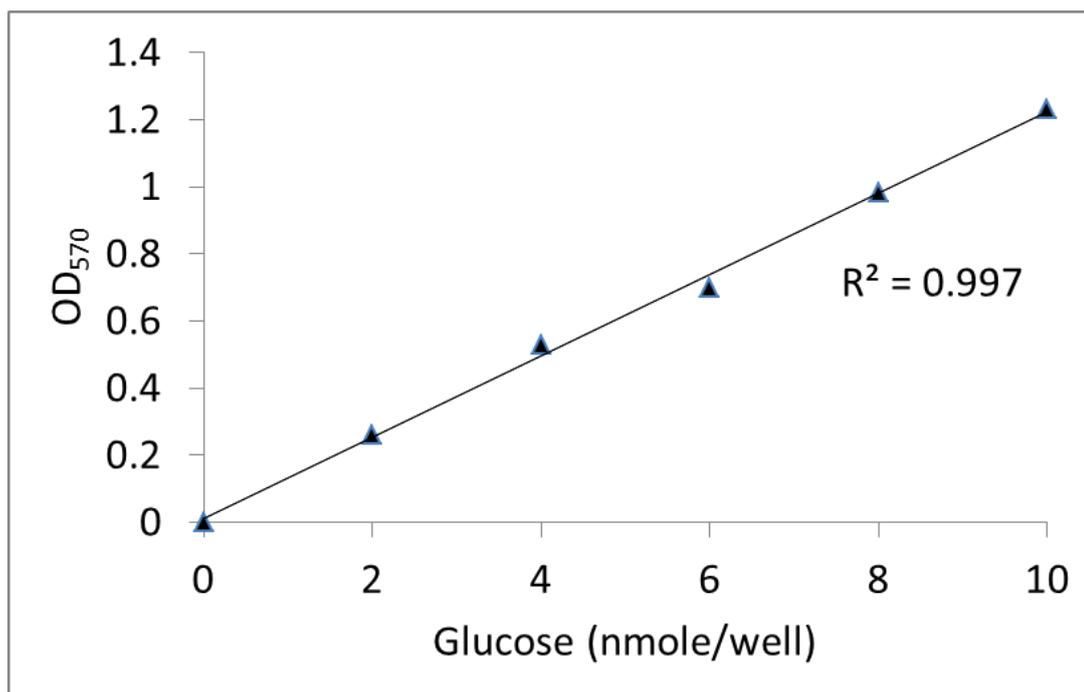
Figure 2.2: Standard curve for monitoring glucose. The 0, 2, 4, 6, 8, 10 μl of the Glucose Standard (1nmol/ml) were added into each well individually. The wells were adjusted to 50μl by adding the glucose assay buffer to generate 0, 2, 4, 6, 8, 10 nmol/well of the Glucose Standard. A 50μl of a reaction mix was added into each well. The reaction mix was prepared from 46μl of the glucose assay buffer, 2μl of glucose probe and 2μl of glucose enzyme in advance. The plate was incubated in an incubator for 30 min at 37°C (protected from light) and finally, the Spectra MAX 340pc spectrophotometer was used to measure the absorbance at 570nm.

2.3.2.2 High Pressure Anion Chromatography (HPAC)

D-glucose (Sigma-Aldrich) was used to make up glucose standards in the following concentrations: 0.27 mM (50mg/l), 0.11 mM (20mg/l), 0.027 mM (5mg/l) and 0.0027 mM (0.5mg/l). Diluted sample (20 μl within the range of 0.27 – 0.027 mM glucose) and standards were injected onto a Carbopac PA10 column (Dionex) at RT using an autosampler. An isocratic program was set up using 125mM NaOH at a flow rate of 1 mL/min. A pulsed amperometric detector was used with the following settings: 420 ms at 0.05 V, 180 ms at 0.75 V, and 420 ms at −0.15 V, while the sensitivity was set at 3K. Glucose in the samples and prepared standards were monitored in the Department of Food & Nutritional Sciences, University of Reading using an 8220i Dionex ion chromatography system (Dionex Corp., Sunnyvale, CA).

2.3.3   Biolog phenotypic microarray

Biolog phenotypic microarray (Biolog) was performed as a collaborative study with Professor Roberto La Ragione, Veterinary Pathology Centre, School of Veterinary, Surrey University. A single colony of each *C. jejuni* strain was grown on BA for 36h at 37°C, then sub-cultured onto an MHA plate for a further 48h. Approximately half of a plate was suspended in IF-0a inoculating fluid (Biolog, Hayward, CA, USA) and adjusted to 16% transmittance (Turbidimeter, Biolog 3531), equivalent to an OD_{600} of 0.8. Four mililliter of the adjusted cells was mixed with 0.88ml sterile DW, 120µl tetrazolium violet (Redox dye D) as a redox indicator dye (100x) (Biolog 74224), 6ml of IFOa liquid, and 1ml of additive solution (X12). The additive solution was prepared in advance from 5ml of 6% (w/v) albumin, (bovine fraction V powder BSA, Sigma A7906), 5ml of 150mM sodium bicarbonate NaHCO3 (Sigma S6014), and 40ml of DW. A final volume of 100µl of the mixture was transferred into each well of a 96-well PM1 Biolog plate. The plate was placed in a plastic Biolog bag with a Campy Gen Compact Sachet (Oxoid, code CN0020C). An impulse sealer was used to seal the end of the bag. The plates were finally placed in an OmniLog PM automatic plate reader (Biolog 91171) and were incubated at 37°C. Metabolism of different carbon sources was monitored every 15min for 48h by quantitating production of a purple color. The color occurs due to the reduction of tetrazolium violet dye. It was known that the following sugars in PM1 give a false positive reaction; L-arabinose (A2), D-xylose (B8), D-ribose (C4) and L-lyxose (H6). The data were analysed by examining the kinetic plot curves of the different strains, the ability of the strains to reach the maximum point at either 24h or 48h, and finding differences in the carbon source utilisation.

2.4 Biofilm assays

2.4.1 Floating biofilm formation

Floating biofilm, as opposed to formation of biofilm on a solid surface, was performed as described by Vegge et al. (2016). The *Campylobacter* strains were grown on BA plates, incubated for 24-36h at 37°C. Biofilm was then assessed in 5cm petri dishes containing 10ml of glucose free tryptic soy broth (TSB, Becton, Dickinson and Company, code 286220), with and without 100mM glucose, inoculated with $0.002OD_{600}$ cells, and supplemented with 1% (v/v) Bolton broth supplement antibiotic. The plates were incubated for 4 days at 37°C under microaerobic conditions without replacing the gas generating system. After incubation, the plates were shaken very gently on the bench. Extensive floating biofilm was formed by the control strain, *C. coli* Dg349 and *C. jejuni* Dg153. A Sony α 65, 24.3 Mpx camera was used to take pictures of the phenotypic properties that were then saved in tif format.

2.4.2 Microtiter plate biofilm -crystal violet (CV) staining assay

The CV based biofilm assay was essentially performed as described by Hanning and Slavik (2009). Duplicate sterile polystyrene 96-well plates (Corning) were prepared to monitor both growth and biofilm. The *C. jejuni* strains were grown on BA plates and incubated for 30-42h at 37°C. Triplicate wells of each plate, containing 180 µl MHB or BHI were inoculated with 20µl 0.02OD standardized cells and incubated according to the microtiter plate growth assay (**see section 2.2.6**). Plates were incubated for 72h without replacing the gas generating system. The OD was read using one plate as described in the microtiter plate growth assay. The second plate, which was used to assess biofilm, was rocked for 2min slowly at 40 rocks/min on a Biometra WT12 rocker. All medium was gently removed using a multichannel pipettor and used to check for absence of contamination. The plate was dried for 30min at 55°C in a GENLAB oven. After drying, 200µl of filtered 0.16% w/v crystal violet was added to each well. The plate was then rocked for 15 min at 40 rocks/min. The dye was removed from the wells that were then washed with 200µl sterile DW, rocking for 15min. The DW was removed from the wells and the plate was again dried at 55°C for 15min in the oven. This washing process was repeated twice. The washed wells were then filled

with 200µl of 80:20 ethanol: acetone to extract bound CV and rocked for a further 15min. Finally, the absorbance at 590nm was measured by using the Spectra MAX 340pc spectrophotometer microplate reader. Control wells contained only MHB. Growth, as measured by OD_{600} was considered a measure of total bacteria (sessile and planktonic). CV reading, OD_{590} was a measure of sessile bacteria (bound to the plate) or biofilm.

2.4.3 Scanning Electron Microscopy (SEM) analysis of biofilms on glass fiber filter

Bacterial samples were prepared for scanning electron microscopy (SEM, FEI Quanta 600 FEG SEM microscope) as described by Kalmokoff et al. (2006). Bacteria were grown on 0.5cm x 1.5cm sterile Whatman glass microfiber membranes (Sigma-Aldrich, code, WHA1825-025) in Corning 24-well plates (Corning, code CLS3526-50EA). Each well contained 900µl of MHB inoculated with 100µl *Campylobacter* (0.025 OD_{600}), and one membrane standing in an upright position. Cultures were incubated without shaking at 37°C in the gas pack generating system for the desired time. All medium was removed approximately every 24h (or as indicated), and replaced with 1ml fresh sterile MHB medium. After ensuring absence of contamination by Gram staining, the membranes were prepared for SEM.

To fix samples for SEM, incubated membranes were placed into Eppendorf tubes using sterile forceps and washed with 1ml of sterile PBS (pH 7.3) (three times for 1min). The PBS was replaced with fixation solution, 2.5% glutaraldeyde (v/v) in 0.1M sodium cacodylate buffer (v/v, PH 7.0). Samples were incubated at either: A, room temperature for 3h or B, overnight at 4°C (or longer when necessary). The membranes were washed with the PBS buffer twice more for 15min. Finally, the samples were subjected to an acetone dehydration series of 30%, 50%, 70%, 80%, 100% (3x), v/v acetone. At the 70% acetone step, samples were left overnight at 5°C, if necessary; otherwise, each step was for 15-20min at RT.

Once dehydrated, samples were dried in a critical point dryer (BALZERS, CPD 030 critical point dryer). The membranes were placed in a sample holder, which was submerged in liquid CO_2. The sample holder was put in a chamber. This chamber was then drained and refilled

with the liquid CO_2 several times. This process was repeated 7-8 times to ensure that the dehydration medium was completely replaced by the liquid CO_2. The specimens were then heated to several degrees above the critical temperature of CO_2, at 40°C, and 80-85 bar of pressure. The gaseous CO_2 was released and the dried specimens then mounted and coated with gold immediately.

To coat with gold, the dried membranes were transferred onto a sample holder of Edwards sputter coater. The pressure of the chamber was decreased to approximately 1×10^{-1} Atm and then increased to just below 4×10^{-1} Atm for 15 -20min, during which time the sample was coated with gold. Gold coated samples were examined under SEM (FEI, QUANTA 600F). Initially, samples were examined at 600x to 1500x magnification to assess biofilm structure and extent of spread versus clustering of cells. In the next step, the membranes were then gradually focused from 15000x to 30000x magnification to examine areas of microcolony/biofilm formation. The files were saved as tif files.

2.5 Negative staining and Transmission Electron Microscope (TEM)

C. jejuni strains (RM-1 stocks, **section 2.2.4**) were subcultured on 2% BA, incubated for 44-48h at 37°C. A single colony was streaked on a 1.2% BA and incubated for 24h at 37°C. MHB (10 ml) was then inoculated with a loopful of bacteria and incubated for 22-24h with shaking (150rpm). One milliliter of the culture was centrifuged at 3000rpm for 3min. Approximately $1mm^3$ of the pellet was re-suspended in 500µl of the following fixative solution; 2.5% (v/v) EM grade glutaraldehyde in 0.1M sodium cacodylate buffer (pH 7.0) and stored overnight at 4ºC. The fixed cells were then centrifuged at 3000rpm for 5min. The supernatant was discarded and the pellet washed (2x) in 0.1ml nH_2O and resuspended in 0.1ml nH_2O. For TEM, a glow discharged copper grid (Carbon film, 300 mesh copper100) was placed on a piece of sterile filter paper and then 30µl bacterial culture was gently pipetted onto the grid and incubated for 10min at RT. The grid was then stained with 30µl of 1% uranyl acetate, left for 1 min, then washed 3 consecutive times (10sec) in 1% (w/v) aqueous ammonium acetate (pH 4.3). The grids were then focused from 2000x to 8000x by TEM at 200KV using a JEM 2100 Plus TEM, product version 1.7.10.1750, with camera AMT XR401 and camera software version: 7.0.0.184 for image capture. Files were saved as

tif files. Cell dimensions and amplitude were measured using ImageJ software (Wayne Rasband, NIH http://imagej.nih.gov/ij) (Rueden et al., 2017). Generally, data from 10-20 cells/strain was averaged.

2.6 Animal studies

2.6.1 Monitoring impact of rat mucin on growth

Intestinal mucin was recovered from available lab rats to be sacrificed, 4 white transgenic GFP (~250g) and a Norway (500g) lab rat. Rats were killed by euthanasia in a CO_2 chamber, by a qualified technician. Sterile forceps were used to cut the abdominal skin of the rats and separately dissect out the small intestine and caecum. Intestinal segments were slit open using sterile scissors and the contents removed by gently rinsing twice with cold PBS into a Petri dish. The mucosal surfaces of the small intestine or caecum were then scraped with a sterile plastic scraper (SARSTEDT) to harvest secreted mucus into a sterile Petri dish. The collected mucin was homogenized using a lysing matrix tube (MP Biomedical, code 116912050) and a MP shaker (FastPrep-24). Approximately 1ml of the mucin was homogenised for 10sec at 4m/s (5x). The MP tubes were then centrifuged for 90sec at 8000 rpm and the supernatant used either directly for supplementation in a growth assay or stored at -20°C. During isolation and homogenisation of mucin, samples were kept on ice, as far as possible. To monitor the impact of the mucin on growth, 5% (v/v) of the mucin and 1% (v/v) Bolton broth supplement antibiotic was used in MHB. *C. jejuni* strains were grown on BA plates for 30-42h. MHB medium was supplemented with 5% of the mucin and inoculated with 0.002OD of the *C. jejuni* cells. The medium was incubated at 37°C for up to 72h under the microtiter plate assay condition (**see section 2.2.6**). Both OD and CFU were monitored every 24h for up to 72h.

2.6.2 *Galleria mellonella* infection model

Galleria mellonella larvae were purchased from Live Foods UK (code number, 5060446440704), stored on wood chips at 5°C without feeding and used from 1-3 days, after arrival. Bacterial strains were routinely subcultured on BA plates from glycerol stocks. A loopful of 24h culture was used to inoculate 10ml of MHB, which was incubated for 18- 26h at 37°C under microaerophilic with shaking at 150rpm. The cells were then washed with PBS (pH 7.3) after spinning down at 8000rpm for 2min. The bacterial cells were concentrated, adjusted to 1OD$_{600}$ in PBS and further diluted in PBS to give the following 10µl doses: 1OD, ~1.23 x 10^{7}CFU; 0.25OD, ~4.5 x 10^6 CFU; 0.062OD, ~9.6 x 10^5 CFU and 0.015OD, ~2.2 x 10^5 CFU. For inoculation, a disposable glove was worn over a protective puncture glove (The Safety Supply Company, code, SA-Q3984LG) and the disposable glove was replaced between each dose and strain. A 0.3ml disposable insulin syringe was used to inject 10µl aliquots of different doses of the bacterial cells adjacent to the right proleg of the larvae. Ten larvae were inoculated for each strain and dose. For controls, 10 larvae were injected with PBS and 10 were incubated without injection. Each treatment group of larvae were incubated on top of filter paper in a separate Petri dish aerobically at 37°C for 24h. The macroscopic appearance (unhealthy black and orange colors or healthy creamy color), and mortality defined by Champion et al. (2010) were recorded at 24h.

2.6.3 Colonisation of chickens with *Campylobacter*

2.6.3.1 Chicken trial

The chicken trial was performed at Nottingham University, Sutton Bonington Campus, School of Biosciences, Division of Food Sciences, in collaboration with Professor Ian Connerton. Prior to start of the trial, chickens were screened for absence of *Salmonella* and *Campylobacter*. Bacterial cells (1 ml, ~71og$_{10}$ CFU) were delivered into the stomach of each of the 16-day-old *Campylobacter*-free male 308 Ross broiler chickens by oral gavage (n= 7/ bacterial strain for three days post infection (dpi) and n= 7/bacterial strain for seven dpi). PBS (1ml) was delivered into the stomach of each of the 7 control chickens for three and seven dpi. The following *C. jejuni* strains were tested; RG-1 (n=1, Dg147), RG-2 (n=2, Dg95

and Dg275), ST 45 clonal complex (n=1, Dg200), ST 42 clonal complex (n=1, Dg153), ST 21 clonal complex (n=2, Dg194 and Dg370), and the standard chicken trial control HPC5 ST 356 (n=1) plus PBS control (n=1). At day 16, the chickens were separated into 9 groups of 7, x2 (a and b). For "b"s, each bird was housed in an individual cage and "a"s remained in co-housing for each group of 7. The average weight of the 16-day-old chickens was 534 ± 6g. Weight and health was monitored throughout the trial. At day 19, all "a" birds (n= 63) were sacrificed and intestinal segments surgically removed. Aliquots of caecal, ileum and jejunum contents were transferred to preweighed labeled tubes and stored on ice. Extra-intestinal organs (heart, spleen, kidney, breast, liver) were aseptically removed, placed in individual sterile homogenisor bags and stored on ice. At the same time, samples of caecal content and intestinal tissue were recovered and stored at -80°C for future metagenomics/cytokine analysis. This was done by following published procedures, Loc Carrillo et al. (2005). At day 23 all "b" birds (n= 63) were sacrificed and processed in the same way.

2.6.3.2   Processing of samples

The Miles-Misra microplate method was used to count *Campylobacter* CFU in the caecal, ileal and jejunal contents essentially as described in **section 2.3.3**. Intestinal content was suspended in Maximum Recovery Diluent (MRD) to 1g /ml making the first sample a 10^{-1} dilution. The contents were then vortexed vigorously for 15-30 seconds, to resuspend before tenfold serial dilution in triplicate to a concentration of 10^{-7}. Five aliquots (10µl) of each triplicate of the 10^{-2} to 10^{-7} dilutions were dispensed onto a 2% mCCDA plate for selective isolation of *Campylobacter*. The plates were incubated at 42°C in a microaerobic atmosphere for 48h in a Modular Atmospheric Controlled System (MACS) cabinet (Don Whitley Scientific) with a gas mixture of 88% (v/v) N_2, 5% (v/v) CO_2, 2% (v/v) H_2 and 5% (v/v) O_2 (BOC Limited). After incubation, the colonies were visually inspected for typical shiny, round, and gray colonies (Chon et al., 2012) of *Campylobacter*. Any suspect samples were confirmed by Gram staining. An occasional contaminant observed with the 10^{-2} dilution from chickens not colonised by the *C. jejuni* strain tested formed a flat white or creamy colony and had an oval cocci shaped cell morphology. Confirmed *Campylobacter* colonies were counted to calculate the CFU of *Campylobacter*/g of intestinal contents.

The extra-intestinal organs were enriched to assess presence or absence of *Campylobacter*. Each organ was homogenized in 10ml of MRD by shaking in a Stomacher Biomaster Lab system for 10sec. The samples (1ml) were then transferred into a universal containing 9ml Bolton Selective Enrichment Broth and incubated for 4h at 37°C, followed by further incubation at 42°C for 20h. Following incubation, 10µl from below the Bolton Broth meniscus was spotted onto a mCCDA plate. The plates were incubated for 48h under microaerophilic conditions at 42°C and any growth confirmed as *Campylobacter* from colony and cell morphology, as needed. *Campylobacter* growth was scored as positive or negative for each chicken.

A Miles and Misra mCCDA plate from each 7 dpi chicken with *Campylobacter* in the caecal content was brought back to the Reading University lab in 3.5L jars with CampyGen Sachets. Five single colonies from the highest dilution of the RG-2 strains and at least one single colony of each of the other strains were streaked on to 2% BA plates, incubated for 40-44h at 37°C and glycerol stocked. Recovered colonies were stocked as, for example, Ch275 – 99 indicating inoculated *C. jejuni* strain Dg275, recovered from chicken number 99. Confirmation of strain identity was performed by amplification and sequencing key MLST loci: *aspA* 288 for RG-1 and *aspA gltA* and *glyA* for other strains (**see Table 2.3 for primers**). Recovered colonies from chickens infected with the RG-2 strains, Ch95 and Ch275, were also screened by PCR and growth in DMEMf plus/minus glucose for possession of a functional *glc* locus. Ten of the infected and recovered strains were sent for WGS.

2.7 DNA Techniques

2.7.1 Preparation of genomic DNA

2.7.1.1 Colony lysis

Lysed colonies were used as a source of genomic DNA for amplification of shorter PCR fragments (~200 bp to 4-6 kbp). For this, a few colonies were re-suspended in 100µl of TE (100mM Tris, 10mM EDTA, pH 8.0). The suspended cells were boiled for 5min and then centrifuged for 3min at 13,000rpm. The supernatant was stored in a new sterile tube at -20°C for later use. Routinely, a 1/50 dilution was used for PCR reactions.

2.7.1.2 Isolation of genomic DNA

GeneJET Genomic DNA purification kit (Thermo Scientific, K0721) was used to prepare genomic DNA from fresh cultures subbed from single colonies onto BA plates and grown at 37 °C for 24-28h BA. Approximately half of a culture plate containing ~2 x 10^9 bacterial cells was harvested and then treated in 180µl of digestion solution. Proteinase K and RNase A solutions were added into the digested solution to remove protein and RNA in the solution. To wash the genomic DNA, ethanol and wash buffer I and wash buffer II were used. Routinely, the purified DNA was eluted in 80-100µl of nH_2O. As an exception, elution buffer of GeneJET PCR Purification Kit (Thermo Scientific, code K0701) was used to elute the DNA of selected chicken strains for WGS as MicrobesNG recommend use of an elution buffer that does not contain EDTA.

2.7.2 DNA quantification

2.7.2.1 Nanodrop spectrophotometer DNA quantification

A 2 µl sample was routinely used to estimate DNA concentration and purity using a Nanodrop spectrophotometer (ND1000, Thermo Scientific). Nanodrop water or elution buffer

was used to blank the Nanodrop reader. GeneTool was used to quantify genomic DNA, following AGE on 0.6 % agarose gels for WGS.

2.7.2.2 Agarose gel electrophoresis (AGE)

AGE was also used to confirm the quality and concentration of DNA using 0.6-0.75% (w/v) agarose gel, prepared in TAE buffer (40mM Tris-HCl, 1mM EDTA, pH 8.0). GelRed nucleic acid stain (1 µl) was mixed with 50ml of melted gel at 45-50°C and then poured into a gel caster (BioRad). DNA loading buffer was added to DNA samples, which were pipetted into the gel along with an appropriate ladder; 1µl of DNA Hyper Ladder 1 (Bioline 1kb) or 2-4µl of High DNA Mass Ladder (**Figure 2.3**). The gel was run at 80 volts for 60-90min. Gel pictures were taken using G-Box (SynGene) and both file formats tif and sgd were saved. Where necessary, DNA quantification was calculated using GeneTools software; a rectangle was drawn around gel bands and then automatic background correction was used to subtract background. DNA in either genomic DNA or PCR products were quantified through comparison with the desired band from the High DNA Mass Ladder (**Figure 2.3**).
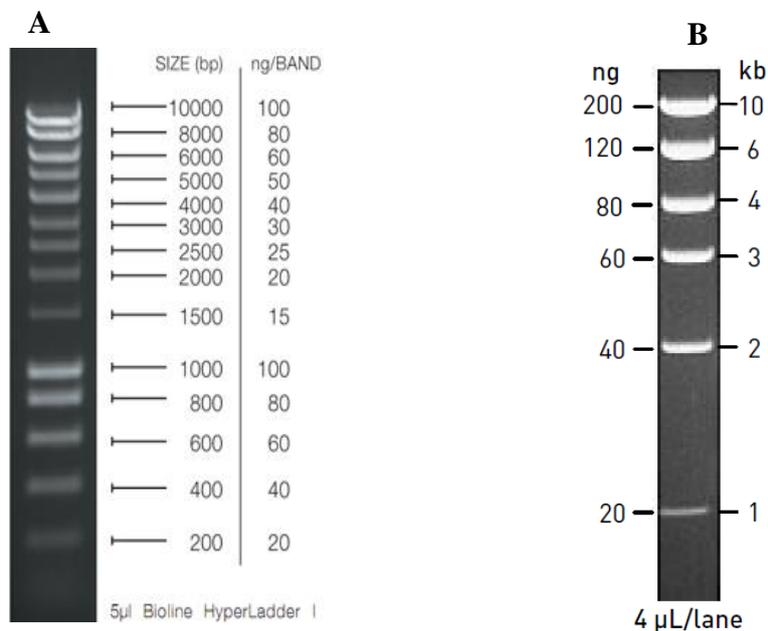
Figure 2.3: DNA ladder. (**A**) DNA Hyper Ladder 1 and (**B**) High DNA Mass Ladder. Molecular size (bp) and respective molecular mass (ng) of each band is shown from 5µl of A and 4µl of B.

2.7.3   Oligonucleotide primers

Oligonucleotide primers were purchased from Eurofin–Genomics (Germany). They were routinely suspended in sterile nH_2O to make stock concentrations of 200 pmol/µl, diluted to a working concentration of 10 pmol/µl, and stored at -20°C. Primers were routinely designed using APE program (A Plasmid Editore, V2.0.51) and/or TM Calculator of Thermo Fisher. Typical conditions for designing primers are stated in **Table** 2.2 except the first three primers of **Table 2.3.** The best pair of forward and reverse primers were used for amplification.

Table 2.2: Primer setting conditions.

	Minimum	Maximum	Salt (mM)	50
Length	18	22	Primer (nM)	250
Tm	54	62	Orientation	5' --> 3'
%GC	38	60	Sort by	5'

Primers that were used to amplify and sequence specific genes to confirm strains are listed in **Table 2.3.** Primers that were used to close the genome sequences of Dg147 (RG-1) and Dg275 (RG-2) are listed in **Table 2.4** (**A & B**) and **Table 2.4** (**C & D**), respectively. Primers that were used to fill big gaps of Dg147 and Dg275 using the primer walking method are listed in **Table 2.4** (**B & D**), respectively. Primers that were used to amplify and primer walking of *C. coli* Dg349 *rrnA* are listed in **Table 2.4 (E)**.

Table 2.3: Amplification and sequencing primers used to confirm isolates.

Primer	Sequence (5′-->3′)	Product size (bp), paired with	GC (%)	Tm (°C)	Comment and/ or used for
aspA9f *	AGTACTAATGATGCTTATCC	941, aspAr20	35	48	RG-1, RG-2, CC-42, CC-45,CC-21
aspAr20	ATTTCATCAATTTGCTCTTTGC	941, aspA9f	52	52	RG-1
aspA r21*	ATTTCATCAATTTGTTCTTTAC	941, aspA r21	23	47	CC-45, CC-42, CC-22
gltA-A1f	GGGCTTGACTTCTACAGCTACTTG	1112, gltA-A2r	50	59	RG-2
gltA-A2r	CCAAATAAAGTTGTCTTGGACG	1112, gltA-A1f	41	56	RG-2
glyA-A1f	GAGTTAGAGCGTCAATGTGAAGG	1052, glyA-A3r	48	58	RG-1, RG-2
glyA-A3r	AAACCTCTGGCAGTAAGAGC	1052, glyA-A1f	50	60	RG-1, RG-2
edaF	TTGGACTTAACCACGAACCACC	3739, zwfR	50	59	RG-1
zwfR	ACGCCTGATTGCCTTCTATTGC	3739, edaF	50	59	RG-2
pgiF	AAGAGTGGGTTCTTGGCTTGG	2462, glcPR	52	59	RG-2
glcPR	AAGACACCATCGCAACCTTTCC	2462, pgiF	50	59	RG-2
zwfF	ACCTTGGGAGTTGAAAGCAGAG	3400, pgiR	50	59	RG-2
pgiR	AACGCCCACCTACCCAATTC	3400, zwfF	55	59	RG-2
hipOF	GAAGAGGGTTTGGGTGGTG	735, hipOR	57	57.7	Specific for *C. jejuni* (Lawsonl et al., 1998)
hipOR	AGCTAGCTTCGCATAATAACTTG	735, hipOF	39	54	Specific for *C. jejuni*,
aspkF	AGCTTGCGGTAAGAGCCTTGC	400, aspkR	57	61.8	Specific for *C. coli* (Lawsonl et al., 1998)
aspkR	ATAAAAGACTATCGTCGCGTG	400, aspkF	43	55.9	Specific for *C. coli*
edaF	TTGGACTTAACCACGAACCACC	3739, zwfR	50	59	RG-2

zwfR	ACGCCTGATTGCCTTCTATTGC	3739, edaF	50	59	RG-2
pgiF	AAGAGTGGGTTCTTGGCTTGG	2462, glcPR	52	59	RG-2
glcPR	AAGACACCATCGCAACCTTTCC	2462, pgiF	50	59.6	RG-2
zwfF	ACCTTGGGAGTTGAAAGCAGAG	3416, pgiR	50	59	RG-2
pgiR	AACGCCCACCTACCCAATTC	3416, zwfF	55	59	RG-2
glcPF1	CAGCTTCTACAGCTTCTTATCC	829, glcPR1	45	54	RG-2
glcPR1	GCAACCAAACCTTGAACTGGAG	829, glcPF1	50	58	RG-2
eddF	ACACCCGCTATGTGTGATGG	956, eddR	55	58	RG-2
eddR	ACTGCTCTGCCTATGTTTCCTC	956, eddF	50	58	RG-2
pgiF	AAGAGTGGGTTCTTGGCTTGG	927, pgiR1	52	59	RG-2
pgiR1	TCTTGCACTTCTTCGAGTGTTC	927, pgiF	45	56	RG-2

* PuBMLST recommended primers

Table 2.4: Primers used for closing genomes of *C. jejuni* Dg147 (RG-1) and Dg275 (RG-2), and to localise the *glc* locus in *C. coli* Dg349

Table 2.4: A- Dg147 amplification and sequencing primers.

Gaps	Primer Name & Primer (5'-->3')	% GC	Tm (°C)	Flanking contigs	Distance from contig end (bp)	Expected Product Size (bp)
*A	Dg147-A-F AGGTTGTTGTAAGTTCAAGGGTTGTAG	40	59	6, 9	371	6485
	Dg147-A-R AGAATTAACAAGTCCGCAATGAGCTAC	40	59		85	
B	Dg147-B-F GTGGTGAAAGATTTGCTGTGTGTG	45	59	9, 4	304	1528
	Dg147-B-R AACCTTTGCCTCATAACCTTGTGG	45	59		391	
	Dg147-B-FA ATCGCTCAACCAAATGGCGAAC	50	60		747	2527
	Dg147-B-RA TAGCTTCCTATCACTCCACCATC	47	57		951	
C	Dg147-C-F TGTAAAGTCCAAACTCCAGTTCCAC	44	59		328	1651
	Dg147-C-R	45	59		516	

	AGGACACTCATCAACCACTTATGC			4, 2			
	Dg147-C-FA TGTCCTTTGGACTTTGCGTG	50	57		741		2633
	Dg147-C-RA CAGTTGGAGTTTCAGCCAAGC	52	58		1085		
*D	Dg147-D-F ATGGCGATGGAAATGCTAATTTAGG	40	58	2, 3	532		7201
	Dg147-D-R TGATACTAAAGCCACAAATGAAAGTGC	37	58		472		
*E	Dg147-E-F ATCAACTCTAAGAACGCTTAACTCAC	38	57	3,7	519		6932
	Dg147-E-R AGCTTGCTGAGAAGATTTAGATATGG	38	57		216		
F	Dg147-F-F TCATGCAGGAATTGGCGTACAG	50	59	7, 10	363		443
	Dg147-F-R GCGTTCCATTCATCAGGGACTATC	50	59		80		
G	Dg147-G-F CTCAGCCACCTATGAAAGTCAAGC	50	59	10, 8	246		1222
	Dg147-G-R AACTCAGCGCCATCATATCTGC	50	59		605		
H	Dg147-H-F TAGGTTATGGGCCAAGTGTTGC	50	59	8,17	379		4244
	Dg147-H-R ACTAAGGCAACTCAAGCAGCTC	50	59		483		
I	Dg147-I-F GCGTTTAATGCACCGATGTTGG	50	59	17, 18	517		1623
	Dg147-I-R AAGCGGCTCAAGATGGACAAAG	50	59		470		
	Dg147-I-FA ATCGCCATCCCTGAAGCATC	55	59		628		1894
	Dg147-I-RA GATGCTTCAGGGATGGCGATAG	54	59		634		
J	Dg147-J-F ATCCCTGAAGCATCATCTGCTG	50	59	18, 15	495		1077
	Dg147-J-R AGCCCTTTACCAACGCTTCAATC	47	59		90		

K	Dg147-K-F TTTCTAGTATGGTCATTGGGTTTCC	40	57		283	938
	Dg147-K-R ATGTGCTTTCATGTGCTACTCATC	41	57	15, 23	163	
	Dg147-K-FA TTTGGCTAAACTAGGTCCTGC	47	56		997	1776
	Dg147-K-RA TTGTGCTGATGGAGCTTTGAG	47	57		289	
L	Dg147-L-F CTTCCACGACTTGTGATCTCCATG	50	59		302	5051
	Dg147-L-R GGCTTGGAGAGGAAGCAACATAC	52	59	23, 1	529	
	Dg147-L-FA CATATTCACCCTCATAAGCCCAC	47	57		511	5852
	Dg147-L-RA AAACCCAGGGATGCCTTATG	50	56		1123	
M	Dg147-M-F CCACTTACAGCTATCCAAACAGG	45	54		123	1242
	Dg147-M-R CAGATGAAGATAGAGCAAGAGC	45	54	1, 6	290	
	Dg147-M-FA GATGCAATCGCTAGAACTCCAGG	52	59		468	2805
	Dg147-M-RA GAGCAGCTATGAGTTTGGCAAAG	47	58		1508	

*A, D, E contain the *rrnABC* operons, respectively

Table 2.4: B- Dg147 primers used in primer walking of big gaps.

Gaps	Primer Name & Primer (5'-->3')	% GC	Tm (°C)	Location on unmapped contigs, alignment manner
A, D, E	glcp3RCC AGTATCATCACCCACGATGTGC	50	58	47bp far from the beginning of unmapped 5, aligns complementary
	glcP3RD TTCCAACCGTTCTGAGCCAACC	54	61	976bp far from the beginning of unmapped 5, aligns complementary
	glcpFCCC AGGTGTGGTTAGCTTCGTACTAGG	50	59	1559bp far from the beginning of unmapped 5, aligns reverse complementary
	glcP3RE ACGACCTTAGACTAGCACTTCC	50	57	1815bp far from the beginning of unmapped 5, aligns complementary
	glcPFC ACCCAGACTACCAGCTAAGG	55	57	2434bp far from the beginning of unmapped 5, aligns reverse

				complementary
	glcp3RF ACGCCTATGGGTTCAGTCCTC	57	60	2689bp far from the beginning of unmapped 5, aligns complementary
	glcp3RG AGTCAAGGCATCCACCATTCG	52	59	3420bp far from the beginning of unmapped 5, aligns complementary
	rnn1 TCCAACCGCAGGTTCTCCTAC	57	60	4212bp far from the beginning of unmapped 5, aligns complementary
	rnn2 GTCAGTTAAGTTCCAGCAGATCG	47	57	4986bp far from the beginning of unmapped 5, aligns complementary
H	Dg147-H-Fa TGCACTGAACCAATATCAGC	45	54	2245bp far from the beginning of unmapped 20, aligns complementary
L	Dg147-LRB ATTACTAGGAGCACCATCTCG	57	55	317bp far from the beginning of unmapped 13, aligns reverse complementary
	Dg147-LRC GAGATACCCGTAAAGCACTGG	52	56	153bp far from the beginning of unmapped 14, aligns reverse complementary
	Dg147-LFC ATCTACCATCCACCAAGGAC	50	55	1723bp far from the beginning of unmapped 14, aligns complementary
	Dg147-LFD AGCTCGGCAGTTTGATTCAG	50	57	1125bp far from the beginning of unmapped 29, aligns complementary
	Dg147-LRE AGAGTTTGCTGTTGGTTAGG	45	54	66bp far from the beginning of unmapped 25, aligns reverse complementary
	Dg147-LFE CCCTAACCAACAGCAAACTC	47	57	65bp far from the beginning of unmapped 25, aligns complementary

Table 2.4: C- Dg275 amplification and sequencing primers.

Gaps	Primer Name & Primer (5'-->3'), or hybridized gene	% GC	Tm (°C)	Flanking contigs	Distance from contig end (bp)	Expected Product Size (bp)
1	Dg275-1F AAGTGTTATCGCCAGTTACG	45	54		263	
	Dg275-1R GAAGCACCACAAGATTGCTC	50	55	9,16	72	4120
	Dg275-1FA CCCGTTTGTCTATCAAGTAGTGC	47	57		375	4532
	Dg275-1RA	55	50		372	

	CCATGACTTGGATCAAGCAC					
2	Dg275-2F CGTATCAGTTGTAGAAGGTGG	47	54	16, 10	335	1021
	Dg275-2R CCATAGAACCTAAAGCTGAACC	45	54		201	
3	Dg275-3F CCAGTTCCACATCCTATGTC	50	53	10, 7	307	869
	Dg275-3R ATGCCAATGGAGCTGTAGAG	50	56		141	
*4	Dg275-4F TTGATGAAATGGGCAGAGTG, 230bp far from in stop codon of encoding a hypothetical protein gene	45	54	7, 3	373	15827
	Dg275-4R AGTTCAACTCTTACGCAGAGC, 59bp far from in stop codon of *murD*	47	56		182	
	Dg275-4FA AGAGTTGATGAGTTGGCTATC G, 567bp far from in stop codon of encoding a hypothetical protein gene	45	56		710	16774
	Dg275-4RA GCATTGATGCTGTTATGGGTG G, 869bp far from in stop codon of *murD*	45	56		992	
	glcPF TTTGGTTGGCGGAAAGGTTGC, 992 bp far from in stop codon of *glcP*	52	60	From *glcP* *murD*	1029	5424
	Dg275-4RA GCATTGATGCTGTTATGGGTG G, 869bp far from in stop codon of *murD*	45	56		992	
5	Dg275-5F TCCAAGGACATCTATGCCAG	50	55	3, 18	133	6798
	Dg275-5R TCTCAGCGCATTCTTTACCAC	47	56		447	
	Dg275-5FA TACTCTTTGCGGCAATACAGG	50	55		1104	7769
	Dg275-5R TCTCAGCGCATTCTTTACCAC	47	56		447	

6	Dg275-6F TGTGTGGAGTTCTCGTATGTG	47	56	18, 35	296	500
	Dg275-6R AGTGTTTGGGTTTGCTGTTC	45	55		243	
7	Dg275-7F TGGTAGGCTATGGAATTTGTGG	45	56	35, 14	261	5732
	Dg275-7R TAAAGCCACACTCTGATTGC	45	54		318	
*8	Dg275-8F ATGGAGAACATCGAGCAAGAG, 50bp far from in stop codon of encoding a hypothetical protein gene	47	56	14, 11	98	6852
	Dg275-8R ACTTAAAGCACTACGAACGAG, 351bp far from in stop codon of encoding a hypothetical protein gene	45	56		489	
9	Dg275-9F TGGACTTGGCTTGTCTCTTTC	47	56	11, 6	211	693
	Dg275-9R GCCTAAATCGTTTGCAGTGTG	47	56		139	
10	Dg275-10F AAGTTCTGCGCTTGAAAGAG	45	55	6, 5	247	901
	Dg275-10R TCCATTGCGTTCCACTTACAC	47	57		311	
11	Dg275-11F TCAAGGACATCTATGCCAGC	50	56	5, 1	131	404
	Dg275-11R TCGGCTGTATTAACACCATG	45	54		363	
	Dg275-11FA TCCACCCGCATCAGTATCTTC	52	58		474	747
	Dg275-11R TCGGCTGTATTAACACCATG	45	54		363	
12	Dg275-12F GCTTTGGACTTAGAATGGCTAC	45	55	1, 13	277	404
	Dg275-12R AGCACTTTCAACCTTGCCTG	50	57		217	
13	Dg275-13F TCCGCTTCCAGCTAAACCAG	55	58		151	389

	Dg275-13R GCTTTGGAGTAAACGCTTTGC	47	57	13, 15	328	
14	Dg275-14F TGGCATCCAAACATTTCTCC	45	54	15, 9	249	338
	Dg275-14R CATCGTCCAGATATAGTAGC	45	50		179	
15	Dg275-15F TAGCCACTCCAAGACATACG	50	55	9, 2	168	6693
	Dg275-15R GTGTCTGTGATGAGTTTGAG	45	52		431	
*16	Dg275-16F CGTGTTGCTGTGAGTGATTC, 147bp far from in stop codon of *ansB*	50	55	2, 19	413	15657
	Dg275-16R GACTTCGGCAATAGAAAGT, 219 bp far from end of node 19 insertion between 16SRNA and *ggt*	45	52		219	
	Dg275-16FA GGCGTGTTGCTGTGAGTGATT C, 149bp far from in stop codon of *ansB*	54	60		415	16244
	Dg275-16RA CAAGATTTCACATGCAGTGGC TAC, 197bp far from in stop codon of *ggt*	45	58		804	
	zwfF ACCTTGGGAGTTGAAAGCAGA G, 755bp far from in stop codon of zwf	59	59	Part of unmappe d 17 + unmappe d 4 + part of 19	6214	15151
	Dg275-16RB TTGCGTATGAGTTTGGCGGTG, 712bp far from in start codon of complementary putative type II restriction endonuclease	52	60		5581	

* *rrnA* and *glc* locus, and *rrnB* and *glc* locus were located in 16 and 4 gaps, respectively, *rrnC* was located in the gap 8, same for D

Table 2.4: D- Dg275 primers used in primer walking of big gaps.

Gaps	Primer Name & Primer (5'-->3')	% GC	Tm (°C)	Location on unmapped or mapped contigs, alignment
1	Dg275-1Fb ACCCATTCTTGCTTTGAGTGC	47	57	1108bp far from the beginning of unmapped 23, aligns complementary
	Dg275-1FC TACCGACAGATTGAAACGCACC	50	59	142bp far from the beginning of unmapped 27, aligns reverse complementary
	Dg275-1RC TGGTGGTGGTTATCTCATTGAGG	47	58	988bp far from the beginning of unmapped 27, aligns complementary
*4	ed1FCc CGATCCCGCTATTCTCCACC	59	61	2080bp far from the beginning of unmapped 22, aligns complementary
	edaF TTGGACTTAACCACGAACCACC	59	59	316bp far from the beginning of unmapped 17, aligns complementary
	edaF1 ACTTGCGGAACTTCGTTTGC	50	58	681bp far from the beginning of unmapped 17, aligns complementary
	eddR ACTGCTCTGCCTATGTTTCCTC	58	58	1403bp far from the beginning of unmapped 17, aligns complementary
	eddF2 AGTCATAGGTCCTGAAGGCAC	52	57	2134bp far from the beginning of unmapped 17, aligns complementary
	eddF3 TTGTGATTTGCTCCAGCTTC	45	55	2649bp far from the beginning of unmapped 17, aligns complementary
	zwfF ACCTTGGGAGTTGAAAGCAGAG	58	59	3448bp far from the beginning of unmapped 17, aligns complementary
	zwfF1 GGGTCCTAAGGAAGTCTTTG	50	53	4164bp far from the beginning of unmapped 17, aligns complementary
	pglF1 AAGCTGAAGAGTTCGAGTGG	50	56	4817bp far from the beginning of unmapped 17, aligns complementary
	glkF3 CAGGAGAGGGTGGGCATACTAG	59	59	5538bp far from the beginning of unmapped 17, aligns complementary
	pgiR AACGCCCACCTACCCAATTC	55	60	6844bp far from the beginning of unmapped 17, aligns reverse complementary
	pgiF AAGAGTGGGTTCTTGGCTTGG	52	60	6451bp far from the beginning of unmapped 17, aligns complementary
	pgiF1 TGACACGGGTCCTATTATTTGG	45	56	7162bp far from the beginning of unmapped 17, aligns complementary
	glcPF TTTGGTTGGCGGAAAGGTTGC	52	61	8634bp far from the beginning of unmapped 17, aligns complementary
	glcPR1 GCAACCAAACCTTGAACTGGAG	50	58	8797bp far from the beginning of unmapped 17, aligns reverse complementary
	glcPF2	47	57	8798bp far from the beginning of unmapped

	TCCAGTTCAAGGTTTGGTTGC			17, aligns complementary
	glcP3RG AGTCAAGGCATCCACCATTCG	52	61	9593bp far from the beginning of unmapped 17, aligns reverse complementary
	Dg275-4-8Ra TAGCTTATCGCAGTCTAGTACG	45	55	60bp far from the beginning of unmapped 4, aligns reverse complementary
5	Dg275-Fa AGAAGCAGGTAGCAGTGTAGAG	50	57	726bp far from the beginning of mapped 1, aligns complementary
	Dg275-5Fb CAAAGCATTACCGATGCAAG	45	53	1657bp far from the beginning of mapped 1, aligns complementary
	Dg275-5Fc ACACCCTCAAAGAGCAATACC	47	56	2313bp far from the beginning of mapped 1, aligns complementary
	Dg275-5Fd AGCGAGATGAAGGAGTTTGC	50	56	3062bp far from the beginning of mapped 1, aligns complementary
	Dg275-5Fe ATGCTGATACCACTCTTGCAGG	50	58	3831bp far from the beginning of mapped 1, aligns complementary
	Dg275-5Ff CAGGAAATCCTTGCTCAAGATG	45	55	4290bp far from the beginning of mapped 1, aligns complementary
	Dg275-5Ra CCACCTACTTTGCTTGATCC	50	54	5249bp far from the beginning of mapped 1, aligns reverse complementary
7	Dg275-7Fb GCTTTGGAGTAAACGCTTTGC	47	57	307bp far from the beginning of mapped 15, aligns reverse complementary
	Dg275-7Fa CCTGCTGCTAAGATAATTGC	45	52	715bp far from the beginning of mapped 15, aligns complementary
	Dg275-7Fc TCCAAGTTCTGAGGCACGAATG	50	59	1437bp far from the beginning of mapped 15, aligns complementary
	Dg275-7Fd GCCCATGTTCTCTAAGTAAAGC	45	55	1970bp far from the beginning of mapped 15, aligns complementary
	Dg275-7Fe TATCTCTTTGATACGCCCAG	45	52	2795bp far from the beginning of mapped 15, aligns complementary
	Dg275-7Ff GCACTCTTACTACTTCCATCTG	45	54	3211bp far from the beginning of mapped 15, aligns complementary
	Dg275-Ra AGAGCGTTATCTTGGAGCTTGG	50	58	4642bp far from the beginning of mapped 15, aligns reverse complementary
	Dg275-7Rb TCTTCACCACTTCCTTTGTGC	47	57	4985bp far from the beginning of mapped 15, aligns complementary
*8	Dg275-8Fa TTCGATCCCGCTATTCTCCACC	54	60	2078bp far from the beginning of unmapped 22, aligns complementary
	Dg275-4-8Ra TAGCTTATCGCAGTCTAGTACG	45	55	60bp far from the beginning of unmapped 4, aligns reverse complementary

	glcpFB GGGCAACCCAATGTATAGAG	50	57	124bp far from the beginning of unmapped 4, aligns complementary
	glcP3RE ACGACCTTAGACTAGCACTTCC	50	57	1626bp far from the beginning of unmapped 4, aligns reverse complementary
	glcpFCCC AGGTGTGGTTAGCTTCGTACTAGG	50	59	1880bp far from the beginning of unmapped 4, aligns complementary
	glcPFdd TTAAAGCGGTACGCGAGCTGG	57	60	2738bp far from the beginning of unmapped 4, aligns complementary
*16	Dg275-8Fa TTCGATCCCGCTATTCTCCACC	54	60	2078bp far from the beginning of unmapped 22, aligns complementary
	EdaF TTGGACTTAACCACGAACCACC	59	59	316bp far from the beginning of unmapped 17, aligns complementary
	edaF1 ACTTGCGGAACTTCGTTTGC	50	58	681bp far from the beginning of unmapped 17, aligns complementary
	edaR1 TAGCAAACGAAGTTCCGCAAG	47	57	682bp far from the beginning of unmapped 17, aligns reverse complementary
	eddF ACACCCGCTATGTGTGATGG	58	58	2339bp far from the beginning of unmapped 17, aligns reverse complementary
	zwfR1 TGGATTGGTTAAGATGCACGAG	45	57	3199bp far from the beginning of unmapped 17, aligns reverse complementary
	zwfR ACGCCTGATTGCCTTCTATTGC	59	59	4033bp far from the beginning of unmapped 17, aligns reverse complementary
	zwfF1 GGGTCCTAAGGAAGTCTTTG	50	53	4164bp far from the beginning of unmapped 17, aligns complementary
	pglR AATGAAGCAACCGCCTTAGC	50	57	4444bp far from the beginning of unmapped 17, aligns reverse complementary
	pglF1 AAGCTGAAGAGTTCGAGTGG	50	56	4817bp far from the beginning of unmapped 17, aligns complementary
	glkR1 TTTGTCATCTGCACCCAATCC	47	57	5278bp far from the beginning of unmapped 17, aligns reverse complementary
	pgiF AAGAGTGGGTTCTTGGCTTGG	52	60	6451bp far from the beginning of unmapped 17, aligns complementary
	pgiR AACGCCCACCTACCCAATTC	55	60	6844bp far from the beginning of unmapped 17, aligns reverse complementary
	pgiR1 TCTTGCACTTCTTCGAGTGTTC	45	57	7356bp far from the beginning of unmapped 17, aligns reverse complementary
	glcPR2 TGGACCGATTGTGCTTCTTG	50	57	8220bp far from the beginning of unmapped 17, aligns reverse complementary
	glcPR1	50	58	8797bp far from the beginning of unmapped

	GCAACCAAACCTTGAACTGGAG			17, aligns reverse complementary
	glcP3RG AGTCAAGGCATCCACCATTCG	52	61	9593bp far from the beginning of unmapped 17, aligns reverse complementary
	Dg275-4-8Ra TAGCTTATCGCAGTCTAGTACG	45	55	60bp far from the beginning of unmapped 4, aligns reverse complementary

*rrnA+ glc locus and *rrnB+ glc* locus locate in gaps 16 and 4 respectively*, rrnC* locates in gap 8

Table 2.4: E- Dg349 primers used in amplification and primer walking. The primers covered from *ansB* to *eda* and from *glcP* to encoding putative type II restriction endonuclease.

Site of amplification	Primer Name & Primer (5'-->3')	% GC	Tm (°C)	Used for sequencing forward (F) or reverse (R) strand	Used for amplification or sequencing (primer walking)
ansB to *eda*	ed1F * AGTTGTTGCAGGTAGTGGTGCAG	52	61	F	Both
	ed1R TCTGCTGGCGGAATCACTATGC	54	61	R	Both
	ed1Fa ATTAGTGGCGCACGGGTGAG	60	61	F	Sequencing
	ed1FBb GGTAGTCCACGCCCTAAACG	60	58	F	Sequencing
	ed1FCc CGATCCCGCTATTCTCCACC	60	58	F	Sequencing
	ed1RBb CCGTGTCTCAGTTCCAGTGTG	57	59	R	Sequencing
	ed1Rc CCCAATTTCTTATCCGCATCAC	45	55	R	Sequencing
glcP to encoding putative type II restriction endonuclease	glcPF TTTGGTTGGCGGAAAGGTTGC	52	61	F	Both
	glcp3R** TACCGCTGTTGCCTTTGTGTATGC	50	62	R	Both
	glcPFb CAAGAGGGAAACAACCCAGAC	52	57	F	Sequencing
	glcpFc CACAGCACTCTGCTAACTCG	55	56	F	Sequencing
	glcPFd TTCGGTCCCTATCTGCCGTG	60	60	F	Sequencing
	glcpFee AGAAGCTAAGCGCATCGTGG	55	59	F	Sequencing
	glcPFdd TTAAAGCGGTACGCGAGCTGG	57	61	F	Sequencing
	glcP3Ra AGAGTAAATCGTGGTGGAAC	45	53	R	Sequencing

	glcp3Rb TCACTGGAACAAACTCGTGC	50	57	R	Sequencing
	glcp3Rc TTCACTTGTGCTTGATGATG	40	52	R	Sequencing
	glcp3Rdd CTACTTGGGTTGTCACGATAGG	50	56	R	Sequencing
	glcp3Ree AGCAAGGCGTTTCCACATCTG	52	59	R	Sequencing
	glcp3Rf TCCAACCGTTCTGAGCCAACC	57	61	R	Sequencing
	glcP3Rg ACTAACCCTACGATGACGAGC	52	57	R	Sequencing

* & ** were initially designed for the Dg275, they worked with Dg349 as well.

2.7.4　PCR amplification

Genes and the *rrn*/*glc* locus were amplified to confirm bacterial species, as preliminary identification of non-sequenced isolates or to confirm the presence of the *glc* locus. For amplification of the entire *rrn/*glc locus, CloneAMP ™ HiFi (Clontech) was used and for screening of the genes, DreamTaq DNA Polymerase (Thermo Scientific) was routinely used. PCR components and conditions of the enzymes are listed in **Table 2.5.**

Table 2.5: Components and conditions PCR reaction.

*	
PCR components	**PCR condition**
CloneAMP ™ HiFi Premix, 12.5μl	Initiation: 95°C
Forward primer, 7.5 pmol/μl, 0.75μl each of forward and reverse primers	Denaturation: 98°C, 10 sec
DNA template, 2.5μl of 20ng/μl	Annealing: primer TM -2 to 5, 15 sec
nH$_2$O, 8μl	Primer extension: 72°C, 1min/1kb
Total: 25 μl	30 cycles

**	
PCR components	**PCR condition**
nH$_2$O, 15.32μl	Initiation: 95°C, 2min
10X DreamTaq Buffer, 2.5μl	Denaturation: 95°C, 30 sec
dNTP Mix, 2 mM, 2.5μl	Annealing: primer TM -2 to 5, 30 sec
Forward primer,1 μl + reverse primer, 1μl	Primer extension: 72°C, 1min/1kb
DreamTaq DNA Polymerase, 0.18μl	Final, 72°C, 10min
DNA template, 2.5μl of 20ng/μl or 1/50 of lysed cells	30 cycles
Total: 25 μl	

* CloneAMP ™ HiFi Premix DNA Polymerase. ** DreamTaq DNA Polymerase

2.7.5 Purification of PCR products

In the initial studies, the QIAquick PCR Purification Kit (Qiagen) was used for purification of PCR products. GeneJET PCR purification kit (Therno Scientific) was subsequently used for all PCR products of 0.1-18 kb. The samples were purified based on the manufacturer's protocol with the exception that the elution buffer or nH_2O was preincubated at 55°C for 20min and 30-40µl of this was used to elute DNA. The products were routinely quantified by a Nanodrop spectrophotometer and were checked on an agarose gel. The samples were then stored at -20°C for DNA sequencing, if required.

2.7.6 DNA sequencing

DNA sequencing of the PCR products was performed commercially using MWG Eurofins (Germany) Sanger sequencing tube service. A total of 15µl of 5ng/µl PCR product (300-1000bp), or 10ng/µl for DNA samples >1,000 bp was sent with 15µl of 10pmol/µl of an appropriate sequencing primer. The primer walking method was followed in order to sequence PCR products more than 1kbp. For WGS, genomic DNA was purified and quantitated as above (**sections 2.7.1** and **2.7.2**). Purified DNA, (50-100µl at a concentration 30ng/µl, except Dg275, 19.18ng/µl), was sent to MicrobesNG, Birmingham University for Illumina genome sequencing with 2 x 250bp paired-end reads and >50 x coverage. Reads were trimmed with Trimmomatic and assembly was performed *de novo* using SPAdes and with *C. jejuni* NCTC11168 genome, as reference using BWAmem to assess quality. Contigs were reordered and reoriented according to the reference genome based on a MUMmer whole-genome alignment. Annotation of the draft genomes were performed using Prokka. The variant calling was performed using VarScan using the Dg275 closed and annotated genome, as reference.

2.8 Bioinformatics analysis

2.8.1 Applications using the PubMLST database and phylogenetic analysis

The draft genomes of the 137 Norway rat strains are publically available via BIGSdb and were accessed via the *Campylobacter jejuni/coli* PubMLST database (Jolley and Maiden, 2010) using ID or Dg numbers, as listed in **Table 2.1**. Illumina sequencing and assembly of all strains were performed on 2013-14 by the Sanger sequencing centre, Cambridge, as described by Cody et al. (2013). For all strains, total nucleotides ranged from 1446185 to 1816779 bp with an average size of 1662728 bp. The median number of contigs was 35, N50 contig value was 4 and N50 length 172554 bp. With the exception of Dg268 (71), Dg16 (162), Dg201 (68), Dg233 (61), Dg381 (77), Dg84 (104), Dg161 (168) and Dg234 (64) the number of contigs ranged from 18 to 56. Annotated sequence data was submitted to Bacterial Isolate Genome Sequence Database (BIGSdb) and curated by Dr Alison Cody and Helen Wimalarathna, Dept of Zoology, University of Oxford.

The BIGSdb database is a free accessible website containing a large number of bacterial sequences from draft to closed genome (Jolley and Maiden, 2010). The most important functions of the BIGSdb and PubMLST/campylobacter database that have been used throughout this study are described below.Gene-by-gene comparisons between strains, of individual selected genes, rMLST, wgMLST,cgMLST and ED types were performed using the genome comparator function with default settings of 70% identity over 50% of the alignment, a BLASTN word size 20. NCTC11168 was routinely used as reference strain. To identify and analyse core genomes of RG-1 and RG-2 groups, the closed, annotated gb files of Dg147 (Closed genome Dg147.gb) and Dg275 (Closed genome Dg275.gb) were uploaded as reference. Output files, recording presence, absence or truncation of genes, were stored in excel format for detailed analysis. This also recorded recorded allelic profile of each gene. To construct a phylogenetic network, data was downloaded as an archive file. The Nexus file was extracted from the archive file and imported into SplitsTree4 programme (version 4.13.1) to visualise and annotate a phylogeny network. This programme permitted editing of the ID numbers and colour-coding strains or clusters.

To construct a 16S rRNA phylogenetic tree, DNA sequence of NCTC11168 (AL111168.1) 16S rRNA (1513bp) was downloaded from NCBI database. This DNA was blasted against the draft genome of all required isolates in the PubMLST database using the BLAST function. The V3-5 region of each 16S rRNA gene was identified, in silico, using 16S rRNA gene primers: F, 5′-ACTCCTACGGGAGGCAGCAG-3′, and R, 5′-CCGTCAATTCCTTTGAGT-3′ and DNADynamo software (Kennedy, N. A. et al., 2014). The extracted 16S rRNA sequences of each isolates was saved with related ID as a text file and then uploaded into New Algorithms and Methods software to to make the phylogenetic tree (Chevenet, F. et al., 2006; Guindon, S. et al., 2010).

Single gene sequences from selected strains, or the translated product were downloaded from BIGSdb via PubMLST using the XMFA / concatenated FASTA format function, with alignment when required. To investigate surrounding sequence, ends of contigs and study the *glc* locus, DNA sequence were downloading with flanking sequence. Fas files containing all contigs for selected strains were exported from BIGSdb using Contigs function. These were used for assembly of Dg275 and Dg147 as below and genome comparisons with Blast atlas.

2.8.2   Contig assembly genome closure of Dg147 and Dg275 draft genomes

All draft genome contigs of the Dg147 (RG-1) and the Dg275 (RG-2) strains were downloaded from the BIGSdb database as an archive file contained in a FASTA file. The FASTA and gb files of the reference genome were downloaded from the freely accessible NCBI database (https://www.ncbi.nlm.nih.gov/genome/ 149). The draft genome of the test strain was aligned with the FASTA file of the reference genome using the CONTIGuator web server (http://combo.dbe.unifi.it/contiguator) with default parameters. The CONTIGuator output files were then aligned with the gb file of the reference strain using Artemis Comparison Tool (ACT) (Carver et al., 2008). This alignment gave a visual representation of the reference genome alignment with the test draft contigs (mapped contigs). Those contigs that did not align were named as unmapped contigs. The genome alignment helped to find gaps between the mapped contigs. Forward and reverse primers were designed from 72-1508 bp away from the end of both contigs. Initially, DreamTaq DNA Polymerase was used to amplify the gaps. If the gap was too large to amplify, the CloneAMP HiFi Premix DNA

polymerase was then used. The primer walking technique was followed for filling these gaps. DNAdynamo was used to assemble all sequenced DNA to cover each gap and DNAstar was then used to assemble the DNA sequence of the gaps with the mapped contigs to make a closed genome. The closed genomes were saved as fas files. To annotate the closed genome, the fas file was uploaded via PATRIC 3.5.11for automatic annotation of the bacterial genome with RASTtk programme (Aziz et al., 2008, Brettin et al., 2015), using taxonomy name *Campylobacter jejuni* taxonomy ID 197. The output gb file was used directly for genome comparisons (**see section 2.8.3 belo**w).

2.8.3   RG-1 or RG-2 genome comparisons

A core gene set for RG1 strains was identified using genome comparator of PubMLST/*Campylobacter* database. The annotated closed genome of Dg147 (Closed genome Dg147.gb) was uploaded as reference for gene-by-gene comparison with the draft genomes of the other 29 RG-1 strains in the database (**Table 2.1**), using the default parameters, a minimum of 70% identity with a minimum of 50% alignment. Presence and absence of genes were recorded in an excel file (**see supplemetary data**). Any genes that were absent or incomplete in at least one of the RG-1 strains were counted as accessory genes. To confirm additional accessory genes of RG-1 group that were absent in the Dg147 genome, the draft genomes of all RG-1 strains were compared using NCTC11168 as reference.

To compare cgMLST genes within RG-1, IDs of all RG-1 strains were pasted in the scheme box of Genome Comparator of the PubMLST database and option *C. jejuni/ C. coli* cgMLST v1.0 selected. The output for all 1343 cgMLST genes were extracted as an excel file. cgMLST genes are identified with CAMP number. These were correlated with the Dg147 annotated locus numbers by combining both in the excel file. This also double checked annotation of the Dg147 genome with a manual comparison of each gene between RG-1 cgMLST and Dg147. Genes were identified as cgMLST, or accessory genes if incomplete or absent in one or more strain. To identify any RG-1 core genes, that are absent from cgMLST defined genes but present in some generalist strains, the closed genome of Dg147 was aligned with the draft genomes of selected strains of sequence types ST21CC, ST45CC, ST42CC using genome comparator. The RG1 core genes were identified within the Dg147 genes. The

cgMLST defined genes were then excluded. The remaining 247 RG-1 core genes were then compared gene by gene, with conserved genes in the ST21CC, ST45CC, ST42CC genomes. The identified 35 common genes were considered to be potentially important in colonization. GView Server https://server.gview.ca (Petkau et al., 2010) was used for a visual comparison of genomes to create Blast atlas figures.

2.9 Statistical analysis

The average of technical replicates was routinely performed and used as one mean. At least three biological replicates were routinely used to obtain average with standard error for example in growth studies and nutrient analysis. Comparison of the means of two groups was made using the Independent Samples T test (Two-samples T test) with the current version of SPSS 21 programme. Results from the chicken trial were analysed using the Mann-Whitney U Test (One- tailed) (http://www.socscistatistics.com/tests/mannwhitney/Default2.aspx).

Chapter 3

Characterisation of the novel

RG-1 Group

3.1 Introduction

Over 3000 sequence types of *Campylobacter* deposited in the BIGSdb database were analysed by Colles and Maiden (2012). The predominant clonal complexes in the database were ST-21, ST-45, ST-257, ST-48, and ST-828. The isolates were from different sources including human, chicken, beef, lamb, bird, environmental waters and pig sources. A six year study on 3300 clinical *Campylobacter* isolates from Oxfordshire, where the farm-associated rat strains investigated in this study were isolated, provided evidence that some clonal complexes showed seasonal variation. The clonal complexes (ST 353 and ST 403) were predominant during winter, whilst clonal complexes (ST 45, ST 283, and ST 42) were predominant during summer (Cody et al., 2012). Strains belonging to the generalist ST 21 and ST 45 clonal complexes have been identified in a wide range of hosts and these CC groups together with *C. coli* ST 828 clonal complex are considered to be comprised of generalist *Campylobacter* strains. The ability of these strains to transmit from one host to another indicates possession of genetic material enhancing adaptation to different hosts and environments (Dearlove et al., 2016, Kwan et al., 2008, Gripp et al., 2011). Other clonal complexes have been identified as more closely associated with specific hosts. For example, *C. jejuni* strains of ST 61 and ST 42 clonal complexes are frequently isolated from cattle and considered to have genotypes adapted for cattle colonization (Jolley and Maiden, 2010, French et al., 2009), whereas ST 177 clonal complex is recovered from wild birds (Colles and Maiden, 2012, French et al., 2009). In addition, as identified from phylogenetic analysis of both MLST and rMLST, a clearly distinct clade of 30 *C. jejuni* strains was readily identifiable in the 137 sequenced genomes of Norway rat isolates (**see Figure 1.11** in literature review, MacIntyre, Prescott and McCartney, unpublished data). This group had been identified as *C. jejuni* by Matrix Assisted Laser Desorption Ionization - Time of Flight mass spectrometry (MALDI-TOF) and was designated *C. jejuni* 'rat group 1', RG-1.

ST numbers of most of RG-1 group have been assigned as 6562, 6564, 5129, 7278 and 6561, except Dgs (289, 347, 348, 295, 164, 105, 157) of cluster 2 and Dgs (36, 301, 179, 276) of cluster 3 (**Figure 3.17** at below). ST numbers of these have not been assigned because of

variability of the *pgm* (*glmM*) MLST locus. Based on MLST typing, and phylogenetic analysis there was no record of this clade of *C. jejuni* in the extensive BIGSdb database. This could be related to geographical niche evolution, although these strains were isolated from 8 different farms in SE England over 3 years. An alternative possibility might be that the RG-1 group has evolved as environmental or host adapted strains. The BIGSdb database is heavily biased towards human and chicken isolates, indicating that RG-1 strains are at least not commonly associated with these hosts. The RG-1 STs are also not present within the large collection of 1413 wild bird isolates. As the RG-1 isolates were all from farm associated Norway rats, it has been considered that these strains may have evolved optimal colonisation of Norway rats.

Therefore, the aim of this first chapter was to initiate characterization of this novel clade of *C. jejuni* strains. Laboratory based studies focused on several common attributes of *C. jejuni*, cell morphology, motility, properties for biofilm formation and basic growth and metabolic properties. As an initial step in identifying genes that might be linked to host specificity of the RG-1 strains, a set of core and accessory RG-1 genes were identified by characterising their genomic contents, and comparing this genomic data with a set of 1343 human campylobacteriosis cgMLST core genes (Cody et al., 2017). Linking this information with the RG-1 phenotypic properties will assist in selecting the RG-1 strains most appropriate for *Galleria mellonella* and chicken *in vivo* studies.

Results and Discussion

3.2   Confirmation of RG-1 isolates as *Campylobacter jejuni*

As RG-1 isolates were identified as a very clearly distinct clade based on rMLST and WGS analysis, it was important to firstly determine how this relates to other *Campylobacter* species. The 16S rRNA sequence analysis is a classical and important method for the identification of various bacteria to genus and species level, including *Campylobacter* spp. (Gorkiewicz et al., 2003, Hansson et al., 2008). However, the method cannot differentiate two main species *C. jejuni* and *C. coli* strains.   There are six different variable regions on the sequence. V3-5 region has been used to differentiate *C.jejuni* and *C. coli* from other species (Gorkiewicz et al., 2003). This region of the 16S rRNA gene has also been used in microbiome analysis (Kennedy et al., 2014). *C. jejuni* has three *rrn* loci (*rrn* A-C) (Taylor et al., 1992). It has been demonstrated that all three copies of 16S rRNA genes are identical in the NCTC11168 reference strain (AL111168), and in each of 45 additional *C. jejuni* strains studied by Hansson et al. (2008).

In this study, the whole DNA sequence of 16S rRNA from strain NCTC11168 was Blasted with the RG-1 and RG-2 groups in the BIGSdb database to extract DNA from the draft genome (Jolley and Maiden, 2010). DNAdynamo was used to find any polymorphisms in V3-5 region of the 16S rRNA and to construct a phylogenetic tree based on the V3-V5 region. Phylogeny of the 16S rRNA shows that RG-1 strains cluster on a different branch to *C. lari*, *C. hyointestinalis* and *C. fectus* species (**Figure 3.1**, **A**), but this analysis is not able to differentiate any of the RG-1 or RG-2 strains from *C. jejuni* subsp. *jejuni, C. jejuni* subsp. *doylei* and *C. coli.* Representative strains of RG-1 and RG-2 isolates are grouped together in the major cluster of the phylogeny. Sequence analysis showed that there is no polymorphism between the V3-V5 sequences of NCTC11168 and that of the *C. coli* strains tested (**Table 3.1**). On the other hand, there is a single mismatch (position 814) compared to *C. jejuni* subsp *jejuni* 81116 and 2 mismatches compared to *C. jejuni* subsp. *jejuni* 81-176. *C. jejuni* subsp. *doylei* (strain RM4099) had 1-3 mismatches compared to the three *C. jejuni* subsp. *jejuni* reference strains. One group of RG-1 strains, Dg61, Dg62, Dg301, Dg179 and Dg276, had the identical V3-V5 sequence to *C. jejuni* subsp *jejuni* 81116, but two mismatches compared

to all other RG-1 strains. Distinct geographic niche of this smaller group of five RG-1 strains might explain the different 16S rRNA sequence, as these were isolated from farms K and X, whereas the others were isolated from farms C, E, G, H, I and X. The other RG-1 strains (83.3% of strains) were closest to NCTC11168 with only a single mismatch, at position 723, in the V3-V5 region. The V3-V5 region in all 23 of the RG-2 strains analysed, including Dg95 and Dg275, was identical to NCTC11168.

These data confirm assignment of the RG-1 and RG-2 strains to the *C. jejuni/coli* group and is consistent with reports that phylogenetic analysis based on 16S rRNA sequence is not always reliable for differentiation between *C. jejuni* and *C. coli* (Hansson et al., 2008, Muralidharan et al., 2017). In addition to 16S rRNA, WGS phylogeny was also analysed to compare RG-1 and RG-2 strains to related and reference strains (**Figure 3.1, B**). This phylogeny highlights the clustering of RG-1 strains on a single distinct phylogenic clade, but the branch point of this clade is closer to the three *C. jejuni* reference strains than *C. jejuni* subsp. *doylei.* These data are all consistent with the MALDI-TOF MS assignment of the RG-1 strains to *C. jejuni* species.
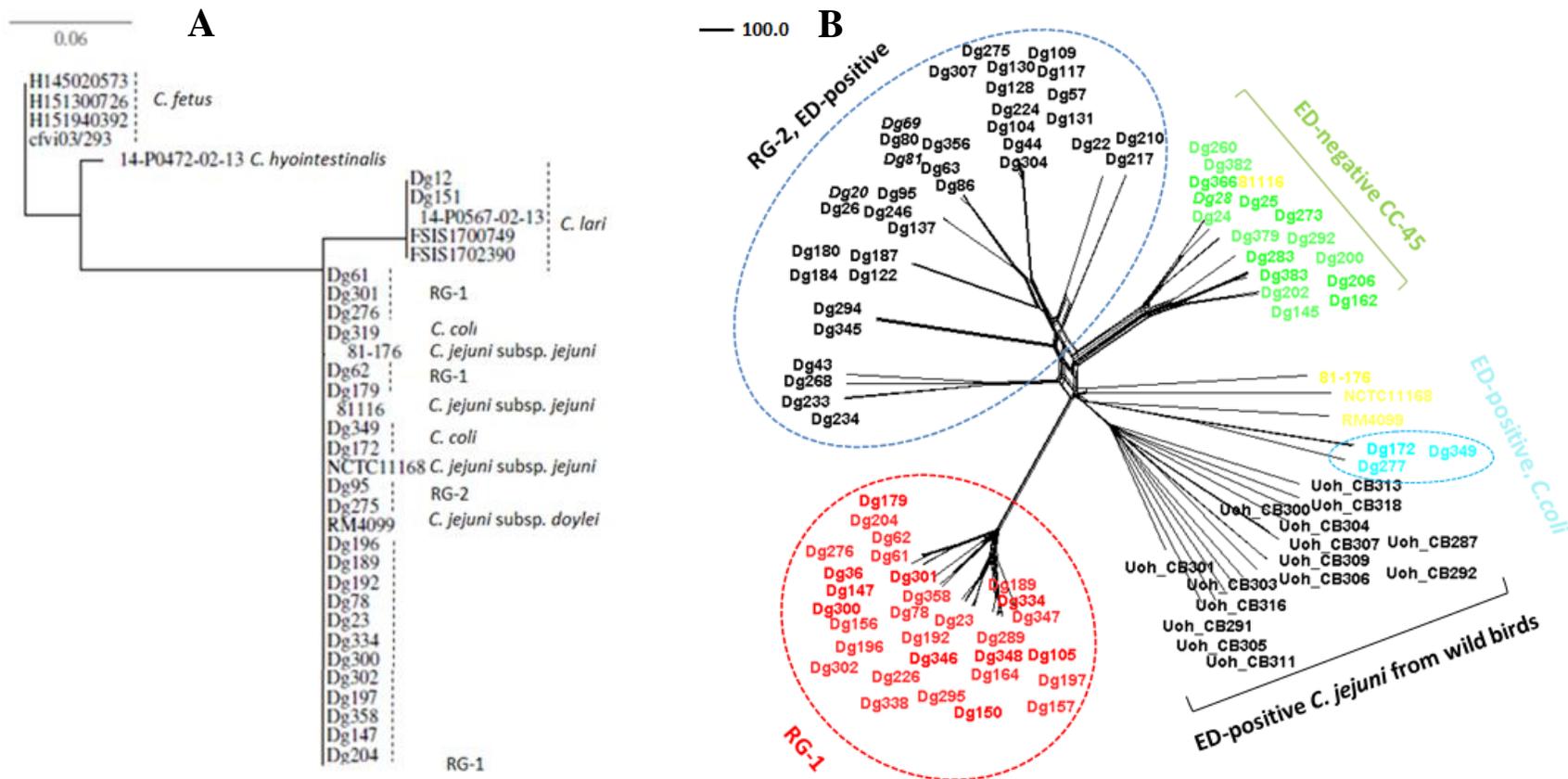
Figure 3.1: Phylogenetic comparison of new *C. jejuni* clonal complexes from Norway rats. (**A**) 16S rRNA dendogram. The complete 16S rRNA for all strains was downloaded from the BIGSdb database using the 16S rRNA of NCTC11168 as BLAST. After that, the V3-5 region of each 16S rRNA gene was identified, *in silico* (Kennedy et al., 2014) and then used to make a phylogenetic tree (Guindon et al., 2010, Chevenet et al., 2006). (**B**) WGS phylogenetic network of RG-1 with different strains of *C. jejuni* and *C. coli*. The WGS phylogenetic network of the genome comparator of BIGSdb were compared among selected strains. NCTC11168, 81-176 and 81116 *C. jejuni* subsp. *jejuni*, and RM4099 *C.jejuni* subsp. *doylei* (yellow). SplitsTree4 was used to visualise the results of the phylogenetic network (Huson and Bryant, 2006). The branch length of the tree is proportional to the number of substitutions per site.

Table 3.1: Polymorphisms in part of the 16S rRNA sequence of RG-1 and RG-2 compared to other *Campylobacter spp.* Extracted V3-5 sequences of 16S rRNA from the genome data of BIGSdb rat associated *Campylobacters* (Dg isolates) and selected reference strains were aligned to identify positions of polymorphism

Group	Species	* Nucleotide positions			
		644	723	814	821
NCTC11168	*C. jejuni* subsp. *jejuni*	C	C	G	T
81116	*C. jejuni* subsp. *jejuni*	C	C	A	T
81-176	*C. jejuni* subsp. *jejuni*	T	C	A	T
All RG-1 except Dg61, Dg62, Dg301, Dg179, Dg276	*C. jejuni* subsp. *jejuni*	C	T	G	T
Dg61, Dg62, Dg301, Dg179 and Dg276, RG-1	*C. jejuni* subsp. *jejuni*	C	C	A	T
RG-2 (main branch 23 strains)	*C. jejuni* subsp. *jejuni*	C	C	G	T
RM4099	*C. jejuni* subsp. *doylei*	C	C	G	C
Dg319, Dg349, Dg172	*C. coli*	C	C	G	T
Dg151, Dg12, FSIS1700749, FSIS1700749	*C.lari*	12 polymorphisms			
H151940392, H145020573, cfvi03/293	*C.fetus*	28 polymorphisms			
14-P0567-02-13	*C. hyointestinalis*	12 polymorphisms			

* All nucleotide mismatches, compared to NCTC11168 16S rRNA sequence, for *C. jejuni* and *C coli* strains are shown.

To further investigate the relationship of RG-1 isolates to other classic *C. jejuni* strains, both Analytical Profile Index (API) and PCR for classic target genes was performed. *Campylobacter* does not normally utilise sugars and these isolates grew poorly. API was used to specifically assess hippurate hydrolysis (HIP) and nitrate reduction (NIT). Eight of the RG-1 isolates (Dg192, Dg189, Dg196, Dg78, Dg334, Dg23, Dg346, Dg358) and two of the classic isolates Dg145 (ST-45 clonal complex) and Dg153 (ST-42 clonal complex) were tested (examples are shown in (**Figure 3.2, A**). The positive NIT reaction, with all RG-1 and Dg153 strains is consistent with positive reactions of *C. jejuni* subsp. *jejuni* and was distinct from the NIT negative reaction for *C. jejuni* subsp. *doylei* strains (Analytical Profile Index, (Martiny et al., 2011). Similarly HIP, which is used as a *C. jejuni* species marker (Wang et al., 2002) was positive for all strains. Genomic information from the draft genomes confirmed that all RG-1 strains harbour the *hipO* gene (hippurate hydrolysis) and the *nap* gene, a subspecies *jejuni* marker (**Figure 3.2, B**) (Miller et al., 2007, Steele and Owen, 1988). In contrast, the draft genome sequences confirm that RG-1 strains do not have propionate-

CoA ligase (*prpE*) nor 2-methyl-citrate synthase (*prpC*), two markers that have been used in *C. coli* identification (Wagley et al., 2014). The absence of these genes in RG-1 was also confirmed experimentally in a Biolog phenotypic microarray (**see below Figure 3.14**). Neither Dg147 nor Dg300 RG-1 strains are able to use propionic acid as a sole carbon source. The above genotypic and phenotypic markers provide significant evidence to conclude that RG-1 strains are *C. jejuni* with many properties of subsp. *jejuni*, and are clearly distinct from *C. jejuni* subsp *doylei* and *C. coli.*

A

Dg192 + +   Dg153 + +

B

Locus	RM4099	Dg196	Dg189	Dg153	Dg192	Dg78	Dg23	Dg334	Dg349
CAMP0715 (*napA*)	X	799	799	21	799	799	799	801	275
CAMP0908 (*hipO*)	85	151	151	6	151	151	151	151	X
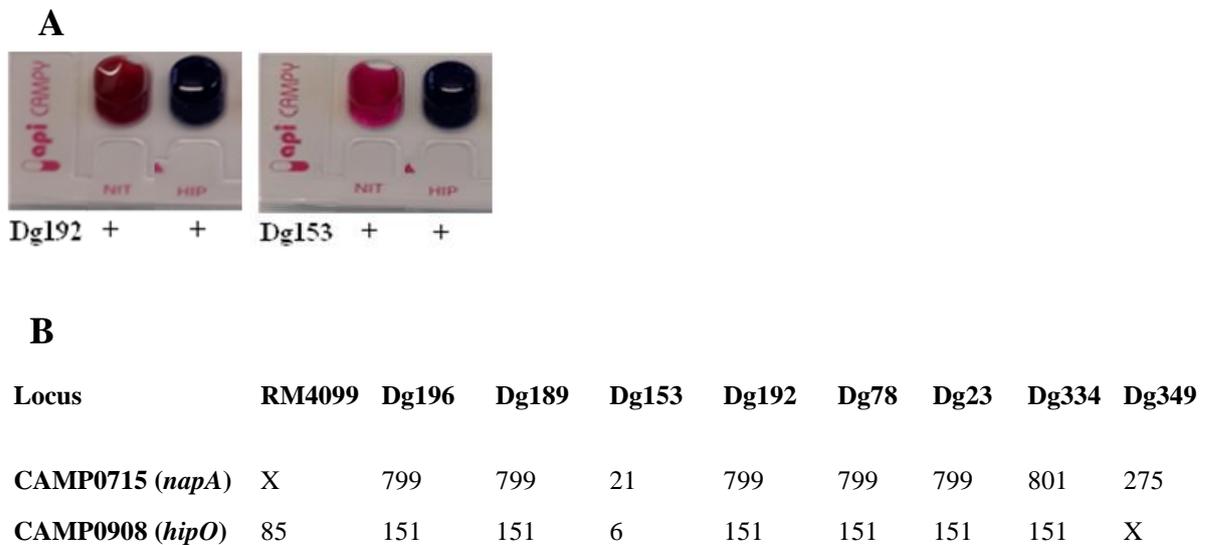
Figure 3.2: API CAMPY test and *napA* and *hipO* gene comparison. (A) API, NIT: potassium nitrate, positive (+) colour, pink/red, an indicator of subsp. *jejuni*. HIP: sodium hippurate, positive (+) colour is violet, an indicator of *C. jejuni* subsp. *jejuni* 1, *jejuni* 2 and subsp. *doylei* (Wang et al., 2002). Dg192 (RG-1) and Dg153 (ST-42 clonal complex). **(B)** Genome comparison of *napA* (nitrate reductase) and *hipO* (hippuricase) of strains as indicated. RM4099, *C. jejuni* subsp. *doylei*, ST 1845 from human blood culture (Jolley and Maiden, 2010); Dg349, *C. coli* (*hipO* negative); all other isolates are RG-1 strains. Numbers, allele numbers for *napA* and *hipO*; X, missing allele.

3.1   Motility and biofilm characteristics of the RG-1 strains

3.1.1   Motility of the RG-1 strains

Motility has been shown to be critical to pathogenicity, has a basic role in host cell invasion and adhesion, and helps cells to survive in different environments (Ramos et al., 2004). This contributes substantially as one factor that contributes to *C. jejuni* colonization of animals, and subsequent invasion of the intestinal mucosa and disease in susceptible hosts (Yao et al., 1994). Motility genes may be switched off, when bacterial samples undergo subculture and are stored for long periods. Therefore, to enhance growth and characterise motility, all strains were monitored for motility following a round of growth on 0.4% MHA plates, followed by subculturing onto BA. From the BA plates, the strains were stocked as Round Motility 1 (RM-1) stocks (**Figure 3.3, A**). An additional round of motility was completed and stocked as Round Motility 2 (RM-2) (**Figure 3.3, B**). RM-1 glycerol stocks were used as starting inoculum for all subsequent studies.

A

B
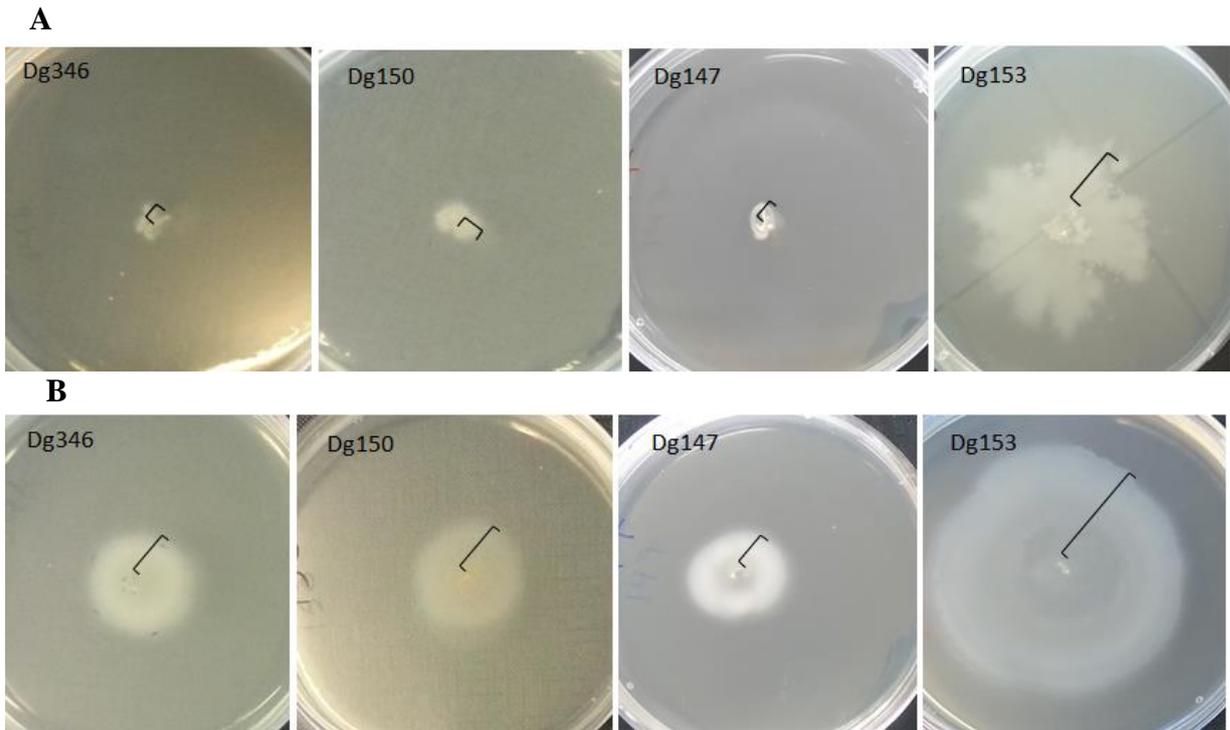
Figure 3.3: Motility on MHA plates. A standardised culture (5 µl) was stabbed into the center of 0.4% MHA of (**A**) the first round motility and (**B**) the second round motility. Plates were incubated for 72h at 37°C. The radius of motility was measured from outside the general growth area as indicated by the black brackets. Stabs of motile culture close to the rim of motility were taken for glycerol stocking of RM-1 cultures and the preparation of inoculum for second round motility plates. Dg346, Dg150 and Dg147 belong to RG-1 and Dg153 belongs to ST-42 clonal complex.

Approximately half of the RG-1 strains' motility had been enhanced after exposure to MHA motility medium, typically increasing from essentially non-motile (4mm) to 14mm. However, the level of motility in all RG-1 strains was significantly less than that observed in Dg153, examples included Dg147 (P≤0.001), Dg300 (P=0.003) and Dg226 (P=0.039) (**Figure 3.4, A**). The other 13 strains of RG-1 (43%) initially showed negligible motility, with no substantial enhancement of motility compared to RM-2. Despite the fact that the classic isolates were initially mostly quite motile in the first round, producing a motility radius from 6mm for Dg209 (ST 42 clonal complex) and 23mm for Dg162, most of these isolates also improved their motility in the second round, producing motility radii of 18mm for NCTC11168 to 34mm for Dg161. Only one classic isolate showed negligible motility, Dg361 and for two others Dg16 and Dg162 motility did not improve in the second round (**Figure 3.4, B**).
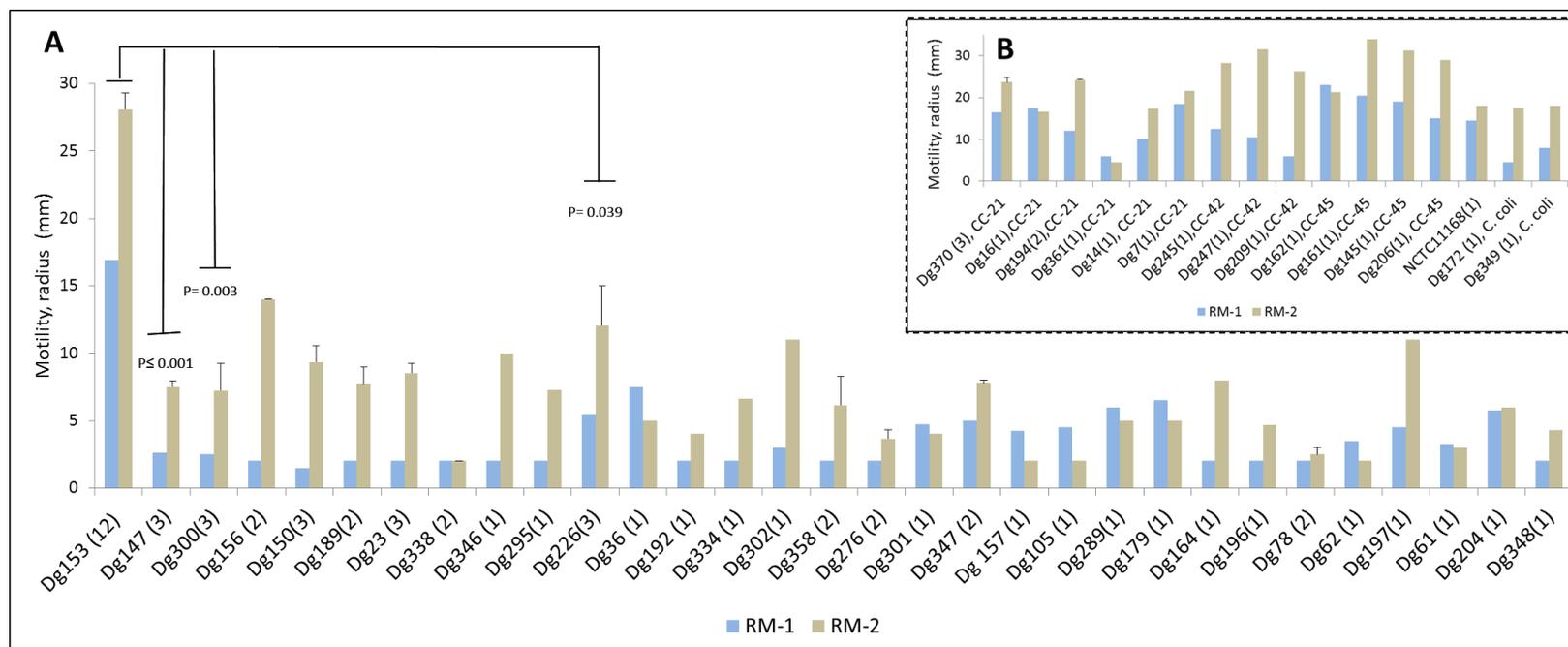
Figure 3.4: The first (RM-1) and second (RM-2) round motility. (**A**) RG-1 and Dg153, and (**B**) *C. jejuni* and *C. coli* Norway rat isolates belonging to classic clonal complexes. The rounds of motility testing were performed as described in **Figure 3.3** and in methods chapter, **section 2.2.4**. Triplicate technical repeats were performed on all second round tests. Numbers in brackets indicate number of biological repeats for RM-2 tests.

To confirm the cell morphology and possession of flagella, the cells were tested under TEM. Seven RG-1 strains (Dg147, Dg61, D164, Dg179, Dg300, Dg276 , Dg347), two RG-2 strains (Dg95, Dg275) and 4 strains belonging to classic ST groups (Dg153, Dg200, Dg194 and Dg370) were grown in liquid culture (MHB) with slow shaking for 22h at 37°C, before cells were fixed in glutaraldehyde and stained with uranyl acetate, as described in **section 2.5**. The TEM images showed that the cell morphology of RG-1 and RG-2 strains were similar to typical *Campylobacter* cell morphology, although the size of the cells varied and cells of RG-2 were straighter (**Figure 3.5**). The cell size of the *C. jejuni* strains was 0.4μm wide in Dg147, Dg164 and Dg275, 1μm wide in Dg179, and 2.3μm long in Dg147 to 5.7μm long in Dg276. The amplitude of cells also varied from 0.45μm in Dg275 to 1.1μm in Dg179 (**Table 3.2**). The bacterial cell morphology, the TEM images have confirmed that both groups possess bipolar flagella, as is found in typical *C. jejuni* cells (Balaban and Hendrixson, 2011). Thus, these images confirm typical *C. jejuni* cell morphology of the RG-1 strains, similar to that of strains belonging to classic CC groups.
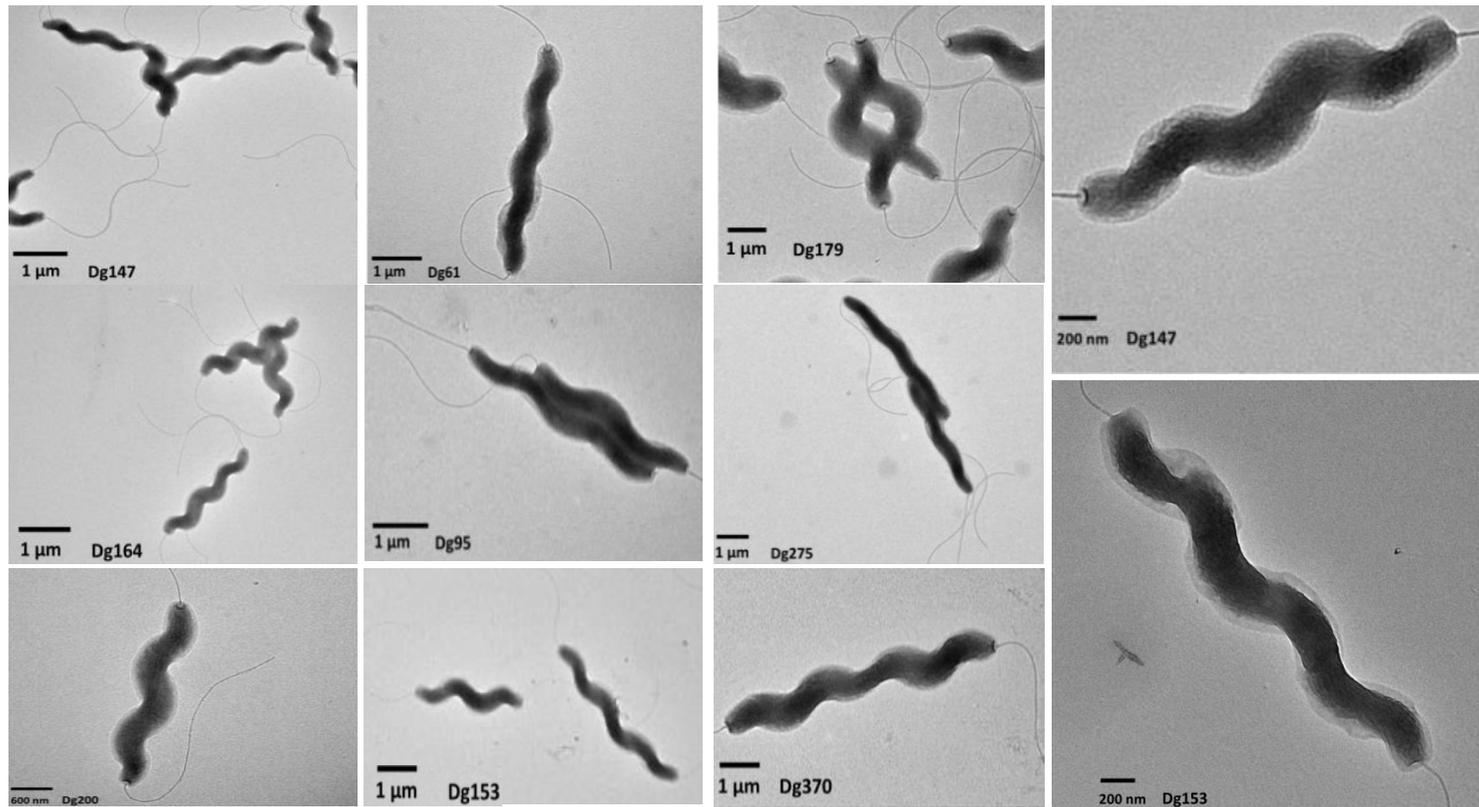
Figure 3.5: TEM micrographs of negatively stained rat-associated *C. jejuni* cells. Each *Campylobacter* strain was grown in 10ml MHB at 37°C for 22h with shaking at 150rpm. Cells were then recovered and fixed in 2.5% (v/v) glutaraldehyde in 0.1M cacodylate buffer pH 7.0, and stained with 1% uranyl acetate. The images were taken by TEM. Strains belong to groups as follows: RG-1 (Dg147, Dg61, Dg164, Dg179), RG-2 (Dg275, Dg95), Dg200 ST-45, Dg153 ST-42 clonal complexes, and ST-50 Dg370 ST-21 clonal complex. Note, single bipolar flagella.

Table 3.2: Morphological characteristics of *Campylobacter jejuni* cells.

Isolates	ST, group	* Cell length (µm)	* Cell width (µm)	* Amplitude of spiral (µm)	n
Dg153	42	3.5±0.4	0.5±0.06	0.7±0.01	10
Dg370	50	4.1±0.2	0.7±0.04	0.8±0.06	10
Dg147	6265, RG-1	2.3±0.1	0.4±0.01	0.48±0.01	30
Dg164	-, RG-1	2.4±0.2	0.4±0.03	0.6±0.05	18
Dg179	-, RG-1	5±0.3	1±0.03	1.1±0.02	7
Dg300	6562, RG-1	4.4±0.2	0.6±0.02	0.8±0.01	10
Dg276	-, RG-1	5.4±0.3	0.67±0.03	0.94±0.02	10
Dg200	45	3.6±0.3	0.6±0.03	0.7±0.03	4
Dg275	7259, RG-2	4.4±0.2	0.4±0.02	0.45±0.02	16

* Mean cell length, width and spiral amplitude ± SE (standard error) of cells imaged as shown in **Figure 3.5**, were measured using ImageJ software (Rueden et al., 2017). (–) indicates unassigned ST (sequence type), n, number of cells measured.

3.1.2 Biofilm formation

Biofilm formation is considered to be one of the contributing factors to successful transmission of *C. jejuni* through the food chain (Bronowski et al., 2014). It has been reasoned that the flagella are required to form biofilm (Kim et al., 2015b, Kalmokoff et al., 2006). Therefore, the RM-1 stocked cultures were used to test the ability of the RG-1 isolates to form biofilm in both a microtitre plate assay and on glass fibre filters monitored by SEM.

3.1.2.1 Biofilm formation-crystal violet based assay

Crystal violet binding is often used as a rapid assay to quantitate bacteria bound to polystyrene plates as the first indicator of microcolony and biofilm formation (Kalmokoff et al., 2006). This assay was used as an intial screening step for the biofilm potential of the bank of *C. jejuni* isolates. The level of crystal violet binding varied substantially among the tested strains, with no grouping of isolates based on either the ST number, motility or growth property. The standard error between biological repeats for CV binding was much higher than compared to growth. Despite this, some isolates gave a strong positive result. RG-1 strains Dg192, Dg334, Dg358, and Dg276 consistently resulted in high levels of CV binding

indicating strong sessile structures >OD_{590} 0.8, while Dg150 and Dg23 appeared to be non-biofilm formers (**Figure 3.6**). This data showed no linear relationship with bacterial growth (total bacterial cells) and motility to biofilm. Many of the poor growers (OD_{600} 0.12 or less) showed a high level of CV binding (0.4-1.2 OD_{590}). In addition, the very poor grower, Dg197 strain, with an OD_{600} of 0.06 still appeared to produce a high level of biofilm, with CV binding levels around 0.7 OD_{590}.
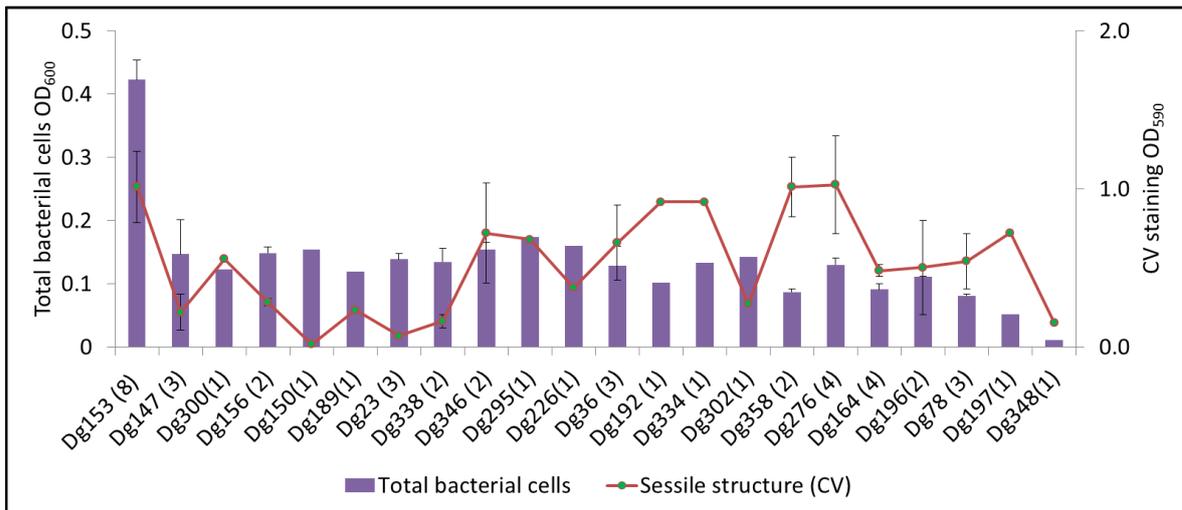
Figure 3.6: Relation of sessile structure (CV staining biofilm) with growth of RG-1. RM-1 stocks were used to prepare a standard inoculum. Two polystyrene 96-well plates containing MHB were inoculated and grown for 72h at 37 °C. One plate was used to monitor growth under microtiter growth assay conditions following re-suspension of cells (**see section 2.2.6**), and the other plate was used for CV staining with only gentle rocking to remove settled bacteria (**see section 2.4.2**). OD_{600}, growth and OD_{590}, CV absorbance. RG-1 strains and Dg153 ST-42 clonal complex were tested. Numbers in brackets refer to the number of independent experiments, with triplicate wells inoculated in each experiment/strain.

3.1.2.2   Biofilm property on glass fiber filters

To further investigate the ability of *C. jejuni* strains to form different biofilm structures, selected strains were grown on glass fiber filters in MHB medium and then analysed under SEM. MHB medium was inoculated with the proposed good sessile structure forming bacteria Dg153 (ST-42) or Dg36 (RG-1), or with poor sessile structure formers Dg189 or Dg62 (RG-1 strains) (**Figure 3.7**). The MHB broth was exchanged with fresh medium at 2, 3, and 4 days. At 4 days, filters were removed and processed for SEM as described in **section 2.4.3**. Filters were initially examined at low magnification to assess biofilm structure and the extent of random spread versus clustering of cells, prior to more detailed examination of microcolony/biofilm formation at higher magnification. Dg36, the representative RG-1 strain that gave a positive reaction in the CV binding assay produced no microcolony of biofilm structures in this assay. Flagellated cells associated with the filter were all randomly distributed. One single cluster of cells were observed in Dg189, but many of the cells were transformed to coccoid form (**Figure 3.7, A, B, C**). This transformation, between coccoid and filamentous morphology of *C. jejuni* has been shown to be due to incubation for prolonged periods under unfavorable conditions (Griffiths, 1993). However, Dg153 covered much of the filter in a typical dense microcolony/ biofilm (**Figure 3.7 D, E**) (Buck et al., 1983). At higher magnification, it was evident that these bacterial cells were still growing and healthy, as most had retained spiral morphology. This is consistent with the ability of Dg153 to form sessile structures in the CV assay and to clump at the bottom of the 96-well plates in MHB cultures. Formation of a microcolony or biofilm structure was convincing only with Dg153, none of the tested RG-1 strains appeared to form biofilm under these conditions. Conditions were not ideal for growth of RG-1 strains as indicated by coccoid formation and low OD_{600}. Optimal growth conditions might enhance the ability of these strains to form biofilm.
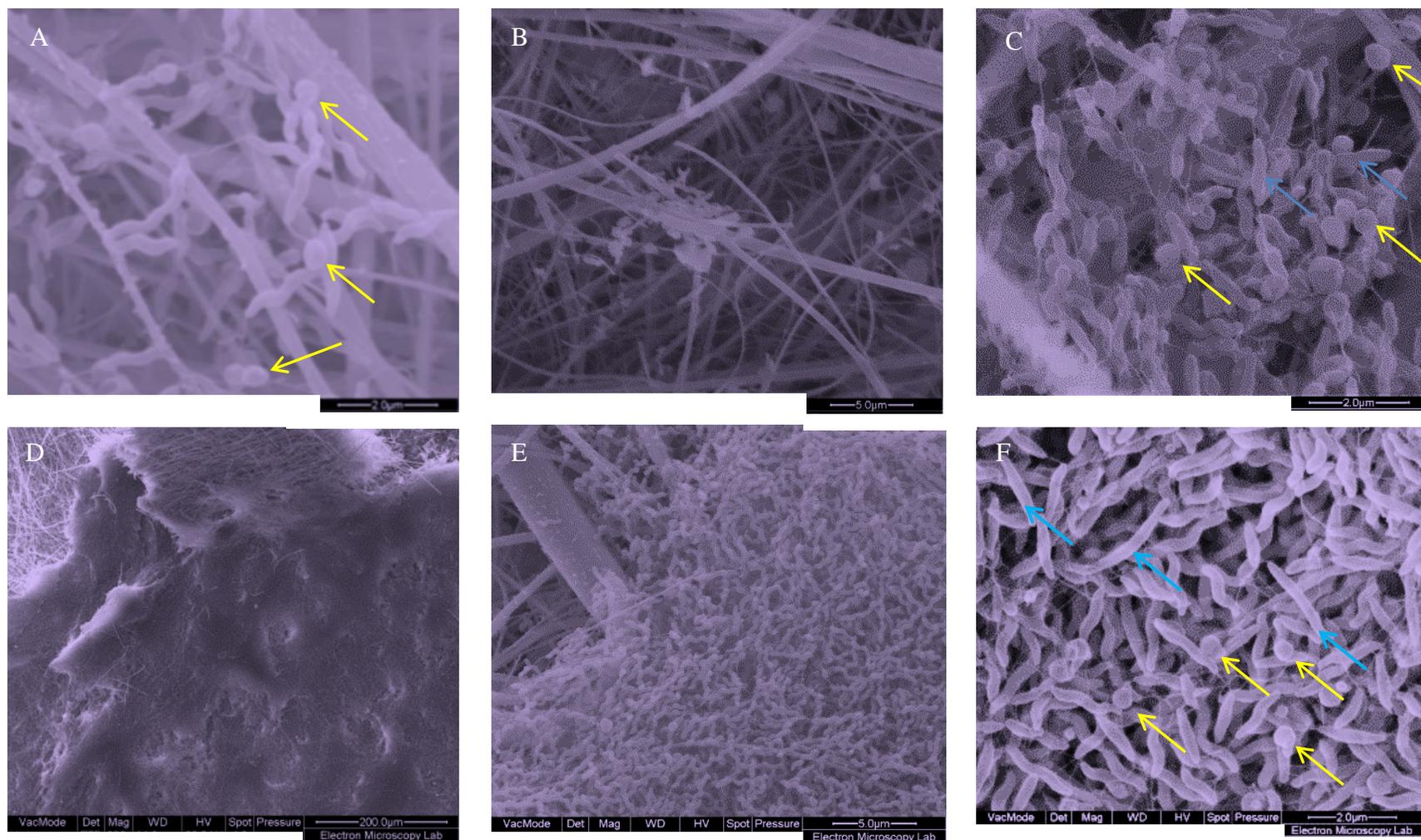
Figure 3.7: SEM of biofilm formation by Dg153 but not RG-1 strains. Cells were grown in MHB with a glass fibre filter for 4 days, with the medium refreshed after 2, 3, and 4 days. Filters were processed for SEM as described in **section 2.4.3**. **A;** Dg36, spiral and coccoid cells, low numbers, randomly distrubuted (30,000x) filter, **B;** Dg62, very low numbers, randomly distributed (10,000x) **C;** Dg189 one aggregate of cells found on filter (30,000x) **D-F; Dg153** D, (1981x), **E;** (10,000x) and **F** (40,000x): microcolony over much of filter. Yellow arrows indicate coccoid cells; blue arrows indicate rod or filamentous shape cells. The images are representative of three independent experiments.

3.1.3   Growth adaptation of RG-1 strains in laboratory conditions

Freshly isolated bacterial strains, which have adapted to available nutritional compounds in their hosts, do not always grow well initially in laboratory media and need to adapt to growth in this new environment. Strains used in this study, were isolated from Norway rats between 2010 to 2012 and had been grown primarily in mCCDA and BA prior to stocking. Growth of these original isolates in MHB was compared to growth of the RM-1 stock cultures, which had already been exposed to MH nutrients in the first round of the motility assay (**Figure 3.3**, **A** at above). A microtitre plate growth assay was used, as described in **section 2.4.6**. The results, as shown in **Figure 3.8**, show extremely poor growth for most of the original stock cultures, especially RG-1 strains. On the contrary, RM-1 results show a dramatic enhancement of growth, many folds higher at both 48h and 72h time points. Thus, even a single round of growth on MHA led to adaptation of the RG-1 strains to this medium. For instance, the cell density of Dg147 when inoculated from the original stock was negligible, 0.01 and 0.02 OD_{600} at 48h and 72h, respectively, but increased to 0.26 OD_{600} by 48h, when inoculated with the RM-1 stock. While not as marked, the cell density of strains belonging to classic CC groups also increased with pre-subculture. Thus, Dg153 inoculated from the original culture grew to 0.2 and 0.33 OD_{600} at 48h and 72h, respectively, but to 0.5 and 0.41 OD_{600} at 48h and 72h, respectively, when inoculated with the RM-1 stock. This growth improvement highlighted adaptation of strains to MH medium following a single round of growth on a motility plate. Therefore, for the subsequent assays RM-1 stocks were used.
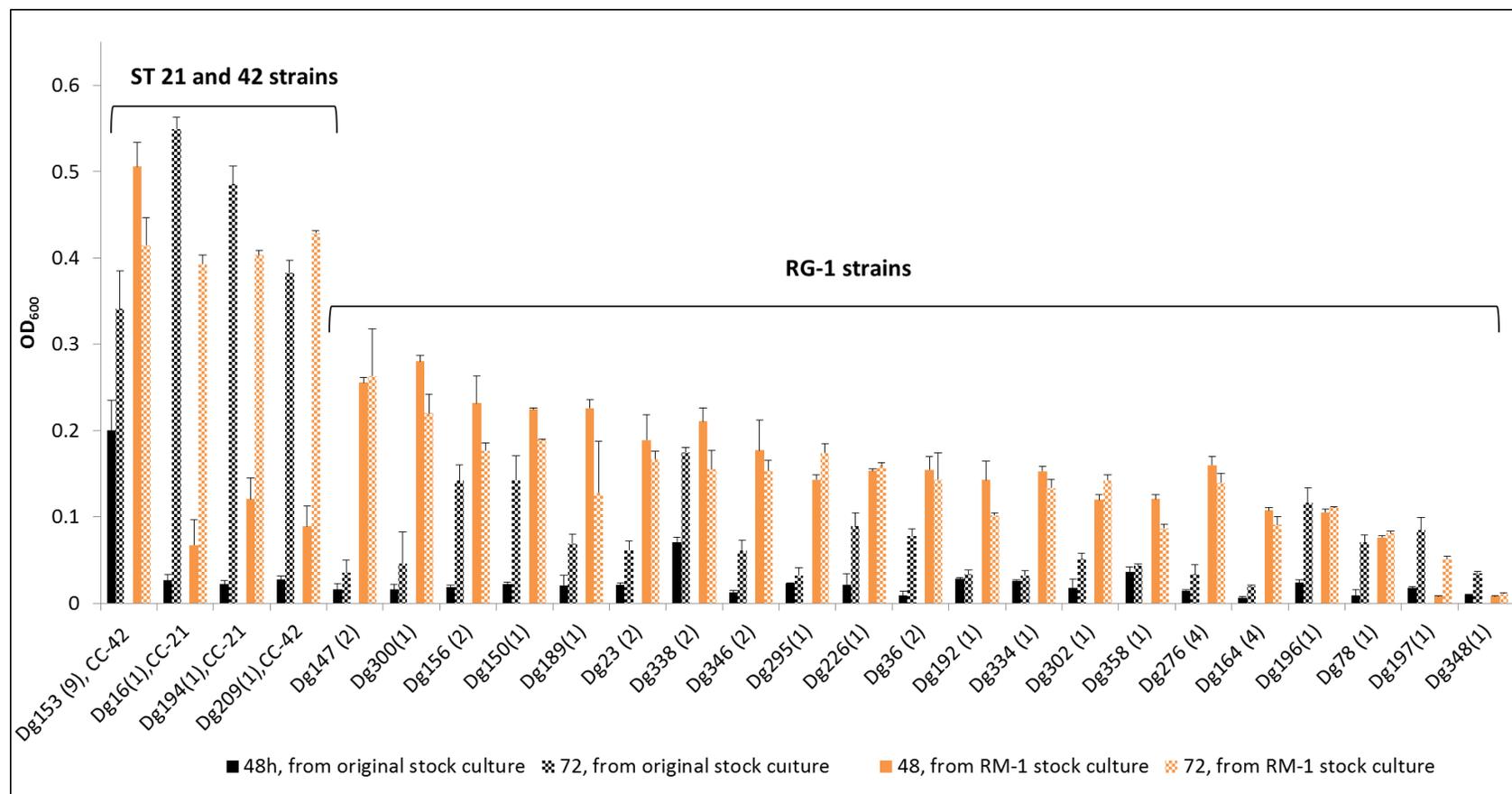
Figure 3.8: Growth enhancement following exposure to MHA medium. Cells were subcultured from either the original stocked cultures (from fresh bacterial strains) or from the RM-1 stocks (**Figure 3.3** at above), onto BA plates and processed for the microtiter plate growth assay (**see section 2.2.6**). Growth at OD_{600} was measured for triplicate wells at 48h and 72h. Brackets indicate the number of independent experiments for RM-1 stocks. Bacterial strains are grouped according to RG-1 and sequence types. Solid bars are growth at 48h, while dotted bars are growth at 72h for both the original and RM-1 stock cultures.

Additional adaptation of RG-1 isolates to MH medium was also tested by subculturing 5µl directly from a 48h MHB culture of RM-1 into fresh MHB (**Figure 3.9, A**). Although direct comparison between the original and the subcultured cultures is perhaps not valid, there was no evidence of growth enhancement in any of the RG-1 and other sequence types. For instance, growth levels in Dgs (300, 150, 226) were not much different between the original and subcultured cultures, only ranging between 0.05-0.26 OD_{600}. Furthermore, the very poor growers Dg348 and Dg36 from RG-1 strains (**Figure 3.8** at above) plus Dg153 were subcultured consecutively 6 times by direct streaking onto MHA plates. Bacterial cells from the last streaking and also from RM-1 stocks were grown in MHB under microtiter growth assay conditions. The OD_{600} was monitored at 48h as shown in **Figure 3.9, B**. Subculturing 6 times in total did not enhance growth of the strains remarkably. This suggests that a single round of growth of the freshly isolated bacterial strains to MH medium was enough to encourage the isolates to adapt to the laboratory condition.

The optimum growth for *C. jejuni* is between 37°C and 42°C. Strains from animal sources grow better at the typical body temperature, 37°C, of most animals. In contrast, strains from poultry have been reported to grow better at 42°C, the body temperature of birds (Duffy and Dykes, 2006). The possibility that some of the RG-1 strains may grow better at higher temperature was considered and growth at both 37°C and 42°C was tested. No significantly enhanced growth was observed at 42°C degree for any strains tested (**supplementary Figure 3.1**). Therefore, a growth temperature of 37°C and inoculum from RM-1 stocks was used for subsequent experiments.
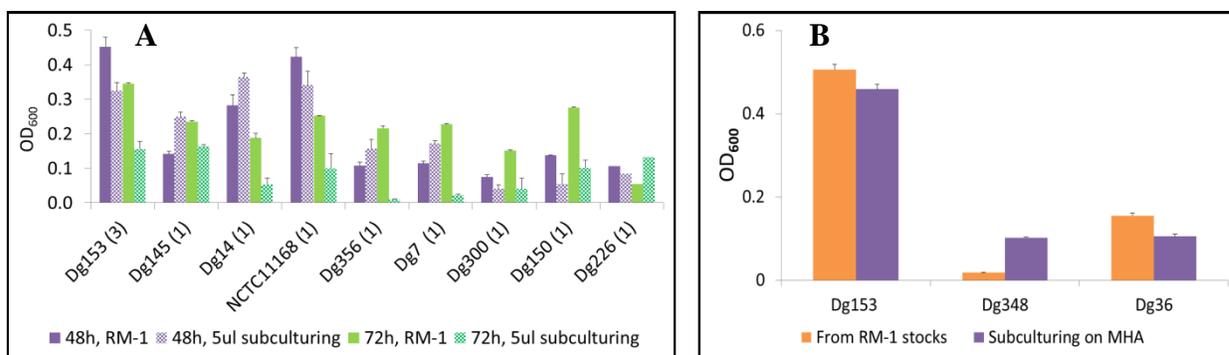
Figure 3.9: Consistent growth of RG-1 strains. (**A**). RM-1 stocks were used to grow cells in MHB, in a microtitre plate growth assay. Cell density was read at 48h and 72h (RM-1). A small volume of 5µl from 48h and 72h cultures was then subcultured into 195µl of fresh MHB and grown for further 48h and 72h (labeled as 5ul subculturing). Brackets indicate the number of independent experiments. Solid bars are RM-1 cultures at 48h and 72h, and dotted bars are 5µl subcultured cultures grown at 48h and 72h. (**B**) Consistent growth properties following 6 times subculturing of RG-1 strains on MHA plate. Bacterial cells were subcultured from the RM-1 stock onto MHA and then resubcultured consecutively onto 6 new MHA plates at each of the following time points; 72h, 47h, 47h, 43h, 24h and 26h. Cells from the final subculture plate (26h culture plate) were harvested and used in the preparation of a standardised inoculum for the microtiter plate growth assay in MHB. The RM-1 stocks were also used directly in the standard microtiter growth assay in MHB. Cell density was read at 48h. Averages and standard errors for RM-1 stock is from repeat experiments, each with triplicate wells. Dg300, Dg150, Dg226, Dg348, and Dg36 are RG-1 strains,

3.1.4   Comparison of growth of RG-1 strains in MHB and BHI

After establishing the growth assay, growth properties of all strains from the novel RG-1 clade (30 strains) and different sequence type groups of rat-associated *C. jejuni* strains were tested. The standard microtiter plate growth assay was followed, subculturing RM-1 glycerol stocks on BA plates to prepare a standard inoculum. All RG-1 isolates grew poorly in MHB as compared to control Dg153 strain, with a significant difference (P≤0.001) (**Figure 3.10, A**). For example, the best grower strain Dg147 among RG-1 strains grew to a bacterial density of 0.19 OD after 48h, whilst the other sequence types grew better to 0.32-0.49 OD (**Figure 3.10, B**). The majority of RG-1 isolates behaved as poor growers in MHB, with cell densities, at 48h, down to 0.08 OD for Dg78, and at the extreme end there were 5 very poor growers, Dg62, Dg61, Dg204, Dg179 and Dg348, which grew to an OD of just 0.02- 0.008 at 48h.

BHI, a richer medium, was also used to characterise the growth of the RG-1 strains. The strains behaved similarly as in the MHB (**Figure 3.11**). The RG-1 strains grew poorly in BHI compared to the other sequence types. They grew to 0.18OD on average by 48h, but the other strains grew to 0.36OD on average by 48h. The majority of the RG-1 strains grew in BHI slightly better than in MHB. The average OD_{600} in MHB was 0.12, but in BHI was 0.18 at 48h. This might be because BHI is a richer medium than MHB. There were some exceptions, for example, Dg62, Dg61, and Dg204 grew substantially better in BHI than MHB, but both Dg197 and Dg348 hardly grew in BHI compared to in MHB. To summarise, despite all being treated similarly on isolation from Norway rats and stocking, the classic isolates grew consistently better than the RG-1 group. It is possible that the RG-1 strains may require additional or alternative growth factors.
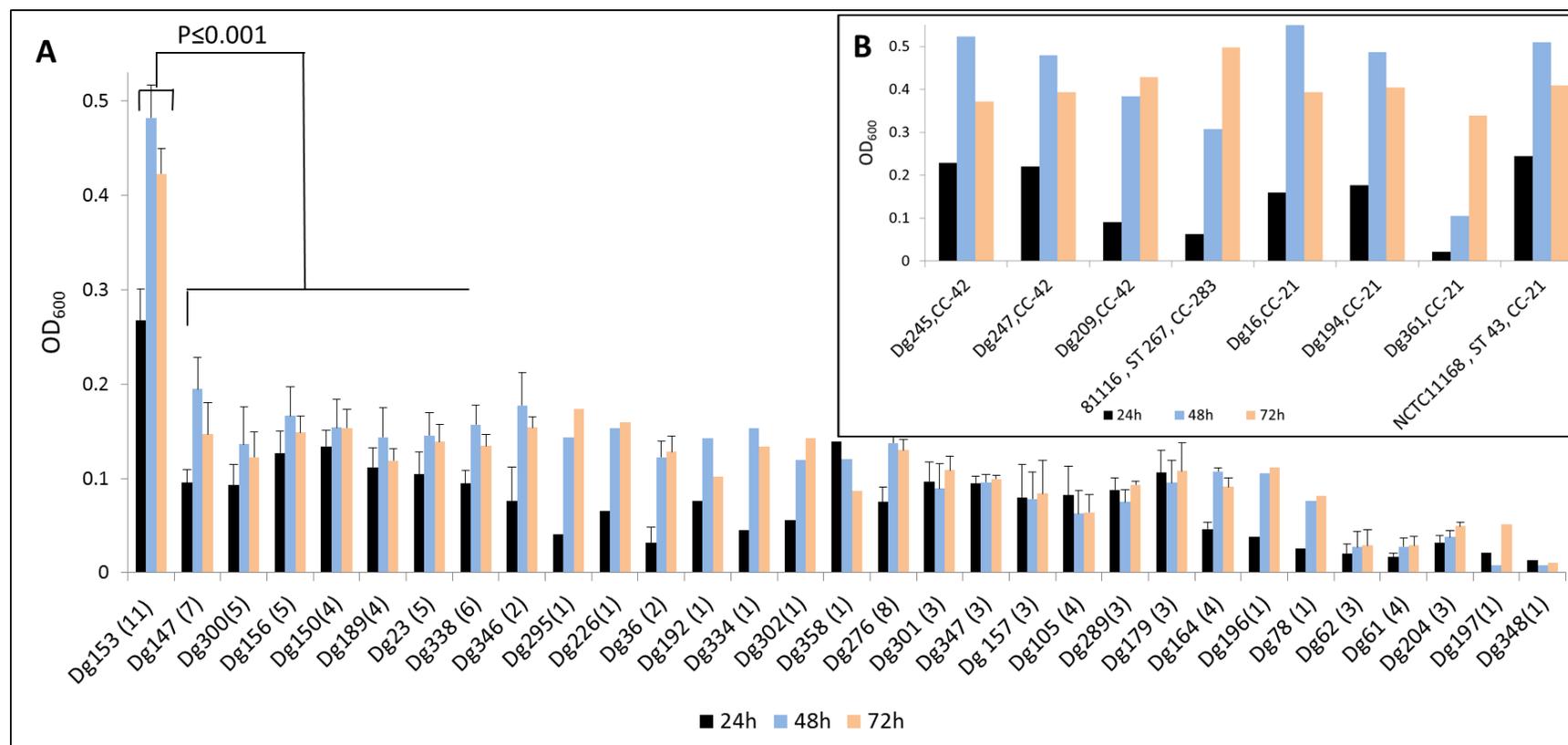
Figure 3.10: Growth of RG-1 in MHB. RM-1 stocks of (**A**) RG-1 strains and Dg153, and (**B**) strains of different sequence types and CC were subcultured on BA, and incubated for 36-42h. The wells contained 20μl standardised inoculum (0.02OD) in MHB. The OD was read every 24h. Brackets indicate the number of biological repeats; SE was made from more than a biological replicate. Three wells were inoculated for each biological replicate.
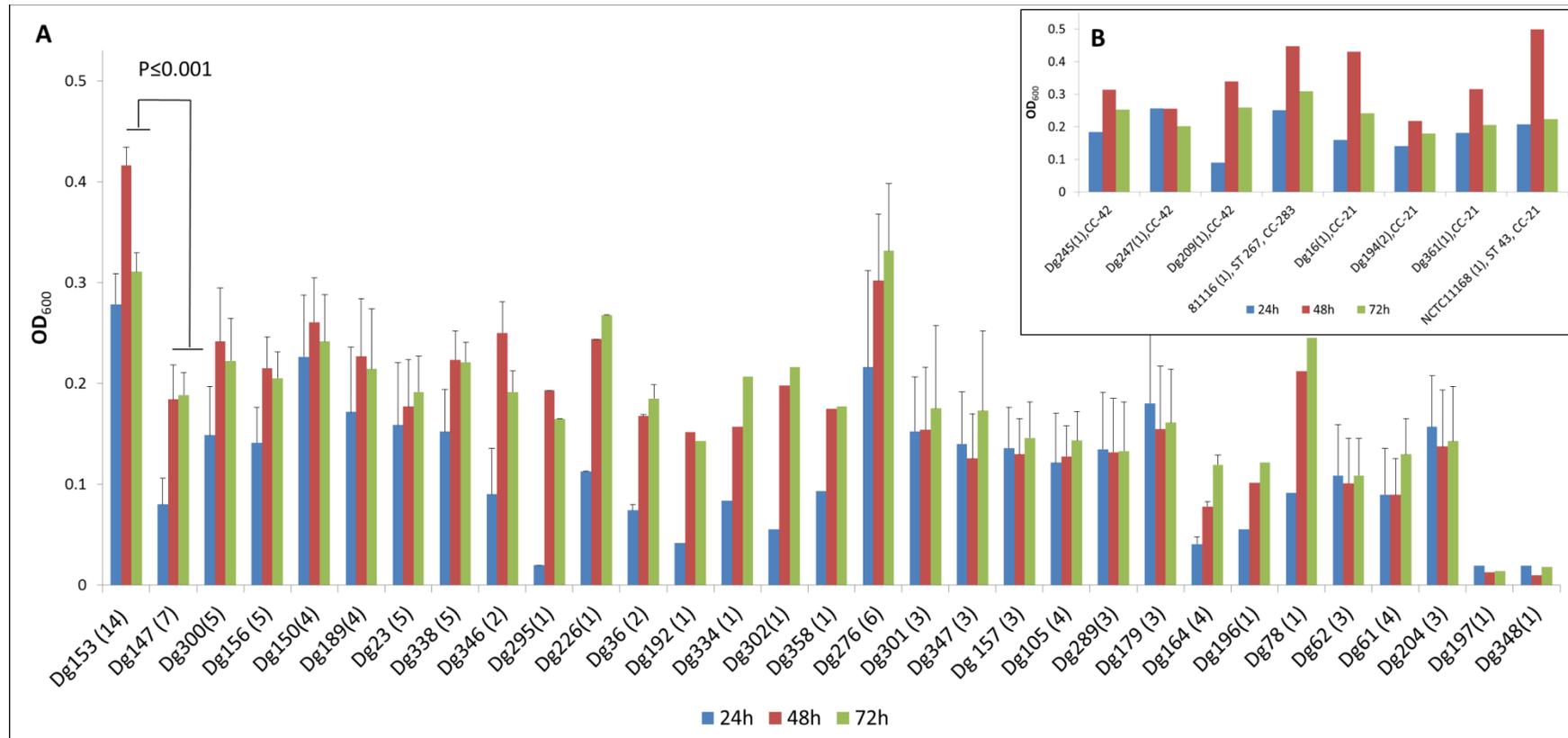
Figure 3.11: Growth of RG-1 in BHI. RM-1 stocks of (**A**) RG-1 strains and Dg153, and (**B**) strains of different sequence types and CC were subcultured on BA, and incubated for 36-42h. The wells contained 20µl standardised inoculum (0.02OD) in BHI. The OD was read every 24h. Brackets indicate the number of biological repeats; SE was made from more than a biological replicate. Three wells were inoculated for each biological replicate.

3.1.5   Growth versus motility

Motility (**Figure 3.4** at above) appeared to correlate with growth of the isolates in MHB (**Figure 3.10** at above). To examine the relationship more closely, growth was plotted against motility in **Figure 3.12**. This highlighted two distinct groups: good growers (primarily classic strains) with a large radius of motility (0.39 to 0.48 OD_{600} and 23 to 29 mm radius of motility) and poor growers. Poor growers includes all RG-1 strains (<0.2 OD_{600} and a much smaller radius of motility) and the 2 RG-2 strains tested, Dg95 and Dg275, (0.02 to 0.19 OD_{600} and motility from 2 to 14 mm radius). This is logical, since bacteria need energy to be motile. However, there were some exceptions. For instance, Dg197, in particular, was a very poor grower, but migrated to a reasonable distance (12 mm). Similarly, Dg370 ST 50 CC-21 the only classic strain among the tested strains that grew poorly with an OD_{600} of 0.15 also showed a high level of motility (24mm) (see circled isolates, **Figure 3.12**). In addition, as shown in **Figure 3.5** at above, all strains tested have bipolar flagellum. This excludes the possibility of their low motility being due to missing flagellum. Genome data confirms that the cgMLST (Cody et al., 2017) *fliA*, *flgR* and *rpoN* potential flagellar expression regulators (Jagannathan et al., 2001) are present in RG-1, RG-2 and reference strains Other factors such as gene expression (Tu et al., 2008) might be necessary to enhance the motility of the poor motile RG-1 and RG-2 strains.
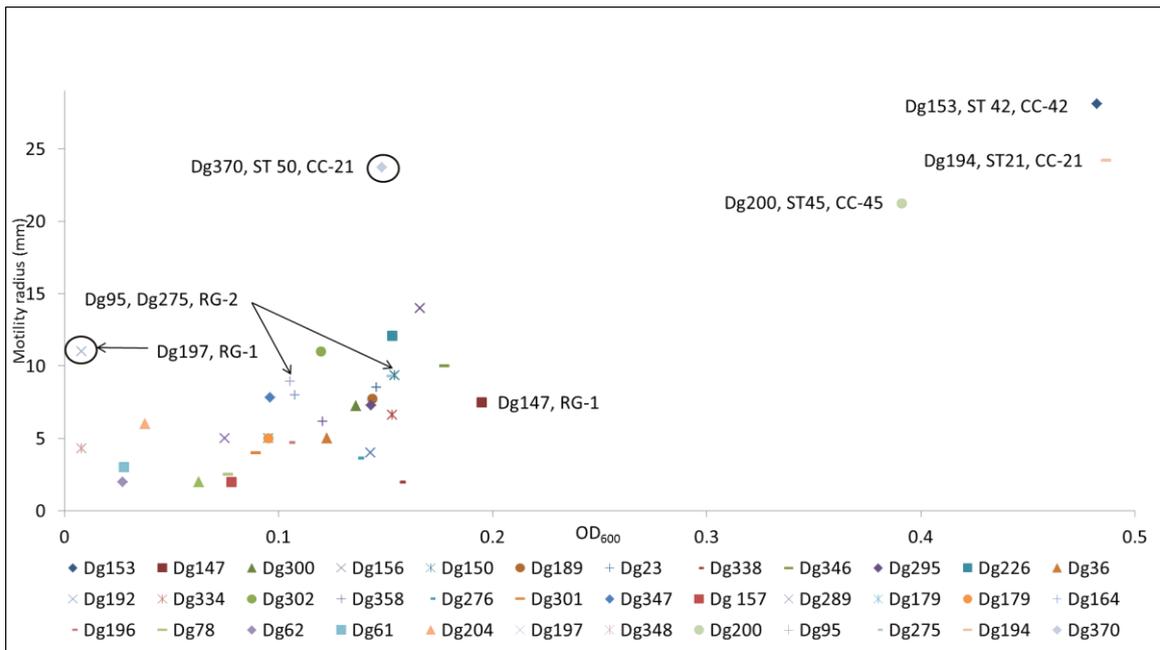
Figure 3.12: Relation between growth and motility. Motility (RM-2) at 72h and growth in MHB at 48h. Circles indicate outstanding strains Dg197 and Dg370 poor growth, high motility.

3.1.6 Rat mucin enhanced RG-1 growth

RG-1 strains grew poorly in MHB and BHI. The possibility that some host factor might enhance growth was tested by growing these strains, together with the reference strain NCTC11168, in MHB supplemented with lab rat intestinal and caecal mucin. Rats were initially confirmed as *Campylobacter*–free by culturing faecal samples from each rat in Bolton enrichment broth with selective antibiotic. Samples were then streaked on selective mCCDA medium and incubated under microaerophilic conditions at 42°C. All rats were free of *Campylobacter* spp. and therefore they could be used to extract their intestinal mucin, as described in **section 2.6.1**. Homogenised mucin was added to MHB with antibiotic supplement to 5% final volume, in a standard microtitre plate assay. Growth was monitored by following viable count, in two separate assays with different mucin preparations.

There was no immediate difference in NCTC11168 reference cell numbers in MHB supplemented with or without small intestine or caecum mucin at 24h (~1x 10^9 CFU/ml) (**Figure 3.13**). However, NCTC11168 cell numbers decreased 10-fold from 1.2 x 10^9 to 1.3x 10^8, and from 3.6 x 10^8 to 1.6 x 10^6 at 48h and 72h, respectively, in the presence of prefrozen

mucin (**Figure 3.13, A**). This is consistent with an experiment that was conducted by two undergraduate students under my supervision (**Appendix 3**). In contrast, the growth of RG-1 strains was enhanced in the presence of both prefrozen and fresh small intestine and caecum mucin. At 24h, Dg300 viable cells increased 6-fold, from $1x10^8$ to $6x10^8$ CFU/ml, and Dg147 viable cells increased 2 fold from $6x10^8$ to $1.3x10^9$ in the presence of small intestine mucin (**Figure 3.13, A**). The same pattern can be seen for both 48h and 72h. Fresh mucin also enhanced the growth of all tested RG-1 strains. At 24h, $1.6 x10^8$ CFU/ml of Dg300 increased to $6.2 x 10^8$ and $9.3 x 10^8$ CFU/ml in the presence of small intestine and caecum mucins, respectively (**Figure 3.13, B**). The same effect occurred in the other Dg276 RG-1 strain, at 45h, whereby the viable count of $4 x10^8$ CFU/ml was increased to $9x10^8$ and $7x10^8$ with small intestine and caecum mucins, respectively (**Appendix 3**). This preliminary data suggests that the lab rat mucin might contain some growth factors required by RG-1 strains as well as inhibitory factors that acts on the NCTC11168 control strains, but to which RG-1 strains are more resistant.

C. jejuni utilises human mucin MUC2 for modulation of expression of the following genes: mucin-degrading enzyme and *Campylobacter* invasion antigen genes: *cj1055c* and *cj0256* putative sulfatase family protein, *cj0843c* putative secreted transglycosylase, *cj1344c* putative glycoprotease, and *cj0914c Campylobacter* invasion antigen (*ciaB*) (Tu et al., 2008). The main role of these loci relates to helping *C. jejuni* colonisation and ability to cause disease. However, these genes are all defined as core genes (Cody et al., 2017), and are present in all strains tested including NCTC11168. The strain dependent growth was consistent with other studies. For example, Stahl and Vallance (2015) show a relationship between the intestinal mucus layer and *C. jejuni*.
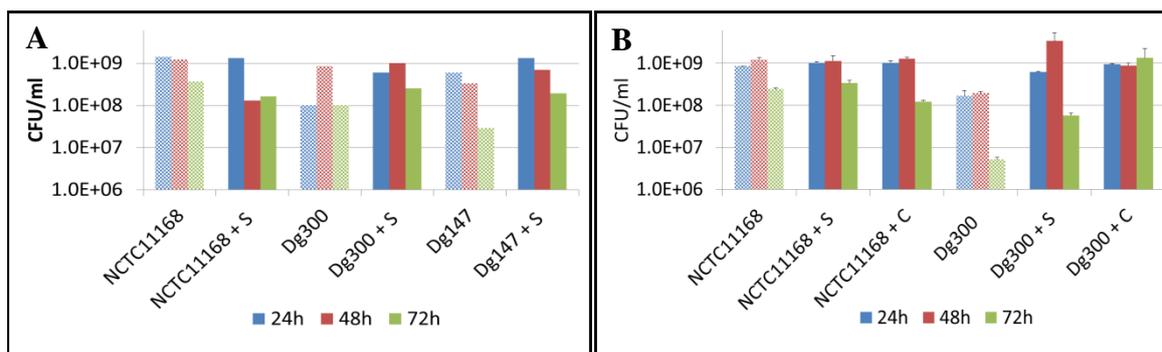
Figure 3.13: Impact of rat intestinal mucins on growth of RG-1 strains Crude preparations of small intestine (S) and caecum (C) mucin, as indicated, were recovered by scraping and then homogenizing in a lysing matrix tube and used directly to supplement MHB + antibiotic supplement in a microtitre plate growth assay. (**A**) The first study was with small intestine (S) mucin from 4 white lab rats (~200g/rat), whereby the mucin was frozen for 5 months prior to use. (**B**) The second study was with small intestine (S) and caecum (C) mucin preparations from a Walesh Norway Rat, which was used fresh. MHB was supplemented with 5% (v/v) mucin, 1% (v/v) antibiotic and 20μl of standard inoculum (0.002OD, 2.9×10^4 CFU), determined by Miles and Misra methods on 2%BA plates. SE is from 2 biological replicates. Patterned bars are without mucin.

3.1.7 Preliminary analysis of metabolic properties of RG-1 using Biolog

To find any unusual metabolic traits of RG-1 strains, which may be important for colonisation of Norway rats, initial screens were performed using a Biolog phenotypic microarray. This method indirectly measures bacterial cell respiration through reduction of the tetrazolium violet dye (Bochner, 2003). The PM 1 Biolog plate that includes 91 single carbon sources (Smalla et al., 1998), was tested for two different RG-1 strains and the NCTC11168 reference strain over a period of 48h.

The respiratory activity of Dg147 and Dg300 RG-1 strains and the NCTC11168 reference strain confirmed that many substrates were utilised by Dg147 and NCTC11168 strains. These carbon sources include succinic acid (A5), Asp (A7), Pro (A8), formic acid (B10), D,L-malic acid (C3), α-keto-glutaric acid (D6), Gln (E1), m-tartaric acid (E2), α-hydroxy butyric acid (E7), glycyl-L-aspartic acid (F1), citric acid (F2), glycyl-L-glutamic acid (G1), L-malic acid (G12), D-psicose (H5), glucuronamide (H7), and pyruvic acid (H8). False positives due to auto-reduction of the tetrazolium violet dye were documented for L-arabinose (A2), D-xylose (B8), D-ribose (C4), and L-lyxose (H6) (Wagley et al., 2014). However, 11 substrates were

utilised by only one or two of the isolates (**Figure 3.14** & **Table 3.3**). Among these 11 substrates, L-fucose (B4) and Glu (B12) were utilised only by NCTC11168. The inability of RG-1 strains to utilise L-fucose is due to the absence of the fucose utilisation operon, *Cj0480c - Cj0490* (Stahl et al., 2011, Dwivedi et al., 2016). This pathway is present in only 50% of genome-sequenced *C. jejuni* and *C. coli* strains, and is reported to be absent in livestock-associated isolates (Dwivedi et al., 2016). Dg300 showed no or very poor respiration on most nutrients tested, apart from the citric acid cycle (TCA) intermediate acids. This might be due to absence of some essential growth factor for this strain.

Three shapes of the kinetic plot curves were identified (**Figure 3.14**). With the first type, A, the respiration level peaks quickly, increasing up to about 24h before falling. In the second type, B, respiration also quickly increases but then remains at a constant level. With the last type, C, appears as a linear curve which slowly increases up to the final reading at 48h. This can be explained by the poor growth of Dg300 and should be excluded. Type A was predominant for NCTC11168. The NCTC1168 strain utilised the carbon sources in a short time, whilst the Dg147 RG-1 strain behaved mostly as type B, (**Figure 3.14 & Table 3.3**). This suggests that the RG-1 strains need longer periods to adapt and use the nutrients and is consistent with the poor growth of RG-1 strains (**Figure 3.10** and **Figure 3.11** at above) under *in vitro* conditions. The basic carbon sources Ser, Arg, Pro, Asp, and Glu were identified for confirmation of utilisation in *in vitro* growth experiments.
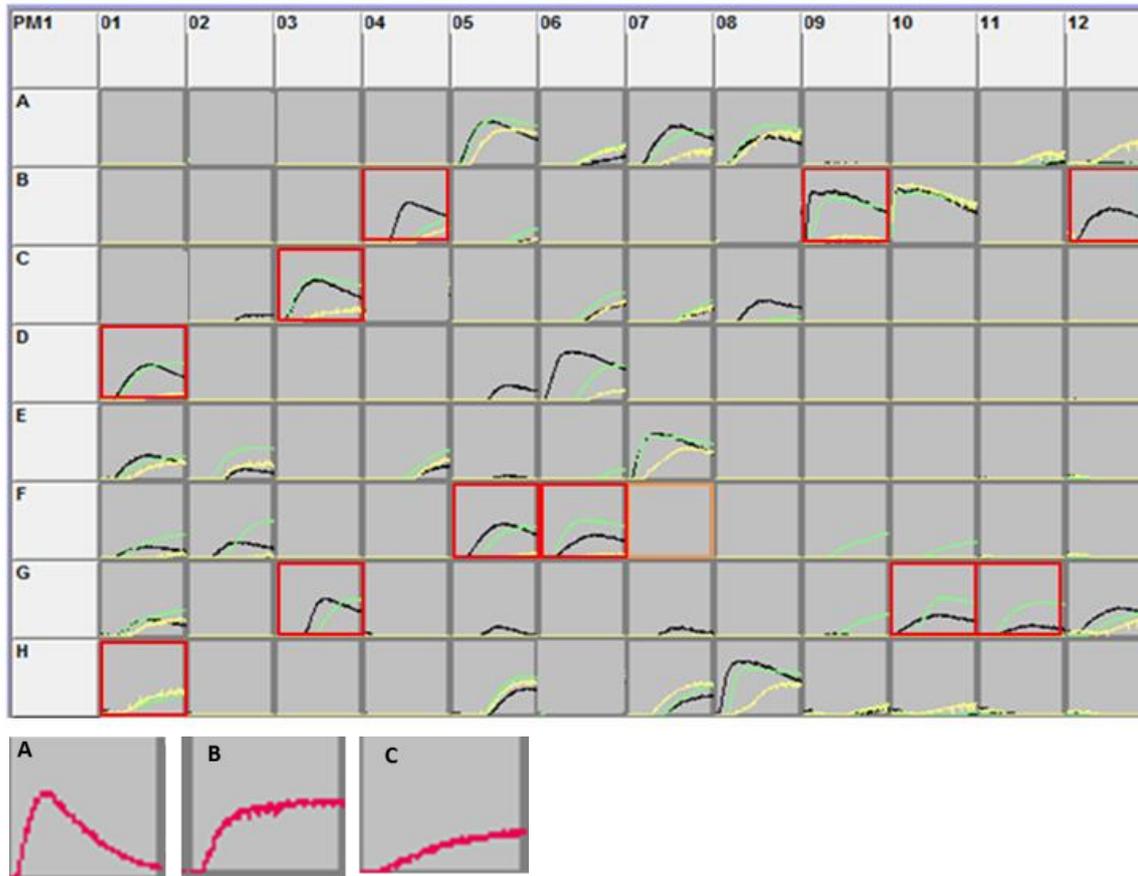
Figure 3.14: Kinetic plot curves of Biolog PM 1 plate. NCTC11168 reference strain (black), Dg147 (green) and Dg300 (yellow) were incubated for 2 days at 37°C under microaerophilic conditions in Biolog PM1 96-well plate, as described in **section 2.2.3**. Redox dye D density was monitored every 15min. The curves were performed by uploading an OFA file into microbial FM analyser software. The orange rectangular well (F7) is propionic acid, used only by *C. coli* (Wagley et al., 2014). A2, B8, C4 and H6 are false positive substrates. Red rectangular wells denote substrates not utilised by some of the strains. Bottom row, defined shapes of representative curves A, B, and C.

Table 3.3: Peak data at 24h and 48h incubation of PM-1 Biolog plate.

*PM1 wells	Substrates	NCTC11168**		Dg300**		Dg147**	
		24h	48h	24h	48h	24h	48h
A5	Succinic Acid	143, A	83	108, B	107	156, B	133
A7	L-Aspartic Acid	129, A	82	ND	ND	84	114, B
A8	L-Proline	92, A	69	85, B	80	125, B	118
B4	L-Fucose	135, A	87	ND	ND	ND	ND
B9	L-Lactic Acid	165, A	98	ND	ND	146, A	116
B10	Formic Acid	168, A	99	169, A	108	159, A	136
B12	L- Glutamic Acid	116, A	71, A	ND	ND	ND	ND
C3	D,L-Malic Acid	137, A	80	ND	51, C	149, A	123
D1	L-Asparagine	117, A	76	ND	ND	101	122, B
D6	α-Keto-Glutaric Acid	158, A	102	ND	ND	ND	115, B
E1	L- Glutamine	76, A	55	ND	62, C	55	79, C
E2	m-Tartaric Acid	ND	ND	ND	55, C	87	103, B
E7	α-HydroxyButyric Acid	144, A	84	81	84, A	140, B	118
F1	Glycyl-L-Aspartic Acid	ND	ND	ND	ND	ND	79, C
F2	Citric Acid	ND	ND	ND	ND	ND	122, B
F5	Fumaric Acid	107, A	75	ND	ND	64	105, B
F6	Bromo Succinic Acid	68, A	56	ND	ND	125, B	105
G1	Glycyl-LlutamicAcid,A	55	ND	ND	57, B	56	89, C
G3	L-Serine,	118, A	83	ND	ND	ND	123, B
G10	Methyl Pyruvate, A	71	ND	ND	ND	118, B	112
G11	D-Malic Acid	ND	ND	ND	ND	100	108, B
G12	L-Malic Acid	90, A	68	ND	ND	ND	80, C
H1	Glycyl-L-Proline	ND	ND	ND	85, B	ND	71, C
H5	D-Psicose	ND	83, A	ND	109, B	62	125, B
H7	Glucuronamide	ND	61, A	65	99, B	ND	101, B
H8	Pyruvic Acid	172, A	109	59	90, B	151	136, B

*PM1 wells A5-H8 contain substrates utilised by at least one of the strains, L-Fucose and Glu were not utilised by RG-1 strains and Ser was not utilized by Dg300.

A, B and C refer to the shapes of kinetic plot curve as defined in (Figure 3.14**). Numbers are max numbers at 24h and 48h of incubation. ND < 50.

3.1.8   Amino acid catabolism by RG-1 and reference NCTC11168 strains

In addition to TCA cycle intermediates, the amino acids Ser and Asp are preferably utilized by *C. jejuni* NCTC11168 and to lesser extent Glu and Pro (Leach et al., 1997). To reinforce the Biolog data, the ability to grow on different amino acids as carbon source were tested by adding them to MEM-FBS medium. Dramatic enhancement of growth was observed in the medium supplemented with Ser for NCTC11168, at both 48h and 72h, and to a lesser extent with Asp and Pro (**Figure 3.15**). A similar pattern was seen with Dg153. RG-1 growth started at a later stage of growth (72h). The OD_{600} of NCTC11168 was 0.37 in Ser, which was much greater than 0.19 in Pro at 48h (**Figure 3.15**) (Leach et al., 1997). In contrast, Dg147 grew to similar lavel in both Pro and Ser at 72h. All three *serABC* genes, are imperative for Ser biosynthesis from 3-phosphoglycerate, are encoded by *sdaC* (*Cj1625c*), a low-affinity and high-capacity Ser system. The *sdaA* (*Cj1624c*) gene for catabolism of Ser to pyruvate and ammonia, are present in the draft genome of NCTC11168 and RG-1 strains (Jolley and Maiden, 2010, Velayudhan et al., 2004, Gundogdu et al., 2007). These data contradict with the Biolog result observed here (**Figure 3.14 and Figure 3.15**). The Biolog data showed that strain Dg300 did not respire with Ser. This might relate to slow growth of the strain in this system. To confirm this, plus the results of other important amino acids, quantification of the nutrients was done during particular time points of the growth curve as described in the next paragraph.
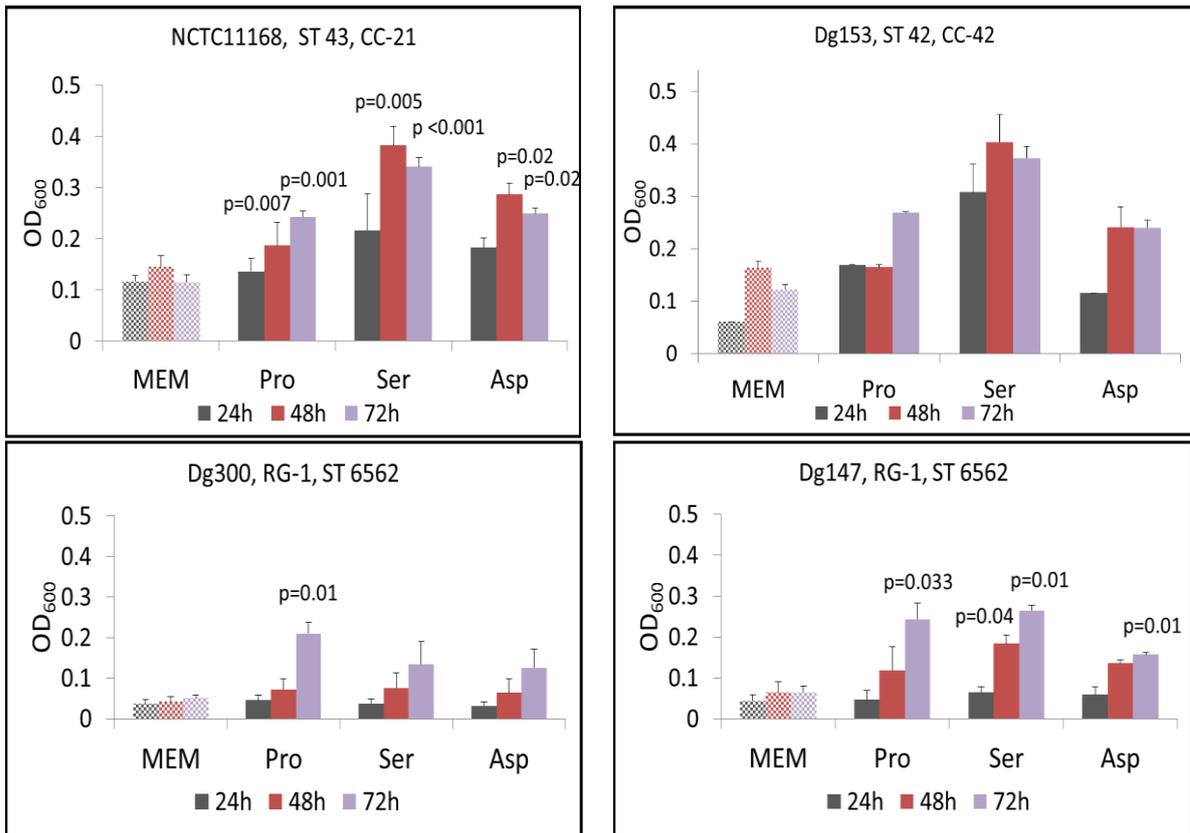
Figure 3.15: Growth of Dg300 and Dg147 in MEM-FBS supplemented with amino acids. The isolates were subcultured on BA from RM-1 stocks. The strains were then streaked on MHA plates and incubated under the microtiter plate growth assay condition (**see section 2.2.6**). The strains were grown in MEM-FBS, inoculated with 0.002OD of bacterial cells, 10% FBS and, with or without 20 mM Pro, Ser or Asp. Values are derived from the mean + SE of three biological replicates, except Dg153 which is from 3 wells on a biological replicate. ST- sequence type, CC-clonal complex. Dotted bars (MEM-FBS) represent cells grown in the medium without 20mM carbon sources (control), p = statistical differences between growth in the MEM-FBS without 20mM amino acids and MEM-FBS with 20mM of the amino acids to the same time point of growth.

The Biolog data suggested that there are some exceptions in the nutrient catabolism of RG-1 compared to NCTC11168. However the RG-1 strain, Dg300 grew slightly better in the MEM-FBS that contained the amino acids (**Figure 3.16, A**). To confirm this, all amino acids were quantified in the MEM that was supplemented with 10% of FBS and 20mM of Pro or Ser. It was clearly seen that Dg300 RG-1 strain had utilised 20% of the Pro and almost all of the Ser by 72h (**Figure 3.16**, **B**). Dg300 utilised more than 50% of 0.4mM glutamate (**Appendix 4**). This is in agreement with the presence of glutamate transporter genes *peb1A* (*cj0921*), *cj0919*, *cj0920*, and *pebC* (*cj0922*) within the Dg300 genome (Guccione et al., 2008, Stahl et al., 2012). The negative result of Dg300 in Glu for the Biolog assay might be due to that the strain cannot respire in these amino acids under this system (**Figure 3.14** at above). In addition to Pro, Ser, and Asp amino acids, the amount of arginine, methionine, and Gln amino acids were also depleted by at least 60% for both the reference and RG-1 strains (**Appendix 4**). This result is in agreement with previously published data conducted in different culture media for Ser, Asp, Arg, Glu, and Gln use by RG-1 strains (Wright et al., 2009). The concentration of the rest of the amino acids had either increased or decreased slightly. This may be due to cell lysis. To summarise, these data confirm that the RG-1 and NCTC11168 strains can use the same amino acids without exception. They have a preference for Ser first, and then in decreasing preference for Asp, Glu, and Pro (Leach et al., 1997, Guccione et al., 2008).
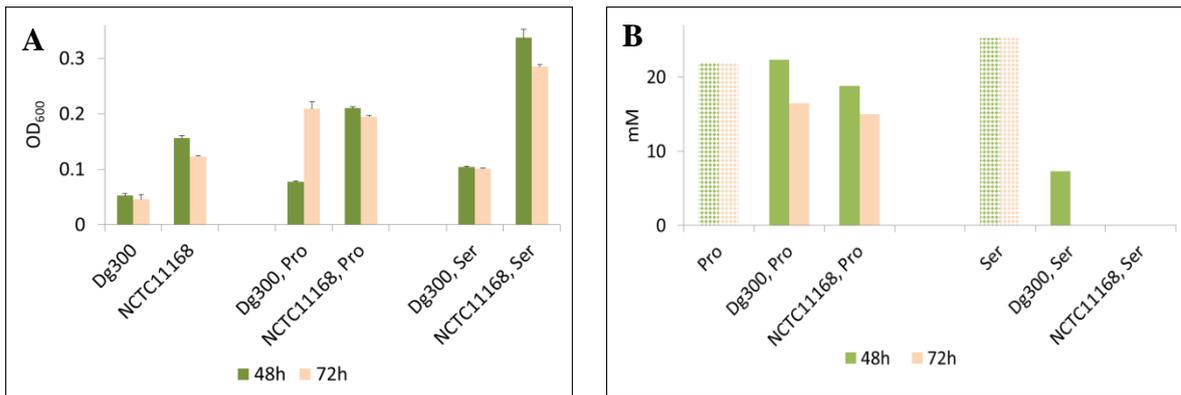
Figure 3.16: Monitoring growth *C. jejuni* strains and quantification of amino acids. (**A**) The isolates were sub-cultured from RM-1 stock cultures onto BA plates, and after incubation the strains were streaked on MHA plates. The MEM-FBS was then inoculated with $0.002OD_{600}$ of the bacterial cells and supplemented with 10% FBS and 20 mM of either Pro or Ser. The isolates were grown under microtiter growth assay conditions. Growth of the Dg300 RG-1 and NCTC11168 were monitored at 48h and 72h. (**B**) The amino acids were quantified at 48h and 72h. Dotted bars represent the concentration of Pro and Ser in non-inoculated medium with the bacterial cells at 48h and 72h. Concentration of the other amino acids are shown in **Appendix 4**. The amino acids were quantified using EZ-Faast-kit and GC-MS.

3.2 Closing Dg147 draft genome, a representative of RG-1 group

Due to the unique nature of the RG-1 clade, the potential of specific host association of these strains, and the excellent draft genome sequences available (an average of 32.6 contigs for this group), it was decided that closing the genome of one representative strain would be worthwhile for detailed comparison with other *Campylobacter* isolates. Dg147 was chosen as a representative of the RG-1 group. The draft WGS of this strain contained only 26 contigs, the smallest number of contigs among all RG-1 strains, and Dg147 belongs to the largest cluster identified by WGS phylogeny of the RG-1 group (**Figure 3.17**). This cluster 1 includes ten isolates, which is in close proximity to cluster 2, which contains eight isolates. Clusters 1 and 2 together comprise 60% of RG-1 and are quite distant from cluster 3 (7 isolates).
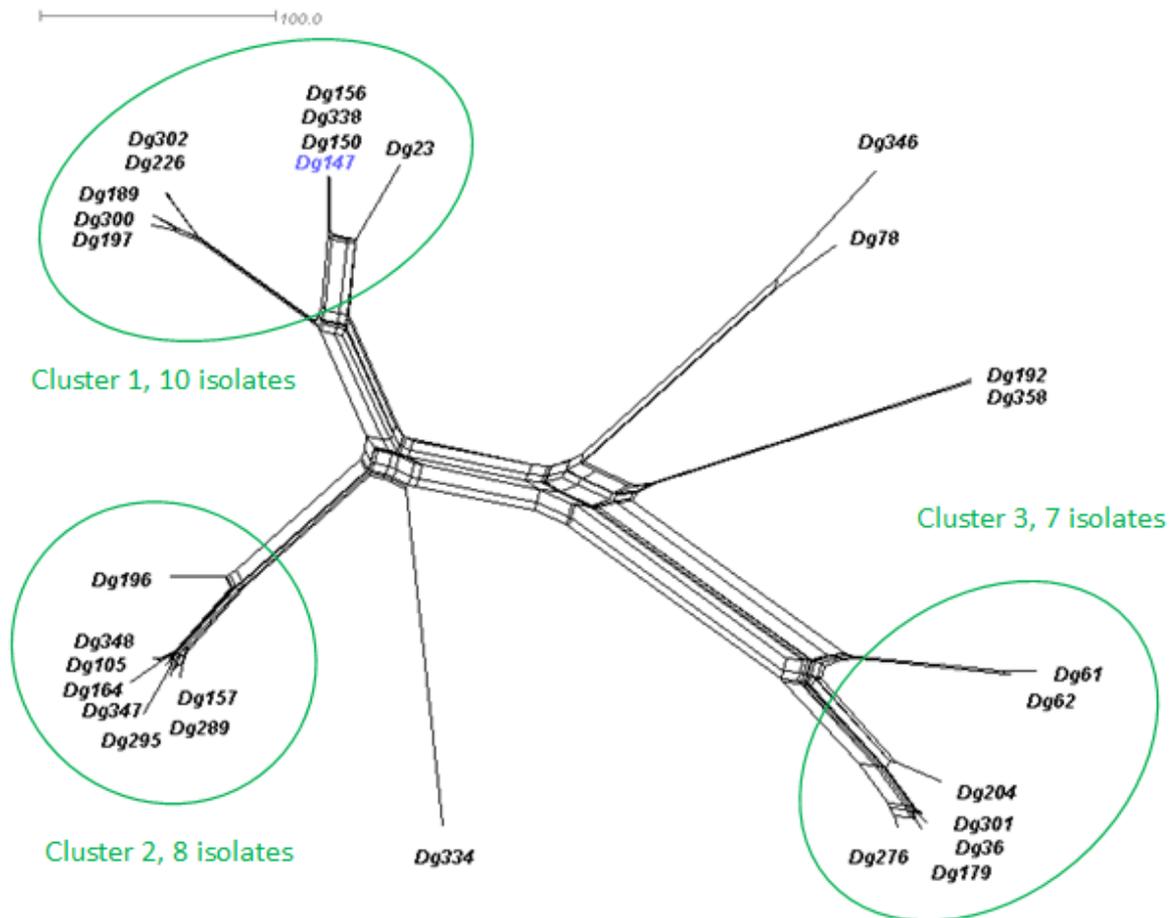
Figure 3.17: WGS phylogenetic network of RG-1 group. Draft genomes of all 30 RG-1 strains were compared using genome comparator of the PubMLST/*Campylobacter* database. SplitsTree4 was then used to visualise and edit the results of the phylogenetic network (Huson and Bryant, 2006). Blue highlighted Dg147 was selected for closing the genome, as a representative RG-1 strain.

3.2.1   Genomic assembly of Dg147 using NCTC11168 as a reference strain

The strategy of genome mapping, sequence assembly and closure are shown in the flow chart in **Figure** 3.18. The FAS file of Dg147 from the PubMLST/*Campylobacter* database and FAS files of the sequenced published genome of *C. jejuni* subsp. *jejuni* NCTC11168 (AL111168, 06-FEB-2015) (Gundogdu et al., 2007), 81116 (Pearson et al., 2007) and 81-176 (Accession number CP000538.1) were downloaded from the NCBI database. The Dg147 draft genome was assembled with each reference strain separately using CONTIGuator software (Galardini et al., 2011). The output Archive file contents from CONTIGuator were read by the Artemis

Comparison Tool (ACT) (Carver et al., 2008) to create a display comparison. Archive files of genome alignment between Dg147 and the three reference strains are provided in the **supplementary data**. Only 9 or 6 Dg147 contigs were mapped with genome sequences of 81116 and 81-176 respectively, while 13 contigs were mapped with the NCTC11168 genome sequence, leaving 13 unmapped contigs. Mapped contigs are shown **Figure 3.19**. The number and size of unmapped contigs were as follows; 5 (6195bp), 11 (210bp), 12 (975bp), 13 (337bp), 14 (1981bp), 19 (284bp), 20 (2377bp), 21 (207bp), 24 (658bp), 25 (251bp), 26 (9537bp), 29 (1562bp), and 37 (802bp).
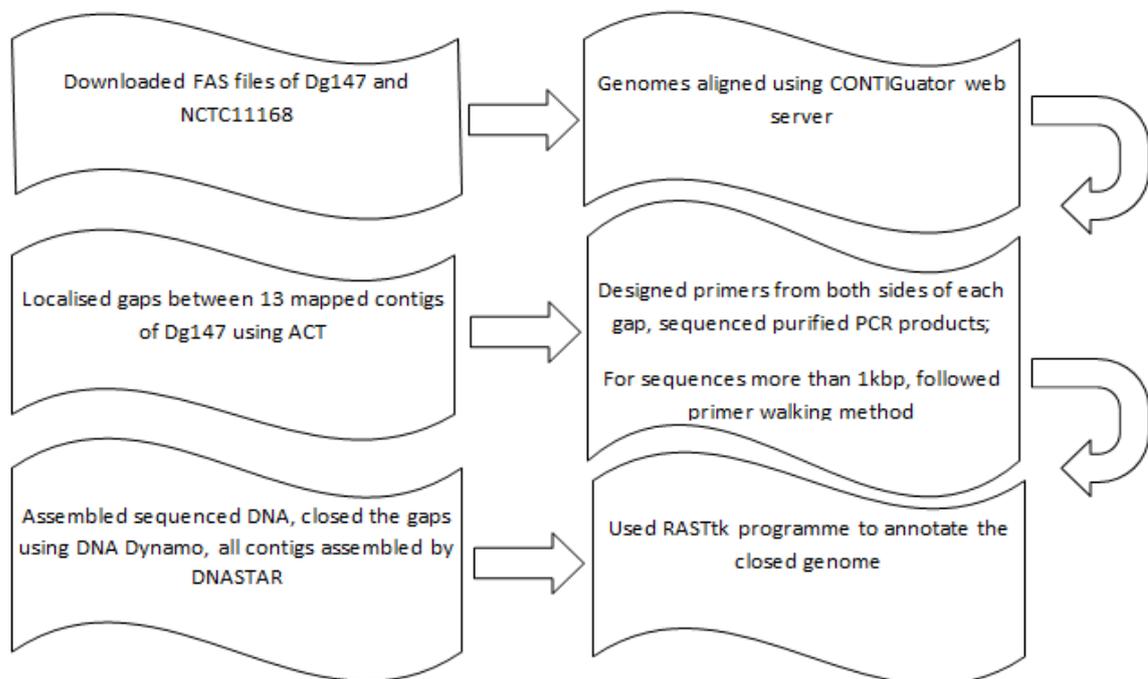
Figure 3.18: Flow chart of genome mapping, closure and annotation strategy for Dg147 RG-1 strain.
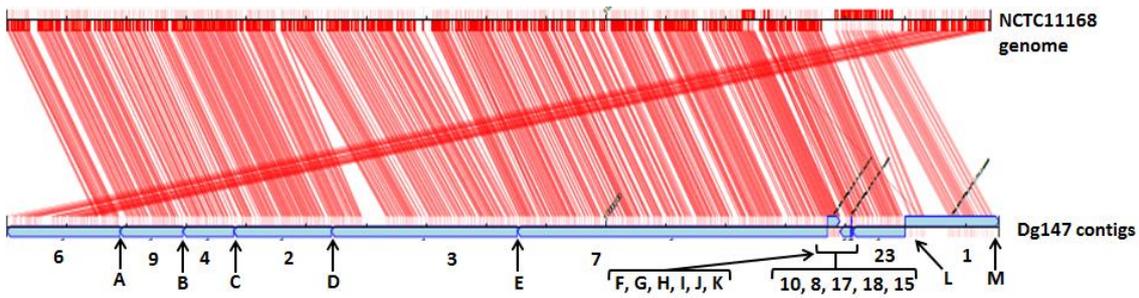
Figure 3.19: Contig mapping of the Dg147 draft genome against the reference genome NCTC11168. The alignment was made using the CONTIGuator web server (Galardini et al., 2011). The archive file was then imported into the ACT programme to visualise the gaps. The top horizontal line represents the reference genome and the bottom horizontal line, segmented into multiple blue contigs, represents the Dg147 strain mapped contigs. Red lines represent the region at which both the reference genome and Dg147 strain share similar DNA sequence, with a minimum of 80 percent identity over ≥80 bp. White regions represent differences between the compared DNA sequences. Mapped contigs 6 (188613bp), 9 (104898bp), 4 (84761bp), 2 (162727bp), 3 (310321bp), 7 (517600bp), 10 (20901bp), 8 (17333bp), 17 (1240bp), 18 (1116bp), 15 (1339bp), 23 (87260bp), and 1 (156631bp), and gaps A to M were subsequently labelled accordingly. Gap M is between contigs number 1 and 6. Sheng Foo is acknowledged with respect to his contribution to closing the Dg147 genome during his Final year project.

3.2.2 PCR amplification and closing gaps A-M of Dg147

Forward and reverse primers were designed from 85 to 1508bp from the end of contigs (**Table 2.4, A**) and used to amplify DNA within each gap using the isolated Dg147 genome as a template. Products varied in size from 443bp for gap F to 7201bp for gap D (**Figure 3.20**). Purified PCR products were sequenced using the amplification and sequencing primers. For larger products, from 1-7kb, primer walking was followed (**Table 2.4, B**). The size of the sequenced DNA for the gaps correlated with sizes of the gel bands. Initially, it was not possible to sequence the PCR product of gap L. Therefore, the other gaps were closed first. Unmapped contigs 14 and 29 were not located in these other gaps. Primers (Dg147-LRC and Dg147-LFC for unmapped 14, and Dg147-LRD and Dg147-LFD for unmapped 29) were designed at both ends of these 2 unmapped contigs. Sequencing of gap L (5,852 bp) PCR product with these primers identified within this gap the following organization within this gap: mapped contig 23, unmapped contig 13, unmapped contig 29, unmapped contig 13, unmapped contig 14, unmapped contig 13 and mapped contig 1 (**Table 3.4**). The three copies of unmapped contig 13 had caused problems with sequencing by primer walking.
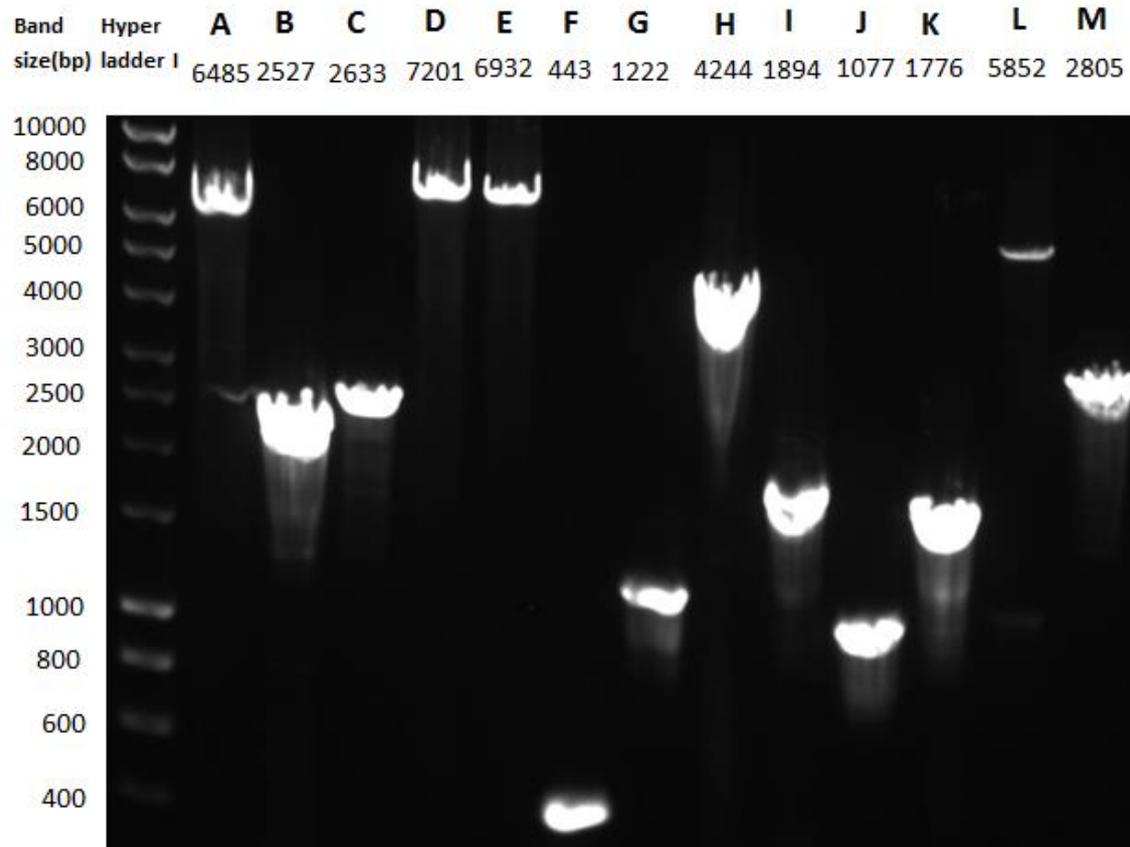
Figure 3.20: AGE of PCR products from amplified gaps between assembled contigs of Dg147. Gaps (A to M) are as shown in **Figure 3.19** at above on assembly of Dg147 contigs with NCTC11168 genome. PCR products were obtained using CloneAmp HiFi, isolated genomic DNA and primers as listed in **Table 2.4 A**. The following set of primers were used to amplify gaps B, C, I, L, K and M; Dg147-B-FA, Dg147-B-RA; Dg147-C-FA, Dg147-C-RA; Dg147-I-FA, Dg147-I-RA; Dg147-K-FA, Dg147-K-RA; Dg147-L-FA, Dg147-L-RA and Dg147-M-FA, Dg147-M-RA. Numbers below the letter for each gap indicate the length of each amplified product following confirmation by DNA sequencing.

The closed genome of Dg147 was saved as FAS, EMBL, and GBK files using DNASTAR software. The closed Dg147 genome was aligned with the NCTC11168 genome using ACT (**Figure 3.21**). Identical regions between both genomes were labelled with vertical red lines. In addition to these, there were some red diagonal and horizontal red lines between the aligned genomes. The main diagonal lines were between the three copies of the ribosomal RNA loci (*rrnABC*) (**rectangles in Figure 3.21**). The diagonal red line denoted with a circle number 1 links two copies of genes encoding putative methyl-accepting chemotaxis signal transduction proteins, *cj0144* and *cj1564* genes (Marchant et al., 2002). The C-termini of both proteins are identical (72.6% total identity). The third diagonal red line (circle number 2) is between *cj1307* (encoding putative amino acid activating enzyme) of NCTC11168 and locus number 1440 of the Dg147 genome. The above repeated gene/s caused difficulty in assembling DNA sequences following Illumina sequencing.
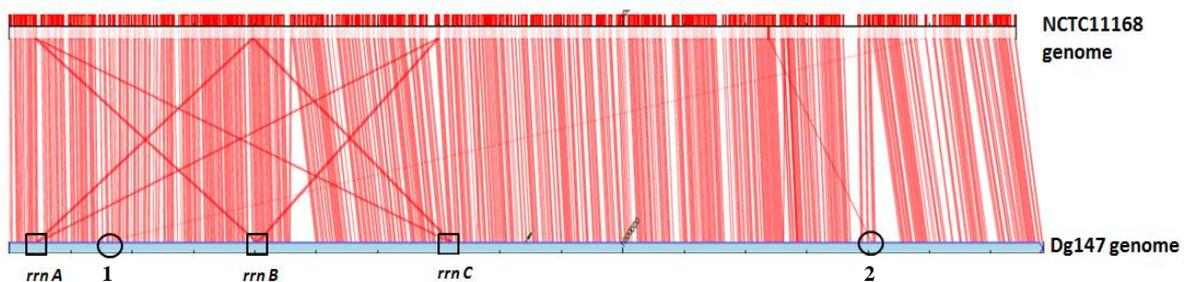
Figure 3.21: Mapped closed genome of Dg147 with NCTC11168 reference genome. Top and bottom horizontal lines represent the reference and Dg147 genomes, respectively. Red lines represent the region at which both the reference genome and Dg147 genome share the same genome similarities. Squares are placed on the 3 *rrn*s in Dg147, and circles 1 and 2 are repeat regions. Details of the comparison is able to be seen by zooming in to the level of genes or DNA bases.

Rapid Annotation using the Subsystem Technology tool kit (RASTtk) (Brettin et al., 2015, Aziz et al., 2008) was used for genome annotation of the closed genome of Dg147. The genome consists of a 1,685,204bp circular chromosome with a GC content of 30.43%, 1746 predicted coding sequences (CDSs), and 44 tRNA and 9 rRNA. The closed genome Dg147 strain does not contain plasmids. The closed, annotated Dg147 genome was uploaded into the ACT program to identify gene/s located in the filled gaps. The three identical copies of RNA operons (5709bp), denoted as A, B, and C (Kim et al., 1993) were located in the gaps A, D, and E, respectively (**Table 3.4**). B, C, and M gaps were filled with genes encoding methyl-accepting chemotaxis proteins, Cj0144, Cj0262c, and Cj1564c, respectively. The C-termini

of these three encoded proteins contains an identical repeat (Marchant et al., 2002, Rahman et al., 2014). The following *maf* family, containing *maf1*, *maf4*, *maf6*, *pseE* and *flaB,* which encode chemotaxis and motility proteins is located in G, H, and K gaps. It is known that these *maf* family genes are variable genes and cause phase variation in *C. jejuni* by a slipped-strand mispairing mechanism (Karlyshev et al., 2002). Both *maf1* and *maf4* are identical and both contained homopolymeric G tracts. This would contribute to problems in assembling sequence data of this region.

Three small-unmapped contigs were not placed in any gap. Unmapped contig 11 (210bp) was identical to a duplicated sequence within contig 7. This sequence encodes tRNA-Asp and tRNA-Val. The second unmapped contig 19 (284bp) is identical to sequences within *flaA* (1719bp) and *flaB* (1719bp) genes which are located on mapped contigs 17 and 18, respectively. The last unmapped contig 21 (207bp), was identical to a sequence within two different genes, *cj0122* and *cj0814*. Both encode hypothetical proteins and are located on contigs 7 and 9, respectively. Each of the unmapped contigs 11, 19 and 21 was Blasted with each of the closed genomes of the three strains NCTC11168, 81116, and 81-176. As with the closed Dg147 genome, there were two copies and only two copies, of each of these contigs in all 3-reference strains. Thus, these three unlocalised contigs can be ignored.

Table 3.4: Identified genes within gaps of partially assembled Dg147 genome.

Gaps	Flanking mapped contigs	Size of gaps *(bp)	Unmapped contig number(s) identified in the gap, state of sequencing	Gene or operon located within gap **
A	6, 9	6027	5, sequenced	*rrnA* operon
B	9,4	829	12, sequenced	Cj0144, methyl-accepting chemotaxis protein
C	4,2	807	12, sequenced	Cj0262, methyl-accepting chemotaxis protein
D	2,3	6195	5, sequenced	*rrnB* operon
E	3,7	6195	5 and 25, sequenced	*rrnC* operon on 5, intergenic DNA on 25
F	7,10	3	No DNA in the gap, sequenced	The break was in a polymeric tract of 19 G.
G	10,8	371	26, sequenced	Part of *maf1* (Cj1318), motility accessory factor
H	8,17	3382	20, 26 and 37, the gap was sequenced except 2085bp was taken from contig 20	Part of *maf4* (Cj1335), motility accessory factor *pseE* (Cj1337) and end of *flaB* (Cj1338)
I	17,18	632	37, sequenced	Part of *flaA* (Cj1339c)
J	18,15	492	24, sequenced	Part of Cj1340c (hypothetical protein), 83.4% identity to Cj1341
K	15,23	490	24, sequenced	Part of *maf6* (Cj1341c), motility accessory factor
L	23,1	4218	Three copies of 13, a copy of 14 and a copy of 29, the gap was sequenced except 1433bp and 672bp were taken from contigs 14 and 29, respectively	Two uncharacterised loci,1797 and 2223bp, part of capsular polysaccharide biosynthesis gene (1647bp) in *C. jejuni* strain HS05, all genes absent in NCTC11168
M	1,6	829	12, sequenced	Cj1564 putative methyl-accepting chemotaxis signal transduction

* Determined following sequencing of PCR product obtained as in

Figure 3.20 ** identified through annotation.

3.3   Genome comparison of RG-1 group

3.3.1   Core and accessory genes of RG-1 group

WGS phylogeny (**Figure 1.11** and **Figure 3.1, B** at above) highlighted that the RG-1 group is a distinct clade of *C. jejuni* isolates, distinct from not only those *Campylobacter* isolates from humans, the environment, and animals, but also from other rat-associated *Campylobacter* strains. Therefore, the genomic content of this group was analysed to identify shared genes with generalist strains, and to highlight genes unique to this group, which might support the concept of adaptation to a rat host. The completed and annotated Dg147 genome was compared to draft genomes of other strains of RG-1 using genome comparator of the PubMLST/*Campylobacter* database after uploading gb file, Dg147.gb used for variant calling, of the closed genome (**see supplementary data**). Importantly, RG-1 strains shared a high number of 1561 core genes and possessed a low number of 269 accessory genes (**Figure 3.22**). Among the accessory genes, 185 genes were present in Dg147, but were absent in at least one of the other RG-1 strains. Accessory genes were grouped, and listed from 1 to 10, according to location in the genome in **supplementary Table 3.1**. Position of each group in the RG-1 genome and relative size of each accessory region is indicated in the Blast atlas shown in **Figure 3.22**. In addition to the 185 accessory genes, present in Dg147 and absent in at least one other strain, 84 additional accessory genes were identified. These were absent in the Dg147 genome, but were present in at least one of the other 29 RG-1 strains (**supplementary Table 3.2**). The RG-1 core genes comprised 85.3% of the pan-genome. This is slightly higher compared to the core genes which are found in 12 *C. jejuni* tested strains, including three common reference strains NCTC11168, 81-176, and 81116 (79%) (Dorrell et al., 2001). This can be explained by the close evolutionary relatedness of this group of isolates, despite the fact that they were isolated from 7 different farm areas over 3 years.

Approximately one third of the 269 accessory genes were from two large phage insertions. The first insertion was located in the first quarter of the Dg147 genome between locus480 to locus538. This big insertion (40.3kb) was present in all RG-1 strains except for Dg334 (**Figure 3.22**, group 3). The insertion consisted of 58 loci and had a GC content of 28%, approximately 2% percent lower than that of the Dg147 genome. The majority of the genes in

this region were uncharacterised encoding hypothetical genes plus five recognized phage genes; loci number 480 (Phage integrase family protein), 489 (Phage antirepressor protein), 505 (phage repressor protein, putative), 524 (Phage antirepressor protein), and 533 (hypothetical protein, phage tail fiber-like) (**see group 3, supplementary Table 3.1**). The second phage insertion (38.8kbp, 30% GC content) was located between locus1448 to locus1502, assigned to group 9. In contrast to the first phage insertion, only Dg147, Dg23, Dg150, Dg156 and Dg336 (16.6% of the RG-1 strains) harboured this phage DNA. Among the 54 loci 18 loci were characterised as encoding phage-related proteins; including locus 1451 (bacteriophage DNA transposition protein A, putative), locus 1470 (Phage (Mu-like) virion morphogenesis protein) locus 1471 (Mu-like prophage FluMu protein gp29), locus 1488 (tail fiber protein H, putative) and locus 1497 (phage virion morphogenesis protein, putative). The other large variable insertion, group 8 encodes capsular biosynthesis genes (**supplementary Table 3.1**). The majority of the remaining accessory genes encode phase variable flagella-mediated motility (Karlyshev et al., 2002) and chemotaxis factors, membrane and membrane transport proteins.

Figure 3.22 : Blast atlas of RG-1 genome sequences. Dg147 The closed and annotated genome file ofDg147 Dg147_closed genome, Supplementary data) was blasted with draft genomes of all other RG-1 strains using the BIGSdb database. The figure highlights the core and accessory genome areas of RG-1 isolates based on the Dg147 genome. The main groups of accessory genes, visible as white gaps, were labelled from 1-10, as listed in **supplementary Table 3.1**. The large areas 3 and 9 correspond to phage DNA. The first and second inner-most circles display the GC-skew and GC content, respectively. The third inner-most green coloured circle displays the gb file of the closed genome of Dg147. The Blast atlas was made using GViewServer https://server.gview.ca (Petkau et al., 2010)

3.3.2   Analysis of cgMLST defined core genes of RG1 strains

A set of core genes present in more than 95% of the draft genomes of 2472 clinical *C. jejuni* and *C. coli* isolates from campylobacteriosis, from Oxfordshire have recently been defined as cgMLST genes (Cody et al., 2017). This is accessible as a cgMLST from genome comparator- *C. coli/ C. jejuni* cgMLST v1.0 via PubMLST/*Campylobacter*. This set of cgMLST genes was compared directly gene-by-gene to RG-1 core genes as described in 2.8.3, to identify any cgMLST genes absent from the set of core RG-1 genes and also any core RG-1 genes absent from cgMLST Only 2.1% (29 genes) of the 1343 cgMLST genes were absent or incomplete in one or more of the RG-1 strains. Of these 16 genes were incomplete or absent in a single RG-1 strain, a further 10 were absent 20-40 % of strains and only 3 cgMLST genes were missing in all 30 RG-1 strains (**Appendix 5**). These three unlinked 'missing' genes include *cj0145*, which encodes a TAT-dependent alkaline phosphatase (PhoA) (van Mourik et al., 2008). As this is the only alkaline phosphatase in *C. jejuni* NCTC11168 and 81168, absence of this enzyme in RG-1 strains may reflect absence of requirement of a phosphatase under alkaline conditions by these strains. The second 'missing' gene *cfbpB* (*cj0174c*) encodes a putative iron-uptake ABC transport system permease of the ferri- transferrin uptake system and the third *cj1721c* encodes an uncharacterised beta-barrel outer membrane protein (**Appendix 5**). Interestingly, high number (247) RG-1 core genes were not included in the 1343 defined cgMLST (**Appendix 6**). This can be attributed to the phylogenetic distance of this closely related group of RG-1 strains from other *C. jejuni* isolates. Among these genes, 27 (18.2%) genes were uncharacterised, 23 (15.5%) genes encoded putative periplasmic proteins, 16 (10.8%) genes encoded putative membrane proteins, and 5 (3.3%) genes encoded putative lipoproteins. The rest of the 78 genes had different roles including cell metabolism, flagellar assembly, glycosylation and chemotaxis. In order to identify genes amongst such a high number of genes, which might be more specific to this group, the genomes of the RG-1 strains were compared to different sequence types as explained in the next section.

3.3.3   Potential host specificity genes within RG-1

Any genes absent from more than 5 % of strains would be excluded from the cgMLST-defined group of core genes (Cody et al., 2017). As these are likely to include some genes contributing to niche survival and specific host colonization, the 247 RG-1 core genes, that were absent from the 1343 set of cgMLST genes, were compared to genomes of selected generalist strains. These included reference strains as well as strains of human and rat source belonging to generalist ST-21 and ST-45 clonal complexes (**Table 3.5**). Clinical strains selected were primarily isolates collected in Oxfordshire between 2012 and 2016, the same general geographic area as the Norway rat isolates.   The closed genome of Dg147 and FASTA files of the draft genomes of all strains listed in (**Table 3.5**) were compared using the GView server to create a BLAST atlas (**Figure 3.23**). The Blast atlas showed genomic differences between Dg147 and other sequence types. All genes present in Dg147 and absent in all other strains were identified and checked using Genome comparator (PubMLST/campylobacter) for presence in all RG-1 strains. A list of these unique RG-1 genes is displayed in **Table 3.6**. Only 35 of this subgroup of 247 RG-1 core genes were identified as absent in all generalist strains tested (**Table 3.6**). Therefore, these 35 genes could potentially be RG-1 specific genes related to specific niche adaptation.

Interestingly, nine of these 35 genes encode a putative iron uptake system. The location of these genes in the Dg147 genome was between locus 254 to 262 (7423bp). The NCBI database showed that 92% (7278bp max query) of this DNA segment was identical to the *C. jejuni* subsp. *doylei* 269.97 genome (Accession no: CP000768.1), and 82% (2981bp max query) to the *Campylobacter helveticus* strain ATCC 51209 genome (Accession no: CP020478.1). In addition to this putative iron uptake system, previously characterized transport system for ferrichrome, ferri-rhodotorulic acid, haem and and ferrous iron appear to be complete in all RG-1 strains (**Appendix 7**). However, the key components of the ferri-enterochelin uptake system; Cj0752, TonB3 (Cj0753c), and CfrA (Cj0755), and components of the ferri-transferrins iron uptake system; CfbpB (Cj0174c), Cj0177, Cj0178, ExbB1 (Cj0179), ExbD1 (Cj0180), and TonB1 (Cj0181) were absent in all RG-1 isolates (**Appendix 7**) (Naikare et al., 2006, Miller et al., 2009). The additional annotated TonB dependant outer membrane gene product Cj0444 is present in all.

Furthermore, a *cdtABC* operon was one of the other operons in the 35 genes identified (locus00283-285 of **Table 3.6**). Only 64% identity was found between the CdtB amino acid sequence of Dg147 and NCTC11168; whereas, it shared 81 % identity (100 % query cover) with *C. coli* ATCC43478 and 88 % identity (94% query cover) with *C. lari.* RM16701CdtA and CdtC from Dg147 were found to be more distantly related, sharing only 44% (99% query cover) and 54 % (88 % query cover) identity with the corresponding gene products in the NCTC11168 genome, 68 % (94% query cover) and 67 % (91% query) respectively with *C. coli* ATCC43478 and 77 % (99 % cover) and 90 % (91 % query) identity respectively with *C. lari.* While the *C. lari ctdABC* locus appears closest to the RG-1 *cdtABC* locus, *C. lari* RM16701 CtdA has in-frame deletions compared to CdtA of Dg147 (**Appendix 8**). Functional studies would be interesting to compare cell specificity and toxicity of CdtABC toxin from these different species. Another potential virulence determinant, classified as RG-1 specific belongs to the two-partner TypeVb autotransporter family (Guerin, J. et al., 2017). Locus 01659 was annotated as belonging to the ShlA/HecA/FhaA family of adhesion/heme binding proteins and locus 01658 as a helper TpsB subunit. Another protein was a putative acetyltransferase. Activity of this enzyme may have implications for the survival of RG-1 in diverse environments.

Table 3.5: Sequence types of *C. jejuni* from human and rat sources used to compare with the Dg147 genome.

Isolate	Year of isolation	ID	Source	**Sequence type (ST)	**Clonal complex (CC)
* NCTC11168	1977	48	human stool	43	21
* 81116 (NCTC11828)	1981	1220	human stool	267	283
* 81-176 (CP000538)	1981	1481	human stool	604	42
Dg370	2012	26041	rat	50	21
Dg7	2011	25965	rat	21	21
Dg16	2011	25983	rat	21	21
Dg361	2012	25994	rat	21	21
Dg375	2012	26048	rat	19	21
Dg194	2012	26061	rat	21	21
Dg200	2012	25985	rat	45	45
Dg153	2012	24195	rat	42	42
OXC6895	2012	22090	human stool	45	45
OXC6953	2012	21127	human stool	21	21
OXC6949	2012	21123	human stool	50	21
OXC7793	2013	24596	human stool	45	45
H135060411	2013	2667	human stool	21	21
OXC8586	2013	28922	human stool	50	21
OXC8678	2014	29010	human stool	21	21
H140620012	2014	2670	human stool	50	21
H143840383	2014	31012	human stool	45	45
OXC9424	2015	31243	human stool	21	21
OXC9640	2015	31459	human stool	45	45
OXC9440	2015	31259	human stool	50	21
OXC10113	2016	43862	human stool	21	21
NWC182	2016	43880	human stool	45	45
NWC325	2016	47897	human stool	50	21

* NCTC11168, 81116 and 81-176 are well characterised closed genomes strains. ** ST-42 and ST-45clonal complexes, ST-50 ST-21 clonal complex and ST 21 ST-21 clonal complex of rat and human associated strains.
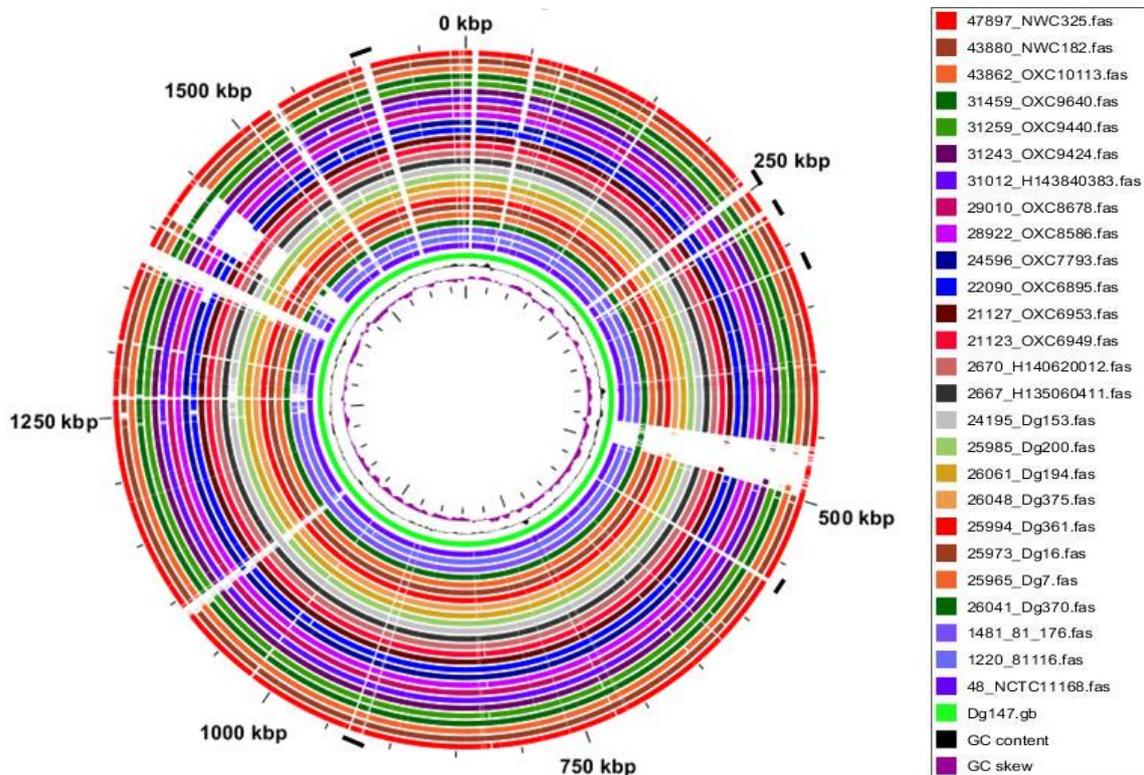
Figure 3.23: Blast atlas of genome sequences from the *C. jejuni* Dg147 reference strain compared to different clonal complex strains. The annotated WGS of Dg147 was blasted with draft genomes of other strains listed in **Table 3.5** using GView Server https://server.gview.ca (Petkau, A. et al., 2010).. White regions in the Blast atlas represent genes absent in some or all of the selected strains.. The location of Dg147 genes missing in all other strains is shown as a dash in this figure.

Table 3.6: Potential RG-1 specific genes identified using GView Server (Figure 3.23)

Locus number in Dg147 genome	Product	Sequence length	Genome position
locus00056	hypothetical protein	777	74640
locus00057	hypothetical protein	585	75437
* locus00254	hypothetical protein	846	243310
* locus00255	Pseudoazurin	438	244155
* locus00256	ABC transporter, ATP-binding protein (cluster 8, B12/iron complex)	780	244576
* locus00257	heme/hemin ABC transporter, permease protein	855	245545
* locus00258	periplasmic binding protein	1095	246389
* locus00259	periplasmic binding protein	1107	247480
* locus00260	Major outer membrane protein	327	248597
* locus00261	Major outer membrane protein	915	249001
* locus00262	Probable tautomerase	207	250072
locus00280	hypothetical protein	576	267862
locus00281	hypothetical protein	1386	268439
** locus00283	Cytolethal distending toxin subunit C	573	270370
** locus00284	Cytolethal distending toxin subunit B, DNase I-like	804	270952
** locus00285	Cytolethal distending toxin subunit A	795	271752
locus00290	putative acetyltransferase	228	275966
locus00335	putative membrane protein	597	314989
locus00336	Membrane protein	714	315566
locus00337	Membrane protein	552	316299
locus00607	hypothetical protein	930	565549
locus00608	hypothetical protein	1671	566520
locus00609	hypothetical protein	423	568206
locus00961	hypothetical protein	312	922834
locus00974	YhbX/YhjW/YijP/YjdB family protein	1188	932326
locus01436	CMP-N-acetylneuraminate-beta-galactosamide- alpha-2,3-sialyltransferase	1413	1402740
locus01437	hypothetical protein	315	1404133
locus01438	Putative acyl carrier protein	219	1404444

locus01576	aminoglycoside 6-adenylyltransferase	192	1527074
locus01577	aminoglycoside 6-adenylyltransferase	576	1527351
locus01578	hypothetical protein	1278	1528079
locus01643	Hemerythrin domain protein	642	1591887
locus01658	Channel-forming transporter/cytolysins activator of TpsB family	1653	1603992
locus01659	Putative large exoprotein involved in heme utilisation or adhesion of ShlA/HecA/FhaA family	5577	1605653
locus01743	hypothetical protein	963	1680263

* putative iron uptake system. ** *cdt* genes.

Another comparison was done to find genes which are absent in all RG-1 strains, but present in all of the strains that are listed in **Table 3.5**. A low number of 23 genes were absent in all RG-1. Among these genes, there was a group of genes between *cj0727* to *cj0736* encoding a putative ABC transport system (**Appendix 9**). Genes of this group might have an essential role in supporting the isolates ability to colonise a host. To investigate this hypothesis, *in vitro* chicken and *Galleria mellonella* trials were performed.

The Venn diagram below (**Figure 3.24**) was made to present a summary of the genomic comparison of RG-1 strains with cgMLST and different sequence types of *C. jejuni* isolates as listed in **Table 3.5**. The majority of RG-1 genes are also cgMLST (1314 genes). In addition, 269 genes were accessory genes in RG-1, but importantly 247 of the RG-1 core genes were not cgMLST and only 35 genes were not in the different sequence types of *Campylobacter* isolates. The last two groups of genes are more important and should be considered with further analysis.
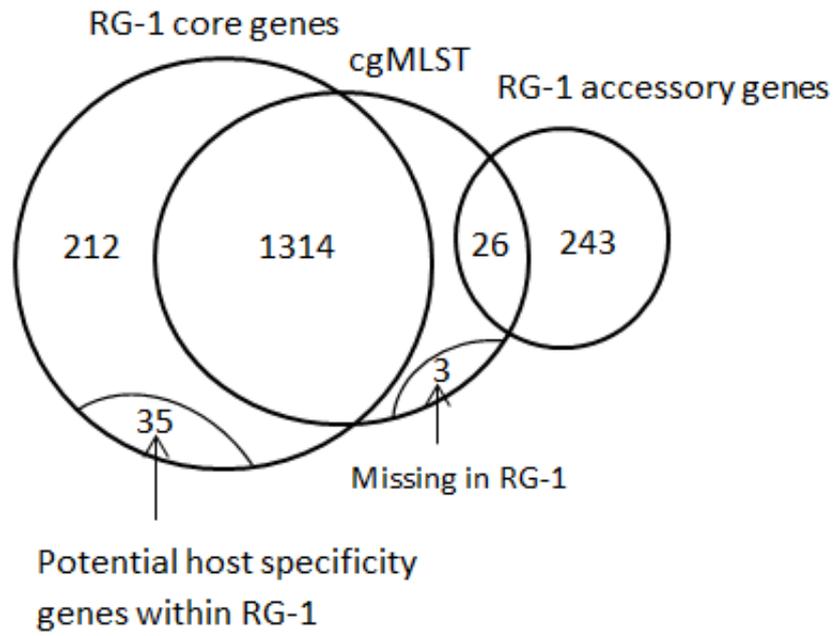
Figure 3.24: Venn diagram summarising RG-1 potentially rat specific, core, and accessory genes.

3.4 Conclusion

WGS and rMLST phylogenetic analysis had previously identified 30 RG-1 *Campylobacter* strains as a closely related phylogenetic clade (**Figure 1.11** at above). This group exhibited typical *Campylobacter* characteristics of microaerophilic growth requirements, spiral morphology and motility with bipolar flagella. However, on phylogenetic analysis this group is clearly distinct from the generalist ST-21 and ST-45 clonal complexes of *C. jejuni*. The 16S rRNA analysis confirmed that the RG-1 strains cluster with the *C. jejuni/ C. coli* group and are clearly distinct from other species. The presence of the *hipO* and *napA* gene markers further distinguishes the RG-1 group from both *C. coli* and *C. jejuni* subsp. *doylei*. The evidence assigns RG-1 group within *C. jejuni* species. However, 3 genes of the cgMLST defined genes of *C. jejuni/ C. coli* are missing from all of the RG-1 strains and 2.1% missing or incomplete in one or more strains. More detailed genetic and phenotypic analysis is required to clearly assign this new clade of rat associated bacteria to *C. jejuni* subsp. *jejuni* or to a new subspecies.

RG-1 strains were treated with the motility MHA medium in order to assist adaptation to the laboratory conditions. This treatment enhanced the bacterial growth by ~5-7 times and in many cases doubled motility. Even though the RG-1 cells had adapted to the laboratory conditions, the growth and motility properties were still 2 or 3 times less in comparison to the reference strains. Nevertheless, RG-1 cells grew as well as reference strains when they were grown in the MHB rich medium supplemented with lab rat mucin. This is likely due to a particular nutrient/growth factor within the rat mucin and deserves further investigation. Biolog in combination with i*n vitro* growth studies confirmed that the RG-1 strains utilised the key nutrients Ser, Asp, Asn, Glu, Gln, Pro, succinic acid, formic acid, and pyruvic acid, rather than carbohydrates, typical of *C. jejuni* catabolism (Guccione et al., 2008, Leach et al., 1997). Ser and Asp were used first followed by Glu and lastly Pro (Velayudhan et al., 2004). This order of amino acid preference is similar to the well-studied NCTC11168. One of the clear differences between RG-1 and the NCTC11168 was that RG-1 did not utilise L-fucose due to absence of the genomic island (*cj0480c-cj0480*), which is required for uptake and metabolism of this carbon source. It is known that this

pathway is absent in approximately 50% of the sequenced *C. jejuni* and *C. coli* strains, including both human gastroenteritis causing *C. jejuni* 81-176 and 81116 strains (Dwivedi et al., 2016).

RG-1 strains share the majority of their genome content (1561 core genes, 85.3%). Meanwhile, 41.6% of accessory genes (112/269) were from two large uncharacterised phage insertions. The remainder of the accessory genes was largely related to motility, chemotaxis, membrane proteins and transport proteins. These genes could be important for evolution of the RG-1 strains to a specific host environment. Only 2.2% of the RG-1 core genes (35 genes) might be rat specific genes. These were absent from the published cgMLST defined core genes of clinically associated *C. jejuni/coli* and were also absent from rat strains belonging to classic clonal complexes, CC-21, CC-45, CC42. A group of 9 of these genes are related to the uptake of iron. This pathway was not identical to the identified 7 class of iron uptake systems in *Campylobacter* (**Appendix 7**). However, based on percentage identity, this group of genes might have been acquired from *C. jejuni* subsp. *doylei* 269.97 (92% identity) or *C. helveticus* (82% identity). On the other hand, RG-1 strains are missing three genes of the ferric-enterochelin uptake system including the *cfrA* gene and five genes of the ferri-transferrin uptake system, including the *cj0178* gene. Different *in vivo* studies have demonstrated that *C. jejuni* NCTC11168 requires both *cfrA* and *cj0178* to be able to survive and colonise the intestines of animals (Palyada et al., 2004, Zeng et al., 2009). In addition to differences in iron uptake, the *cdtABC* of RG-1 strains is much closer to *C. lari* than to the *C. jejuni cdt* locus. It is well known that *Campylobacter* spp. are commensal bacteria to poultry. Even though many *C. jejuni* strains from broiler chickens have *cdt* genes, they do not cause disease in the birds (Bang et al., 2001).

On the other hand, RG-1 lacks some important genes. Vitamin B_5 biosynthesis has been identified as a host specificity factor (Sheppard et al., 2013). Interestingly, *panBCD* genes were absent in all RG-1 and RG-2 strains, but present in the other ST-21, ST-22 and ST-45 clonal complexes found in farm associated rats. Sheppard et al. (2013) found that *Campylobacter* isolates from cattle have the required pathway, but the operon was not found in chicken associated strains. They proposed that chickens could obtain the nutrients from rich vitamin B_5 cereals and grain, and therefore they do not need to synthesise it

themselves. This may be true for the RG-1 strains. They could get vitamin B_5 from the typical grain diet of rats. Lastly, another implicated host specificity factor, the fucose utilisation operon was not found in the RG-1 genomes. Mutation in the fucose transporter inhibits uptake of fucose and reduce colonisation levels in a piglet model, but without presenting a significant difference in a chicken model (Stahl et al., 2011). In a different case study, the diet of chicken was supplemented with fucose, which increased the level of colonisation by the wild type strain. The *fuc*-positive NCTC1168 also formed a more sessile structure of biofilm than *fuc*-negative 81116 (Dwivedi et al., 2016). Therefore the ability to use fucose might allow the bacterial isolates to survive better in pigs over other hosts. Similarly, RG-1 strains might not be able to colonise pigs, suggesting that they have adapted to other specific hosts.

Analysis of the draft genomes of both biofilm and non-biofilm formers for selected genes have shown that the keys in biofilm formation did not provide any information as to why Dg153 was such an efficient biofilm former. Flagellar complex *flaA*, *flaB*, *flaG*, *fliA*, *fliD*, *flgG*, *flgG2*, *maf5,* chemotactic c*heA*, oxidative stress responses *groEL, groES*, *tpx*, capsular polysaccharide *kpsM*, and adhesins *peb1*, and *flaC* genes, were all present (Kalmokoff et al., 2006). The inability of RG-1 strains to form a biofilm microcolony may limit their prevalence amongst different hosts (Joshua et al., 2006).

To conclude, draft genomes, mass spectrometry, and 16S rRNA have confirmed that the RG-1 strains are *C. jejuni* and presumably subsp. *jejuni*. Approximately 15% of the RG-1 core genes (247) were not cgMLST. Among these genes there are some operons or genes that might contribute to host specificity such as an alternate putative iron uptake system, more distantly related *cdt* genes, gene encoding ShlA/HecA/FhaA exoproteins, putative acetyltransferase and cytolysins. RG-1 strains grew poorly in rich medium compared to other differing sequence types, were motile, but under the conditions tested did not form biofilm on the glass membrane filter. Growth of poorly growing RG-1 strains was encouraged through the addition of rat mucins to the rich medium. The loss of some host specific operons and the improved growth of these RG-1 isolates in the rat mucin suggests that this group of *C. jejuni* strains may not be able to survive in different hosts and may be restricted to specific hosts. To further investigate this, selected RG-1 strains were studied

in both *Galleria melonella* and Ross chicken *in vivo* models, as shown in chapter 5. The next chapter will be on the ability of RG-2 strains to utilise glucose.

Chapter 4

Characterisation of the glucose utilising RG-2 Group

4.1 Introduction

The main carbon and energy sources of *C. jejuni* are amino acids, pyruvate, lactate and TCA cycle intermediates (Hofreuter et al., 2012). In rich media, NCTC11168 utilised only the following amino acids, Ser (Velayudhan et al., 2004) Asp, Glu (Guccione et al., 2008, Leon-Kempis et al., 2006), and Pro (Leach et al., 1997). The strain 81116 was shown to preferentially use Ser and Asp when it is grown in rich medium and only later shift to Glu and Pro utilisation (Leach et al., 1997). In addition, some strains such as 81-176 can grow in a medium containing only Asn and Gln as carbon source (Hofreuter et al., 2008). Most *Campylobacter* cannot use glucose and are defined as non-glycolytic due to lack of the glycolytic enzyme 6-phosphofructokinase and incomplete Embden-Meyerhof-Parnas pathway (EMP) (Velayudhan and Kelly, 2002, Szymanski, 2015). In contrast to this dogma, more than 50% of strains of *C. jejuni*, isolated from farmed animals, have been shown to harbour a fucose utilisation pathway (Dwivedi et al., 2016, Muraoka and Zhang, 2011) and approximately ten years ago, genes encoding enzymes of the Entner-Doudoroff pathway were identified in the genome of *C. jejuni* subsp. *doylei* 269.97 (Miller, 2008).

With the exponential increase in WGS data, more widespread possession of the ED pathway by *Campylobacter* has been very recently identified (Vorwerk et al., 2015, Vegge et al., 2016). In 2015, functionality of the ED pathway in a human isolate of *C. coli*, strain CHW470, was demonstrated using ^{13}C-labelled glucose (Vorwerk et al., 2015). The *glc* locus (encodes enzymes of the ED pathway) was shown to be inserted into a ribosomal RNA gene cluster of *C. coli* CHW470 and could be transferred to an ED-negative strain of *C. coli* by natural transformation. In 2016, Vegge and coworkers (Vegge et al., 2016) published a survey of presence and analysis of the *glc* locus in draft *C. jejuni/coli* genomes submitted to the BIGSdb database. They identified the *glc* locus in only 1.7% of more than 6000 genomes from various environmental, clinical and animal sources. Alleles of each gene in the *glc* locus were classified according to their DNA sequence in the same way as MLST sequence types were defined (Vegge et al., 2016). Any new sequence, which was not identical to an existing sequence in the database, was defined as a new allele and was given a new number. ED types were then identified from the allelic profiles of all seven

different digit numbers of the ED pathway genes. Thirty-six types of the ED pathway were identified in *C. jejuni* and ten in *C. coli*. The predominant ED type 1 was found in *C. coli*. ED type 1 has the allelic profile (3-5-7-9-5-9-9) corresponding to seven genes of the *glc* locus alleles in the order *glcP* to *eda*. ED type 2 was identified in 18 *C. jejuni* farm-associated rat isolates. It has the allelic profile of 3-4-4-5-4-5-5. There were several additional ED types identified in the rat isolate collection, but most of the other 19 different ED types were from wild birds and were very varied. The benefit of glucose utilisation to *C. coli* has been addressed. In some strains glucose stimulated biofilm production, in others glucose stimulated metabolism. Enhanced survival in stationary phase was common to all strains (Vegge et al., 2016).

The predominant *C. jejuni* ED-positive isolates identified in PubMLST/*Campylobacter* database were from wild birds and Norway rats, with 51% of the total from the Reading strain collection (Vegge et al., 2016). None were from human disease or chickens suggesting that glucose utilisation may have some advantage to environmental rather than clinical/ agricultural strains. Thus, the benefit of this locus to the bacterium and its correlation with niche and or host association remained unclear. Importantly, whilst two studies had demonstrated glucose metabolism by *C. coli,* none had demonstrated a functional ED pathway in the more important pathogen, *C. jejuni*. Hence, this chapter focuses on functionality of the *glc* locus in *C. jejuni*. It provides the first demonstration of glucose utilisation by *C. jejuni*, highlights the preference of these strains for amino acids over glucose as carbon source, and highlights the interesting phenomenon that the *glc* locus is located within two of the three *rrn* operons in the two RG-2 strains studied in detail.

Results and Discussion

4.2 RG-2 strain selection for *in vitro* study

A high percentage of the farm-associated rat *C. jejuni* isolates harbour a *glc* locus; 39 of the 116 sequenced strains (32.7%). The majority of these strains (27) have been assigned as CC-45, but to differentiate these strains from other closely related ED-negative CC-45 strains, these ED-positive strains are referred to as RG-2 group. Based on WGS, rMLST, and cgMLST phylogeny data it is evident that the RG-2 strains cluster together, close to the ED-negative CC-45 of generalist strains and clearly distinct from the other unique clade of rat isolates RG-1 (ED-negative, 25.8% of *C. jejuni*) (**Figure 4.1** and **Figure 3.1, B** at above). Only *aspA* and *pgm* (*glmM*) MLST allele numbers of the RG-2 group were different from the MLST profiles of the ED-negative CC-45 strains. Based on the 16S rRNA sequence, RG-2 strains are clustered with *C. jejuni* subsp *jejuni, C. coli*, and *C. jejuni* subsp. *doylei* (**Figure 3.1, A** at above). The V3-V5 region of the 16S rRNA of RG-2 strains was identical to that of NCTC11168 and to the *C. coli* strains Dg349 and Dg172. All strains possess the *hipO* and *nap* genes, which are markers for *C. jejuni* and *C. jejuni* subsp *jejuni,* respectively (Wang et al., 2002, Miller et al., 2007), and lack *prpE* and *prpC*, two *C. coli* markers (Wagley et al., 2014). Interestingly, while TEM analysis of the RG-2 strains revealed a curved biflagellate cell morphology typical of *C. jejuni* strains (Buck et al., 1983, Pead, 1979), these isolates appeared straighter with a lower average amplitude than many of the classic isolates (**Figure 3.5** at above).

RG-2 strains were isolated from seven different farm sites over 3 years (**Table 2.1, B** at above). Twenty-three of these strains harboured either ED type 2 or 3, and were isolated from farm sites (C, E, L, X) and (G, H), respectively (**Figure 4.1**). These ED types shared 6 of the 7 genes of the *glc* locus, with a single allelic difference in *pgl,* ED type 2 has allele number 5 and ED type 3 allele number 6. The other 16 ED-positive strains harboured 8 different ED types from the following farm sites; C, E, G, H, I and X. The ED-negative CC-45 strains were isolated from similar farm sites (G, C, I, X, J). As the purpose of this study, was to demonstrate functionality and characterise the ED pathway of *C. jejuni*, strains carrying the

common ED types 2 and 3 were selected for the study along with the ED-negative CC-45 control strain, Dg200 strain.
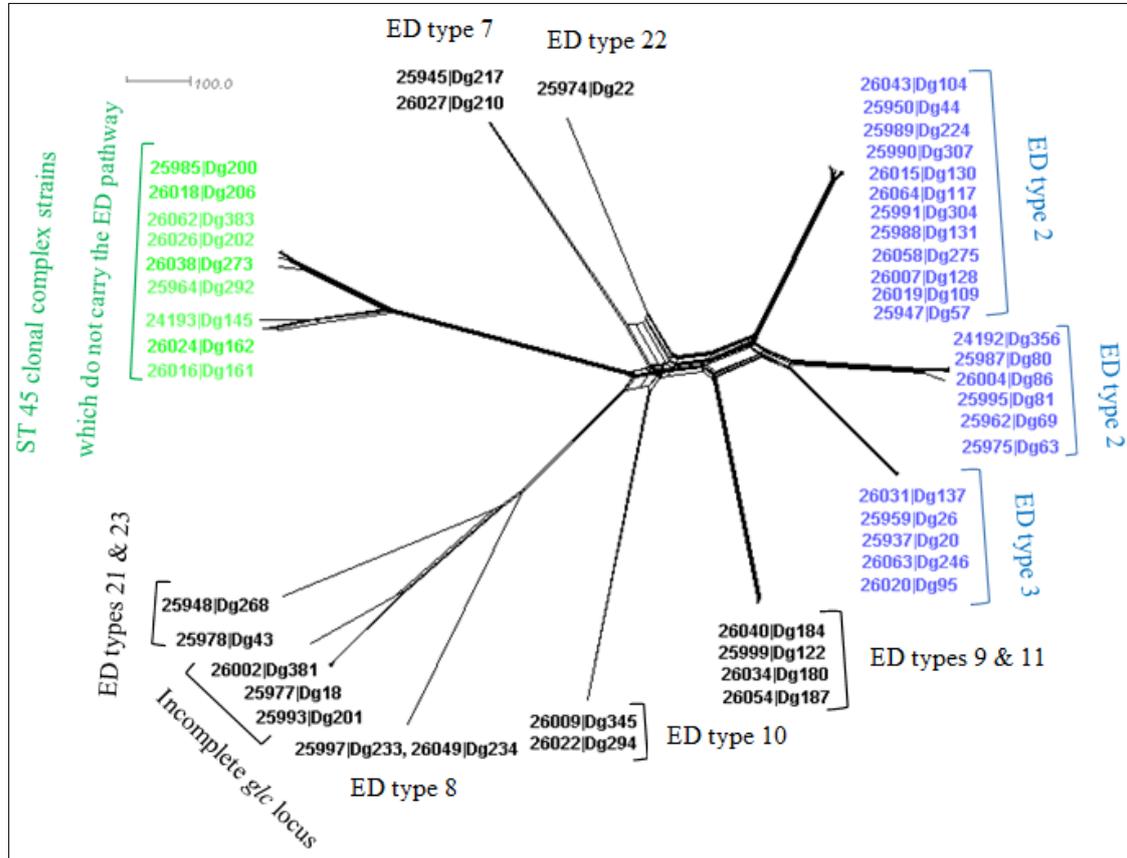
Figure 4.1: Phylogeny of ED groups of *glc* positive farm-associated rat *Campylobacter jejuni* strains. The phylogenetic network was generated based on a cgMLST (95%) comparison of 9 ED-negative *C. jejuni* rat associated strains, CC-45 strains (colour-coded green), plus 39 ED-positive *C. jejuni* rat associated strains (colour-coded blue and black). The *glk* gene of Dg381, Dg18, and Dg201was incomplete (i.e., at the ends of contigs) as well as *eda* and *edd* genes of Dg381and Dg201. Therefore, the ED types of these strains have not been assigned. Dg161 CC-45 had a complete *eda* gene but incomplete *edd*, *pgl*, *glk* and *pgi* genes, and also lacks *zwf* and *glcP* genes. NCTC11168 was used as a reference strain (Cody et al., 2017). SplitsTree4 was used to visualise the results of the phylogenetic network (Huson and Bryant, 2006). ED allele types of the phylogenetic branches are shown by numbers (Vegge et al., 2016).

4.3 Growth and motility properties of RG- 2

As for the RG-1 group, the established motility and growth assays were performed to adapt the RG-2 strains to laboratory growth. RM-1 strains were stocked from MHA motility plates as for RG-1 (**see section 2.4.2**). Growth of RM-1 strains (post one round of growth on MHA motility plates) was compared to growth of the original inoculum in MHB in the microtitre plate growth assay. Interestingly, growth of RG-2 isolates was not enhanced (**Figure 4.2, A**). This was in contrast to RG-1 strains in which growth was enhanced about 3-5 fold following growth on the MHA motility plates. This might be explained by the presence of ~3mM glucose in MHB supporting the growth of the RG-2 strains. RG-2 strains still grew poorly, about 2-3 fold less than the Dg200 strain. Motility of Dg95, Dg275 and Dg356 was enhanced on the second round of motility, but was still much lower than that of the ED-negative strain, Dg200 CC-45 (**Figure 4.2**, **B**). As with RG-1 strains, better motility appeared to correlate with better growth (**Figure 3.12** at above). TEM micrographs confirmed that both Dg275 and Dg95 strains have bipolar flagella (**Figure 3.5** at above). The adapted RM-1 stocks were subsequently used in all assays.
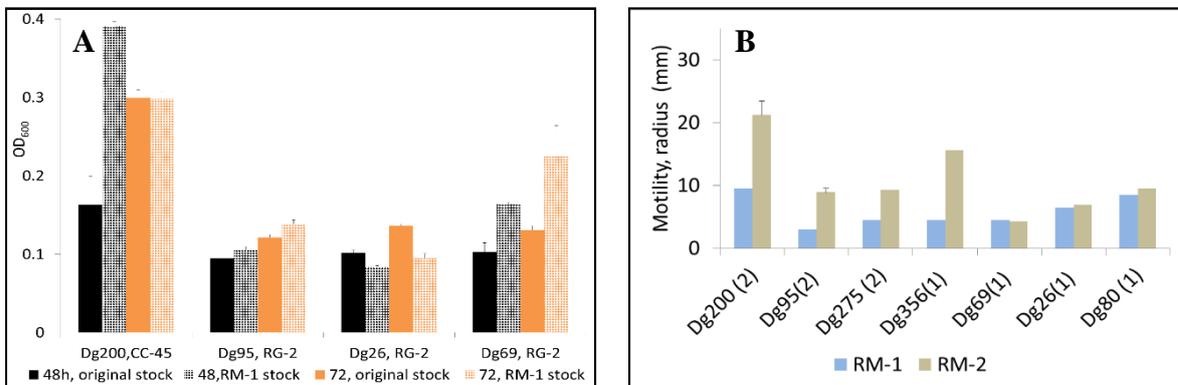
Figure 4.2: Exposure of RG-2 strains to the motility MHA medium. (**A**) Cells were subcultured from either the original stocked cultures (black) or from the RM-1 stocks (orange) onto BA plates, and incubated in MHB in microtiter plates growth assay (**see section 2.2.6**). Growth was measured at 48h and 72h. Solid bars show growth at 48h and dotted bars show growth at 72h. (**B**) Motility of RG-2 strains**.** The assay was performed as explained in **section 2.2.4**. Triplicate technical repeats were performed on all second round tests. Brackets indicate the number of biological repeats for RM-2 tests.

Growth properties of 13 RG-2 strains and the Dg200 control strain were studied in two types of rich media, MHB and BHI, using a microtiter plate growth assay. The RG-2 strains grew poorly in MHB, to an OD_{600} of only 0.08 to 0.23 by 48h, in contrast to the Dg200 control, which grew to approximately twice this density by 48h (**Figure 4.3, A**). Similarly, the control strain grew better in BHI at the early stage of growth but then the RG-2 strains continued to grow, reaching a much higher density than in MHB. (**Figure 4.3, B**). Free amino acids (**Appendix 10**) and glucose were monitored in both MHB and BHI. There was no relevant difference in the concentrations of preferred amino acids; Ser, Asp, Glu and Pro between the two media, and BHI contained only 0.8mM more glucose than MHB ie 3.8 mM. It would seem unlikely that this small difference in glucose concentration alone would have such a large impact on growth of the ED positive RG-2 strains. It is likely that an additional factor in BHI enhances growth of these strains.
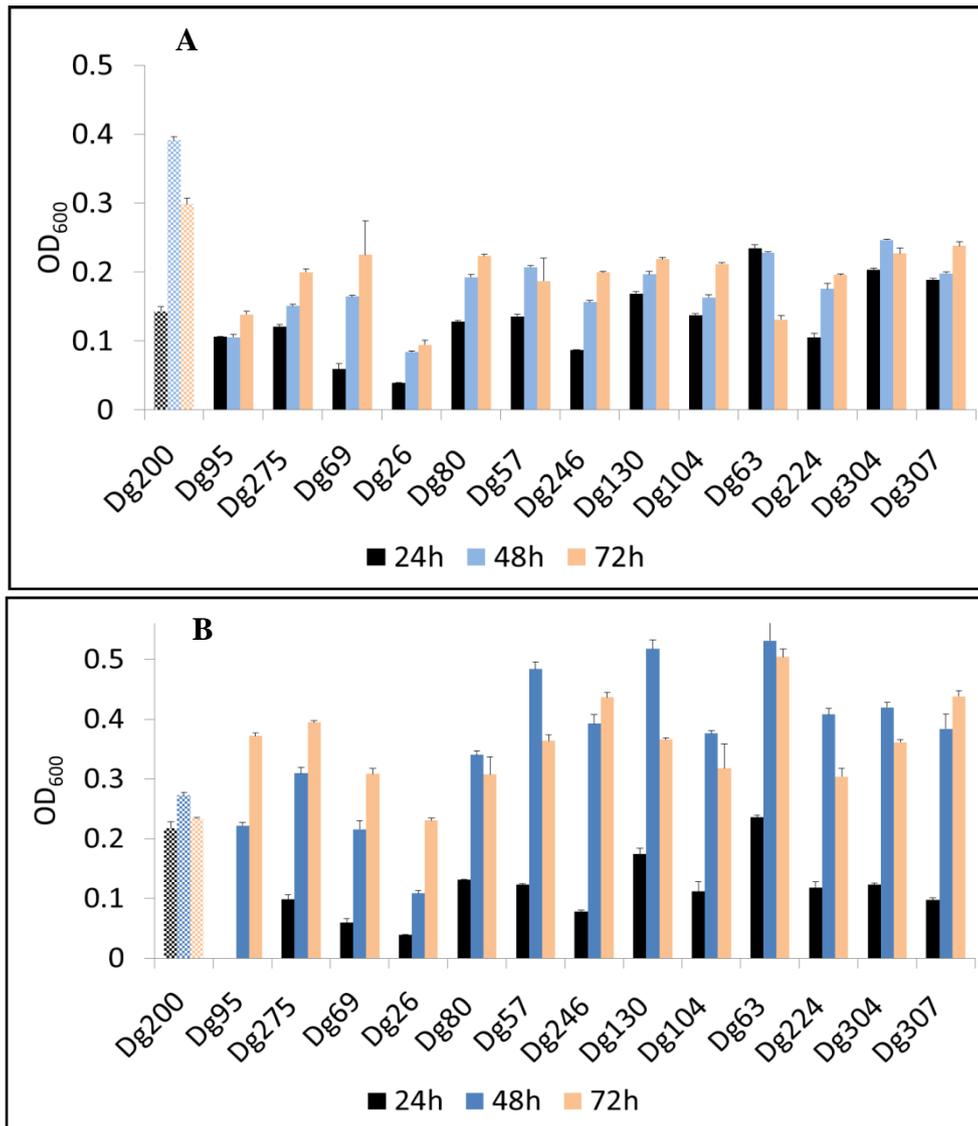
Figure 4.3: Growth of RG-2 strains in rich media. An RM-1 stock was grown on a BA plate for 36-42h. (**A**) MHB and (**B**) BHI media in a microtitre plate assay inoculated with 20ul standard 0.02OD suspension of culture. Cell density at 24h, 48h and 72h are shown. Pattern fill represents Dg200 control ED-negative CC-45; Solid fill represents RG-2 strains. Error bars derived from triplicate technical repeats.

4.4 Growth of RG-2 strains on different amino acids

To test the ability of RG-2 strains to catabolise different amino acids, Dg95 ED type 3 and Dg275, Dg57, and Dg80 ED type 2 of the RG-2 strains were grown in MEM-FBS medium supplemented with single amino acids. The established MEM-FBS medium contained ~3mM glucose, 1.5mM Glu from FBS, and 2mM Gln from the MEM-α itself (**Appendix 10**), which bacteria could use for basal growth. Both ED-positive and negative strains grew to 0.09-0.2 OD by 72h in this basal medium. For all strains, growth was enhanced by inoculation with 20mM of any of the following amino acids; Pro, Ser, Asp, Glu (**Figure 4.4**). Enhanced growth was only evident at 72h and was increased by around 0.5 to 2-fold over growth without amino acids supplement. Interestingly, while Ser is the preferred amino acid (**see Figure 4.7** at below), there is an indication that each RG-2 strain grew as well or better in Pro compared to growth in Ser. This could be related to the fact that Pro uptake by *C. jejuni* increases in stationary phase (Wright et al., 2009). Each RG-2 strain also had the ability to utilise Asn and Gln (**Figure 4.4**), consistent with possession of genes encoding the enzymes asparaginase (AnsB) and γ-glutamyl-transpeptidase (GGT) in their genomes. Taken together this data confirms that RG-2 strains use Ser, Asp, Glu and Pro key amino acids similar to the Dg200 and reference 81-176 strain and also Asn and Gln as does 81-176 (Guccione et al., 2008).
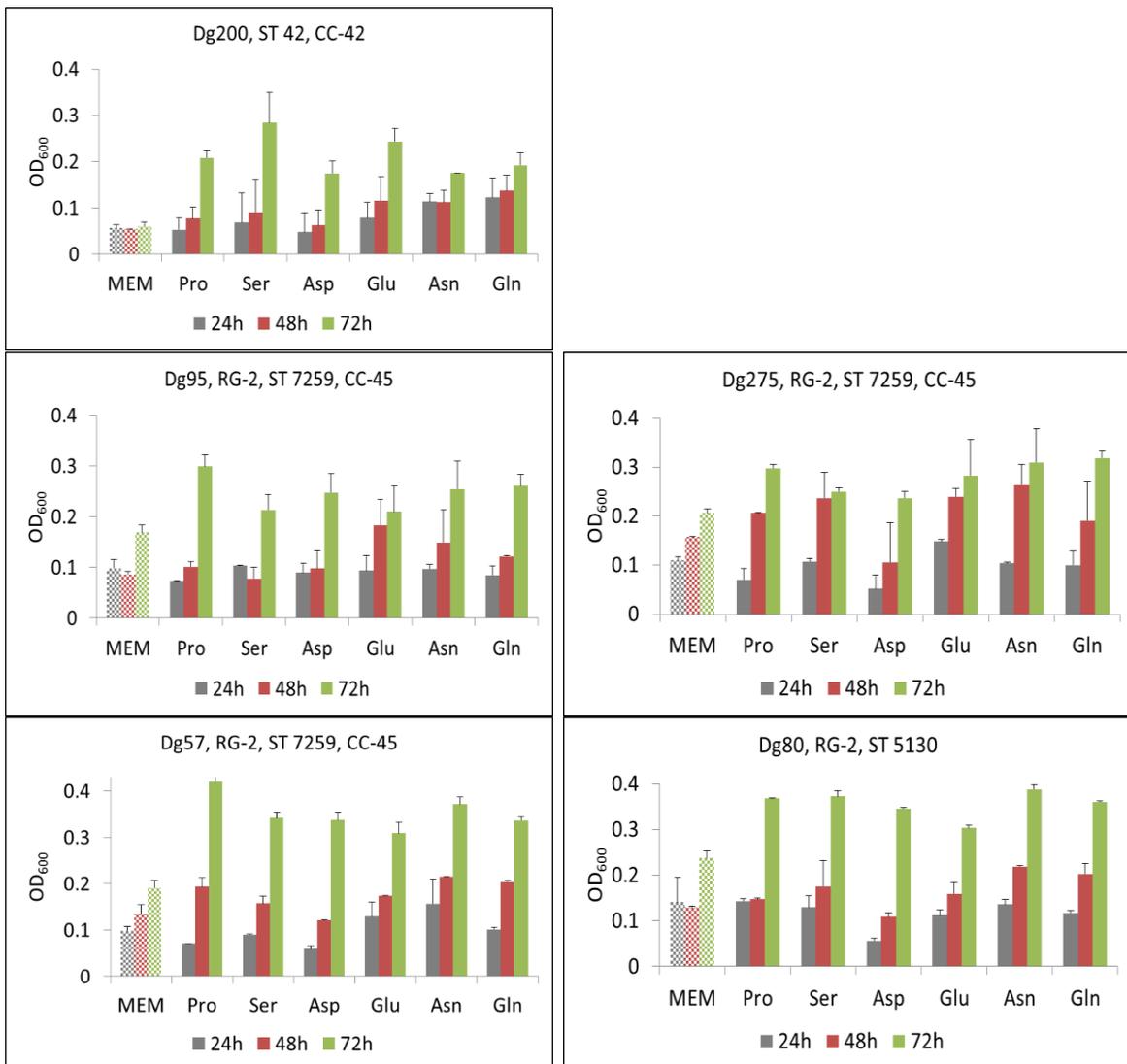
Figure 4.4: Growth of RG-2 in MEM-FBS minimal media supplemented with amino acids.
The RM-1 stocks of strains were subcultured on BA plates, incubated for 30-36h and subcultured on MHA medium for a further 30-36h. Standardised cells, from the MHA plate, were diluted in MEM medium to 0.002OD and 20µl used as inoculum. Microtitre plates contained MEM-FBS medium with 20mM Pro, Ser, Asp, Glu, Asn or Gln, diluted from a 400mM stock. Dotted bars, control MEM-FBS medium without addition of the above carbon sources. Solid bars, growth with amino acid supplement as shown. The error bars of Dg200 and Dg275 are from two biological replicates. Triplicate wells were prepared for each biological replicate.

4.5 Glucose utilisation

4.5.1 Screening *glc* positive strains for glucose utilisation

Fourteen *C. jejuni* strains (ED types 2 and 3) were screened for their ability to grow on glucose. *C. coli* Dg349, which had already been shown to metabolise glucose (Vegge et al., 2016) was included as a positive control. Strains were tested in a microtitre plate assay with DMEMf, with or without 20mM glucose, incubated for 4 days. Growth was enhanced in all RG-2 *C. jejuni* and *C. coli* strains tested (**Figure 4.5**). These initial trials indicated enhanced growth during the later stationary phase of growth. For example, for Dg275 and Dg95 at 72h, the OD_{600} was 0.17 and 0.06 without glucose, respectively, compared to 0.26 and 0.25 with glucose. This was in contrast to the ED-negative strains for which there was no enhanced growth in glucose. With the same glucose dependent stimulation of growth for all 14 RG-2 strains tested, this provides strong support of functionality for the ED pathway in these farm-associated rat *C. jejuni* strains.
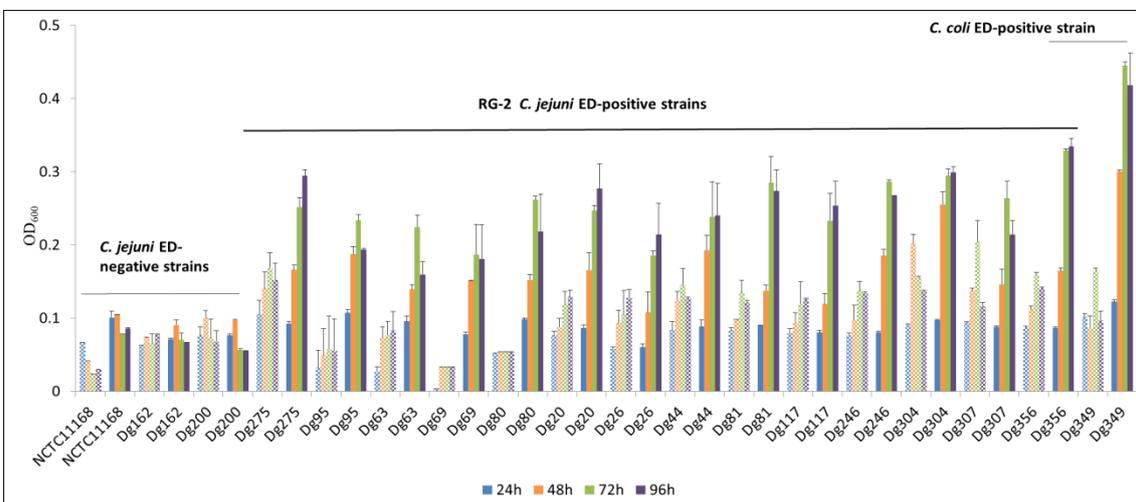
Figure 4.5: Glucose enhanced growth of ED-positive *Campylobacter* spp. Fourteen ED-positive RG-2 *C. jejuni*, one rat *C. coli* ED positive and three ED- negative *C. jejuni* were grown in DMEMf medium with or without 20mM glucose. The cells were grown in a standard microtiter plate growth assay. Stipled bars represent no added glucose, and solid bars represent 20mM glucose. Data for Dg200, Dg275, Dg95, and Dg63 was collected from three biological replicates. Triplicate wells were prepared for each biological replicate.

4.5.2   Preference of ED-positive strains for Ser and Asp over glucose

The preference of the ED-positive *C. jejuni* strain, Dg95, for Ser or glucose was tested by growth in DMEMf plus ~20mM Ser, with or without 20mM glucose, in a flask with slow shaking (150 rpm). Based on optical density, cell numbers increased until 24h, with significant differences between samples grown with and without glucose (P=0.014) (**Figure 4.6, A**). In contrast, viable cell counts indicated entry of cultures into stationary phase around 18h. The continued increase in OD after CFU counts plateau can be attributed to transformation of older/stressed cells to the coccoid form, which continue to contribute to the OD reading. Gram staining confirmed that the majority of cells following growth with Ser but no glucose were coccoid by 36h incubation. This correlated with a drastic reduction in viable cells (2.9 $\log_{10}$ CFU/ml). In contrast, cells grown in the presence of glucose had only approximately 25% coccoid cells and viable counts of around 8.6 $\log_{10}$ CFU/ml at the same time point.

During the exponential phase of growth (18h), approximately 50% of the Ser was utilised in both glucose supplemented and non-supplemented medium (**Figure 4.6, B**). The glucose did not enhance the growth of the bacterial cells during the exponential phase. Based on CFU, the doubling time of Dg95 in the exponential phase of growth was 2.14 ± 0.1h during growth with Ser and glucose, and 2.11 ± 0.2h with Ser alone without significant differences (P value = 0.89). It was not until almost all the Ser and Gln present in DMEMf had been utilised, that this strain started to use glucose. Glucose utilisation began at some point between 18h and 24h, by 24h growth, only 12% of the free glucose had been used (2.5mM). Aside from ~0.15mM of Asp and Glu, Dg95 did not utilise any other amino acid, which were in the medium, prior to using glucose (**Figure 4.6, C**). From 24h to the last time point 48h, glucose levels continued to decrease down to 10 mM. While there was no return to exponential growth in the presence of glucose, cells remained viable greatly extending the stationary phase of growth. In this experiment, the ED-negative strain Dg200 was used to determine if other factors might influence the glucose concentration in the absence of possession of a *glc* locus. There was no decrease in glucose concentration, which remained at 19.6 – 20mM through the 48h growth (**Figure 4.6, A**). This data shows a clear preference of Dg95 for catabolism of Ser over glucose, implying some diauxic effect.
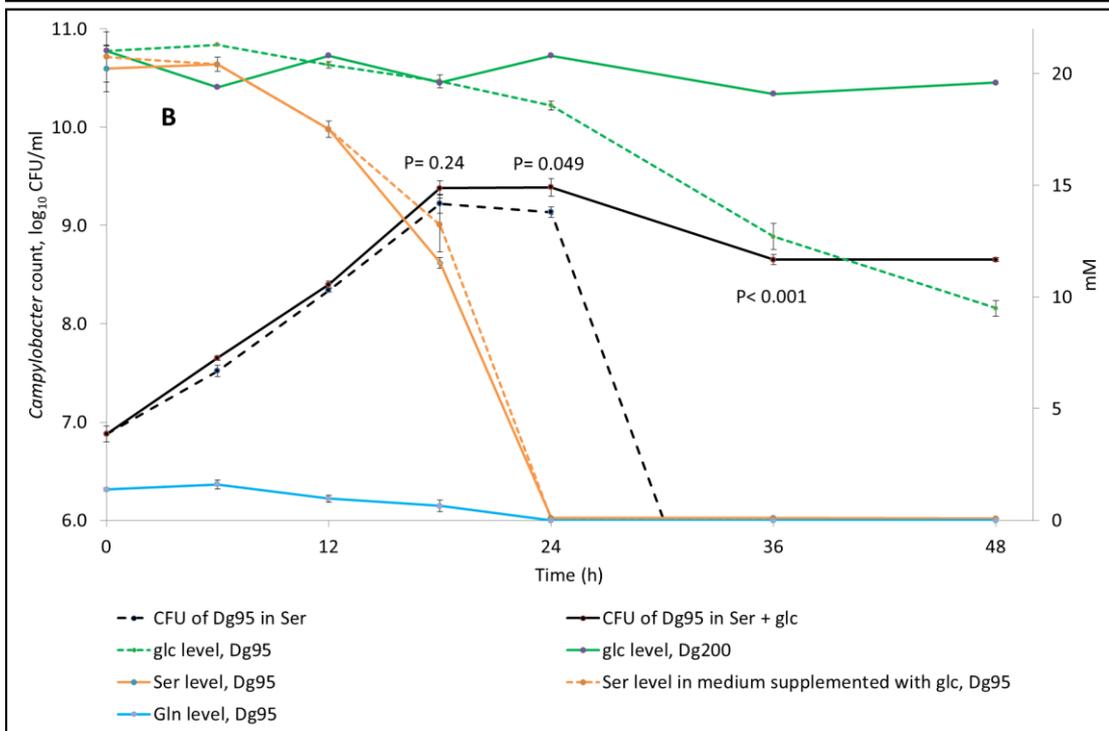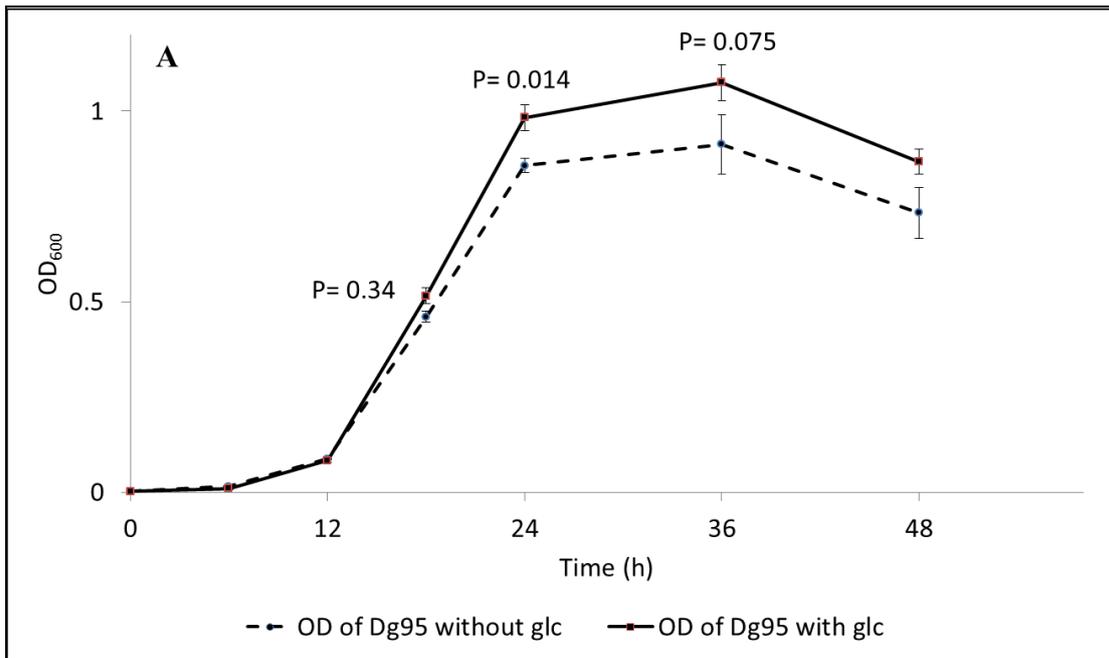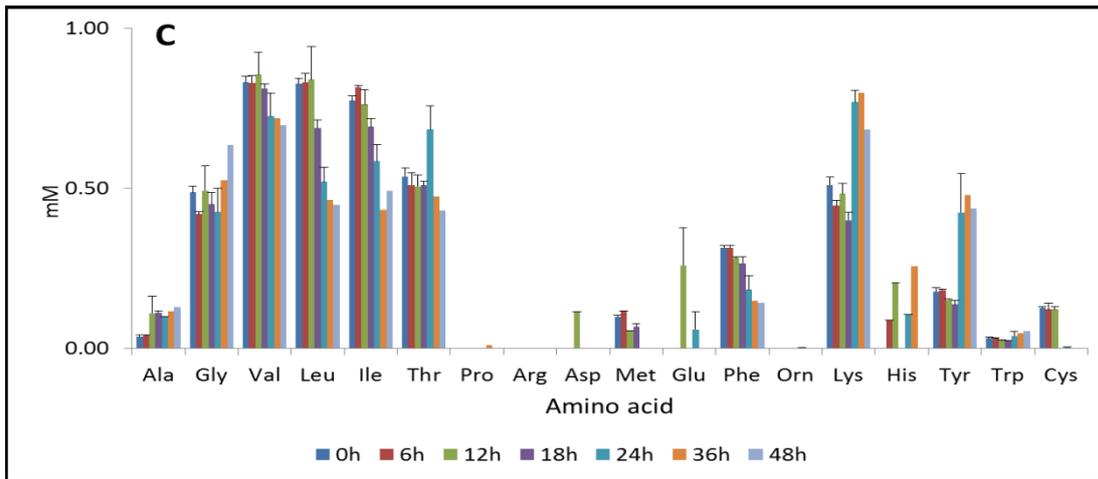
Figure 4.6: Growth of *C. jejuni* Dg95 in DMEMf medium, supplemented with Ser, with or without glucose (glc). *C. jejuni* Dg95 was incubated in 10ml DMEMf medium, 20mM Ser ± 21mM glc, and incubated for 48h in 25ml flasks with shaking at 150rpm. Aliquots were taken during growth, as indicated, and monitored for (**A**) OD and (**B**) viable cells (CFU) plus concentration of glucose, Ser and Gln in culture media. (**C**) Concentrations of other amino acids present in DMEMf. Glucose concentration was quantified by using High-Pressure Anion Chromatography (HPAC) during growth of ED-positive Dg95 and for the ED-negative Dg200 strain (grown as for Dg95, but monitored only for glc levels). Amino acids were monitored using an EZ-Faast-kit followed by GC-MS. Means and error bars show standard errors from three independent biological cultures.

To study the preference of RG-2 strains for different amino acids in the presence of glucose in rich medium, Dg95 and Dg275 were grown in MHB medium supplemented with FBS, with or without glucose. The same sample was used to quantify all amino acids and glucose. Based on OD readings and viable counts, as with minimal media, glucose did not affect growth at the exponential and early stationary growth phase of either strain (**Figure 4.7, A**). Only at late stationary phase (72h), there was a significant difference recorded between the OD of cultures with and without glucose, for Dg95 and Dg275, P= 0.002, P= 0.001, respectively, whereas for the control Dg200 it was P=0.22. Gram staining (data not shown) showed that glucose supported the survival of the majority of spiral cells; whereas, most of the remaining cells in the medium without glucose were transformed to coccoid form by 72h.

157

Interestingly, this experiment clearly highlights the preference of these strains for amino acids. Both Dg275 and Dg95 showed the same pattern of amino acid utilisation. Ser and Asp were preferred and rapidly depleted between 12 and 18h (Dg275) or 24h (Dg95) growth. This corresponded to exponential growth with generation times for Dg200, Dg275, and Dg95 calculated from CFU as: 2.5 ± 0.1h, 2.7 ± 0.2h, and 3.2 ± 0.2 in cultures without glucose and 2.7 ± 0.03h, 2.63 ± 0.09h, and 3.15 ± 0.3 in cultures with glucose, respectively. Glu was utilised primarily following depletion of both Ser and Asp and Pro was used much more slowly. Interestingly, the CFU started to decrease in all cultures once the Glu had been depleted, 36h with Dg275, 36-48h with Dg95 and 36h with Dg200 (**Figure 4.7, B, C, D**), irrespective of the presence or absence of glucose. However, while in the absence of glucose there was a steep death phase, the presence of glucose substantially extended the stationary phase allowing cells to survive for a longer period ($P< 0.001$ for Dg95 and $P=0.001$ for Dg275) even up to 72h when glucose was also essentially depleted. Although all of the above nutrients were utilised by this stage of growth, no tested isolates used any other available amino acids (**Figure 4.7, E, F**). The inability to catabolise other amino acids has also been recorded for other *C. jejuni* strains (Guccione et al., 2008, Wright et al., 2009).

For both Dg95 and Dg275, glucose began to be utilised sometime between 24-36h and was reduced by about 20% by 36h. This was only after Ser and Asp were depleted and use of Glu initiated. This diauxic use of these substrates may be a consequence of some form of catabolite repression of the *glc* locus, with Ser and Asp as preferred nutrients over other amino acids and glucose. Glucose, Glu, and Pro did not appear to affect the utilisation of each other when used in the same cultures. An available glucose source allowed cells to be sustained as viable spiral shapes for a longer period. In both minimal and rich media, glucose extended the stationary phase of growth in the *C. jejuni* RG-2 isolates, which are known to be glucose utilising.
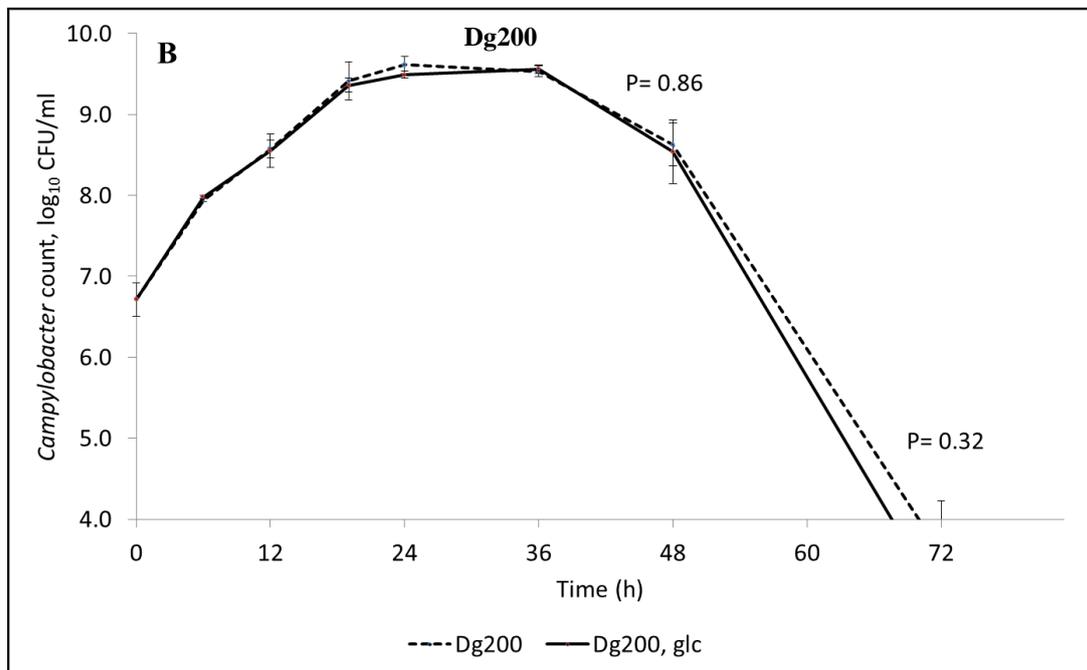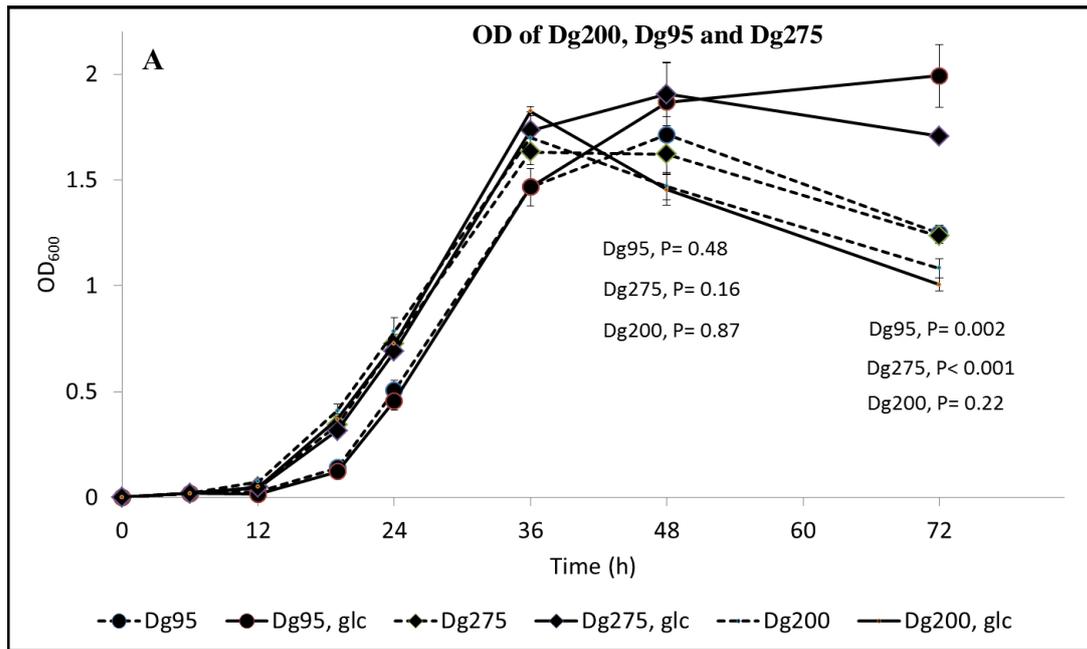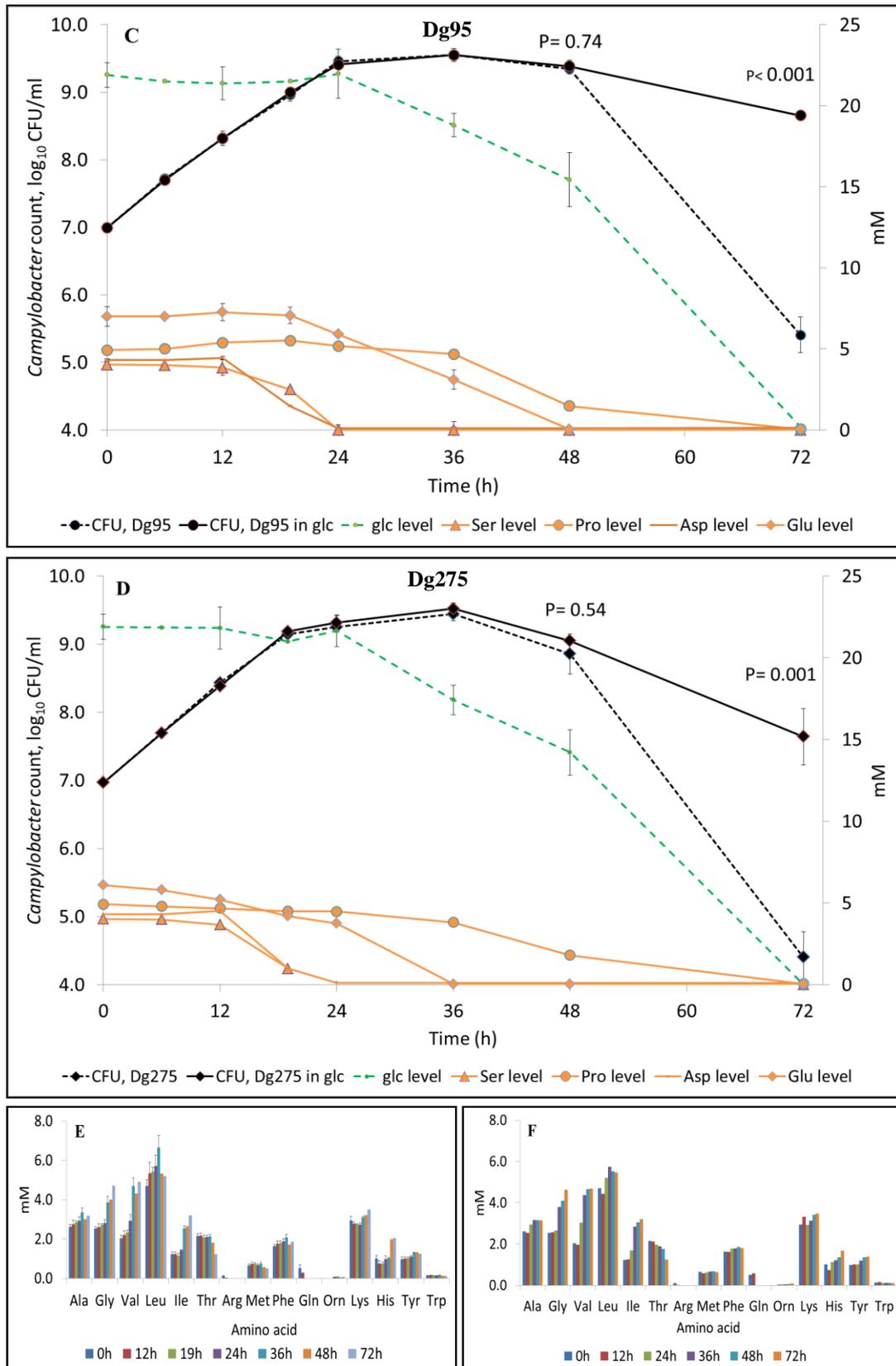
Figure 4.7: Preference of RG-2 strains to some amino acids over glucose. The ED-Positive *C. jejuni* Dg95 and Dg275 strains, and the ED-negative Dg200 strain were supplemented with 10ml MHB, 10% FBS, ± 21.9 mM glucose, and incubated in flasks with shaking (**section**

160

2.2.8). Aliquots were taken during growth for monitoring of OD and viable cell count. (**A**) OD of tested strains, (**B**) CFU of Dg200. (**C**) Dg95 and (**D**) Dg275 CFU, concentration of glucose (glc) and utilised amino acids. Non utilized amino acids in cultures Dg95 and Dg275 are shown in figures (**E**) and (**F**), respectively. Glucose concentration was monitored using a Merk Glucose Assay Kit, and amino acids were quantified using a EZ-Faast-kit followed by GC-MS. Means and error bars for OD, CFU, amino acid, and glucose concentrations show standard errors from three independent cultures except amino acids concentration in the medium was inoculated with Dg275 (**Figure 4.7**, **D & F**) were from a biological replicate. P value compares CFU and OD with or without glucose.

4.5.3   Strain-specific floating biofilm formation

The role of the ED pathway may not solely be to metabolise glucose for energy. For some strains of *C. coli* enhanced floating biofilm formation has been demonstrated (Vegge et al., 2016). Therefore, the ability of two ED-positive strains, Dg95 and Dg275, to form biofilm was studied in Tryptic Soy Broth (TSB) with and without glucose supplementation The two published ED-positive *C. coli* strains Dg172 (biofilm –ve) and Dg349 (biofilm +ve) (Vegge et al., 2016) were included as controls. As shown in **Figure 4.8** *C. coli* Dg349 formed a typical floating biofilm and as expected *C. coli* Dg172 did not. Neither of the *C. jejuni* ED-positive RG-2 strains nor the closely related ED-negative Dg200 strain formed a biofilm. The ST-42 clonal complex *C. jejuni* Dg153 strain was also included in this investigation because only this strain formed microcolonies of biofilm on the glass fiber filters (**Figure 3.7**). Interestingly, the Dg153 strain formed the same sort of floating biofilm (unattached to a solid surface) seen with *C. coli* Dg349 (**Figure 4.8**). Both biofilm-former strains produced biofilm with and without glucose. Dry weight analysis would need to be assessed as performed by Vegge et al. (2016) to establish differences in extent of biofilm. Many genes contribute to the enhancement or reduction of biofilm formation in *C. jejuni*. For example, mutations in *cmeB*, *nusG*, *Cj0268c* (Teh et al., 2017), *fliS,* and *cj0688* have been shown to reduce biofilm formation (Joshua et al., 2006). In contrast, mutation in alkyl hydroperoxide reductase (*ahpC*) was found to increase biofilm formation and also increase susceptibility *C. jejuni* to aerobic stress (Oh and Jeon, 2014). Biofilm formation might not be solely dependent on the ability of a strain to utilise glucose, other factors such as gene expression or enzyme activity and environment will likely affect biofilm formation.

(- glucose)   **(+ glucose)**   **(- glucose)**   **(+ glucose)**

blank   blank

RG-2 ED-positive *C. jejuni* isolates

Dg95   Dg95   Dg275   Dg275

ED-negative *C. jejuni* isolates

Dg200   Dg200   Dg153   Dg153

ED-positive *C. coli* isolates

Dg172   Dg172   Dg349   Dg349

-**

+
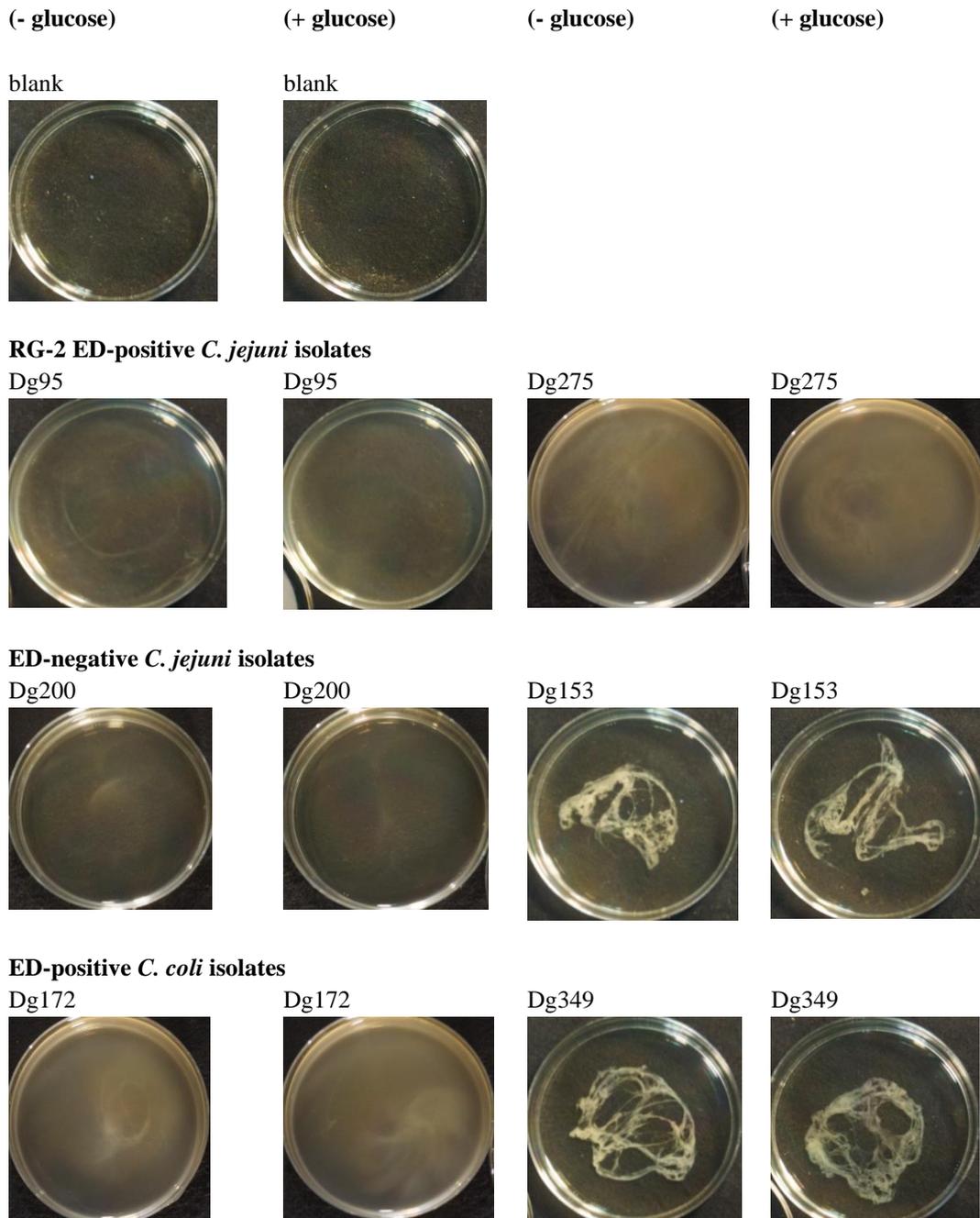
Figure 4.8: Strain dependent floating biofilm formation. *C. jejuni* and *C. coli* of farm-associated rat strains, as indicated, were incubated for four days at 37 °C in glucose free TSB (-glucose) or TSB supplemented with 100mM glucose (+ glucose).

4.6 The closed genome of Dg275 identifies two copies of the *glc* locus

The Dg275 strain belongs to the most common ED type, type 2, among the RG-2 strains studied here (**Figure 4.1**). The draft genome of this strain had a relatively low number of contigs (34), and growth properties of this strain had been previously studied. Previous studies had shown that the *glc* locus was inserted into an *rrn* operon of *C. coli* and *C. jejuni* subsp. *doylei* (Vorwerk et al., 2015). It was difficult to assemble the three copies of the *rrn* operon. Therefore, to localise the *glc* locus within the genome of typical RG-2 strains, the draft genome of Dg275 was closed.

4.6.1 Assembly of Dg275 contigs against the NCTC11168 reference genome

Dg275 genome mapping, assembly, sequencing, and closure was performed as explained in the flow chart of **Figure** 3.18. The output of CONTIGuator based mapping of Dg275 contigs against the NCTC11168 reference strain is shown in **Figure 4.9**. Of the 34 contigs, 16 were mapped to the genome sequence of NCTC11168 (named as mapped contigs). Unmapped contigs were as follows: 4 (3258bp), 8 (24435bp), 12 (5050bp), 17 (9662bp), 20 (523bp), 22 (2187bp), 23 (1234bp), 24 (1509bp), 25 (710bp), 26 (5460bp), 27 (1677bp), 28 (497bp), 29 (203bp), 30 (254bp), 33 (206bp), 34 (605bp), 39 (204bp), and 47 (210).
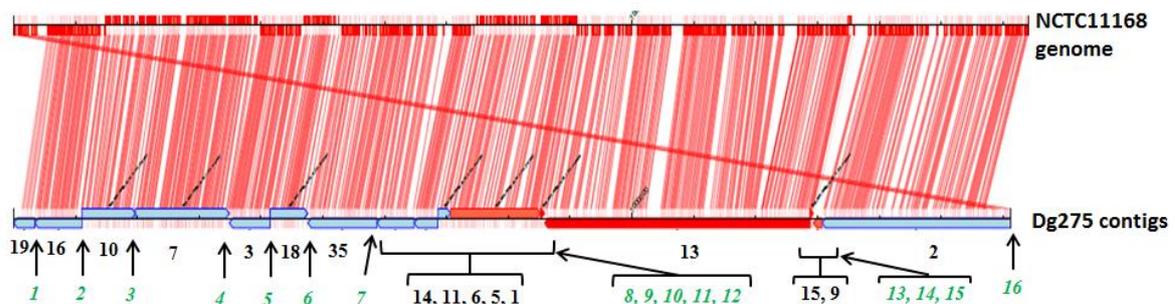
Figure 4.9: Contig mapping of the Dg275 strain draft genome against NCTC11168. The alignment was made using the CONTIGuator web server (Galardini et al., 2011). The archive file was then imported into the ACT programme to find the gaps. The top horizontal line represents the NCTC11168 genome, and the bottom horizontal blue and red segmented lines represent mapped contigs of the Dg275 draft genome. Vertical red lines represent the regions at which both genomes are similar. Mapped contigs 19, 16, 10, 7, 3, 35, 18, 14, 11, 6, 5, 1, 13, 15, 9, and 2 are shown in black, and gaps 1-16 in green italics. Gap 16 is located between contig number 19 and 2.

4.6.2 PCR amplification and closing of the Dg275 genome

Primers used to amplify and sequence all 16 gaps are listed in **Tablse 2.4, C and D** and were designed to amplify from close to the end of flanking contigs. PCR products were obtained from all gaps and ranged in size from 0.33kb to 16.77kb. Each purified PCR product was sequenced. Small gaps were sequenced in both directions and large gaps in one direction. Size of PCR products for each gap, as visualised by AGE, correlated with the size of the sequenced DNA (**Figure 4.11, A**). Interestingly, the *glc* locus was identified, by sequencing, in two gaps, one within *rrnA* locus located in gap 16 and the second in *rrnB* locus located in gap 4. For confirmation of this, an additional pair of primers was designed to amplify (i) from the conserved *murD* gene in gap 4 to an ORF encoding a hypothetical on the other side of the gap and (ii) from *ggt* in gap 16 to *ansB*, both conserved genes. Gap 8, encoding *rrnC* alone was amplified from conserved genes encoding hypothetical proteins. In addition, fragments were amplified from *murD* to the glc locus gene *glcP* (gap4) and from the gene encoding type II restriction endonuclease to *zwf* of the *glc* locus (gap16). In each case the size of the PCR product and sequencing confirmed inclusion of the *glc* locus in *rrnA* and *B* but not *rrnC* in the Dg275 genome (**Figure 4.11, B**). The same results were obtained with PCR amplification for Dg95. The *glc* locus was also inserted into the same *rrnA* and *rrnB* operons in the Dg95 strain (**Figure 4.11, C**).
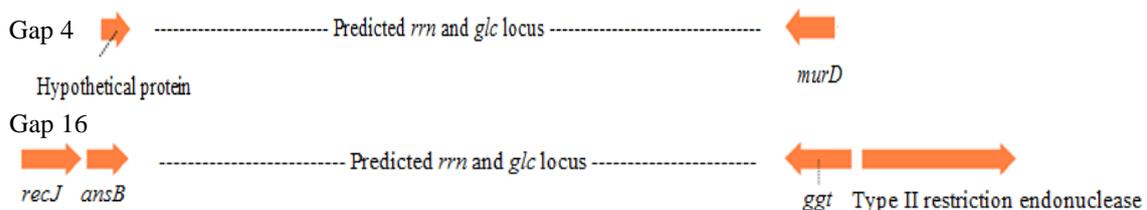
Figure 4.10: Localisation of a pair of primers used to amplify predicted *rrn* and *glc* loci in gaps 4 and 16. A pair of Dg275-4FA and Dg275-4RA primers for gap 4 were used between the conserved regions of the genome. They hybridised to a gene encoding hypothetical protein and *murD*, respectively. glcPF and Dg275-4RA primers were also used to amplify *murD* to *glcP* of the glc locus. A pair of Dg275-16FA and Dg275-16RA primers were used to amplify *ggt* to *ansB* in gap 16. zwfF and Dg275-16RB were also used to amplify putative type II restriction endonuclease gene to *zwf* of the *glc* locus (**Figure 4.11, A, B**). The same primers of Dg275 were used to localise the *rrn/glc* locus in the ED type 3 Dg95 strain.
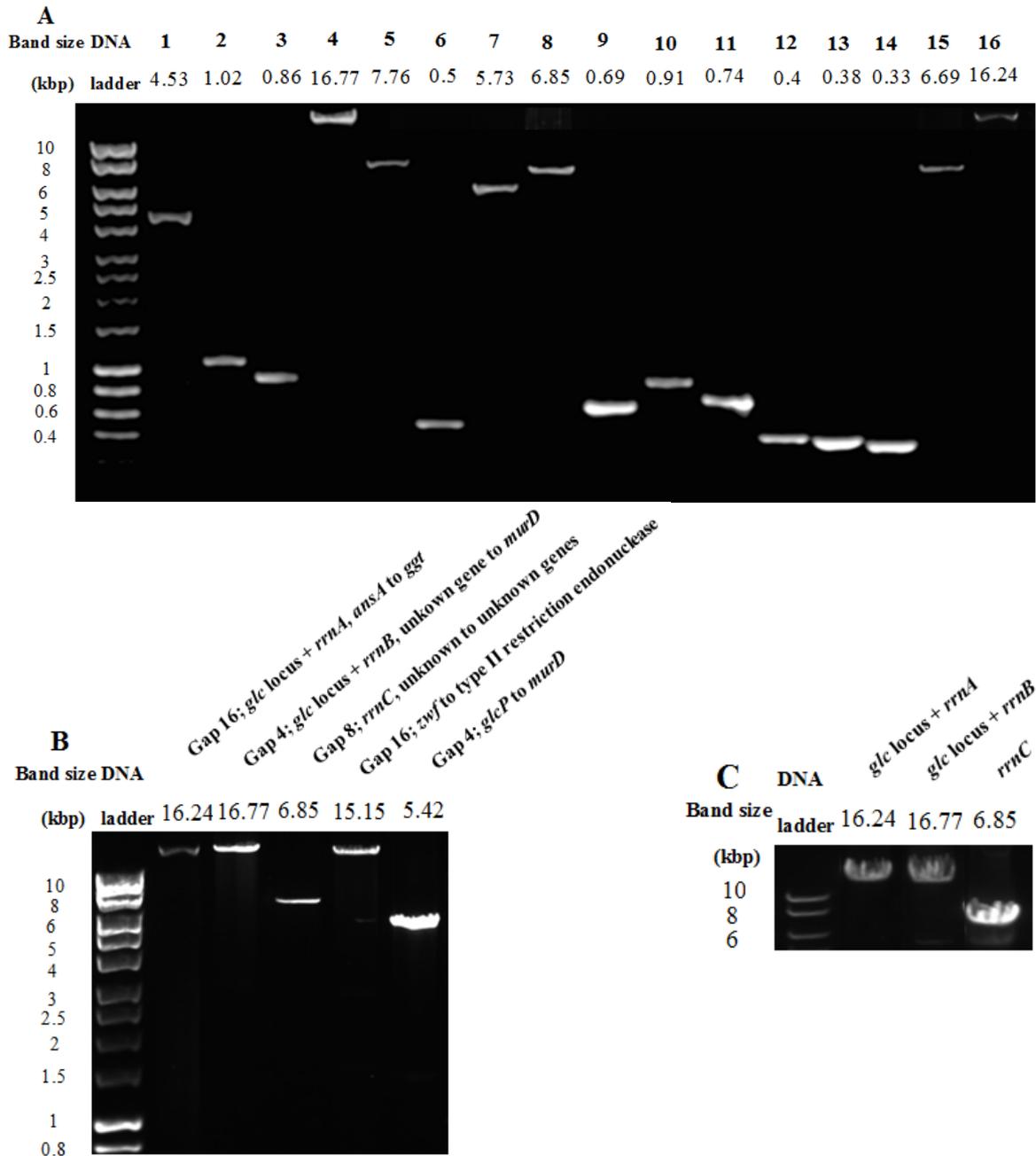
Figure 4.11: PCR products across gaps of the partially assembled Dg275 genome. (A) Dg275 PCR products of the gaps between the mapped contigs. (**B**) *rrn*s/*glc* locus PCR products of Dg275, and (**C**) Dg95 *rrnAB/glc* and *rrnC* PCR products. PCR products were obtained using CloneAmp HiFi, and isolated genomic DNA. Primers and primer characteristics are listed in **Table 2.4, C**. PCR products of the following set of primers were used to amplify these gaps: Dg275-1FA, RA (gap 1); Dg275-4FA, RA (gap 4); Dg275-5FA, R (gap 5); Dg275-8F, R (gap 5); Dg275-11FA, R (gap 11) and Dg275-16FA, RA (gap 16). These are shown in gel A. The same samples of 16, 4, and 8, plus from between *zwf* to type II restriction endonuclease of gap 16, and *glcP* to *murD* of the gap 4 are shown in gel B. Gap numbers 1 to 16 are along the top of gel A in bold. The numbers below, above the lanes, indicate the length of each amplified product following confirmation by DNA sequencing.

While virtually all gaps were readily filled, with initially unmapped contig sequences gap 1 proved more difficult, due to a duplicate short contig 23 in this gap. The same approach was taken as with the Dg147 genome. Primers were designed from all unmapped contigs. They were used in sequencing reactions of Gap1 PCR product (~4.5 kb). The gap 1 PCR product was successfully fully sequenced; identifying two copies of contig 23 and one of contig 27 within this gap with total insert size of 4.53 kbp.

Sequencing of gaps 6, 11, 12, 13, and 14 revealed no additional DNA in these regions (**Table 4.1**). In most cases, breaks occurred where there were long series of G/C or A/T. Gaps 2 and 3 were each filled with a gene encoding a methyl-accepting chemotaxis protein, *cj0144* and *cj0262c* for gaps 2 and 3, respectively. The C-termini of these two encoded proteins have identical sequence (Marchant et al., 2002, Rahman et al., 2014). In addition, to gaps 2 and 3, and gaps 4, 16 and 8 with the *rrn* locus, repeated DNA sequences were found within gaps 9 and 10, 1 and 15, and 5 and 7. All three *rrn* operons were identical aside from the intergenic region between 23S rRNA and 5S rRNA of *rrnB*, which shared only 98.81% identity with the same intergenic region in *rrnA* and *rrnC*.

Three large unmapped contigs; 8 (24.4kbp), 12 (5kbp), and 24 (1.5kb), were not mapped within the closed genome of Dg275. They were Blasted in the NCBI database. This identified a hit with pFORC46.2 plasmid of *C. jejuni* strain FORC_046 (accession number NZ_CP017231.1) with the following percent identity 99, 100, 92, and percentage query cover 83, 97, and 100 for contigs 8, 12 and 24, respectively. The DNA sequences of this plasmid encode the following P-type conjugative transfer proteins; TrbB, TrbE, TrbL, TrbG, and TrbI; IncP-type DNA transfer primase TraC; IncP-type DNA transfer coupling protein TraG; IncP-type DNA relaxase TraI and Transposase.

Only unmapped contigs 33 (206bp) and 46 (210bp) could not be placed within the closed genome of the Dg275. The sequence of contig 33 was identical to a duplicated sequence in the mapped contig 13 of Dg275, encoding motility accessory factor (locus 01301 and locus 01302). The same sequence was duplicated in the 3 main reference strains genes *c8j_1252* and *c8j_1253* in strain 81116 and have a similar position as in the genome of Dg275. As with assembly of Dg147 one short unmapped contig (contig 46 – 210 bp, Dg275) encodes tRNA-Asp and tRNA-Val and could be mapped to duplicate sequences within a mapped contig

(contig 5 for Dg275). However, with the Dg275 sequence there was a single nucleotide mismatch between these copies.

Table 4.1: Identified genes within gaps of partially assembled Dg275 genome.

Gap	Flanking, mapped contigs	Size of gaps (kbp)	Unmapped or mapped contig/s number/s located in gap; state of sequencing	*Gene* or operon locate/s in the gaps
1	19, 16	4.53	two copies of unmapped contigs 23 and one copy of unmapped 27; sequenced except 0.06kbp of 27	Four genes encoding hypothetical proteins, plus putative periplasmic protein (*cj0057*) and a putative peptidase C39 family protein gene (*cj0058*) Contig arrangements as below: 19, 23, 27, 23, 16
2	16,10	1.02	unmapped contig 34; sequenced	putative methyl-accepting chemotaxis protein (*cj0144*)
3	10, 7	0.86	unmapped contig 34; sequenced	putative methyl-accepting chemotaxis protein (*cj0262*)
4	7, 3	16.77	unmapped contigs 22, 17, 4, 29 and 30; sequenced except 0.3kbp of 22 and 0.85kbp of 4	*rrnB* (contigs 22, 4, 29, 39) and *glc* locus (contig 17)
5	3, 18	7.76	mapped 1; sequenced	*cj0967* putative periplasmic protein, a gene encoding hypothetical protein (found in RM1246-ERRC strain) and *cj0975* putative outer-membrane protein
6	18, 35		no DNA; sequenced	Break was at an intergenic region between an encoding hypothetical protein (*cj0563*) and encoding putative integral membrane protein gene (*cj0564*).
7	35, 14	5.73	mapped 15; sequenced	encoding sugar nucleotidyltransferase, putative amidotransferase, phosphoenolpyruvate synthase / pyruvate phosphate dikinase and methyltransferase
8	14, 11	6.85	unmapped contigs 4, 22, 25, 29, 30; sequenced except 0.09kbp of 22 and 0.6kbp of 4	*rrnC*
9	11, 6	0.69	unmapped contig 20; sequenced	encoding a hypothetical protein gene
10	6, 5	0.91	unmapped contig 20; sequenced	encoding a hypothetical protein gene

11	5, 1	0.74	no DNA; sequenced	break was in an encoding putative periplasmic protein (*cj0967*)
12	1, 13	0.4	no DNA; sequenced	break was at an intergenic region between a genes encoding a hypothetical protein and hemerythrin family non-heme iron protein genes. 90bp overlapped between ends of both mapped contigs 1 and 13.
13	13, 15	0.38	no DNA; sequenced	break was within gene encoding putative sugar nucleotidyltransferase gene (*cj1416c*). 109bp overlap between ends of both mapped contigs 13 and 15.
14	15, 9	0.33	no DNA; sequenced	breaks were at an intergenic region between encoding putative methyltransferase (*cj1420c*) and encoding putative sugar transferase genes. 90bp overlapped between ends of both mapped contigs 15 and 9.
15	9, 2	6.69	two copies of unmapped 28 and a copy of 26; sequenced except 4.9kbp of 26	break was in a gene encoding beta-1,3-galactosyltransferase/beta-1,4 galactosyltransferase gene, a capsular polysaccharide biosynthesis gene, four encoding hypothetical protein genes, encoding putative sugar transferase gene. All the above genes are absent in NCTC11168, but present in 81116 except last one which is present in NCTC11168, but absent in 81116.
16	2, 19	1624	unmapped contigs 4, 17, 22, 30, 39; sequenced except 0.39kbp of unmapped 17	*rrnA* and *glc* locus

The NCTC11168 genome was aligned with genomes from each of the ED-positive Dg275 and ED-negative 81-176 and 81116 *C. jejuni* strains (**Figure 4.12**). The diagonal red lines were labelled with rectangles and circles. The diagonal lines (rectangles) were drawn between the three copies of the *rrnABC* operon of the strains to the same operon in the NCTC11168 (Kim et al., 1993). Circle number 1 highlights an encoding putative methyl-accepting chemotaxis gene (*cj0144*) that is located in gap 2. This gene aligned with an encoding putative methyl-accepting chemotaxis gene (*cj1564*) in NCTC11168. The C-termini of both genes were identical (**Table 4.1** at above) (Marchant et al., 2002, Rahman et al., 2014). Circle number 2 located in gap 5 and this region was identical to the mapped contig 1. This region contained a gene, *cj0967*, for a putative periplasmic protein, a gene encoding a

hypothetical protein, and *cj0975* encoding a putative outer-membrane protein. Two copies of the *cj0975* gene are also present in the genomes of 81-176 and 81116 (**Figure 4.12**). Circle number 3, located in gap 7, was identical to mapped contig 15. Gap 7 contained a gene encoding a sugar nucleotidyltransferase, putative amidotransferase, phosphoenolpyruvate synthase/pyruvate phosphate dikinase, and methyltransferase genes (**Table 4.1** at above). A similar order of genes was also seen in the 81116 genome. For confirmation of this arrangement, two sets of primers were used to amplify both gaps 5 and 7 (**Table 2.4, C** at above).
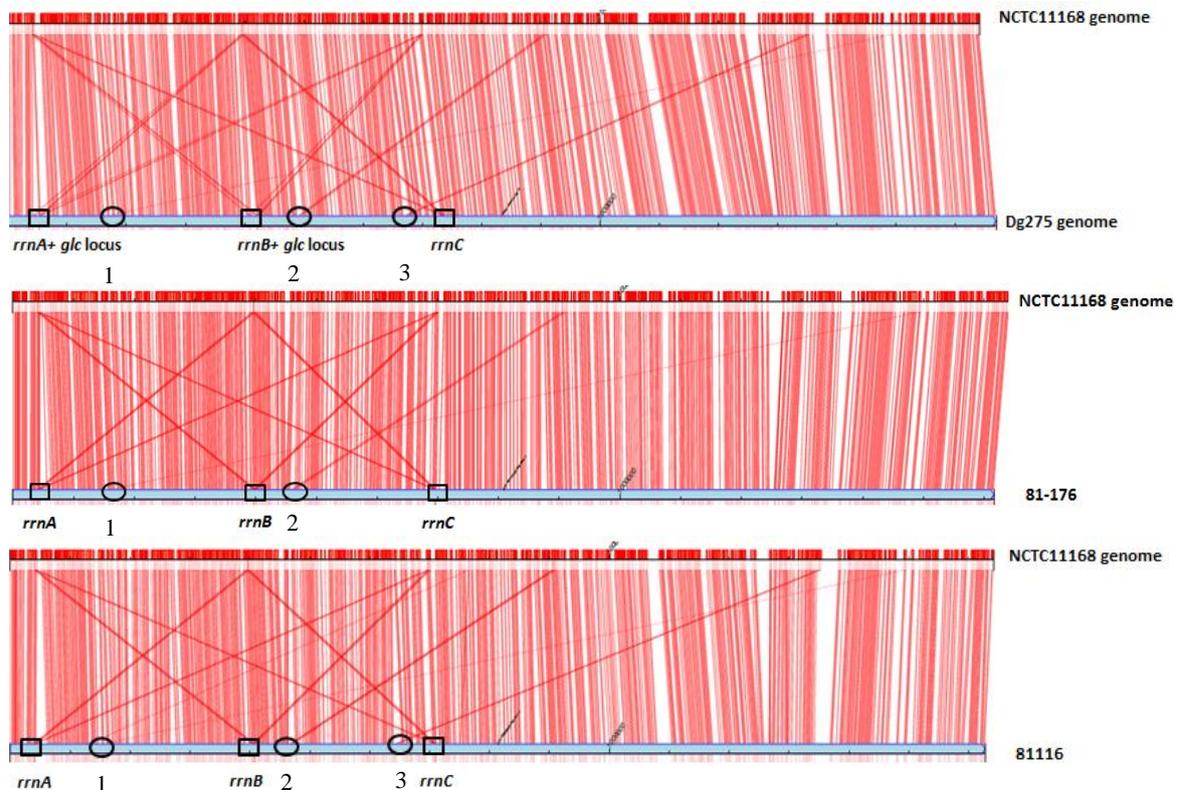
Figure 4.12: Mapping NCTC11168 genome with Dg275, 81-176 and 8116 genomes. Top horizontal line represents the NCTC11168 genome and bottom blue line represents Dg275, 81-176 or 81116 genome. Strains sharing the same genome similarities are represented with vertical red lines. Cross red lines identifing the 3 copies of *rrn* operons are labeled with rectangles. *rrnAB* of Dg275 are each have a *glc* locus inserted. Circles numbered as 1, 2 and 3 are located in gaps 2, 5, and 7 of Dg275, respectively.

4.7   Comparison of *glc* locus insertion in different *Campylobacter* spp.

Only three ribosomal RNA loci (~5.7kb) are found in the two most common *Campylobacter* species; *C. jejuni* and *C. coli* (Kim and Chan, 1989, Vorwerk et al., 2015). Complete assembly of Dg275 into a single closed genome highlighted that the complete *glc* locus was inserted in both the *rrnA* and *rrnB* operons but not in the *rrnC* operon of both *C. jejuni* strains, Dg275 and Dg95 (**Figure 4.11** at above). In addition, the location of a *glc* locus in the *rrnA* operon of the rat *C. coli* strain Dg349 was also confirmed (**Figure 4.13, A**). Organisation of the *glc* locus genes was 5'-16S rRNA- tRNA (Ala)-tRNA (Ile)- *glc* locus-23S rRNA- 5S rRNA -3' (**Figure 4.13**), which was similar to the organisation of these genes as reported for both *C. coli* CHW470 and *C. jejuni* subsp. *doylei* (Vorwerk et al., 2015).

The *glc* locus within the *rrnA* gene cluster was downstream to the *recJ* and *ansB* genes in *C. jejuni* Dg275, as with *C. coli* Dg349, and *C. coli* CHW470 (**Figure 4.13, A**). The *recJ ansB* side was conserved in all of the strains, but the other side was variable although for both Dg275 and *C. jejuni* subsp. *doylei* it was identical. However, The second copy of the *glc* locus was in the *rrnB* gene cluster, flanked by a conserved hypothetical protein gene and by *murD* (**Figure 4.13, B**). Flanking regions of *rrnB* were more conserved than *rrnA* and *rrnC*. These data show that it is possible that each of the three copies of the *rrn* operons might have potential to accommodate an inserted *glc* locus. The *glc* locus of *C. jejuni* subsp. *doylei* has inserted into *rrnC* (**Figure 4.13, C**).

To understand the source of the *glc* locus in Dg275, the DNA percentage identities of the *rrn*/*glc* and intergenic regions of different strains were compared (**Table 4.2** at above). Both *rrnA*/*glc* and *rrnB*/*glc* of Dg275 were identical except for the intergenic region between 5S rRNA and 23S rRNA (0.25kbp, 98.81% identity). Both copies of the *glc* locus and *rrnAB* of Dg275 were closer to *C. coli* Dg349 and CHW470 strains than to *C. jejuni* subsp. *doylei*. Furthermore, the GC content of the *glc* locus in the *C. jejuni* subsp. *doylei* was lower than in the other strains (**Table 4.3** at above). For example, the GC content of *pgl* and *glk* in *C. jejuni* subsp. *doylei* were 26% and 27%, respectively, which is lower than the genome GC content of 31%. Thus, based on the DNA percentage identity and the GC content of the *rrn*s/*glc* as well as flanking regions of the *rrn* operons, it is more likely that the *glc* locus of Dg275 is from *C. coli*, not from *C. jejuni* subsp. *doylei*. Acquiring the locus might have been by natural

horizontal gene transfer (HGT) from live cells of different strains (Dingle et al., 2001, Vorwerk et al., 2015, Ochman et al., 2000).

C

Dg275, *C. jejuni*
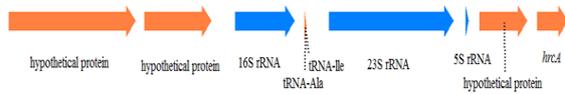
C. jejuni subsp. *doylei* 269.97
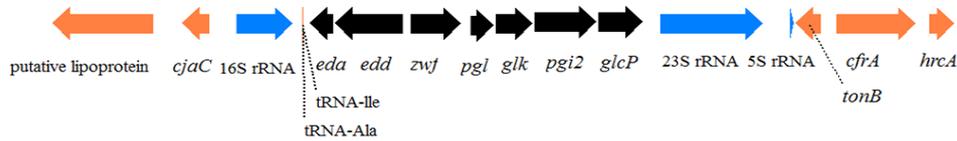
Figure 4.13: Diagram of the ED pathway-encoding the *glc* locus integrated into ribosomal RNA operons. The figure shows a comparison between rRNA operons plus the *glc* locus from several strains with respective rRNA operons in ED-negative strains. rRNA loci (**A**) *rrnA*, (**B**) *rrnB* and (**C**) *rrnC*. Blue color-coded arrows are the *rrn* operon; black arrows, the *glc* locus; orange arrows, genes flanking the *rrn* loci. The *glc* locus is inserted into *rrn* operons adjacent to Ile-tRNA gene and 23S rRNA. The entire *glc* locus of both inserts in Dg275 was confirmed to be identical by Sanger sequencing using the set of primers listed in **Table 2.4, D** and ~15.4 kbp product as a template. Of this, 0.39kb of the copy of *glc* locus inserted into the *rrnA* remains to be confirmed. Primers, listed in **Table 2.4, E**, were used to localise the *glc* locus in the farm-associated rat *C. coli* Dg349 strain. These primers amplified and/or sequenced from *ansB* to *eda* and *glcP* to the gene encoding putative type II restriction endonuclease. Sequences of the other strains were taken from the NCBI database by uploading the following accession numbers; NCTC11168 (NC_002163), 81-176 (NC_008787.1), *C. coli* CHW470 (study; KT001108) and *C. jejuni* subsp. *doylei* 269.97 (NC_009707).

Table 4.2: Comparison of *rrn/glc* loci of Dg275 to *C. coli* and *C. jejuni* subsp. *doylei*.

Gene/Insertion	*Length (bp) based on Dg275 (rrnA/glc)*	*Percentage identity of the rrn/glc locus compared to the Dg275 rrnA/glc locus*				
		Dg275 (*rrnB/glc*)	*C. coli* rat associated Dg349 (*rrn /glc*)	*C. coli* human associated CHW470 (*rrn /glc*)	*C. jejuni* subsp. *doylei* 269.97 (*rrn /glc*)	NCTC11168 (*rrnA*)
16S rRNA	1513	100	99.34	99.41	99.47	99.6
Intergenic region	104	100	99.02	99.02	100	100
tRNA-Ala	76	100	100	100	100	100
Intergenic region	8	100	100	87.5	100	100
tRNA-lle	77	100	100	100	100	100
Intergenic region	153	100	99.35	99.35	53.8	-
eda	624	100	99.68	99	87.34	-
Intergenic region	11	100	100	100	81.81	-
edd	1803	100	99.78	99	93.73	-
Intergenic region	144	100	100	100	69.34	-
zwf	1401	100	99.79	99.79	85.44	-
Intergenic region	165	100	99.39	98.79	67.30	-
*Pgl***	681	100	99.71	99.71	81.64	-
glk	1005	100	99.70	99.70	87.29	-
Intergenic region	3	100	100	100	66.67	-
Pgi2	1644	100	99.88	99.88	84.55	-
Intergenic region	10	100	100	100	100	-
glcP	1200	100	100	100	85.38	-
Intergenic region	665	100	86.80	87.27	73.50	-
23S rRNA	3057	100	99.21	99.52	98.53	99.59
Intergenic region	254	98.81	81.50	82.21	92.46	98.02
5S rRNA	120	100	99.17	100	100	100

* Sequence comparison of *glc* locus of both *rrnA* and *rrnB* operons of *C. jejuni* Dg275 and *rrn* of *C. coli* rat associated strains sequenced (**see Figure 4.13 for detail**). Sequences of other strains were downloaded from the NCBI database. ** There are 19bp overlap between *pgl* and *glk* genes. (-), indicates absent gene or missing data

Table 4.3: GC content comparison of *rrn/glc* locus of *C. jejuni* Dg275 to *C. coli* and *C. jejuni* subsp. *doylei*.

Gene	Dg275 (*rrnA/glc*)	Dg275 (*rrnB/glc*)	*C. coli* rat associated Dg349 (*rrn /glc*)	*C. coli* human associated CHW470 (*rrn /glc*)	*C jejuni* subsp. *doylei* 269.97 (*rrn /glc*)	NCTC11168 (*rrnA*)
16S rRNA	49	49	49	49	49	49
tRNA-Ala	60	60	60	60	60	60
tRNA-lle	53	53	53	53	53	53
eda	33	33	32	32	30	-
edd	35	35	35	35	35	-
zwf	31	31	31	31	26	-
pgl	31	31	31	31	27	-
glk	36	36	36	36	33	-
pgi2	32	32	32	32	31	-
glcP	34	34	34	37	32	-
23S rRNA	46	46	46	47	45	47
5S rRNA	50	50	50	50	50	50

4.8 Genome comparison

Phylogenetic analysis demonstrated that the RG-2 ED-positive rat *C. jejuni* strains cluster together and are close to the ED-negative CC-45 generalist, farm-associated rat strains (**Figure 1.11** and **Figure 4.1** at above). To find any other genes that might be related to glucose metabolism, such as regulators, the annotated genome of Dg275 was blasted with ED-negative CC-45 and RG-2 draft genomes using the genome comparator of the BIGSdb database. This was to help identify any core genes that are present in all 39 RG-2 strains except Dg381, Dg18 and Dg201, but absent in the ED-negative CC-45 strains. Only 32 genes out of 1483 core genes in RG-2 were absent in all of the selected ED-negative CC-45 strains (**Appendix 11**). Among these, locus numbers between 32 to 38 and locus numbers between 431 to 437 were both copies of the *glc* locus. Both locus numbers 72 and 77 encode TolA-like membrane proteins, which are involved in the uptake of colicin (Levengood et al., 1991) and phage infection (Riechmann and Holliger, 1997). In addition, locus number 484 (Transcriptional regulator, IclR family) was interestingly absent in 81-176 and 81116 as well as the ST 45 CC strains screened, yet also present in 65% (17 isolates) of the ED-positive strains from wild birds. A transcriptional regulator of the IclR family was reported to participate in metabolism and regulation of L-gluconate catabolic processes in *Paracoccus* sp. 43P bacterium (Shimizu and Nakamura, 2014). Locus numbers 484 to 493 were for another more common plasticity genomic island, encoding products required for fucose utilisation (Stahl et al., 2012, Muraoka and Zhang, 2011). The ability of RG-2 strains to utilise glucose and fucose (data not shown) might support the strain's ability to adapt to one particular host with greater efficiency.

4.9   Conclusion

Contradictory to the dogma that *C. jejuni* is unable to utilise glucose, this study is the first demonstration of the ability of the most important *Campylobacter* species, *C. jejuni*, to use glucose. All tested *C. jejuni* RG-2 strains utilised the sugar under *in vitro* conditions. Interestingly, it was found that the ED-positive strains still preferentially catabolise Ser and Asp despite the availability of glucose. After depleting stocks of the preferred amino acids, the isolates started to metabolise glucose and Glu, with least preference shown for Pro. Ser was utilised exclusively at the exponential phase of growth, and upon depletion of Ser and Asp, glucose was utilised during the second growth phase. This may relate to the inhibition of glucose uptake or expression of the *glc* locus genes. This alternative and additional pathway for carbon metabolism and/or energy provision may be beneficial for RG-2 cells, enabling them to compete within their habitat and evolve to better survive in a host environment.

The percentage identity and GC content of the plasticity region of the RG-2 strains indicated that the locus might have been horizontally transferred from *C. coli* by a natural transformation mechanism (Pearson et al., 2003, Jeon et al., 2008). A high percentage (11.6%) of the 137 farm-associated rat *Campylobacter* spp. were *C. coli*, none belonged to *C. jejuni* subsp. *doylei,* and 18.7% of the *C. coli* were ED-positive strains. Hence, *C. coli* is a good candidate for origin of this *C. jejuni glc* sequence. The locus was interestingly found in two of the *rrn* operons, which was identified by closing the genome of the RG-2 strain Dg275. The locus was inserted into each of the *rrnA* and *rrnB* operons, but was not found in the *rrnC* operon. The same organisation was confirmed in a second ED positive strain Dg95. At present, it has only been reported that *rrnA* of *C. coli* and *rrnC* of *C. jejuni* subsp. *doylei* harbour the *glc* locus (Vorwerk et al., 2015). This would then appear to cause problems in creating knockout mutants in this pathway for testing of different conditions and *in vivo* modelling. This study provides further evidence to support previous assumptions that the rRNA operon, not its flanking regions, is the area of recombination for acquisition of the plasticity *glc* locus (Vorwerk et al., 2015). Although flanking regions of the *rrnA* and *rrnC* operons are genetically diverse, the loci have inserted into *rrnA* and *rrnB*, which has more highly conserved flanking regions.

The genomes of all RG-2 strains also contained the fucose pathway. This group of isolates can be documented as the first group of *C. jejuni* isolates that are able to use fucose and glucose sugars. The presence of these pathways in a high number of the Norway rats and wild birds might indicate that catabolism of both sugars results in the production of metabolites that aid survival in a specific host or environment and might promote a competitive advantage (Dwivedi et al., 2016). The ability of the glucose utilising RG-2 strains to colonise chickens and the stability of the ED pathway will be shown in the next chapter, *in vivo* chicken assay.

Chapter 5

Galleria mellonella and

Chicken *in vivo* Trials

5.1 Introduction

Chicken and *Galleria mellonella in vivo* models have been used to study colonisation and virulence of *C. jejuni* (Humphrey et al., 2015, Champion et al., 2010, Hendrixson and DiRita, 2004, Loc Carrillo et al., 2005, Senior et al., 2011). Use of the chicken model is important because it is considered that this bird is the main source of human infection by *C. jejuni* (EFSA, 2016, Wingstrand et al., 2006). The bacterium primarily colonises the caeca of the chicken, although there have also been reports of low level colonisation of the small intestine (Beery et al., 1988, Newell and Fearnley, 2003, Humphrey et al., 2015). Although many strains of *C. jejuni* efficiently colonise chickens, reports also include strains showing lower efficiency to poor colonisation. In one case study, 14 day old chickens were infected with different strains *C. jejuni*. Only three of seven isolates of human origin colonised the chickens; whereas the majority of strains of chicken origin (10 of 12) colonised the chickens well (Korolik et al., 1998). Thus, the chicken model is a natural system to compare the ability of different strains to efficiently colonise this farm-yard host. The larvae from *Galleria mellonella* has been shown to succumb to infection by many strains of *C. jejuni.* This insect model has the advantage that it can be used at 37°C, and provides a cheap and rapid method to initially test strains for potential differences in colonisation efficiency and virulence.

The ED-positive strains used in this study (**Chapter 4**) were all isolated from Norway rats. Among the ED-positive strains identified in the BIGSdb database, seventy C. *jejuni* strains from different sources harbour the *glc* locus. Thirty-nine strains (51.4%) belong to the ED-positive RG-2 *C. jejuni* strains isolated from Norway rats, 30 isolates (42.9%) are from wild birds, 2 isolates (2.9%) are from farm environments, and 2 (2.9%) from an unknown source (Vegge et al., 2016). None of the ED-positive strains identified were from chicken or clinical samples. The large and comprehensive BIGSdb database contains no isolates belonging to, or closely related to, the distinct group of RG-1 isolates (ST 6562, ST 6561, ST 6564, ST 5129, ST 7279, ST 7278) (**see Figure 1.11** at above).   Strains phylogenetically related to RG-1 and ED positive strains appear to be excluded from banks of human and chicken isolates. While this might be attributed to recent evolution of

these strains in a localised geographic niche, it could also be explained by adaptation to survival in the rat host or some other environment and an inability to colonise chickens. Enhanced growth of tested RG-1 strains in the presence of rat mucin would also be consistent with Norway rats as host for these strains. For these reasons, the ability of selected well characterised RG-1 (**Chapter 3**) and RG-2 strains (**Chapter 4**) to colonise chickens was compared to colonisation ability of several generalist strains. Initially, strains were tested using a *Galleria* virulence assay. Strains were then tested, for their ability to colonise 16 day old Ross chickens. In this Chapter, several questions were addressed. (i) How efficiently did RG-1 and RG-2 strains colonise chickens? (ii) Was there any evidence from WGS analysis of recovered strains of adaptation for enhanced survival in chickens? How stable was the *glc* locus in ED-positive strains during chicken colonisation? (iii) Could strains be recovered from other sites apart from the caecum? (iv) What was the impact of growth of different strains on the caceal microbiota (samples saved for future study).

Results and Discussion

For animal trials, Dg147 was selected as a representative of the clade RG-1 as the genome had been closed and analysed, and the strain studied in depth. Dg147 represents one main branch of RG-1 strains and possesses two insertions of phage. Similarly Dg275 was the best characterised example of RG-2, with closed genome and represents the main cluster of RG-2, possessing the common ED type 2. Dg275 was isolated from a sheep/ pig farm at 42°C. Dg95 (RG-2), represents strains with ED type 3 and was isolated from a turkey farm at 37°C. The generalist strain, Dg200 ST45CC, was selected because of its close phylogenetic position to the RG-2 clade. The other generalist strain selected for the studies was Dg194, which belongs to the classic human and chicken colonisation group ST21CC. Dg153 ST42CC was chosen as it had consistently been shown to be a good grower and biofilm former and interesting was isolated on the same day and same farm as Dg147.

5.2 Virulence of strains in the *Galleria mellonella* model

In initial trials with *Galleria* each strain was tested at an inoculum level of 7.1, 6.6, 6 and 5.3 $\log_{10}$ CFU, and melanisation or death was recorded at 24h (**Figure 5.1, A**), as described in **section 2.6.2**. A dose-response was observed for each strain. At the highest level, 7.1 log10 CFU, all larvae died except for one survivor for each of strains Dg147 and Dg153. At the lowest inoculum level, 5.3 $\log_{10}$ CFU, most larvae survived varying from 3 mortalities with Dg194 down to none with Dg153. **Figure 5.1, B** shows the results of repeat experiments with larvae infected with either ~5.9 or 6.6 $\log_{10}$ CFU bacterial cells. The insects showed the same strain-dependent pattern of susceptibility with the two different doses of bacterial cells (**Figure 5.1, B**). The highest number of healthy insects were recovered following infection with RG-1 strains Dg147 and Dg300, or the ST 42 clonal complex strain Dg153. The highest mortality was noted with strains belonging to clonal complexes ST-21, Dg194, and ST-45, Dg200. With an infectious dose of ~5.9 $\log_{10}$ CFU, 60 ± 2% and 55 ± 5% of the larvae survived infection with 11168H and Dg194, respectively, whilst the same dose of the RG-1 strains, Dg147 and 300, resulted in 80 ± 10% and 80 ± 5.7% survival, respectively. Even with approximately five times the infectious dose, 6.6 $\log_{10}$ CFU, over 50% of the insects survived infection with the RG-1 strains. Significantly, surviving larvae, inoculated with Dg194,

Dg200 or Dg275 displayed melanisation, whereas those infected with other strains remained healthy (beige). These data provided a preliminary indication that the *G. mellonella* was more susceptible to the generalist ST-21 and ST-45 clonal complex strains tested than to either the ST-42 clonal complex, or the RG-1 and RG-2 strains.
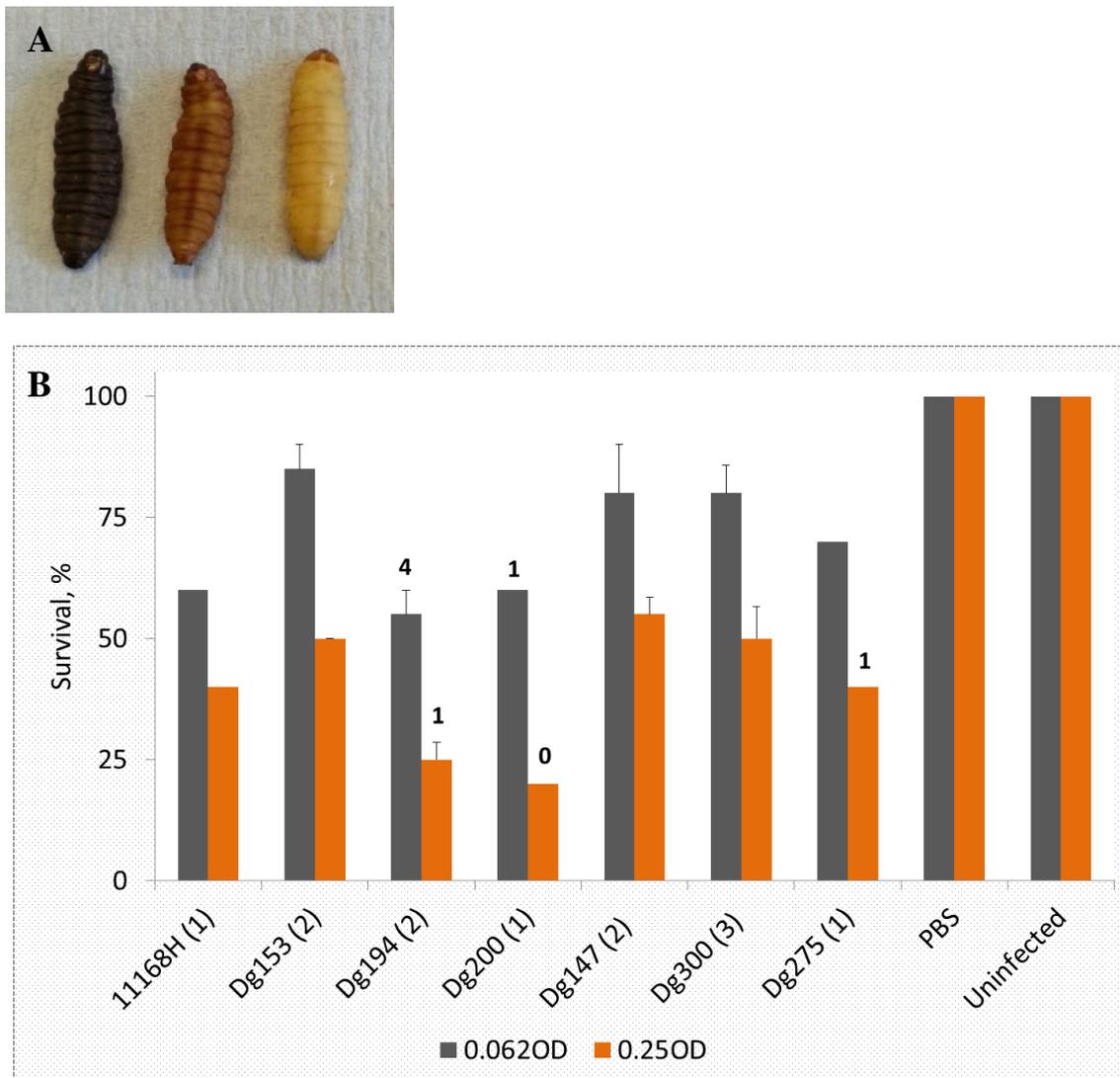
Figure 5.1: *Galleria mellonella* **infection with** *Campylobacter jejuni* **strains.** (**A**) Morphological appearance of the Wax Moth larvae challenged with the *C. jejuni* strains: beige (live healthy insect), brown (live unhealthy or dead insect), and black (dead insect). Dead was confirmed by a lack of movement following stimulation with forceps. (**B**) Groups of ten *G. mellonella* larvae were injected with a 10µl inoculum of standardised *Campylobacter* cells at an infectious doses of 0.062 or 0.25 OD (~9.6 x 10^5 CFU or 4.5 x 10^6 CFU), respectively, with PBS or were uninfected. Melanisation and mortality were recorded 24h after challenge. Numbers over bars are the number of beige (healthy) surviving larvae at 24h; bars without numbers, all survivors were beige. Strain 11168H is a hypermotile variant of the reference NCTC11168 isolate, ST-43/CC 21 (Karlyshev et al., 2002). Other strains were farm-associated rat isolates; Dg153, ST-42; Dg194, ST-21; Dg200, ST-45; RG-1 strains,, Dg147 and Dg300, both ST-6562; RG-2 strain, Dg275, ST-7259/ CC-45 (**see Table 2.1 for detail of strains**). Brackets indicate the number of biological replicates, error bars are made from two or three biological replicates.

5.3 The chicken trial

Two groups of seven 16-day-old Ross chickens, per strain, were housed as follows; the first group (a) was left in the co-housing and the second group (b) were housed in individual cages. Both groups were infected with ~7 $\log_{10}$ CFU of one test strain by the oral gavage method. Groups (a) and (b) were sacrificed at 3 and 7 days post-infection (dpi), respectively and caecal, ileum and jejunum contents individually removed for quantitation of viable *Campylobacter* on mCCDA agar. There has been recent interest in the ability of certain strains of *Campylobacter* to become invasive and infect internal organs of chickens (Firlieyanti et al., 2016, Humphrey et al., 2015). Therefore, the extra-intestinal organs, liver, heart, kidney, spleen and breast meat were also aseptically removed, homogenised and enriched in Bolton broth media, to score for presence/absence of *Campylobacter*. Caecal contents and intestinal tissue were also stored at -80°C for future metagenomics/cytokine analysis (**see 2.6.3 for detail**).

5.3.1 Limited efficiency of chicken colonisation by *Campylobacter jejuni* RG-1 and RG-2 strains

Viable cells recovered from the caecal content after 3 dpi highlighted a low level of colonisation by Dg147 (RG-1), both Dg95 and Dg275 strains (RG-2), and surprisingly Dg370 (ST50 CC-21 clonal complex) (**Figure 5.2, A**). For 5 of the 7 chickens inoculated with Dg147, *Campylobacter* could not be recovered. Similarly, with both Dg95 and Dg370, there was no recovery of *Campylobacter* for 3 chickens. These *C. jejuni* strains had started to die even after just 3 dpi. Moreover, in contrast to the closely related ST 45 strain Dg200 colonisation of chickens with both RG-2 strains resulted in a very broad distribution of levels of colonisation (**Figure 5.2**). The median level of colonisation with Dg95 was 3.6 $\log_{10}$ CFU/g, for Dg275 it was 5.0 log10 CFU/g and for Dg200 8.8 $\log_{10}$ CFU/g. In addition, the other sequence types, Dg153 (ST 42) and Dg194 (ST 21) also colonised the chickens well, 7.6 ± 0.7 $\log_{10}$ CFU/g and 7.3 ± 0.5 $\log_{10}$ CFU/g, respectively, as did the control chicken coloniser HPC5 strain 6.3 ± 0.4 $\log_{10}$ CFU/g (Loc Carrillo et al., 2005). Interestingly, Dg370 ST50 did not colonise the chicken compared to Dg194 ST21 (P= 0.001), although both belong to clonal complex ST21. Chickens were inoculated with a dose of ~7 $\log_{10}$ CFU and

chickens are coprophagic (Newell and Fearnley, 2003). At 3 dpi, some recovery of inoculated *Campylobacter*, without colonisation might be expected. Thus, it can be concluded that both RG-1 the Dg370 strain did not colonise the chickens within the short period of 3 days and the RG-2 strains survived poorly.

The second group of infected chickens were sacrificed after 7 dpi and the caecal content analysed as before for CFU/g. For the 7 day trial, the identity of the recovered strains was confirmed through sequencing appropriate MLST genes, *in vitro* studies, and amplification of the *glc* locus, as explained below in **section 5.3.3**. This confirmed identity of virtually all isolates, aside from those from chicken numbers 89 and 100 that had been inoculated with Dg147 and Dg275, respectively. Both of these chickens were excluded from the results due to contamination. Dg200, Dg194, and Dg153, the strains that had efficiently colonised the caecum in the shorter period of 3 days, retained the same high level of colonisation at 7 dpi. For example, colonisation of the caecum by Dg200 was 7.9 $\log_{10}$ CFU/g at 7 days dpi compared to 8.4 $\log_{10}$ CFU/g at 3 days (no significant difference, P=0.12). These strains appeared to rapidly colonise the chicken caecum and remain colonised for at least 7 days. For the ED-positive strains, Dg95 and Dg275, recovery of viable cells from caecal content at 7 days was much lower, only 5.2 and 5.3$\log_{10}$ CFU/g, respectively. By 7 dpi, without colonisation recovery of *Campylobacter* would be below detectable levels as seen for some other strains (**Figure 5.2**, **B**). Hence, despite the poorer colonisation by the ED-positive strains, these strains showed survival in the chicken caecum and some ability to colonise. The average levels of Dg95 and Dg275 colonisation were only marginally higher at 7 dpi compared to 3 dpi, by 0.68 and 0.41 $\log_{10}$ CFU/g respectively, and there was no significant difference for either, P=0.35 and P=0.26, respectively. However, at both time points there was a broad spectrum in the level of recovery for both Dg95 and Dg275, from the 5-7 chickens infected. These results are consistent with poorer ability of these two ED-positive strains to colonise the chicken caecum. It would be interesting to monitor if colonisation ability improved with time or if the strain eventually disappears from caecal content.
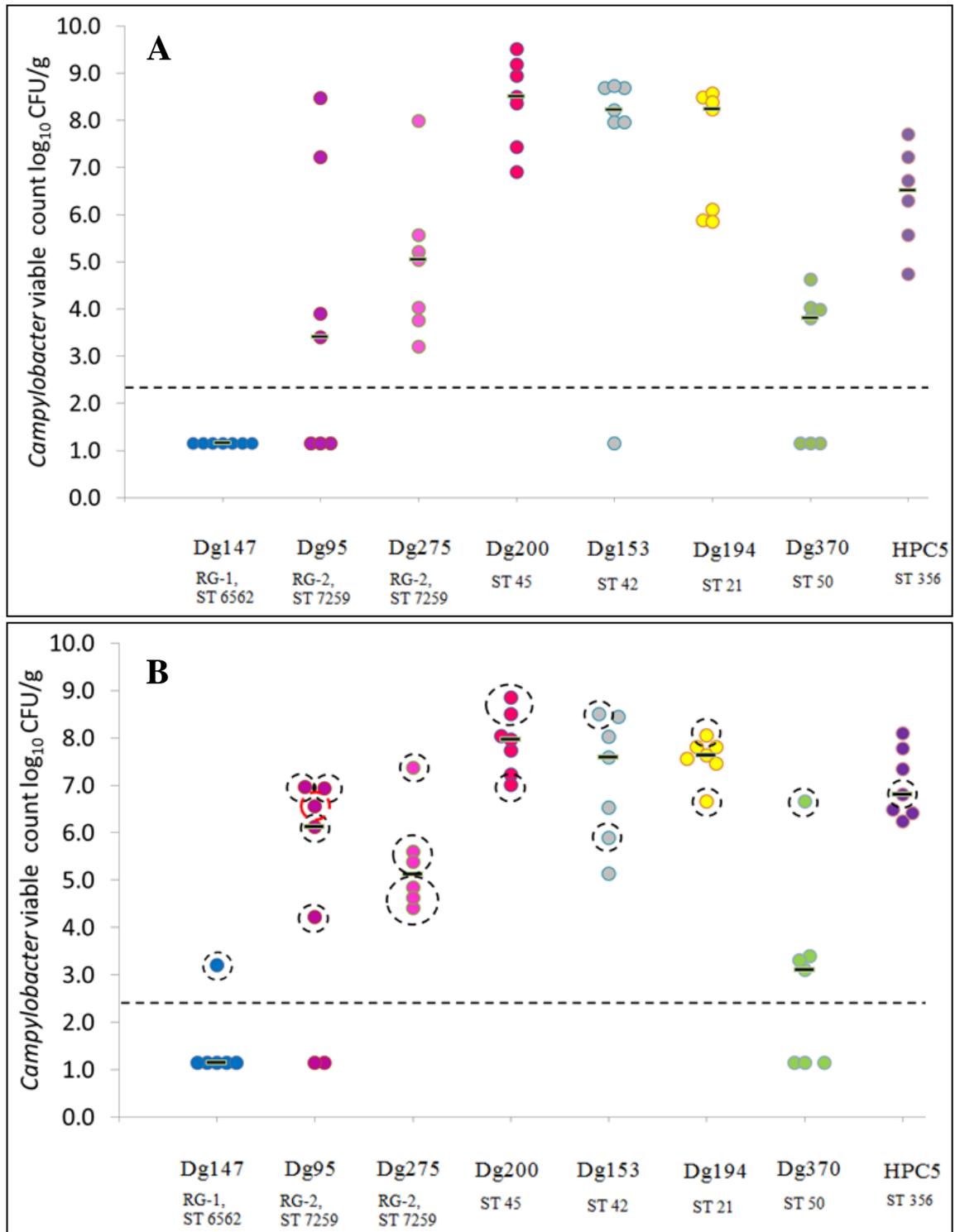
Figure 5.2: Recovery of *Campylobacter jejuni* strains from the caecum of chickens. The 16 day-old chickens were administered with ~7-log$_{10}$ CFU of the bacterial strains by the oral gavage technique as described in **section 2.6.3**. (**A**) At 3 dpi (19 days-old) and (**B**) 7 dpi (23 days-old) the chickens were sacrificed and colonisation of the caecum assessed by Miles and Misra viable count on mCCDA (**see section 2.2.3**). All strains were confirmed by colony morphology. For 7 dpi, recovered colonies were also confirmed by sequencing a target gene/s (dashed black circles), the ability to utilise glucose, and amplification of the *glc* locus (**Table 5.4**). The dashed red circle is strain Ch95-95, recovered from chicken 95. It harbours a naturally mutated *glc* locus and cannot use glucose. Non *Campylobacter* white non creamy and spread colonies were recovered

to high level of 6.07 and 8.1 log10 CFU/g at 3dpi from 2 chickens which were infected by Dg147, and showed no evidence of *Campylobacter* recovery Mann–Whitney U Test (One-tailed) with a significance level of 0.05 was followed for making statistical analysis. Dashed black lines indicate the threshold of detection, 2.3 log10 CFU/ g caecum content. Any data less than the threshold of detection (non-colonised chickens) was counted as LOD/2 (1.15 log10 CFU/g) (Wood et al., 2012), (-) median.

The Dg95 and Dg275 RG-2 isolates carry the *glc* locus permitting metabolism of glucose. There was no evidence from these results that the ability to utilise glucose provided any advantage in colonisation of chickens. The related ED-negative ST-45 clonal complex Dg200 strain was the best coloniser and compared to the RG-2 strains, it colonised to a level more than ~2.6 $\log_{10}$ CFU/g higher (p= 0.007 and P=0.001, respectively). The ED-positive strains were also significantly less efficient colonisers compared to ST-42 clonal complex Dg153 (P=0.043, P=0.023), ST-21 clonal complex Dg194 (P=0.006, P≤ 0.001), and to control ST-356 HPC5 strain (P=0.033, P=0.007), respectively. While comparison of *glc*$^-$ isogenic strain would be helpful, these results indicate that the ability to use glucose is unlikely to confer any advantage in colonising chickens and might even be a disadvantage.

Interestingly, the RG-1 strain, Dg147, behaved according to the hypothesis. It was unable to colonise the caecum of any of the 6 chickens. Viable cell counts were below the threshold of detection, 2.3 $\log_{10}$ CFU/ g, for 5 chickens and *Campylobacter* was detectable in only one of the chickens with a count of 3.2 $\log_{10}$ CFU/g. This strain clearly does not survive well within the chicken caecum. However, this was not the only strain showing no colonisation. Very low levels of recovery were also observed with Dg370 (ST-50 CC-21), although many studies have identified ST-50 CC-21 as a sequence type isolated from human samples (de Haan et al., 2010, Harvala et al., 2016, Ramonaite et al., 2017). A focused chicken trial including more strains of the RG-1 group of strains is required to establish if inability to colonise chickens is a general feature of RG-1 strains.

Motility is recognised as an essential determinant for successful chicken colonisation (Kanji et al., 2015). Therefore, the possibility that loss of motility might be the reason for poor or no chicken colonisation was tested. Both input (Dg strains) and recovered (Ch strains) were tested for motility on both 0.4% MHA and BHI plates (**Figure 5.3**). The data showed that all input and recovered strains are motile in both media without showing large differences

between both types of cells. The level of motility was not related to the level of colonisation in the caecum of the chickens. For example, both Dg153 and Dg200 were strains that successfully colonised chickens, but Dg153 was approximately twice as motile compared to Dg200. In addition, Dg370 was not able to colonise the chicken model, but its level of motility was approximately 1.5 times higher than that of the best chicken coloniser, Dg200. There were no recovered isolates of Dg147 for 7 dpi, but motility of this strain was confirmed in **section 3.1.1**. This confirms that absence of colonisation could not be simply explained by absence of motility. Factors other than motility must explain the ineffective colonisation of the chicken gut by these strains.
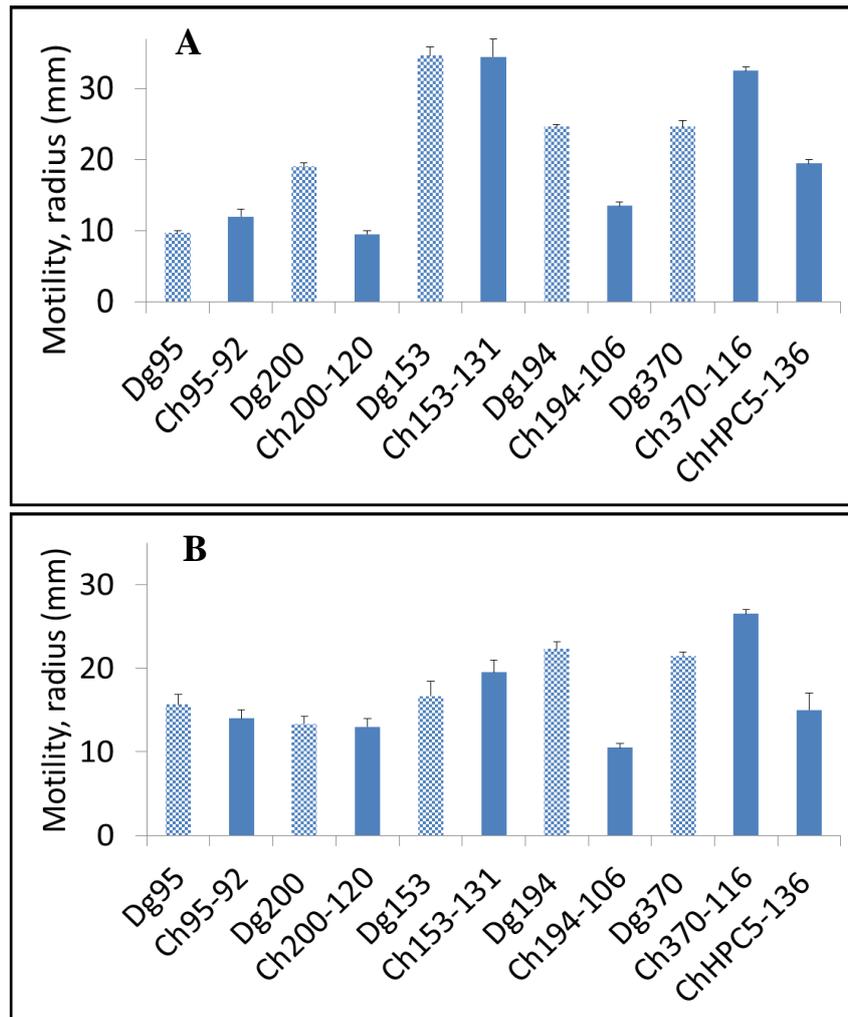
Figure 5.3: Motility property of input and recovered some of the chicken trial strains. Motility properties are shown on (**A**) 0.4% MHA and (**B**) 0.4% BHI motility plates, after 3 days incubation. Dotted histograms are input (Dg) strains and solid histograms are recovered (Ch) strains. Standard errors are from three technical plates.

Viable *Campylobacter* cell numbers present in ileum and jejunum contents were also assessed (**Table 5.1**). For strains most efficient at colonising the caecum, Dg200, Dg194, and HPC5, detectable numbers of *C. jejuni*, between 3.1 to 4.2 $\log_{10}$ CFU/g, were recovered from both the ileum and jejunum, at both 3 and 7 dpi. These values were 3-4 logs lower than the numbers recovered from the caecum and are consistent with published data showing that the caecum of poultry is the primary site of *C. jejuni* multiplication (Hermans et al., 2012). There was essentially no recovery of *Campylobacter* from the ileum and jejunum of chickens infected with RG-1 strain Dg147 and the RG-2 strains Dg95 and Dg275. Hence, these strains are not associated with an alternate site of colonisation of the chicken intestine.

Table 5.1: *Campylobacter* **counts in the ileum and jejunum of chickens post-infection**

Isolate	Average ± SE, log$_{10}$ CFU/g of each *Campylobacter* strain, ND, no detectable cells or below 3 log$_{10}$ CFU/g			
	Ileum, at 3 dpi	Ileum, at 7 dpi	jejunum, at 3 dpi	jejunum, at 7 dpi
Dg147	ND	ND	ND	ND
Dg95	ND	ND	3.34 ± 0.22	ND
Dg275	ND	ND	ND	ND
Dg200	3.01 ± 0.0	3.35 ± 0.35	3.07 ± 0.07	3.31 ± 0.31
Dg153	3.01 ± 0.01	3.14 ± 0.13	ND	ND
Dg194	3.64 ± 0.27	3.89 ± 0.4	3.08 ± 0.07	3.91 ± 0.39
Dg370	3.04 ± 0.0	ND	ND	3.14 ± 0.14
HPC5	3.37 ± 0.32	4.22 ± 0.3	3.27 ± 0.12	3.8 ± 0.19

5.3.2   Recovery of *Campylobacter jejuni* from extra-intestinal organs

As shown in **Table 5.2**, invasion of extra-intestinal organs was a rare event in this study. As might be expected, none of the RG-1, RG-2, and ST-50 Dg370 strains invaded the extra-intestinal organs at either 3 or 7 dpi. In contrast, Dg200 and Dg153 strains were identified in all of the extra-intestinal organs from 1 to 6 of the chickens at day 3, although not at 7 dpi. Dg153 infected the liver in 85% of chickens. The Dg194 chicken coloniser was not recovered from the organs at 3 dpi, but at 7 dpi some of the organs of a couple of chickens were infected with this strain. The good colonising strains showed more evidence of spread to the extra-intestinal organs without any evident symptoms of clinical illness. The only recorded effect of *Campylobacter* infection was a decrease in weight of chickens infected with Dg153 (P=0.02 compared to HPC5 control strain) (**Figure 5.4**). Interestingly, although ST 50 CC-21 Dg370 did not colonise the caecum of chickens well, the body weight of chickens infected with this strain was significantly less compared to other groups; to Dg147 RG-1 strain (p=0.03), and to non-infected control chickens (p=0.02).   Factors other than physical colonisation of the epithelial surface or system infection might be responsible for these results. For example, competition of the poor coloniser Dg370 and good coloniser Dg153

strains with the native microbiota could have an impact on weight gain (Spees et al., 2013, Indikova et al., 2015).

Table 5.2: Recovery of *Campylobacter* from extra-intestinal organs of chicken.

Campylobacter recovery by enrichment at 3 dpi, ND – none recovered

Isolates	Heart	Spleen	Liver	Breast	Kidney
Dg147	ND	ND	ND	ND	ND
Dg95	ND	ND	ND	ND	ND
Dg275	ND	ND)	ND	ND	ND
Dg200	2/7 (28%)	2/7 (28%)	1/7 (14%)	2/7 (28%)	3/7 (42%)
Dg153	4/7 (57%)	2/7 (28%)	6/7 (85%)	3/7 (42%)	3/7 (42%)
Dg194	1/7 (14%)	ND	ND	ND	ND
Dg370	ND	ND	ND	ND	ND
HPCS	1/6 (16%)	ND	1/6 (16%)	1/6 (16%)	ND

Campylobacter recovery by enrichment at 7 dpi, ND – none recovered

Isolates	Heart	Spleen	Liver	Breast	Kidney
Dg147	ND	ND	ND	ND	ND
Dg95	ND	ND	ND	ND	ND
Dg275	ND	ND	ND	ND	ND
Dg200	ND	ND	ND	ND	ND
Dg153	1/7 (14%)	ND	ND	ND	ND
Dg194	2/7 (28%)	1/7 (1 4%)	1/7 (14%)	1/7 (14%)	ND
Dg370	ND	ND	ND	ND	ND
HPCS	ND	ND	ND	ND	ND
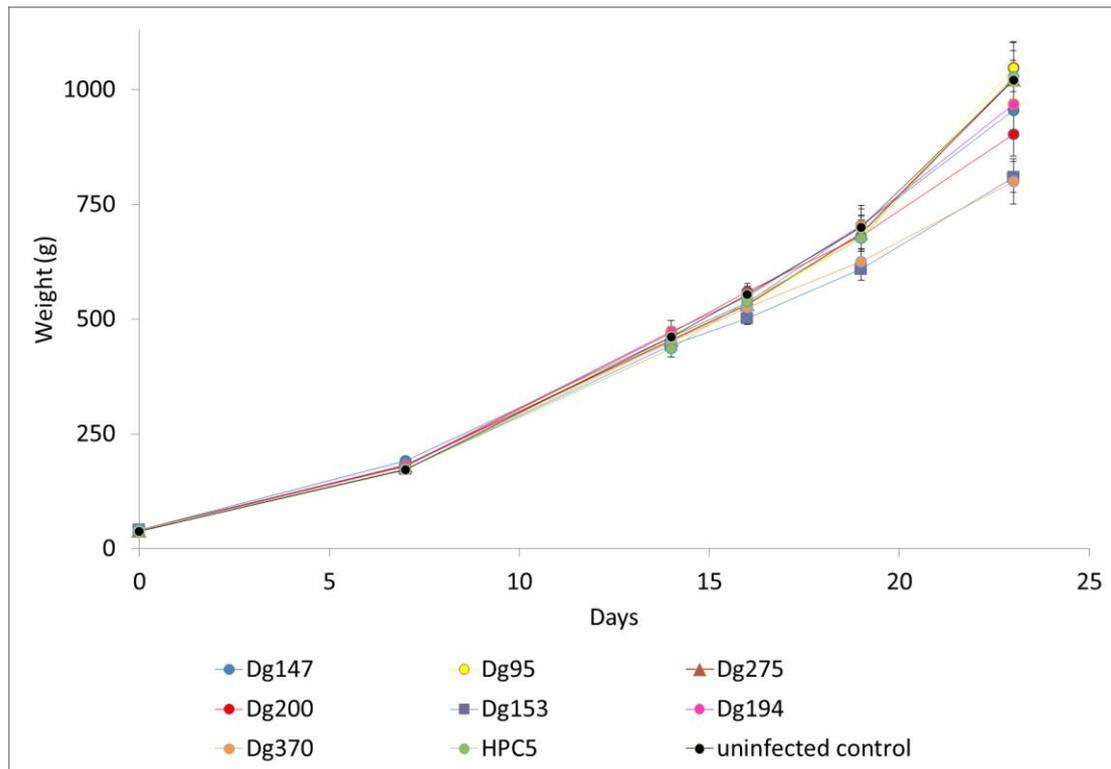
Figure 5.4: Body weight of infected and control chickens over 23 days. This was provided by Veterinarians working in the animal house. Dg275 is overlapped.

5.3.3   Confirmation of *glc* locus and MLST profile of recovered strains

Selective isolation and high numbers facilitated specific recovery of *Campylobacter* from the caecum. Colonies recovered from the caecum of each chicken, 7 dpi, and stocked in Reading, were analysed for confirmation of identity of each strain. The 7 MLST loci of strains used in the trial were screened for loci that would differentiate strains. The allele *aspA* 288 is unique to RG-1 strains and the three genes *aspA*, *gltA*, and *glyA* were identified as suitable to differentiate other strains (**Table 5.3** and **Table 5.4**). In addition, three sets of primers were used to amplify the entire *glc* locus in the RG-2 strains. While the correct 3 *glc* PCR products were recovered from 21 out of the 23 Ch95 and Ch275 isolates tested. One chicken, chicken 100, infected with Dg275 was evidently contaminated. No *glc* PCR products were produced by either Ch275-100-1 or Ch275-100-3 colonies. In addition, MLST analysis revealed that the same two isolates were unrelated to Dg275 (and more closely related to Dg194) (**Table 5.4**). Therefore, chicken number 100 was excluded from the trials. One other contaminated chicken was identified. Chicken number 89, which had been inoculated with Dg147,

possessed the allele *aspA* 4 rather than *aspA* 288. Evidence from the other 6 chickens infected with Dg147 was that this strain showed a very poor ability to colonise (**Figure 5.2** at above). Therefore, it was not surprising that the unusual colonization of chicken 89 was due to a contaminant of *C. jejuni*. Thus, chicken 89 was also removed from the study. All other recovered colonies tested gave the correct anticipated MLST PCR product.

Table 5.3: MLST genes used for confirmation of Ch-strains recovered from the chicken trial.

Isolate	*aspA*	*gltA*	*glyA*
Dg95	**1**	**10**	**91**
Dg275	**1**	**10**	**91**
Dg147	**288**	29	28
Dg194	2	**1**	3
Dg370	2	**12**	**3**
Dg200	**4**	10	**4**
Dg153	**1**	3	**4**
HPC5	14	5	**2**

Allele numbers for each strain as assigned in PUBMLST. Underlined and bold alleles were used for the confirmation of identity of strains recovered at 7 dpi from caecum.

Table 5.4: Summary of confirmed ST alleles and *glc* locus.

Strain[1]	Sequenced MLST genes[2]	*glc* locus; *eda-zwf, zwf-pgi, pgi-glcP*[3]
Dg95	*glyA* 91, *gltA* 10	Present
Ch95-92-1	*glyA* 91, *aspA* 1	Present
Ch95-93	*glyA* 91, *aspA* 1	Present
Ch95-94	*glyA* 91	Present
Ch95-94-1	-	Present
Ch95-94-5	-	Present
Ch95-95-1	*glyA* 91, *gltA* 10, *aspA* 1	Present
Ch95-95-2	-	Present
Ch95-95-3	*aspA* 1	Present
Ch95-95-4	-	Present
Ch95-95-5	*glyA* 91, *aspA* 1	Present
Ch95-97-1	*glyA* 91	Present
Ch95-97-2	-	Present
Ch95-97-5	-	Present

Dg275	*glyA* 91	Present
Ch275-99-1	*glyA* 91	Present
Ch275-100-1*	<u>*glyA* 3, *aspA* 2</u>	<u>Absent</u>
Ch275-100-2	*glyA* 91, *aspA* 1	Present
Ch275-100-3*	<u>*glyA* 3, *aspA* 2</u>	<u>Absent</u>
Ch275-101-1	*glyA* 91, *gltA* 10	-
Ch275-101-2	*gltA* 10	-
Ch275-102-1	*glyA* 91	-
Ch275-102-5	-	Present
Ch275-103-1	*glyA* 91	-
Ch275-103-5	-	Present
Ch275-104-1	*glyA* 91	Present
Ch275-104-5	-	Present
Ch275-105-1	*glyA* 91	Present
Ch275-105-5	-	Present
Dg147	*aspA* 288	NA
Ch147-85	*aspA* 288	NA
Ch147-89*	<u>*aspA* 4</u>	NA
Dg194	*gltA* 1	NA
Ch194-106	*gltA* 1	NA
Dg370	*gltA* 12	NA
Ch370-116	*gltA* 12, *glyA* 3	NA
Dg200	*aspA* 4	NA
Ch200-120	*aspA* 4, *glyA* 4	NA
Ch200-122	*glyA* 4	NA
Ch200-124	*aspA* 4, *glyA* 4	NA
Ch153-127	*aspA* 1, *glyA* 4	NA
Ch153-129	*aspA* 1, *glyA* 4	NA

[1]Input strains (Dg numbers), recovered strains (Ch numbers, with strain number-chicken number-recovered isolate number). *Strain with wrong allele, chicken excluded from the trial [2]MLST loci confirmed by amplification from lysed colonies and sequencing PCR product. Underlined loci, wrong indicates contaminant. (-) indicates a gene or an operon that has not been tested. NA, not applicable.

5.3.4   Loss of the ability to utilise glucose during the *in vivo* chicken trial

The ED pathway is found within a plasticity region of the bacterial genome and can be naturally transferred between *C. coli* strains (Vorwerk et al., 2015). The ED positive strains Dg275 and Dg95 colonised ~2.5 $\log_{10}$ CFU/g at 7 dpi less effectively than the related CC-45 strain Dg200. The *glc* locus had not been identified in the genome sequence of any chicken isolate sequence in the database. It was therefore of interest to establish how stable this locus is during *C. jejuni* infection in chickens. PCR analysis had already identified presence of the *glc* locus in all recovered isolates tested (**Table 5.4** at above). Five single colonies had been stocked from the day 7 caecal samples from each of the 6 chickens colonised with Dg275 and the 5 chickens colonised with Dg95. To test for functionality of the *glc* locus, all 55 colonies were screened in a microtire plate growth assay for enhanced growth in DMEMf in the presence of glucose (**Figure 5.5**). For most isolates, growth increased from a maximum OD of around 0.1 in the absence of glucose to around 0.2 or higher in the presence of glucose, comparable to the ED positive input strains, Dg95 and Dg275. Interestingly, for all five Ch95-95 colonies, the presence of glucose in the medium had no impact on growth. These recovered strains behaved in the same way as ED-negative Dg200 and Dg162 strains, despite confirming the presence of the *glc* locus (**Table 5.4** at above). To confirm the loss of glucose utilisation, the recovered Ch95-95 strains, as well as the ED-positive Dg95 and Dg275 strains were grown with shaking in Erlenmeyer flasks (**Figure 5.6**). Viable cells were monitored at 24h and 36h growth in DMEMf, with or without glucose. Ch275-99, Ch275-100, Ch275-101, Ch275-102, Ch275-103, and Ch95-92 all utilised glucose. At 36h, the viable counts had increased from 2 to 4.2 $\log_{10}$ CFU/ml without glucose and from 6 to 8.8 $\log_{10}$ CFU/ml with glucose. In contrast, viable counts for Ch95-95-1 and Ch95-95-5 strains remained at ~2.2 $\log_{10}$ CFU/ml in media with or without glucose. These data were consistent with the data from cells growing in microtiter plates.
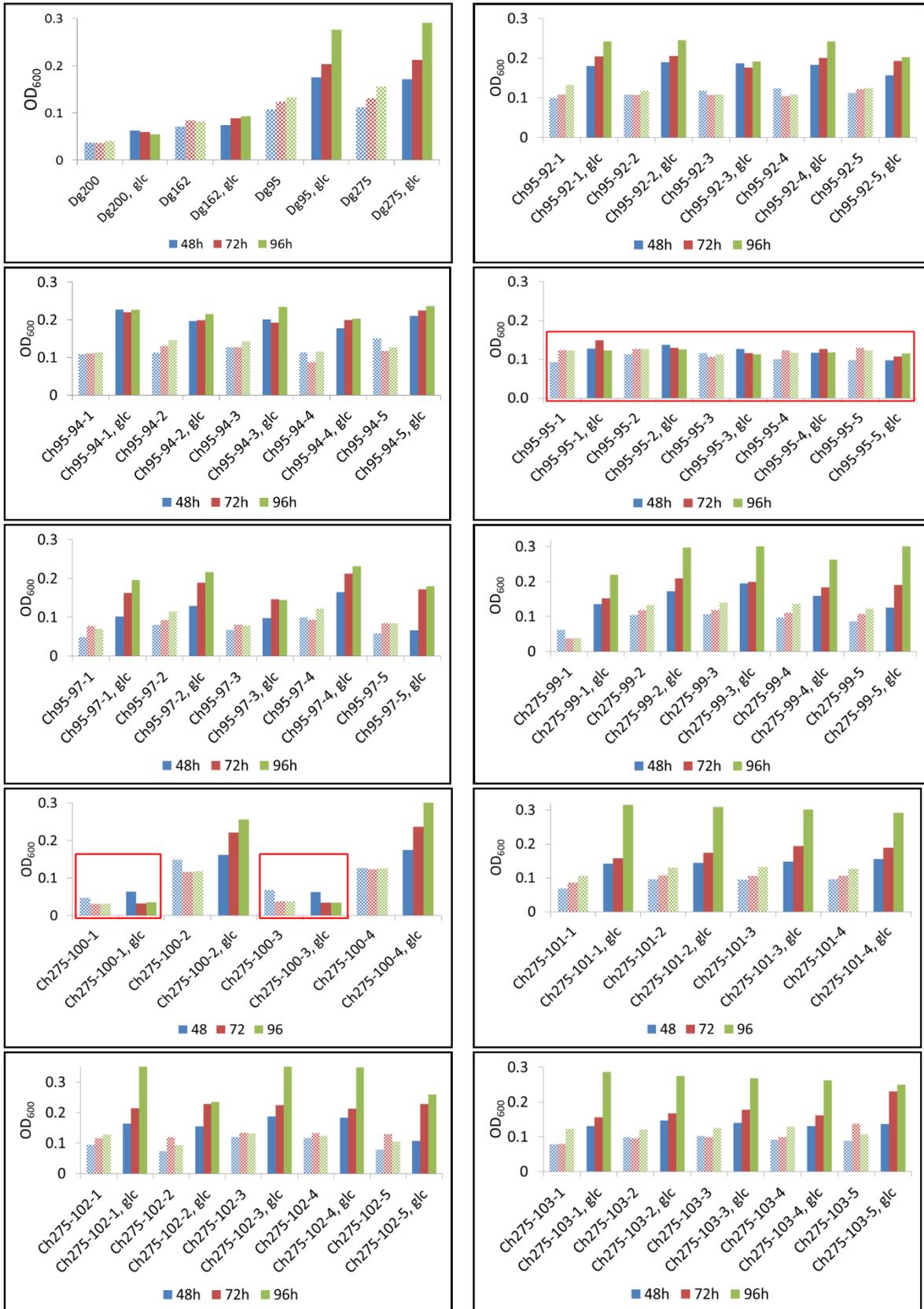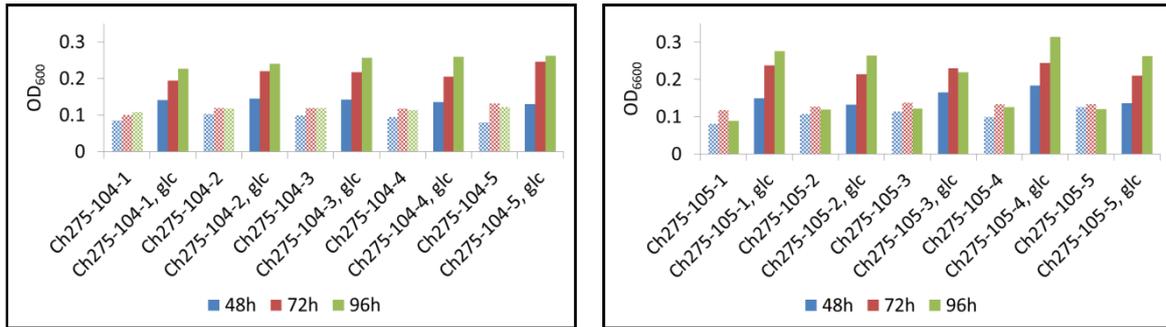
Figure 5.5: Microtitre plate screens for glucose utilisation of the recovered ED-positive strains. Stocked Dg strains (ED positive, Dg95 and Dg275), ED negative strains (all other Dg strains), and isolates from chicken trials (Ch95 or Ch275) were grown in the DMEMf with or without 20mM glucose. Dotted bars, no glucose; bars, 20mM glucose; red rectangles, recovered isolates that were unable to utilise glucose. Ch275-100-1 and -3 were contaminants and this chicken has been excluded from trial results. The data is averaged from three technical wells.
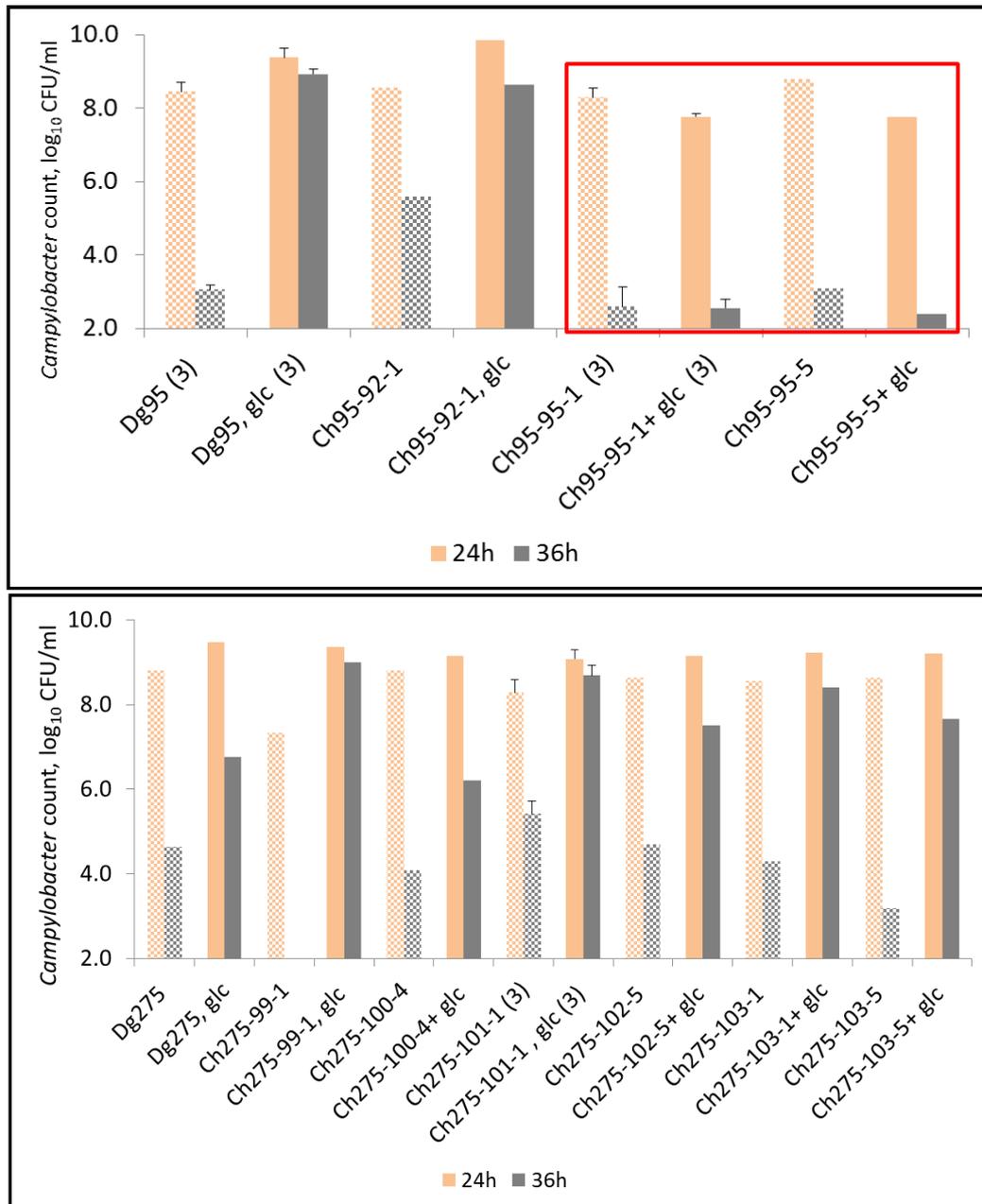
Figure 5.6: Viability of the bacterial cells of the ED- positive strains. DMEMf medium was inoculated with 0.002OD of the input (Dg) or recovered (Ch), isolates and incubated for 36h at 37°C in flasks with shaking at 150rpm. Aliquots were taken at 24h and 36h and monitored for viable cells (CFU). Stippled bars, no glucose; solid bars, 20mM glucose. Red rectangle indicates strains unable to utilise glucose. Error bars are from three biological replicates.

As mentioned above, key MLST genes and presence of the *glc* locus confirmed that the Ch95-95 strain is an RG-2 strain (**Table 5.4** at above). RG-2 strains were shown to possess 2 copies of the *glc* locus, within the *rrnA* and *rrnB* locus respectively (**see section 4. 6**). To confirm that both copies of the *glc* locus are still present in the non-glucose utilisers, the ED pathway was amplified using specific primers flanking each of the *rrnA* and *rrnB* loci (**Figure 4.11** at above). **Figure 5.7** shows that the *glc* locus was inserted into both of the *rrn* operons of the strain Ch95-95, as with the other strains tested. Each PCR product migrated in accord with the expected size of *rrn* plus the *glc* locus (14697bp).
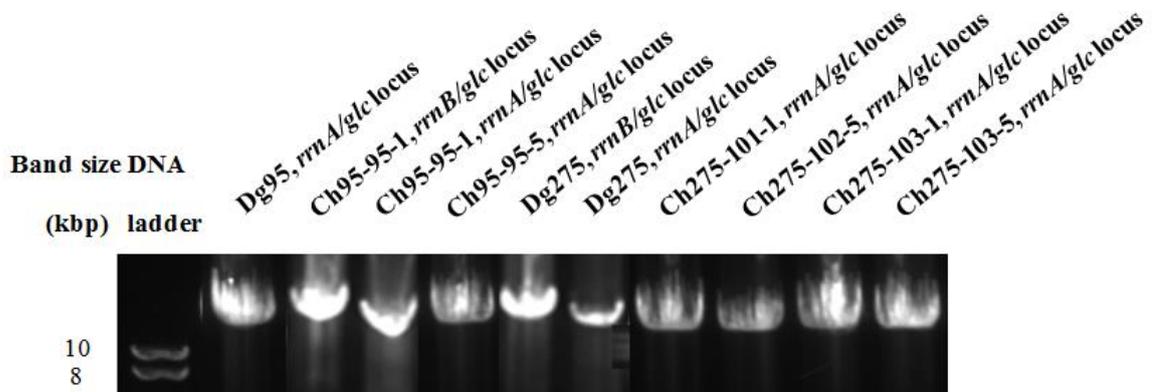
Figure 5.7: Agarose gel of the PCR products of the *rrn/glc* locus. Primers and primer characteristics are listed in methods chapter, **Table 2.4**. The following pairs of primers were used; Dg95, Dg275-16F/R; Ch95-95-1, Dg275-4FA/RA; Ch95-95-1, Dg275-16FA/RA; Ch95-95-5, Dg275-16F/R; Dg275, Dg275-4FA/RA; Dg275, Dg275-16FA/RA; Ch275-101-1, Dg275-16F/R; Ch275-102-5, Dg275-16F/R; Ch275-103-1, Dg275-16F/R; Ch275-103-5, Dg275-16F/R. The gel was run in 0.6% agarose at 80V for 75min. A 1Kbp Hyperladder was used.

5.4 Illumina whole-genome sequencing of Dg95-95 identifies a pseudogene in the *glc* locus

All 5 isolates of Ch95-95 recovered from the chicken trial possessed the *glc* locus but were unable to utilise glucose for growth. Therefore, it was proposed that there must be either an extragenic mutation affecting expression or function of the *glc* locus, or a mutation within one or both of the *glc* loci within this strain. To identify any mutations, both the input and three recovered Ch-95 isolates from the chicken trial were sent for Illumina WGS. Input and output Dg275, and Dg200 strains were also sent for WGS for comparison. **Figure 5.8** shows the genomic DNA sent for sequencing. This had been prepared from growth of a single colony on a BA plate, which had in turn been prepared from a single colony. Some properties of the sequenced DNA are summarised in **Table 5.5**. The number of contigs was low, between 27 to 35 contigs, apart from strain Ch200-122-1 which had 50 contigs. The GC content of the Ch95-95 was 30.46%, which was similar to the input Dg95.
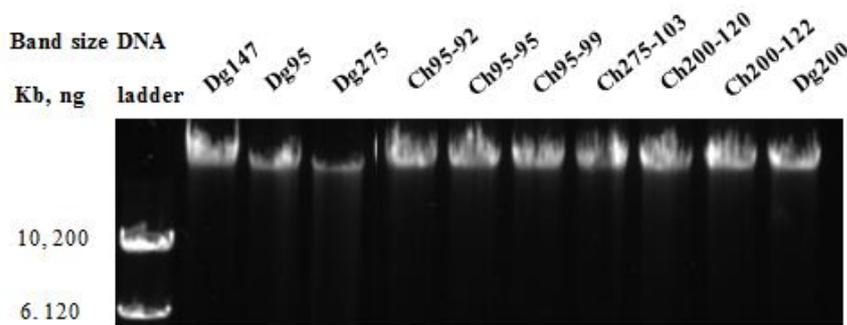
Figure 5.8: Gel electrophoresis of genomic DNA of the sequenced chicken trial strains. 4µl of High DNA Mass Ladder and 3µl of the genomic DNA of the sequenced samples were run for 120min at 80V in 0.75% agarose. The GeneTool programme was used to quantify the DNA in the purified genomic DNA. The DNA concentration was between 200-250ng/3µl except for Dg95 and Dg275, which were 137/3ul and 58ng/3µl, respectively. The samples were then normalised to 30ug/µl in 100µl of Elution buffer from the GeneJet PCR purification kit (10 mM Tris-HCl, pH 8.5).

Table 5.5: Characteristics of the sequenced strains from the chicken trial.

Strain	Total contigs	Largest contig (bp)	Total length (bp)	GC (%)
Dg147	32	562818	1672457	30.43
Dg95	27	429112	1636102	30.46
Dg275	29	429105	1670521	30.45
Ch95-92-1	30	314094	1636730	30.47
Ch95-95-1	32	429111	1636101	30.46
Ch275-103-1	31	429107	1669570	30.44
Ch200-120	35	325750	1652573	30.48
Ch200-122	50	353086	1650113	30.48
Dg200	28	353085	1648271	30.47

To find the *glc* locus in the WGS data for the input and the recovered ED-positive strains from the chicken trial, the draft genome contigs of each recovered strains was opened as a single file in DNADynamo and then searched with a fragment of one gene of the *glc* locus. Identified contigs encoding the ED pathway from Dg95 and Dg275 input strains, and Ch95-92-1, Ch95-95-1, and Ch275-103-1 recovered strains, were then aligned. Alignment of these sequences confirmed that the complete *glc* locus (8358bp) of the Dg95 input and Ch95-92-1 recovered strain, and the Dg275 input and Ch275-103-1 recovered strains are identical, aside from a single nucleotide polymorphism (SNP), GGA to TGA, in the *glk* gene of the *glc* locus in the recovered strain Ch95-95-1 (**Figure 5.9, A**). The SNP occurred in nucleotide number 490 of the *glk* gene (1005bp) and replaced a Gly codon with a TGA stop codon (Povolotskaya et al., 2012, Korkmaz et al., 2014). This results in a drastically truncated protein of 164 instead of 335 amino acids, which could readily explain non-functionality of this *glc* locus (**Figure 5.9, B**).

```
A
Dg268      GCTTTATCTGGGGAAGGTGGGCATACGAGTTTTTCTCCTTTTGATGATACTGAAGTGATA
Dg345      GCTTTAGCAGGAGAGGGTGGGCATACTAGTTTTTCTCCTTTTGATGATACTGAAGTGATA
269.97     GCTTTATCTGGGGAAGGTGGGCATACGAGTTTTTCTCCTTTTGATGATACTGAAGTGATG
Dg217      GCTTTAGCAGGAGAGGGTGGGCATACTAGTTTTTCGCCTTTTGATGATACTGAAGTTATG
Dg22       GCTTTAGCAGGAGAGGGTGGGCATACTAGTTTTTCGCCTTTTGATGATACTGAAGTTATG
Dg349      GCTTTAGCAGGAGAGGGTGGGCATACTAGTTTTTCGCCTTTTGATGATACTGAAGTTATG
Dg95       GCTTTAGCAGGAGAGGGTGGGCATACTAGTTTTTCGCCTTTTGATGATACTGAAGTTATG
Dg275      GCTTTAGCAGGAGAGGGTGGGCATACTAGTTTTTCGCCTTTTGATGATACTGAAGTTATG
CHW470     GCTTTAGCAGGAGAGGGTGGGCATACTAGTTTTTCGCCTTTTGATGATACTGAAGTTATG
Ch95-95    GCTTTAGCATGAGAGGGTGGGCATACTAGTTTTTCGCCTTTTGATGATACTGAAGTTATG
           ****** *   * ** ********** ******* ****************** **

B
>Dg268     ALSGEGGHTSFSPFDDTEVI
>Dg345     ALAGEGGHTSFSPFDDTEVI
>269.97    ALSGEGGHTSFSPFDDTEVM
>Dg217     ALAGEGGHTSFSPFDDTEVM
>Dg22      ALAGEGGHTSFSPFDDTEVM
>Dg349     ALAGEGGHTSFSPFDDTEVM
>Dg95      ALAGEGGHTSFSPFDDTEVM
>Dg275     ALAGEGGHTSFSPFDDTEVM
>CHW470    ALAGEGGHTSFSPFDDTEVM
>Ch95-95   ALA-EGGHTSFSPFDDTEVM
           **   **************
```

Figure 5.9: Multiple alignment of part of the *glk* DNA and protein sequences. (**A**) DNA and (**B**) protein sequences of the *glk* gene of the strains are listed in **Table 5.6** were aligned using Clustal Omega (Sievers et al., 2011). The DNA was extracted from bp number 480 to 540 of the *glk* gene. The same region was translated to create the amino acids sequences. Dg268, Dg345, Dg217, Dg22 and Dg95 were farm-associated *C. jejuni* strains. Dg349 was farm-associated rat *C. coli*. CHW470 was human-associated *C. coli* isolate (Vorwerk et al., 2015) and strain 269.97 *C. jejuni* subsp. *doylei* was isolated from a blood sample of human bacteraemia in South Africa (GeneBank database). Green coloured GGG and GGA codons show the location of the point mutation in the Ch95-95 recovered ED-positive *C. jejuni* strain. The codon has been replaced by stop codon TGA (green and red) in the mutated gene. The stop codon results in a peptide of 164 amino acids (dashed red colour).

Whilst the loss of a functional glucokinase (encoded by *glk*) would explain the loss of glucose metabolism via the ED pathway, there are two copies of the *glc* locus in Dg95. Therefore, both *glc* loci were specifically amplified on a 15kbp fragment together with *rrnA* and *rrnB* as described in **Figure 5.7** at above. Each product was then used as a template to amplify and sequence both copies of *glk*. This confirmed the presence of the same GGA to TGA SNP in *glk* within both the *rrnA* and *rrnB* associated *glc* loci of Ch95-95-1.

There are seven different *glk* alleles among 39 *C. jejuni* and 3 *C. coli* rat-associated ED-positive strains (**Table 5.6**). Both ED type 3 Dg95 and ED type 2 Dg275 have the common *glk* allele number 4, hence 64.2% of these strains. Multiple sequence alignment of the *glk* gene among the different strains showed some variation in this area with both GGG and GGA codons encoding Gly. The mutant stop codon was not observed in any of these strains (**Figure 5.9** at above). The inability of strain Ch95-95-1 to utilise glucose supports the consensus that a non-functional or missing *glk* gene in the typical *C. jejuni* genome is key in the inability of glucose utilisation by *C. jejuni* (Guccione et al., 2008, Hofreuter et al., 2008, Kelly, 2001, Velayudhan and Kelly, 2002).

Table 5.6: The *glk* allele numbers in 39 *C. jejuni* and 3 *C. coli* ED-positive farm associated strains.

Allele number	4	8	6	7	12	13	9
Number of strain (42 in total) (%)	27 (64.2)	4 (9.5)	3 (7.1)	3 (7.1)	2 (4.7)	2 (4.7)	1 (2.3)
Example of strain/s	Dg95, Dg275	Dg18	Dg268	Dg349	Dg217	Dg345	Dg22
***Campylobacter* Species**	*C. jejuni*	*C. jejuni*	*C. jejuni*	*C. coli*	*C. jejuni*	*C. jejuni*	*C. jejuni*

5.5 Conclusion

The data presented in this chapter supports the hypothesis that neither the RG-1 nor the RG-2 strains tested are highly adapted to colonisation of chickens. Initial infection trials with *G. mellonella* larvae provided an initial indication of differences in virulence between strains. The majority of the larvae, approximately 75%, remained healthy when infected with Dg147 or Dg300 RG-1 strains. In contrast, the strains belonging to the generalist complexes ST-21 (11168H and Dg194) and ST-45 (Dg200) clonal complexes killed approximately 50% of the insects at the same dose.

A much greater difference was observed between strains in the chicken caecum colonisation trial. The example strain of RG-1, Dg147, could not colonise the caecum even after 7 dpi. Specific attributes that Dg147 lacks for chicken colonisation are not clear. ST-21 and ST-45 clonal complexes strains are considered generalist strains and are the most commonly recovered strains from clinical infection and are also commonly recovered from chickens (Colles and Maiden, 2012, Mullner et al., 2009, McCarthy et al., 2007, Sheppard et al., 2009, Wilson et al., 2008, de Haan et al., 2010). It was therefore not surprising that strain Dg194, ST-21 clonal complex, was one of the most effective chicken colonisers tested. The reason why Dg370, ST 21 clonal complex, ST 50 colonised so poorly is not known. The average number of viable cells in four of the chickens was 3.64 slog10 CFU/g at both 3 and 7 dpi. Draft genome comparisons of the chicken trial strains, highlighted the absence of genes between *cj1421* to *cj1442* in Dg370 compared to Dg194 (s**upplementary Table 5.1**). This locus is known to be involved in the biosynthesis of capsular polysaccharides (CPS) (Karlyshev et al., 2005). Among these genes, *cj1421, cj1422, cj1426* and *cj1429* are hypervariable sequences (Parkhill et al., 2000). CPS modification enhances protection from bacteriophage in the *in vivo* environment of chickens (Holst Sorensen et al., 2012) and capsule production is a key factor in the invasive properties of ST677CC strains (Skarp et al., 2015). Lack of this locus might contribute to the inability of Dg370 to colonise chickens.

Data thus far suggests that possession of the ED pathway is restricted to environmental rather than clinical strains. There is also no evidence of the *glc* locus in strains isolated from chickens (Vegge et al., 2016). Results from this study could be consistent with this. There is no evidence that the ability of the RG-2 strains to use glucose enhances colonisation of chickens. At both time points, RG-2 strains colonised the caecum of the chickens less efficiently than the closely related ED-negative CC-45 strain Dg200. Average recovery of RG-2 strains from the caecum was at 4.5-5.4 $\log_{10}$ CFU/g, which was at least 1.5 $\log_{10}$ CFU/g less than the level recovered from chickens inoculated with Dg200, Dg153, Dg195, and HPC5 strains. Whether possession of the ED pathway actually hinders colonisation, resulting in negative selection against carriage of this locus, remains to be seen. However, during the trial, Dg95 evolved in one chicken, resulting in recovery of ED-negative strains carrying a pseudo *glk* gene. All five recovered isolates from chicken 95 were negative for growth in glucose, suggesting that the mutation occurred either quite early on or that strains with this mutation rapidly outgrew any ED-positive strains. Irrespective of when the mutation occurred, non-glucose utilisers dominated this environment and this strain, despite losing the ability to use glucose, colonised the caecum to a level of 6.5 $\log_{10}$ CFU/g, which was close to the highest level of colonisation achieved for Dg95. The ability to use glucose is clearly not necessary for colonisation of chickens by ED-positive strains. It remains to be seen what the advantage of this locus is to environmental strains and if there is negative selection against possession of this locus by chickens.

Chapter 6

General Discussion and

Future Directions

General Discussion and Future Directions

C. jejuni is an important food-borne pathogen that causes gastroenteritis worldwide. The bacterium can be carried by many different hosts including farm animals, turkeys, chickens and wild birds (Zautner et al., 2011, Penner, 1988). *C. jejuni* may live as part of the microbiota in both wild animals and livestock, and can even be transferred to humans by pets. Chicken meat is considered a primary source of transmission of *C. jejuni* within the food chain leading to infection in humans. Understanding the mode of transmission of *Campylobacter* around the farmyard and how it enters the food chain via chicken, sheep, pig and cattle colonisation will assist in elimination and control of this food-borne bacterial pathogen. A couple of studies have identified *Campylobacter* spp. in wild rats and shown that these rodent animals have a potential role in contribute to the spread of *Campylobacter* (Meerburg and Kijlstra, 2007a, Meerburg et al., 2006). Through contamination of grain stocks and water with rat faeces, *Campylobacter* can be transmitted to primary reservoirs such as chicken and farmed animals (Nkogwe et al., 2011). A bank of *Campylobacter* strains had been isolated from farm associated Norway rats (*Rattus norvegicus*) and the genomes sequenced (www.pubmlst/campylobacter) as a start to assessing the ability of these animals to act as a reservoir for pathogenic strains of *Campylobacter* in the farmyard. The majority of isolates were *C. jejuni*, but also included *C. coli*, *C. lari*, *C. hyointestinalis* and *C. upsaliensis*. Genomic analysis showed that among the 137 sequenced strains were typical generalist sequence types (ST21CC, ST45CC) commonly associated with farm animals, chickens and human disease, and also some recognised specific host associated sequence types, including ST42CC commonly associated with cattle and ST828CC, associated with high prevelance in pigs (Kwan et al., 2008). Example strains from each of these CCs, Dg194 (ST21), Dg200 (ST45) and Dg153 (ST42), were investigated in some depth in this study and shown to grow well in *vitro* and to very efficiently colonise chickens. Dg153 also consistently formed dense biofilm whether in floating liquid culture or on a solid surface. Hence, it can be concluded that wild Norway rats can harbour both generalist (ST21CC and ST45CC) and the cattle associated (ST42CC) strains of *C. jejuni* that are associated with human disease and that have the ability to be transferred into the food chain via colonised chickens.

This study primarily focused on two additional unusual and dominant groups of rat associated *C. jejuni,* each of which was frequently isolated over the three-year collection period from a variety of different farm sites. The first of these, RG-1 with 30 strains, clustered as a very distinct phylogenetic clade based on WGS and rMLST. The RG-1 group could not be assigned to an existing CC group, due to the absence of at least 4 shared MLST alleles. Comparisons of 16S rRNA sequence, presence and absence of specific gene markers for the identification of species and subspecies and MALDI-TOF MS all confirmed these strains as *C. jejuni.* Because these strains were phylogenetically distinct and there was no record of strains closely related to these in the very extensive BIGSdb database, these were considered to potentially represent host specific strains for rats.

Growth of the RG-1 strains in rich and supplemented minimal media was generally much poorer than that of the generalist sequence types of *C. jejuni* also isolated from Norway rats. The reason why this clade did not grow well is not clear. RG-1 strains possess the genomic information required to utilise the key amino acids Ser, Asp, Glu and Pro (Velayudhan et al., 2004) and ability to use these amino acids was confirmed experimentally. Interestingly, growth properties of this group were similar to the control strains when supplemented with the rat mucin, a fact that would be consistent with adaptation of these strains to rat colonisation. Other *C. jejuni* strains have been identified that show metabolic differences associated with host association. For example, genes required for fucose uptake and metabolism (Stahl et al., 2011) and vitamin B_5 synthesis (Sheppard et al., 2013) are located on genome islands and possession of these genes is variable among *C. jejuni* strains. In recent surveys, operons for the metabolism of fucose have been shown to be absent in approximately half of the *C. jejuni* strains analysed, and more dominant in *C. coli* strains originating from pigs (Dwivedi et al., 2016). Host specific operons for the metabolism of vitamin B_5 are found in *C. jejuni* strains originating from ruminants, but not in those originating from chickens. The reason for this has been attributed to the food given to chickens (grain), which is rich in vitamin B_5, and should compensate for any vitamin deficiency (Sheppard et al., 2013). This is most likely true for rats as well, as grain is a preferred food source (Wang, 1990). Both of these sets of genes, for fucose metabolism and vitamin B_5 synthesis, are absent from the genome of all strains belonging to RG-1 genomes. Whether supplementation with Vitamin B_5 enhances growth of RG-1 strains remains to be tested.

Gene-by-gene comparisons have identified a core set of 247 genes for the RG-1 strains, 3 *C. jejuni* cgMLST genes defined by Cody et al. (2017), but missing from the RG-1 strains and 35 genes present in all RG-1 strains (core genes), but variably present in a selected set of strains known to be pathogenic to humans. Whether any of the 3 defined *C. jejuni* cgMLST genes missing from RG-1 strains are required for colonisation of chickens should be tested experimentally. Similarly, many of the core RG-1 genes unique to RG-1 or variable in other strains represent potential colonisation factors and as such targets for future investigations into specific host specificity of the RG-1 strains. Examples might include the lack that the distantly related *cdtABC* genes of RG-1 strains which are more similar to those of *C. lari* RM16701 and may interact with different host cells.. Iron uptake is known to be critical to virulence. RG-1 possesses four of the characterised iron uptake system, mediating uptake of iron as heme, ferri-rhodotorulic acid, ferrichrome and ferrous iron (Miller et al., 2009). However, the key *cfrA* gene of ferric enterobactin is absent in Dg147. A mutation in the *cfrA* (cj0755) gene in *C. jejuni* has been shown to inhibit the ability of strain NCTC11168 to colonise the gut of chickens (Palyada et al., 2004). A second system, known as the ferri-transferrin iron uptake system, is also absent in Dg147. Lack of these genes might contribute to the inability of Dg147 to colonise chickens. Interestingly, analysis of the annotated Dg147 genome identified a locus nine genes for an additional putative iron uptake system (7423bp). This system shared 92% identity with the *C. jejuni* subsp. *doylei* 269.97 genome and 82% identity to *C. helveticus*. It could be important in alternate host or niche colonisation.

The inability to colonise chickens and enhanced growth in the presence of rat mucin is consistent with the hypothesis that the RG-1 strains do not effectively colonise chickens or humans, but are rat-adapted strains. Confirmation of this hypothesis will require extending chicken trials to include additional RG-1 strains, identification of factors within rat mucin that might enhance growth and animal trials with a Norway rat infection model. Strains have clearly infected wild Norway rats. Strains were recovered from faecal pellets or intestine. Importantly, the ability of RG-1 strains to colonise strains of laboratory rats needs to be tested. A laboratory rat colonisation model would permit design of experiments to identify factors involved in adaptation of these strains to the rat intestine. If RG-1 strains can efficiently colonise rats, they could represent an excellent animal model for *Campylobacter*-host interaction. Trials also including the potentially more virulent generalist strains isolated from Norway rats may also be useful in defining virulence properties as well as colonisation.

Glucose utilising (ED positive) RG-2 strains were the second unusual group of Norway rat associated strains investigated in this study. These strains belong to ST45 clonal complex, sharing 4 of the 7 MLST alleles with the central ST45CC strain, but in addition to possessing the *glc* locus represent a phylogenetic group distinct from the ED-negative CC-45 rat-associated strains. While two previous studies have demonstrated catabolism of glucose by *C. coli* strains (Vegge et al., 2016, Vorwerk et al., 2015), this is the first demonstration of utilisation of glucose by the important pathogen *C. jejuni*. However, among strains in the BIGSdb database, only a small number of *C. jejuni* strains harbour the *glc* locus (Jolley and Maiden, 2010, Vegge et al., 2016). This locus is predominant in strains that originate from Norway rats (39 strains), wild birds (30 strains), farm environments (2 strains) and environmental waters (1 strain) (Jolley and Maiden, 2010, Vegge et al., 2016). From over 6000 isolates within the database, none of the ED-positive *C. jejuni* subsp *jejuni* strains are from humans or chickens. Thus, it was suggested that this pathway is mainly restricted to strains that are from the environment and rats. For this reason, the probability of transmission of these strains via the food chain to humans or chickens appears low. Another possibility is that ED-positive *C. jejuni* strains are transmitted to chickens, but that possession of the *glc* locus is a disadvantage in the chicken host and these strains are selected against or the *glc* locus itself is lost. In accord with this, during the short time of the chicken trial one of the recovered strains, Ch95-95, had lost the ability to utilise glucose through a single SNP creating a pseudogene in *glk* (**Chapter 5**). This ED negative strain dominated in chicken 95 and colonised at the same or higher rate in comparison to glucose utilising strains. Reasons why possession of the ED pathway might help strains to survive in the environment demands further investigation. ED pathway products are different slightly from those of the EMP pathway. In addition to ATP and NADH molecules, the pathway results in production of one molecule NADPH per molecule of glucose (Flamholz et al., 2013). NADPH has reductive characteristics and can be utilised for oxidative stress protection, a vital requirement for *Campylobacter*, as well as having a role in normal physiology and cell survival (Stanton, 2012). Thus, the ED pathway may be beneficial to these ED-positive strains through protection from oxidative stress when exposed to high levels of oxygen in the environment.

The *glc* locus is inserted within an *rrn* operon (Vorwerk et al., 2015). This study shows the first demonstration of the insertion into 2 of the 3 *rrn* operons in *C. jejuni* (**Chapter 4**). Whether this is specific to *C. jejuni* remains to be established. Genomic DNA of both Dg275

and Dg95 was made from a population of cells that originated from a single colony. ED pathway sequencing of the mutated Ch95-95 strain showed that both copies of the *glc* locus, inserted into *rrnA* and *rrnB* operons, have the same SNP within the *glk* gene. This provides evidence that the ED pathway rapidly recombines between *rrnA* and *rrnB*. Why it was not recombined into *rrnC* locus in any of the *C. jejuni* strains analysed is not clear. The copy number of the *rrn* operon varies from 1 to 15 per bacterial genome (Rainey et al., 1996), with between 10 to 15 rRNA operons in *Clostridium paradoxum* (Rainey et al., 1996) and *Bacillus subtilis* (Loughney et al., 1983) and 7 copies in the enteric bacterium *Escherichia coli* (Ellwood and Nomura, 1980). However, there are some bacteria which possess only a single copy RNA operon, such as *Rickettsia prowazekii* (Andersson et al., 1995) and *Mycoplasma pneumonia* (Bercovier et al., 1986). It has been demonstrated that deletion of one of the *rrn* operons in *E. coli* did not have a significant negative effect on its growth (Condon et al., 1995), and only one *rrn* copy is required for *E. coli* to synthesise the required number of ribosomes for a high growth rate (Bremer, 1975). However, inactivation of one or two of the *rrn*s in *C. jejuni* may have a negative effect and it is important to investigate if insertions within the *rrn* loci have a negative impact on growth *in vitro* and competitive survival *in vivo*.

Another important result is the demonstration of preference for Ser and Asp over glucose. Only after depleting these amino acids, was glucose used suggesting some regulatory mechanism for this diauxic effect. Although it is possible that the *glc* locus is expressed only in stationary phase, the data on Ser and glucose utilisation suggests that the presence of serine inhibited glucose uptake or expression of the *glc* locus genes. Examples of 2 different diauxic phenomena are those found in *Azotobacter vinelandii* (George et al., 1985), and *E. coli* (Loomis and Magasanik, 1967). *A. vinelandii* can uptake acetate before glucose and acetate inhibits glucose uptake. On the contrary to *C. jejuni* and *A. vinelandii*, *E. coli* utilises glucose before lactose via catabolite repression. Similar behaviour is found in *Pseudomonas,* which uses organic acids before glucose during the first stage of growth (Lynch and Franklin, 1978). Proline is the least preferred amino acid by *C. jejuni* (Wright et al., 2009, Guccione et al., 2008). This might relate to the prediction that *putP* (*cj1502c*) and *putA* (*cj1503c*) genes that encode proline permease and the enzyme that oxidises proline to glutamate, respectively (Cairney et al., 1984), are upregulated at the stationary phase of growth. In contrast, it has been shown that *pebC* (*cj0922c*) and *peb1A* (*cj0921c*) are upregulated throughout the whole

growth cycle. *peb1A* encodes for aspartate and glutamate periplasmic binding and transport protein, and the *pebC* gene encodes for an ATP binding protein (Leon-Kempis et al., 2006). Upregulation of genes relating to aspartate and glutamate before genes of proline might result in a reduced preference of *C. jejuni* to proline. It has been shown that the ability to use glucose has helped the ED-positive *C. coli* strains to make more biofilm in a liquid medium, however this behaviour has not been identified in the RG-2 *C. jejuni* strains. The benefit of glucose utilisation in these strains might be to make energy and NADPH, as is the case in *Pseudomonas putida* KT2440. This strain can convert glucose to gluconate, which then enters into the central carbon metabolism mechanism resulting in NADPH synthesis (Nikel et al., 2015). If the RG-2 strain makes NADPH after metabolising glucose, the bacteria may be more resistant to environmental stresses such as oxidative stress.

In summary, farm-associated rats carry two unusual groups of *C. jejuni* that may prove to be host specific strains. However, these rats also act as a reservoir for virulent generalist *C. jejuni* strains and might have a significant role in the dissemination of these strains in the farmyard. The effect of RG-2 compared to virulent generalist strains on the chicken microbiome will be investigated with stored samples. Laboratory rats as a model for colonisation and virulence by these strains is an important aspect to address. The mutated ED negative Ch95-95 RG-2 strain is already available, for future comparisons on the advantages and disadvantages of glucose utilisation by *C jejuni*.

Appendices

Appendix 1: Corresponding CFU to OD for different strains. WPA used for OD reading.

Cells grown on routine BA, then grown on 2%BA for counting CFU, viable cells in 0.2OD		Cells grown on routine MHA, then grown on 2%BA for counting CFU, viable cells in 0.02OD	
Isolate	CFU/ml	Isolate	CFU/ml
Dg153	7×10^8	Dg153	9×10^7
Dg164	5×10^8	Dg370	8×10^7
Dg276	5.7×10^8	Dg200	5×10^7
Dg36	9×10^8	Dg275	6.8×10^7
Cells grown on routine BA, then grown on 2%BA for counting CFU, viable cells in 0.02OD		Dg95	9×10^7
Isolate	CFU/ml	Dg80	5.6×10^7
Dg153	5.4×10^7	Dg307	6.6×10^7
Dg370	8×10^7	Dg14	7×10^7
Dg200	6×10^7	Dg246	6.3×10^7
Dg95	9×10^7	Dg81	7.5×10^7
Dg275	6×10^7	Dg44	7.2×10^7
Cells grown on routine BA, then grown on 2%mCCDA for counting CFU, viable cells in 0.02O		Dg194	9×10^7
Isolate	CFU/ml	Dg224	4.3×10^7
Dg200	5.8×10^7	Dg57	5.7×10^7
Dg95	8×10^7	NCTC11168	7.7×10^7
Dg275	7.8×10^7	81116	6.3×10^7

Appendix 2: List of detailed chemical constituents of different media or solution used.

<u>**Blood agar (BA) base No.2, Oxoid, code: CM0271**</u>

Typical Formula	g/litre
Proteose peptone	15.0
Liver digest	2.5
Yeast extract	5.0
Sodium chloride	5.0
Agar	12.0

pH 7.4 ± 0.2 at 25°C

A 1 litre of DW contained 40g of BA powder, sterilised by autoclaving at 121°C for 15 min, cool to 45-50°C and added 7% defifbinated blood (v/v)

<u>*Campylobacter* **blood-free selective agar (mCCDA), Oxoid code: CM739B**</u>

For counting CFU, mCCDA was supplemeted with bacteriological agar No. 1 (Oxoid, CM0271B to a concentration of 2% (w/v). While the media was used for isolation of *Campylobacter*, a vial of CCDA selective freeze-dried supplement (Oxoid, SR0155E) was added into 500ml of the sterilised mCCDA medium before pouring into Petri dishes. This was prepared according to manufacturer's instructions.

Typical Formula	gm/litre
Nutrient Broth No.2	25.0
Bacteriological charcoal	4.0
Casein hydrolysate	3.0
Sodium desoxycholate	1.0
Ferrous sulphate	0.25
Sodium pyruvate	0.25
Agar	12.0

pH 7.4 ± 0.2 at 25°C

22.5g of CCDA was suspended in 500ml of DW and brought to boil in a water bath. It was then sterilised by autoclaving at 121°C for 15 min and cooled to 45-50°C. A vial of CCDA selective supplement SR0155 was reconstituted in 2ml of sterile DW, mixed thoroughly with the medium and then poured into sterile Petri dishes.

<u>**CCDA Selective supplement, Oxoid, Code: SR0155**</u>

Vial contents	per vial
Cefoperazone	16mg
Amphotericin B	5mg

Each vial is sufficient for 500ml of CCDA medium

Mueller-Hinton Agar (MHA), Oxoid, code: CM0337

Typical Formula	gm/litre
Beef, dehydrated infusion from	300
Casein hydrolysate	17.5
Starch	1.5
Agar	17

pH 7.3 ± 0.1 at 25°C

1 litre of DW containing 38g of MHA powder was sterilised by autoclaving at 121°C for 15 min, cooled to 45-50°C, and poured into sterile Petri dishes.

Campylobacter motility agar

Mueller-Hinton (MH) or BHI broth was supplemented with bacteriological agar No. 1 to a final concentration of 0.4% (w/v) and prepared according to manufacturer's instructions. 1 litre of motility agar contained 21g of the medium and 4g of the agar No. 1

Brain heart infusion (BHI), Oxoid, code: CM1135B

Following sterilization, BHI was stored at RT for a maximum of 8 weeks

Typical Formula	gm/litre
Proteose peptone	10.0
Brain infusion solids	12.5
Beef heart infusion solids	5.0
Glucose	2.0
Sodium chloride	5.0
Disodium phosphate	2.5

pH 7.4 ± 0.2 at 25°C

1 litre of DW containing 37g of BHI powder was poured into a universal, sterilised by autoclaving at 121°C for 15 min, before being cooled to 45-50°C.

Peptone water Oxoid, code: CM0009

Typical Formula	gm/litre
Peptone	10.0
Sodium chloride	5.0

pH 7.2 ± 0.2

1 litre of DW containing 15g of peptone powder was poured into a universal, sterilised by autoclaving at 121°C for 15 min, then cooled to 45-50°C.

Bolton broth, Oxoid, code: CM0983

Typical Formula	gm/litre
Meat peptone	10.0
Lactalbumin hydrolysate	5.0
Yeast Extract	5.0
Sodium chloride	5.0
Alpha-ketoglutaric acid	1.0
Sodium pyruvate	0.5
Sodium metabisulphite	0.5
Sodium carbonate	0.6
Haemin	0.01

pH 7.4 ± 0.2 at 25°C

13.8g of Bolton broth was suspended in 500ml of DW, sterilised by autoclaving at 121°C for 15 min, and cooled to 50°C. In advance, a vial of Bolton broth selective supplement (SR0183) was resuspended with 5ml of 50% ethanol (v/v). Aseptically, 25ml of lysed defibrinated horse blood (HB034) and a vial of Bolton broth selective supplement (SR0183) were mixed thoroughly with the sterilised 500ml medium. 10ml of this medium was aliquoted into a sterile universal in a safety cabinet.

Bolton broth selective supplement, Oxoid, code: SR0183

Vial contents (each vial is sufficient for 500ml of medium)	per vial
Cefoperazone	10.0mg
Vancomycin	10.0mg
Trimethoprim.	10.0mg
Cycloheximide	25.0mg

Lysed horse blood

20ml of defibrinated horse blood (HB034) was added to a sterile falcon tube, frozen overnight and then thawed at RT. This freeze-thaw process was repeated 4-5 times. 20ml of nH2O was added to the tube, vortexed, centrifuged at 5000 rpm for 30min. The supernatant was stored at -20°C.

Maximal recovery diluents (MRD) , (peptone saline diluent), Oxoid, code: CM0733B

Following sterilisation, MRD was stored at RT for a maximum of 8 weeks

Typical Formula	gm/litre
Peptone	1.0
Sodium chloride	8.5

pH 7.0 ± 0.2 @ 25°C

Bactot^m trypic soy broth without dextrose (TSB, BD & Company, code 286220

Typical Formula	gm/litre
Pancreatic digest casein	17.0
Enzymatic digest of soybean meal	3.0
Sodium chloride	5.0
Dippotassium phosphate	2.5

pH 7.3 ± 0.2

1 litre of DW containing 27.5g of TSB was warmed up until the solution was complete, before being sterilised by autoclaving at 121°C for 15 min.

Phosphate buffer saline (PBS), (PH7.3)

Formula	gm/litre
Sodium chloride	8
Potassium chloride	0.2
Disodium hydrogen phosphate	1.15
Potassium dihydrogen phosphate	0.2

pH 7.3

Dulbecco's Modified Eagle Medium (DMEM), Gibco, code: 11966025

Components	Molecular Weight	Concentration (mg/L)	mM
Amino Acids			
L-Glutamine	146	584	4
L-Serine	105	42	0.4
L-Methionine	149	30	0.2
Glycine	75	30	0.4
L-Arginine hydrochloride	211	84	0.39
L-Cystine 2HCl	313	63	0.2
L-Histidine hydrochloride-H2O	210	42	0.2
L-Isoleucine	131	105	0.8
L-Leucine	131	105	0.8
L-Lysine hydrochloride	183	146	0.79
L-Phenylalanine	165	66	0.4
L-Threonine	119	95	0.79
L-Tryptophan	204	16	0.07
L-Tyrosine disodium salt dihydrate	261	104	0.39
L-Valine	117	94	0.8
Vitamins			
Folic Acid	441	4	0.009
Choline chloride	140	4	0.02
Folic Acid	441	4	0.009
D-Calcium pantothenate	447	4	0.008
Niacinamide	122	4	0.03
Pyridoxine hydrochloride	206	4	0.019
Riboflavin	376	0.4	0.001
Thiamine hydrochloride	337	4	0.01

i-Inositol	180	7.2	0.04
Inorganic Salts			
Ferric Nitrate (Fe(NO3)3"9H2O)	4.4	0.1	2.4
Calcium Chloride (CaCl2) (anhyd.)	111	200	1.8
Magnesium Sulfate (MgSO4) (anhyd.)	120	97.6	0.8
Potassium Chloride (KCl)	75	400	5.3
Sodium Bicarbonate (NaHCO3)	83	3700	44
Sodium Phosphate monobasic (NaH2PO4-H2O)	138	125	0.9
Sodium Chloride (NaCl)	58	6400	110.3
Other Components			
Phenol Red	376.4	15	0.039

Alpha minimum essential medium (MEM-α, Gibco, code, 11534476)

MEM-α is a modification of minimum essential medium (MEM) that contains adenosine, glycine, guanosine, L-alanine, L-arginine, L-asparagine, Cytidine, Calcium Chloride ($CaCl_2$), 2Deoxycytidine HCl, (anhyd), 2'Deoxyguanosine, 2Deoxyadenosine, L-glutamine, deoxyribonucleosides and ribonucleosides , D-Calcium pantothenate (vitamin B_5), ascorbic acid (vitamin C), vitamin B_{12}, Folic Acid (vitamin B9), lipoic acid, biotin, and sodium pyruvate. The medium does not contain phenol red.

Preparation of MEM-FBS

Phenotypic growth properties were monitored in MEM-α which was supplemented with fresh 2.5% w/v FBP [ferrous sulfate (Acrose), sodium metabisulphite (Fluca), and sodium pyruvate (Acrose)] in order to provide an oxidative stress protectant and an iron source (Shou et al., 1983). The FBP was filter sterilised and then 1:2000 dilutions of this was added into the medium. MEM-α medium was also supplemented with 1% of Bolton broth selective supplement antibiotic (Oxoid) (as indicated) and 10% Foetal Bovine Serum (Gibco).

Preparation of DMEMf

For showing the effect of glucose on growth and monitoring glucose in a culture medium over a growth curve, Broth the DMEM medium was supplemented with 0.2mM iron (II) L-ascorbate (IA) as described by Vorwerk et al. (2015). Bolton supplementary antibiotic (1%) and Foetal Bovine Serum (10%) were added, where indicated, this medium was designated as DMEMf.

Bacterial storage medium (15% glycerol stock medium)

The storage medium was made from sterile BHI with 50% sterile glycerol to give a final concentration of 15% (v/v) glycerol. A full loopful of bacterial cells was suspended into the stock medium.

TE buffer

5mM Tris-HCl, 1 mM disodium ethylene diamine tetraacetic acid (EDTA), pH 8 (TE buffer) was prepared by mixing 0.1214g/l of Tris (Fisher Scientific, code BP1521-1) with 0.3722g/l of disodium ethylene diamine tetraacetic acid (EDTA, Fisher Scientific, code D/0700/53). The pH was adjusted to 8.0 through addition of HCl. Following sterilisation, TE buffer was stored at room temperature, before using a 1/50 used for bacterial cell lysis.

Bacterial cell fixation for SEM or TEM

Campylobacter cells were fixed in 2.5% EM glutaraldehyde (Agar Scientific, code AGR1012) in 0.1M cacodylate buffer (Agar Scientific, AGR1105code). PH 7 of cacodylate buffer was adjusted by adding 0.42ml of 0.2M HCl into 5ml of 0.2M cacodylate buffer. The pH was checked by using pH meter filter paper (Fisher Scientific, code 10642751). After adjusting the pH, 3.58ml of DW and 1ml of 25% glutaraldehyde were added into the fixation solution. This solution was ready to use and kept at 4°C for up to 5 days.

Appendix 3: Growth of *Campylobacter jejuni* in small intestinal and caecal lab rat mucins.
The small intestine (S) and caecum (C) mucins were scraped from 4 white lab rats (~250g/rat) and the mucin was homogenised using a MP shaker. Under my supervision, the growth experiment was done by Nancy, F. and Sarah, J. undergraduate students. MHB was supplemented with 5% mucin, 1% antibiotic and inoculated with $0.002OD_{600}$ of bacterial cells, and incubated for 45h. Cells were then grown under microtiter plate growth assay conditions. CFU values are the mean of triplicate drops of three wells. Pattern filled bars are growth in MHB without mucin.
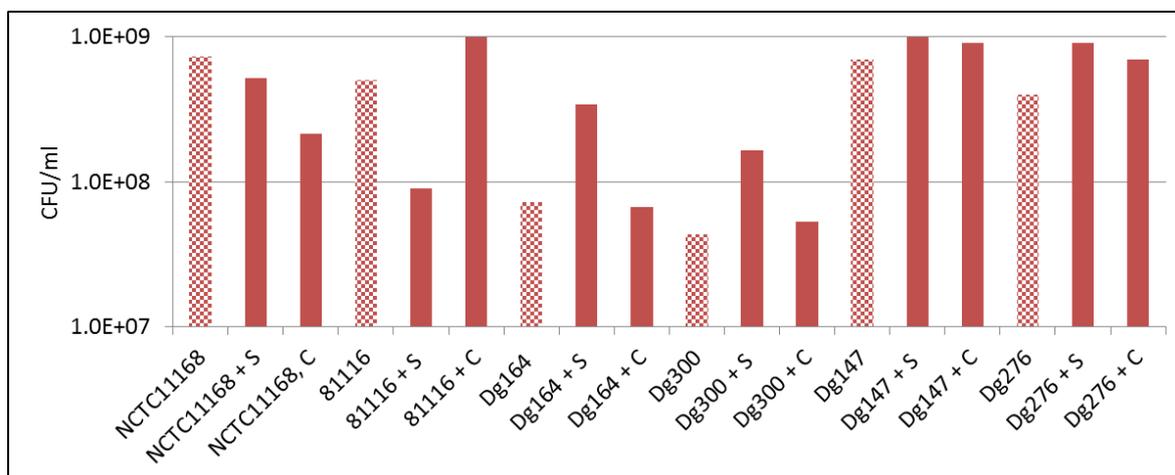
Appendix 4: Concentration of amino acids in MEM-FBS medium. The medium was supplemented with 10% FBS and 20mM (**A**) Pro and (**B**) Ser. Amino acid concentration was measured at 48h and 72h of incubation. The Pro in A and Ser in B data has been shown in **Figure 3.16**, **B**.
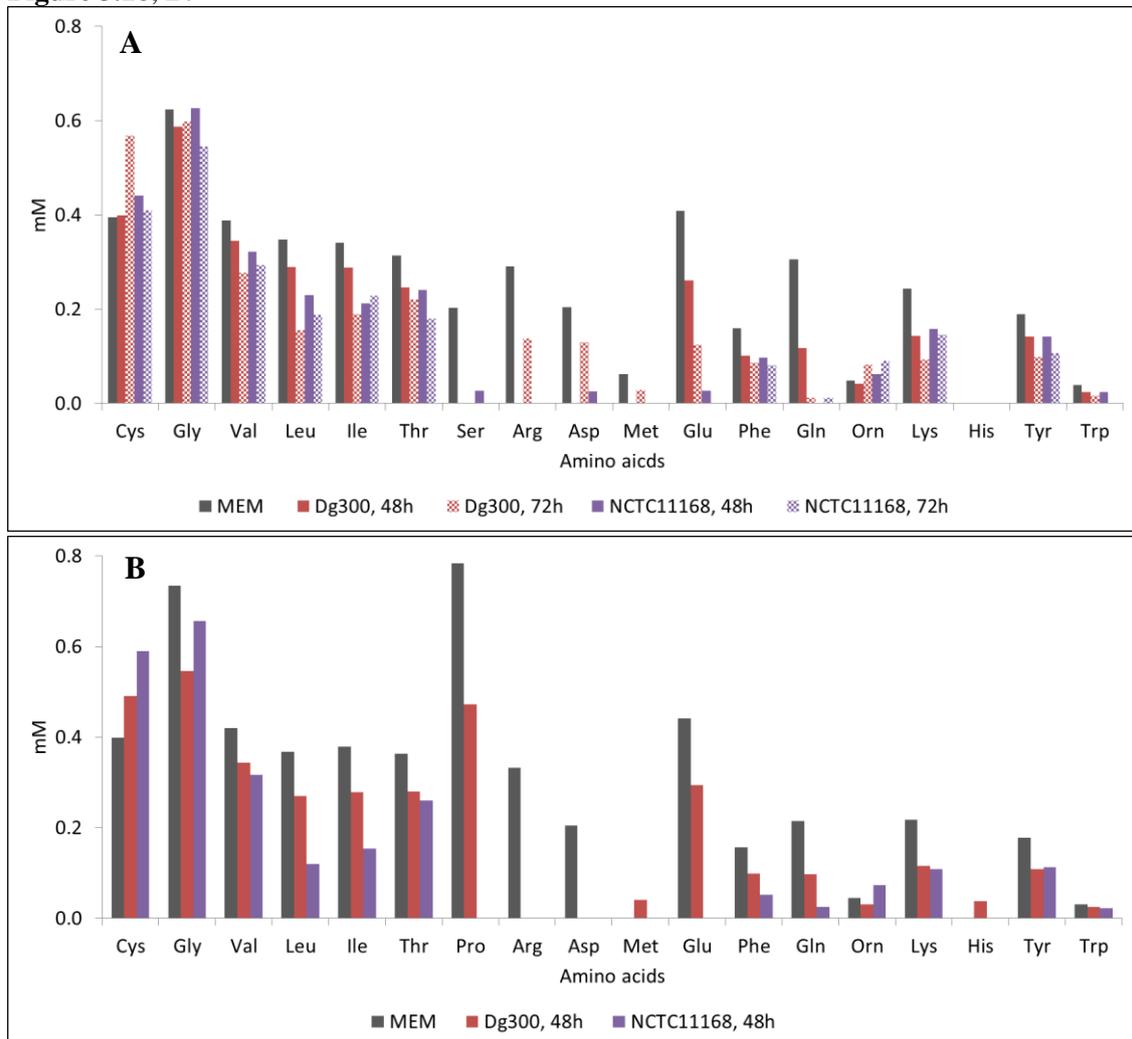
Appendix 5: cgMLST 29 genes are absent (A) or incomplete (I) in at least one of the RG-1 strains.

cj number	Product	Sequence length	Genome position	Absent (A) or Incomplete (I) gene among RG-1
cj0118	conserved hypothetical protein	756	122366	I in Dg197
*cj0145**	TAT (Twin-Arginine Translocation) dependant alkaline phosphatase	1782	148819	A in all
*cj0174c/cfbpB**	putative iron-uptake ABC transport system permease protein	1617	169946	A in all
cj0243c	hypothetical protein	1167	224794	I in Dg150
cj0430	putative integral membrane protein	1227	391711	I in Dg61
cj0484	putative MFS (Major Facilitator Superfamily) transport protein	1233	451046	A in 36.6%
cj0561c	putative periplasmic protein	930	524034	A in 36.7%
cj0763c/cysE	serine acetyltransferase	639	714138	I in Dg156
cj0777	putative ATP-dependent DNA helicase	2031	728500	I in Dg189
cj0799c/ruvA	putative Holliday junction ATP-dependent DNA helicase	552	749307	I in Dg78
cj0800c	putative ATPase	1860	749834	I in Dg289
cj0801	putative integral membrane protein (MviN homolog)	1452	751797	I in Dg189
cj0810/nadE	NH(3)-dependent NAD(+) synthetase	741	761404	I in Dg276
cj0841c/mobB	putative molybdopterin-guanine dinucleotide biosynthesis protein	492	789049	I in Dg189
cj0850c	putative MFS (Major Facilitator Superfamily) transport protein	1188	797653	I in Dg347
cj1004	putative periplasmic protein	417	934201	I in Dg189
cj1014c/livF	branched-chain amino-acid ABC transport system ATP-binding protein	696	947343	I in Dg347
cj1041c	putative periplasmic ATP/GTP-binding protein	852	975230	A in 23.4%
cj1042c	putative transcriptional regulatory protein	891	976144	A in 23.4%
cj1198/luxS	S-ribosylhomocysteine lyase (autoinducer-2 production protein LuxS)	495	1127437	A in 36.7%
cj1064	pseudogene (nitroreductase)	620	1001218	A in 23.4%
cj1199	putative iron/ascorbate-dependent	993	1128243	A in 36.7%

	oxidoreductase			
cj1201/metE	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	2265	1130028	A in 36.7%
cj1202/metF	5,10-methylenetetrahydrofolate reductase	849	1132302	A in 36.7%
cj1295	conserved hypothetical protein	1308	1226978	I in Dg147
cj1296	hypothetical protein	360	1228282	A in 23.4%
cj1346c	1-deoxy-D-xylulose 5-phosphate reductoisomerase	1071	1278851	I in Dg150
cj1411c	putative cytochrome P450	1362	1342550	I in Dg150
cj1721c*	putative outer membrane protein	645	1632901	A in all

* absent in RG-1 including Dg147

Appendix 6: 247 RG-1 core genes, but non-cgMLST based on annotated Dg147 genome. RG-1 strains were Blasted in Genome Comparator (95% threshold) (Cody et al., 2017).

locus number	**Product**	**Sequence length**	**Genome position**
locus00013	Periplasmic dsDNA and ssDNA-binding protein contributing to transformation	240	17883
locus00016	hypothetical protein	525	20683
locus00021	MCP-domain signal transduction protein	1779	25108
locus00022	Cytochrome c551 peroxidase	915	26876
locus00032	hypothetical protein	1398	46507
locus00042	Cytochrome c family protein	1032	59360
locus00046	Flagellar hook-length control protein FliK	1974	63487
locus00050	Hemerythrin domain protein	720	69913
locus00051	C4-dicarboxylate transporter	315	71097
locus00052	C4-dicarboxylate transporter	801	71374
locus00053	C4-dicarboxylate transporter	747	72125
locus00056	hypothetical protein	777	74640
locus00057	hypothetical protein	585	75437
locus00058	FIG00973752: TolA-like membrane protein	585	76316
locus00059	FIG00973752: TolA-like membrane protein	249	76921
locus00060	Bacteriocin resistance protein; peptidase C39	600	77154
locus00075	Predicted L-lactate dehydrogenase, hypothetical protein subunit YkgG	657	89604
locus00076	Predicted L-lactate dehydrogenase, Iron-sulfur cluster-binding subunit YkgF	1440	90253
locus00077	Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE	741	91689
locus00082	hypothetical protein	105	97603
locus00083	Uncharacterized protein YgeA of aspartate/glutamate/hydantoin racemase family	123	97786
locus00089	putative lipoprotein	369	103802
locus00091	Putative periplasmic protein	1338	104900
locus00092	Putative periplasmic protein	1203	106247
locus00138	McrBC restriction endonuclease system, McrB subunit, putative	1812	144212
locus00139	hypothetical protein	1338	146023
locus00144	PhoX, Predicted phosphatase	237	152488
locus00145	PhoX, Predicted phosphatase	120	152725
locus00146	PhoX, Predicted phosphatase	123	152882
locus00147	Gluconate 2-dehydrogenase, membrane-bound, flavoprotein	321	153045

locus00148	Gluconate 2-dehydrogenase, membrane-bound, flavoprotein	363	153843
locus00149	hypothetical protein	411	154238
locus00150	hypothetical protein	405	154763
locus00179	Ferric iron ABC transporter, permease protein	150	175222
locus00180	Ferric iron ABC transporter, iron-binding protein	372	175371
locus00181	Ferric iron ABC transporter, iron-binding protein	600	175739
locus00186	Integral membrane protein TerC	720	180740
locus00199	Putative periplasmic protein	1122	193545
locus00201	Membrane protein	606	195212
locus00202	hypothetical protein	261	195801
locus00203	Citrate/H+ symporter of CitMHS family	1326	196407
locus00209	hypothetical protein	267	204847
locus00210	hypothetical protein	489	205444
locus00211	hypothetical protein	150	206325
locus00212	serine protease pet	183	206609
locus00233	Putative MCP-domain signal transduction protein	1977	224689
locus00237	Highly acidic protein	159	229484
locus00246	Small hydrophobic protein Cj0260c-related protein	222	235650
locus00249	Zinc transporter ZupT	876	238925
locus00254	hypothetical protein	846	243310
locus00255	Pseudoazurin	438	244155
locus00256	ABC transporter, ATP-binding protein (cluster 8, B12/iron complex)	780	244576
locus00257	ABC transporter, permease protein (cluster 8, B12/iron complex)	855	245545
locus00258	ABC transporter, substrate-binding protein (cluster 8, B12/iron complex)	1095	246389
locus00259	ABC transporter, substrate-binding protein (cluster 8, B12/iron complex)	1107	247480
locus00260	Major outer membrane protein	327	248597
locus00261	Major outer membrane protein	915	249001
locus00262	Probable tautomerase cj0270	207	250072
locus00280	hypothetical protein	576	267862
locus00281	hypothetical protein	1386	268439
locus00282	Type III restriction-modification system methylation subunit	327	269834
locus00283	Cytolethal distending toxin subunit C	573	270370
locus00284	Cytolethal distending toxin subunit B, DNase I-like	804	270952
locus00285	Cytolethal distending toxin subunit A	795	271752
locus00290	putative acetyltransferase	228	275966

locus00291	Class D beta-lactamase (EC 3.5.2.6) => OXA-184 family	747	276190
locus00292	CAAX amino terminal protease family protein	819	277051
locus00295	hypothetical protein	402	279417
locus00320	RidA/YER057c/UK114 superfamily, group 2, YoaB-like protein	351	302226
locus00332	Putative transmembrane transport protein	1362	312225
locus00333	Inosine-uridine preferring nucleoside hydrolase	1008	313583
locus00335	putative membrane protein	597	314989
locus00336	Membrane protein	714	315566
locus00337	Membrane protein	552	316299
locus00339	Membrane protein	786	319747
locus00374	hypothetical protein	681	352769
locus00407	Gluconate 2-dehydrogenase, membrane-bound, gamma subunit	729	386473
locus00408	Gluconate 2-dehydrogenase, membrane-bound, flavoprotein	1722	387203
locus00409	hypothetical protein	678	389093
locus00410	hypothetical protein	207	389804
locus00417	Integral membrane protein	423	393568
locus00418	hypothetical protein	489	394108
locus00419	hypothetical protein	468	394646
locus00431	Succinate dehydrogenase flavoprotein subunit	360	412044
locus00432	Succinate dehydrogenase flavoprotein subunit	666	412575
locus00433	Succinate dehydrogenase flavoprotein subunit	279	413483
locus00434	Succinate dehydrogenase iron-sulfur protein	135	413875
locus00435	Succinate dehydrogenase iron-sulfur protein	219	414375
locus00436	Succinate dehydrogenase iron-sulfur protein	171	414660
locus00437	Heterodisulfide reductase subunit B-like protein, Putative succinate dehydrogenase subunit	348	415160
locus00444	Uridine diphosphate glucose pyrophosphatase	597	421194
locus00452	putative membrane protein	528	427043
locus00459	Cyclic dehypoxanthine futalosine synthase	1047	433174
locus00462	hypothetical protein	384	437250
locus00545	Selenophosphate-dependent tRNA 2-selenouridine synthase	999	501807
locus00547	Ammonium transporter	282	503948
locus00594	FIG015373: Membrane protein	693	552041
locus00595	FIG001614: Membrane protein	954	552720
locus00596	hypothetical protein	1005	553683
locus00597	Dicarboxylate carrier protein	1287	554812
locus00598	putative 2-pyrone-4,6-dicarboxylic acid hydrolase	798	556109

locus00605	hypothetical protein	561	564344
locus00606	Membrane protein	207	564895
locus00607	hypothetical protein	930	565549
locus00608	hypothetical protein	1671	566520
locus00609	hypothetical protein	423	568206
locus00631	putative lipoprotein	429	587164
locus00666	Family of unknown function (DUF450) family	1044	623928
locus00691	Di-tripeptide/cation symporter	1572	648532
locus00692	Di-tripeptide/cation symporter	1440	650274
locus00693	Putative periplasmic protein	549	651840
locus00694	Putative transmembrane protein	1107	652381
locus00705	Potassium-transporting ATPase A chain	516	662582
locus00706	Potassium-transporting ATPase A chain	246	663097
locus00707	Potassium-transporting ATPase A chain	147	663339
locus00708	Potassium-transporting ATPase A chain	501	663452
locus00709	Potassium-transporting ATPase A chain	315	663956
locus00710	Potassium-transporting ATPase B chain	2046	664280
locus00711	Potassium-transporting ATPase C chain	579	666327
locus00712	Osmosensitive K+ channel histidine kinase KdpD	1023	667055
locus00713	Osmosensitive K+ channel histidine kinase KdpD	747	668107
locus00719	Possible sugar transferase	303	673715
locus00721	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase (flavodoxin) (EC 1.17.7.3)	1074	675254
locus00726	Membrane protein	939	680402
locus00761	Putative periplasmic protein	396	712857
locus00763	Putative periplasmic protein	318	713839
locus00765	hypothetical protein	234	715612
locus00766	hypothetical protein	375	716036
locus00767	hypothetical protein	1182	723216
locus00770	Chaperone protein DnaK	1872	725946
locus00781	Methionine ABC transporter substrate-binding protein	777	737080
locus00782	Methionine ABC transporter substrate-binding protein	771	737866
locus00783	Methionine ABC transporter substrate-binding protein	774	738685
locus00797	Small hydrophobic protein	174	755142
locus00805	hypothetical protein	1380	761375
locus00825	hypothetical protein	1335	782267
locus00832	Putative processing peptidase	777	789081

locus00836	CoA-binding domain protein	414	792812
locus00837	Uncharacterized protein jhp1395	417	793346
locus00865	hypothetical protein	429	822107
locus00866	Membrane protein	873	822630
locus00870	Periplasmic thiol:disulfide interchange protein, DsbA-like	663	827517
locus00872	Putative arylsulfate sulfotransferase (EC 2.8.2.22)	1863	829073
locus00873	Periplasmic thiol:disulfide interchange protein, DsbA-like	642	831021
locus00874	Cytochrome c family protein	987	831696
locus00875	hypothetical protein	141	832892
locus00896	Small hydrophobic protein	147	857850
locus00960	Putative periplasmic protein	216	922348
locus00961	hypothetical protein	312	922834
locus00962	Putative hemolysin activation/secretion protein	144	923377
locus00963	Putative hemolysin activation/secretion protein	168	924080
locus00964	Putative hemolysin activation/secretion protein	183	924244
locus00973	Surface-exposed lipoprotein JlpA	1119	931175
locus00974	YhbX/YhjW/YijP/YjdB family protein	1188	932326
locus00977	Hippurate hydrolase (EC 3.5.1.32)	1152	934815
locus00978	Integral membrane protein	786	935968
locus00979	Integral membrane protein	411	936777
locus00981	Uncharacterized protein Cj0990c	759	937630
locus01009	High-affinity leucine-specific transport system, periplasmic binding protein LivK	1110	966985
locus01010	Branched-chain amino acid ABC transporter, amino acid-binding protein	1116	968116
locus01012	Putative periplasmic protein	192	969705
locus01027	hypothetical protein	252	986286
locus01032	Uncharacterized MFS-type transporter	702	990744
locus01046	Phosphoglycerol transferase-like protein	1773	1003273
locus01051	Small hydrophobic protein	111	1008821
locus01061	D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)	438	1018648
locus01062	SSU ribosomal protein S6p	378	1019312
locus01067	Flagellar assembly factor FliW	390	1023781
locus01069	Transformation system protein	303	1024898
locus01071	Probable periplasmic protein Cj1079	453	1025856
locus01102	Methyl-accepting chemotaxis sensor/transducer protein	1194	1057382
locus01111	N-linked glycosylation glycosyltransferase PglG	894	1065295

locus01126	Putative two-domain glycosyltransferase	1548	1082255
locus01127	Beta-1,3-galactosyltransferase / Beta-1,4-galactosyltransferase	1116	1083862
locus01139	Phosphoheptose isomerase 1	561	1095098
locus01177	Signal transduction protein CetB, mediates an energy taxis response	498	1130453
locus01179	Putative C4-dicarboxylate transport protein	1377	1132945
locus01208	Hemerythrin domain protein	600	1166132
locus01225	Putative transmembrane transport protein	1200	1182264
locus01227	hypothetical protein	324	1184053
locus01240	Putative isomerase	210	1199816
locus01284	hypothetical protein	894	1243681
locus01289	hypothetical protein	579	1247828
locus01293	Putative acyl carrier protein	228	1251894
locus01294	hypothetical protein	1071	1252118
locus01295	hypothetical protein	510	1253210
locus01296	hypothetical protein	618	1253806
locus01300	Similar to imidazole glycerol phosphate synthase cyclase subunit (LPS cluster)	747	1256443
locus01301	Similar to imidazole glycerol phosphate synthase amidotransferase subunit (LPS cluster)	606	1257190
locus01302	Pseudaminic acid biosynthesis protein PseA, possible Pse5Ac7Ac acetamidino synthase	1137	1257792
locus01306	GDP-N-acetylglucosamine 4,6-dehydratase [NAD+]	993	1262037
locus01307	GDP-2-acetamido-2,6-dideoxy-alpha-D-xylo-hexos- 4-ulose aminotransferase [PLP]	1143	1263022
locus01314	N,N'-diacetyllegionaminic acid synthase (EC 2.5.1.101)	1005	1268472
locus01315	UDP-N,N'-diacetylbacillosamine 2-epimerase (hydrolyzing)	1155	1269469
locus01316	Glucosamine-1-phosphate guanylyltransferase	1026	1270632
locus01317	Glutamine--fructose-6-phosphate transaminase (isomerizing), isomerase subunit (EC 2.6.1.16)	903	1271654
locus01318	CMP-N,N'-diacetyllegionaminic acid synthase (EC 2.7.7.82)	708	1272549
locus01319	Glutamine--fructose-6-phosphate transaminase (isomerizing), glutaminase subunit	771	1273256
locus01321	Motility accessory factor	1881	1275942
locus01346	hypothetical protein	2274	1307689
locus01376	MmgE/PrpD family protein, putative	144	1342684
locus01377	MmgE/PrpD family protein, putative	438	1342818
locus01388	Putative periplasmic protein	351	1352375
locus01397	Adenylylsulfate kinase	513	1361362

locus01398	Putative sugar nucleotidyltransferase	762	1361864
locus01399	Putative amidotransferase (Type 1 glutamine amidotransferase - GATase1)	603	1362627
locus01400	Phosphoenolpyruvate synthase / Pyruvate phosphate dikinase	2340	1363220
locus01401	Methyltransferase, possibly involved in O-methyl phosphoramidate capsule modification	762	1365574
locus01404	D-glycero-D-manno-heptose 1-phosphate guanosyltransferase	675	1369061
locus01405	Phosphoheptose isomerase 2	606	1369723
locus01406	D,D-heptose 7-phosphate kinase	1020	1370316
locus01430	Putative periplasmic protein	315	1398176
locus01436	CMP-N-acetylneuraminate-beta-galactosamide- alpha-2,3-sialyltransferase	1413	1402740
locus01437	hypothetical protein	315	1404133
locus01438	Putative acyl carrier protein	219	1404444
locus01439	Putative amino acid activating enzyme	1509	1404702
locus01502	Type II secretion system protein	843	1449493
locus01504	Transformation system protein	588	1451881
locus01543	Formate dehydrogenase-O, major subunit, selenocysteine-containing	543	1493778
locus01552	CRISPR-associated protein Cas2	432	1501223
locus01553	CRISPR-associated protein Cas1	891	1501647
locus01554	CRISPR-associated endonuclease Cas9	2970	1502527
locus01574	Homolog of BLC protein	450	1525093
locus01576	aminoglycoside 6-adenylyltransferase	192	1527074
locus01577	aminoglycoside 6-adenylyltransferase	576	1527351
locus01578	hypothetical protein	1278	1528079
locus01579	4-carboxymuconolactone decarboxylase (EC 4.1.1.44)	753	1529410
locus01580	hypothetical protein	468	1530257
locus01581	Dienelactone hydrolase and related enzymes	1020	1530758
locus01582	oxidoreductase of aldo/keto reductase family, subgroup 1	801	1531860
locus01583	Putative transporter	1179	1532664
locus01611	MBL-fold metallo-hydrolase superfamily	420	1564437
locus01623	hypothetical protein	957	1572974
locus01641	Alpha-ketoglutarate permease	438	1590181
locus01643	Hemerythrin domain protein	642	1591887
locus01644	Hemerythrin domain protein	600	1592532
locus01649	Serine transporter	1251	1597466
locus01652	hypothetical protein	273	1599938

locus01655	putative TonB-dependent receptor	684	1601175
locus01658	Channel-forming transporter/cytolysins activator of TpsB family	1653	1603992
locus01659	Putative large exoprotein involved in heme utilization or adhesion of ShlA/HecA/FhaA family	5577	1605653
locus01683	hypothetical protein	204	1631990
locus01692	CopG protein	438	1640646
locus01693	Putative periplasmic protein	417	1641412
locus01735	Small hydrophobic protein	99	1672419
locus01736	lipopolysaccharide core biosynthesis protein LpsA	966	1672934
locus01743	hypothetical protein	963	1680263

Appendix 7: Iron uptake groups of genes in *Campylobacter jejuni*. Genes that are missing in RG-1 indicated by underlining in red. Image taken from Miller et al. (2009).

Appendix 8: CdtA of *Campylobacter lari* RM16701and Dg147 alignment. Deletions in amino acid number 14 and between 49 to 54 in the CdtA of Dg147 (red).

```
C.lari RM16701    MQKIKFIMIFSAFVAFFIGCSSKEQFNPLGRSLGDVNDIDPLKIGKNPTPPAKQKTPALI
Dg147             MQKIKFTMVISVLIAFFIGCSSKE-INPLGHSLGAINDSDPLNIGKNPTPPAKQKTPALV
                  ****** *::*.::**********  :****:*** :** ***:****************:


C.lari RM16701    EGKHFPAIPLVPPIITTNTFKGDNPIKGPLPRLTSDNKFTSNALYENLGIASDFVTIMNP
Dg147             EGKKFPAIPLEPPIPKTDVFKKDNAGQT------LSLESSPNGLYGKTGLASDFVTIMNP
                  ***:****** *** .*:.** **  :          . : : *.** : *:********


C.lari RM16701    NGAALTVWALNPGNWIWGYNLFDSKSFGDARVWQLIEFPNNKVMIKNAKTNTCLNAYGNG
Dg147             NGAALTVWALNPGNWIWGYNLFDSKPFGNARVWQLIEIPNNTVMIKNSKTSTCLTAYGNG
                  *********************** **:********:*** .*****:**.***.*****


C.lari RM16701    IVHYPCDQSNFAQYWGFFPMSNGAYQIQNYATQQCIQTTVGNVMEEFYFNYYSIYLTPCL
Dg147             IVHYPCDQGNFSQFWRFHPMSNGAYEIQNFATHKCIQTTVSNVMEEFNFNYYSIYLTSCL
                  ********.**:*:* *.*******:***:**::******.******.****** ********* **


C.lari RM16701    KSKEKNLDRQWSISAPAFQTRSPYAEGGVF
Dg147             KSKEKNLDRQWNVSAPAFETKSPYAYEGES
                  ***********.:*****:*:****  *
```

Appendix 9: Genes absent in RG-1, but present in the different sequence types of
Campylobacter jejuni. Dg147 genome was Blasted against different sequence types of strains
are listed in **Table 3.5**.

* cj number	Product
cj0241c	putative iron-binding protein
cj0264c	molybdopterin containing oxidoreductase
cj0265c	putative cytochrome C-type haem-binding periplasmic protein
cj0270	putative tautomerase family protein
cj0289c/peb3	major antigenic peptide PEB3
cj0570	putative ATP/GTP binding protein
** *cj0727*	putative periplasmic solute-binding protein
** *cj0729*	putative type I phosphodiesterase/nucleotide pyrophosphatase
**cj0730*	putative ABC transport system permease
** *cj0731*	putative ABC transport system permease
** *cj0732*	ABC transport system ATP-binding protein
** *cj0733*	putative HAD-superfamily hydrolase
** *cj0735*	putative periplasmic protein
** *cj0736*	hypothetical protein
cj0741	hypothetical protein
cj0748	hypothetical protein
cj0752	pseudogene (IS element transposase)
cj0967	putative periplasmic protein
cj0969	pseudogene (putative periplasmic protein)
cj0971	hypothetical protein
cj1546	putative transcriptional regulator
cj1553c/hsdM	putative type I restriction enzyme M protein
cj1721c	putative outer membrane protein

* Genes are absent in Dg147 and present in all the other strains were then blasted with draft genome of the
rest of RG-1 strains in the Genome comparator to find unique genes to the reference strains. ** putative
ABC transport system.

Appendix 10: Concentration of free amino acids in fresh MHB, BHI, DMEM, FBS and MEM.
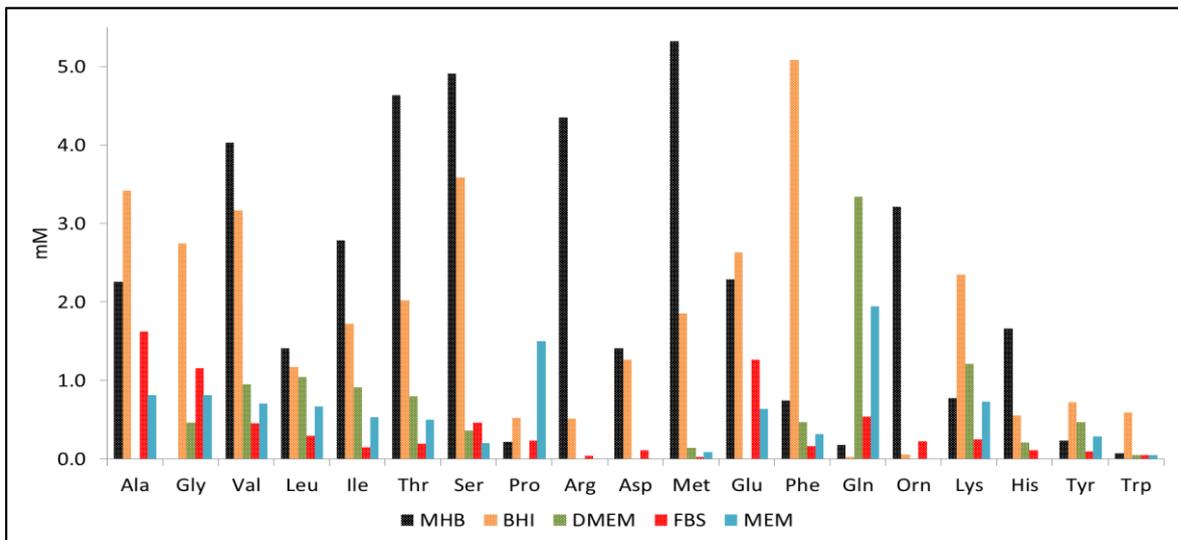
Appendix 11: Genes present in RG-2 strains, but absent in the ED-negative CC-45 farm-associated rat strains. Annotated genome of Dg275 was blasted with the 8 strains of the CC-45, Dgs (200, 245, 292, 206, 162, 202, 273, 383) and other RG-2 strains, separately using the genome comparator of the BIGSdb database. Dg381, Dg18 Dg201 of RG-2 clonal complex strains were excluded from the analysis due to having an incomplete *glc* locus and/or high contig numbers.

Locus number in Dg275 genome	Product	Sequence length	Genome position
locus00032*	4-hydroxy-2-oxoglutarate aldolase, 2-dehydro-3-deoxyphosphogluconate aldolase	624	42660
locus00033*	Phosphogluconate dehydratase	1803	43295
locus00034*	Glucose-6-phosphate 1-dehydrogenase	1401	45242
locus00035*	6-phosphogluconolactonase	681	46808
locus00036*	Glucokinase	1005	47469
locus00037*	Glucose-6-phosphate isomerase	1644	48477
locus00038*	GlcP (Homolog of fucose/glucose/galactose permeases)	1200	50131
locus00072	FIG00973752: TolA-like membrane protein	855	89762
locus00077	FIG00973752: TolA-like membrane protein	855	92493
locus00160	McrBC restriction endonuclease system, McrB subunit	1668	162329
locus00431*	4-hydroxy-2-oxoglutarate aldolase 2-dehydro-3-deoxyphosphogluconate aldolase	624	410802
locus00432*	Phosphogluconate dehydratase	1803	411437
locus00433*	Glucose-6-phosphate 1-dehydrogenase	1401	413384
locus00434*	6-phosphogluconolactonase	681	414950
locus00435*	Glucokinase	1005	415611
locus00436*	Glucose-6-phosphate isomerase	1644	416619
locus00437*	Homolog of fucose/glucose/galactose permeases	1200	418273
locus00484	Transcriptional regulator, IclR family	762	471495
locus00485	Dihydrodipicolinate synthase family	909	472486
locus00486	Altronate dehydratase	264	473391
locus00487	Altronate dehydratase	1167	473657
locus00488	Uncharacterized MFS-type transporter	1287	474835
locus00489	L-fuco-beta-pyranose dehydrogenase, type 2	789	476122
locus00490	Fucose permease	1257	476921

locus00491	L-fuconolactone hydrolase	771	478161
locus00492	L-fucose mutarotase	318	478932
locus00493	Aldehyde dehydrogenase A Glycolaldehyde dehydrogenase	1440	479263
locus01060	Putative Dihydrolipoamide dehydrogenase; Mercuric ion reductase; PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	1356	1055595
locus01061	D-alanyl-D-alanine carboxypeptidase	438	1056965
locus01529	ABC transporter, substrate-binding protein (cluster 5, nickel/peptides/opines)	1536	1538076
locus01571**	Ser transporter	1323	1578189
locus01672	Putative oxidoreductase	963	1660393

* Locus numbers 32 to 38 and 431 to 437 were both copies of the ED pathway genes. Both copies were practically confirmed in Dg275 and Dg95 RG-2 strains (**Figure 4.11**). ** ˜ 60.99% identity to *Cj1625c* (*sdaC*) Ser transporter in NCTC11168.

Bibliography

Bibliography

Albert, M. J., Leach, A., Asche, V., Hennessy, J. & Penner, J. L. (1992). Serotype distribution of *Campylobacter jejuni* and *Campylobacter coli* isolated from hospitalized patients with diarrhea in central Australia. *Journal of Clinical Microbiology* 30**,** 207-10.

Alemka, A., Nothaft, H., Zheng, J. & Szymanski, C. M. (2013). N-Glycosylation of *Campylobacter jejuni* Surface Proteins Promotes Bacterial Fitness. *Infection and Immunity,* 81**,** 1674-1682.

Allen, V. M., Bull, S. A., Corry, J. E., Domingue, G., Jorgensen, F., Frost, J. A., Whyte, R., Gonzalez, A., Elviss, N. & Humphrey, T. J. (2007). *Campylobacter* spp. contamination of chicken carcasses during processing in relation to flock colonisation. *International Journal of Food Microbiology,* 113**,** 54-61.

Andersson, S. G., Zomorodipour, A., Winkler, H. H. & Kurland, C. G. (1995). Unusual organization of the rRNA genes in *Rickettsia prowazekii*. *Journal of Bacteriology,* 177**,** 4171-5.

Ang, C. W. (2002). Structure of *Campylobacter jejuni* Lipopolysaccharides Determines Antiganglioside Specificity and Clinical Features of Guillain-Barre and Miller Fisher Patients. *Infection and Immunity,* 70**,** 1202-1208.

Aspinall, G. O., Lynch, C. M., Pang, H., Shaver, R. T. & Moran, A. P. (1995). Chemical structures of the core region of *Campylobacter jejuni* O:3 lipopolysaccharide and an associated polysaccharide. *European Journal of Biochemistry,* 231**,** 570-8.

Aspinall, G. O., Mcdonald, A. G. & Pang, H. (1992). Structures of the O chains from lipopolysaccharides of *Campylobacter jejuni* serotypes O:23 and O:36. *Carbohydrate Research,* 231**,** 13-30.

Aydin, F., Atabay, H. I. & Akan, M. (2001). The isolation and characterization of *Campylobacter jejuni* subsp. *jejuni* from domestic geese (Anser anser). *Journal of Applied Microbiology,* 90**,** 637-42.

Aziz, R. K., Bartels, D., Best, A. A., Dejongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., Mcneil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics,* 9**,** 75.

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. *Molecular Microbiology,* 40**,** 769-77.

Baker, J., Barton, M. D. & Lanser, J. (1999). *Campylobacter* species in cats and dogs in South Australia. *Australian Veterinary Journal,* 77**,** 662-6.

Balaban, M. & Hendrixson, D. R. (2011). Polar flagellar biosynthesis and a regulator of flagellar number influence spatial parameters of cell division in *Campylobacter jejuni*. *PLOS Pathogens: A Peer-Reviewed Open-Access Journal,* 7**,** e1002420.

Bang, D. D., Scheutz, F., Ahrens, P., Pedersen, K., Blom, J. & Madsen, M. (2001). Prevalence of cytolethal distending toxin (cdt) genes and CDT production in *Campylobacter* spp. isolated from Danish broilers. *Journal of Medical Microbiology,* 50**,** 1087-94.

Beery, J. T., Hugdahl, M. B. & Doyle, M. P. (1988). Colonization of gastrointestinal tracts of chicks by *Campylobacter jejuni*. *Applied and Environmental Microbiology,* 54**,** 2365-70.

Ben, P., Hilary, L., Samuel, K. S. & Hiroshi, A. (2014). *Does Biofilm formation aid colonisation and infection in Campylobacter?,* Norfolk, UK, *Caister Academic Press*.

Bercovier, H., Kafri, O. & Sela, S. (1986). Mycobacteria possess a surprisingly small number of ribosomal RNA genes in relation to the size of their genome. *Biochemical and Biophysical Research Communications,* 136**,** 1136-41.

Bingham-Ramos, L. K. & Hendrixson, D. R. (2008). Characterization of two putative cytochrome c peroxidases of *Campylobacter jejuni* involved in promoting commensal colonization of poultry. *Infection and Immunity,* 76**,** 1105-14.

Blaser, M. J. (1997). Epidemiologic and Clinical Features of *Campylobacter jejuni* Infections. *The Journal of Infectious Diseases,* 176**,** S103-S105.

Bochner, B. R. (2003). New technologies to assess genotype-phenotype relationships. *Nature Reviews Genetics,* 4**,** 309-14.

Bremer, H. (1975). Parameters affecting the rate of synthesis of ribosomes and RNA polymerase in bacteria. *Journal of Theoretical Biology,* 53**,** 115-24.

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., 3rd, Stevens, R., Vonstein, V., Wattam, A. R. & Xia, F. (2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports,* 5**,** 8365.

Bronowski, C., James, C. E. & Winstanley, C. (2014). Role of environmental survival in transmission of *Campylobacter jejuni*. *FEMS Microbiology Letters,* 356**,** 8-19.

Buck, G., Parshall, K. & Davis, C. (1983). spiral to coccoid.-1983-Buck-420-1, Electron Microscopy of the Coccoid Form of *Campylobacter*. *Journal of Cinical Microbiology,* 18**,** 420-421.

Byrne, C. M., Clyne, M. & Bourke, B. (2007). *Campylobacter jejuni* adhere to and invade chicken intestinal epithelial cells in vitro. *Microbiology,* 153**,** 561-9.

Cairney, J., Higgins, C. F. & Booth, I. R. (1984). Proline uptake through the major transport system of *Salmonella typhimurium* is coupled to sodium ions. *Journal of Bacteriology,* 160**,** 22-7.

Carver, T., Berriman, M., Tivey, A., Patel, C., Bohme, U., Barrell, B. G., Parkhill, J. & Rajandream, M. A. (2008). Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics,* 24**,** 2672-6.

Chaisowwong, W., Kusumoto, A., Hashimoto, M., Harada, T., Maklon, K. & Kawamoto, K. (2012). Physiological Characterization of *Campylobacter jejuni* under Cold Stresses Conditions: Its Potential for Public Threat. *Journal of Veterinary Medical Science,* 74**,** 43-50.

Champion, O. L., Karlyshev, A. V., Senior, N. J., Woodward, M., La Ragione, R., Howard, S. L., Wren, B. W. & Titball, R. W. (2010). Insect infection model for *Campylobacter jejuni* reveals that O-methyl phosphoramidate has insecticidal activity. *The Journal of Infectious Diseases,* 201**,** 776-82.

Chevenet, F., Brun, C., Banuls, A. L., Jacq, B. & Christen, R. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics,* 7**,** 439.

Chidgeavadze, Z. G., Beabealashvilli, R. S., Atrazhev, A. M., Kukhanova, M. K., Azhayev, A. V. & Krayevsky, A. A. (1984). 2',3'-Dideoxy-3' aminonucleoside 5'-triphosphates are the terminators of DNA synthesis catalyzed by DNA polymerases. *Nucleic Acids Research,* 12**,** 1671-1686.

Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E. E., Turner, S. W. & Korlach, J. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods,* 10**,** 563-9.

Chmielewski, R. a. N. & Frank, J. F. (2003). Biofilm Formation and Control in Food Processing Facilities. *Comprehensive Reviews in Food Science and Food Safety,* 2**,** 22-32.

Chon, J. W., Hyeon, J. Y., Yim, J. H., Kim, J. H., Song, K. Y. & Seo, K. H. (2012). Improvement of modified charcoal-cefoperazone-deoxycholate agar by supplementation with a high concentration of polymyxin B for detection of *Campylobacter jejuni* and *C. coli* in chicken carcass rinses. *Applied and Environmental Microbiology,* 78**,** 1624-6.

Cody, A. J., Bray, J. E., Jolley, K. A., Mccarthy, N. D. & Maiden, M. C. J. (2017). Core Genome Multilocus Sequence Typing Scheme for Stable, Comparative Analyses of *Campylobacter jejuni* and *C. coli* Human Disease Isolates. *Journal of Clinical Microbiology,* 55**,** 2086-2097.

Cody, A. J., Maiden, M. J. C. & Dingle, K. E. (2009). Genetic diversity and stability of the porA allele as a genetic marker in human *Campylobacter* infection. *Microbiology,* 155**,** 4145-4154.

Cody, A. J., Mccarthy, N. D., Jansen Van Rensburg, M., Isinkaye, T., Bentley, S. D., Parkhill, J., Dingle, K. E., Bowler, I. C., Jolley, K. A. & Maiden, M. C. (2013). Real-time genomic epidemiological evaluation of human Campylobacter isolates by use of whole-genome multilocus sequence typing. *J Clin Microbiol,* 51**,** 2526-34.

Cody, A. J., Mccarthy, N. M., Wimalarathna, H. L., Colles, F. M., Clark, L., Bowler, I. C., Maiden, M. C. & Dingle, K. E. (2012). A longitudinal 6-year study of the molecular epidemiology of clinical *Campylobacter* isolates in Oxfordshire, United kingdom. *Journal of Clinical Microbiology,* 50**,** 3193-201.

Colles, F. M. & Maiden, M. C. (2012). *Campylobacter* sequence typing databases: applications and future prospects. *Microbiology,* 158**,** 2695-709.

Condon, C., Liveris, D., Squires, C., Schwartz, I. & Squires, C. L. (1995). rRNA operon multiplicity in *Escherichia coli* and the physiological implications of rrn inactivation. *Journal of Bacteriology,* 177**,** 4152-4156.

Conway, T. (1992). The Entner-Doudoroff pathway: history, physiology and molecular biology. *FEMS Microbiology Reviews,* 9**,** 1-27.

Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N. & Rosenbusch, J. P. (1992). Crystal structures explain functional properties of two *E. coli* porins. *Nature,* 358**,** 727-33.

Dasti, J. I., Tareen, A. M., Lugert, R., Zautner, A. E. & Gross, U. (2010). *Campylobacter jejuni*: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. *International Journal of Medical Microbiology* 300**,** 205-11.

De Haan, C. P., Kivisto, R., Hakkinen, M., Rautelin, H. & Hanninen, M. L. (2010). Decreasing trend of overlapping multilocus sequence types between human and chicken *Campylobacter jejuni* isolates over a decade in Finland. *Applied and Environmental Microbiology,* 76**,** 5228-36.

De Vries, S. P., Linn, A., Macleod, K., Maccallum, A., Hardy, S. P., Douce, G., Watson, E., Dagleish, M. P., Thompson, H., Stevenson, A., Kennedy, D., Baig, A., Coward, C., Maskell, D. J., Smith, D. G., Grant, A. J. & Everest, P. (2017). Analysis of *Campylobacter jejuni* infection in the gnotobiotic piglet and genome-wide identification of bacterial factors required for infection. *Scientific Reports,* 7**,** 44283.

Dearlove, B. L., Cody, A. J., Pascoe, B., Meric, G., Wilson, D. J. & Sheppard, S. K. (2016). Rapid host switching in generalist *Campylobacter* strains erodes the signal for tracing human infections. *The ISME Journal,* 10**,** 721-9.

Debruyne, L., Gevers, D. & Vandamme, P. (2008). Taxonomy of the Family *Campylobacteraceae. In:* NACHAMKIN, I., AZSZYMANSKI, C. M. & BLASER, M. J. (eds.) *Campylobacter. 3rd ed*. Washington. USA: *ASM Press*.

Dingle, K. E., Colles, F. M., Falush, D. & Maiden, M. C. (2005). Sequence typing and comparison of population biology of *Campylobacter coli* and *Campylobacter jejuni. Journal of Clinical Microbiology,* 43**,** 340-7.

Dingle, K. E., Colles, F. M., Wareing, D. R., Ure, R., Fox, A. J., Bolton, F. E., Bootsma, H. J., Willems, R. J., Urwin, R. & Maiden, M. C. (2001). Multilocus sequence typing system for *Campylobacter jejuni. Journal of Clinical Microbiology,* 39**,** 14-23.

Dorrell, N., Mangan, J. A., Laing, K. G., Hinds, J., Linton, D., Al-Ghusein, H., Barrell, B. G., Parkhill, J., Stoker, N. G., Karlyshev, A. V., Butcher, P. D. & Wren, B. W. (2001). Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. *Genome Research,* 11**,** 1706-15.

Doyle, M. P. & Roman, D. J. (1982). Recovery of *Campylobacter jejuni* and *Campylobacter coli* from inoculated foods by selective enrichment. *Applied and Environmental Microbiology,* 43**,** 1343-53.

Duffy, L. & Dykes, G. A. (2006). Growth temperature of four *Campylobacter jejun*i strains influences their subsequent survival in food and water. *Letters in Applied Microbiology,* 43**,** 596-601.

Dwivedi, R., Nothaft, H., Garber, J., Xin Kin, L., Stahl, M., Flint, A., Van Vliet, A. H. M., Stintzi, A. & Szymanski, C. M. (2016). L-fucose influences chemotaxis and biofilm formation in *Campylobacter jejuni. Molecular Microbiology,* 101**,** 575-589.

Efsa (2016). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2015. *European Food Safety Authority Journal,* 14.

Ellis-Iversen, J., Ridley, A., Morris, V., Sowa, A., Harris, J., Atterbury, R., Sparks, N. & Allen, V. (2012). Persistent environmental reservoirs on farms as risk factors for *Campylobacter* in commercial poultry. *Epidemiology & Infection,* 140**,** 916-24.

Ellwood, M. & Nomura, M. (1980). Deletion of a ribosomal ribonucleic acid operon in *Escherichia coli. Journal of Bacteriology,* 143**,** 1077-80.

English, A. C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J., Qin, X., Muzny, D. M., Reid, J. G., Worley, K. C. & Gibbs, R. A. (2012. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One,* 7**,** e47768.

Esson, D., Gupta, S., Bailey, D., Wigley, P., Wedley, A., Mather, A. E., Meric, G., Mastroeni, P., Sheppard, S. K., Thomson, N. R., Parkhill, J., Maskell, D. J., Christie, G. & Grant, A. J. (2017). Identification and initial characterisation of a protein involved in *Campylobacter jejuni* cell shape. *Microbial Pathogenesis* 104**,** 202-211.

Fauchere, J. L., Rosenau, A., Veron, M., Moyen, E. N., Richard, S. & Pfister, A. (1986). Association with HeLa cells of *Campylobacter jejuni* and *Campylobacter coli* isolated from human feces. *Infection and Immunity,* 54**,** 283-7.

Ferrara, L. G. M., Wallat, G. D., Moynie, L., Dhanasekar, N. N., Aliouane, S., Acosta-Gutierrez, S., Pages, J. M., Bolla, J. M., Winterhalter, M., Ceccarelli, M. & Naismith, J. H. (2016). MOMP from *Campylobacter jejuni* Is a Trimer of 18-Stranded beta-Barrel Monomers with a Ca(2+) Ion Bound at the Constriction Zone. *Journal of Molecular Biology,* 428**,** 4528-4543.

Firlieyanti, A. S., Connerton, P. L. & Connerton, I. F. (2016). *Campylobacters* and their bacteriophages from chicken liver: The prospect for phage biocontrol. *International Journal of Food Microbiology,* 237, 121-127.

Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W. & Milo, R. (2013). Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proceedings of the National Academy of Sciences U S A,* 110, 10039-44.

Flanagan, R. C., Neal-Mckinney, J. M., Dhillon, A. S., Miller, W. G. & Konkel, M. E. (2009). Examination of Campylobacter jejuni putative adhesins leads to the identification of a new protein, designated FlpA, required for chicken colonization. *Infect Immun,* 77, 2399-407.

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.

Fox, J. G., Ackerman, J. I., Taylor, N., Claps, M. & Murphy, J. C. (1987). *Campylobacter jejuni* infection in the ferret: an animal model of human campylobacteriosis. *American Journal of Veterinary Research,* 48, 85-90.

French, N. P., Midwinter, A., Holland, B., Collins-Emerson, J., Pattison, R., Colles, F. & Carter, P. (2009). Molecular epidemiology of *Campylobacter jejuni* isolates from wild-bird fecal material in children's playgrounds. *Applied and Environmental Microbiology,* 75, 779-83.

Friis, C., Wassenaar, T. M., Javed, M. A., Snipen, L., Lagesen, K., Hallin, P. F., Newell, D. G., Toszeghy, M., Ridley, A., Manning, G. & Ussery, D. W. (2010). Genomic characterization of *Campylobacter jejuni* strain M1. *PLoS One,* 5, e12253.

Galardini, M., Biondi, E. G., Bazzicalupo, M. & Mengoni, A. (2011). CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code for Biology and Medicine,* 6, 11.

George, S. E., Costenbader, C. J. & Melton, T. (1985). Diauxic growth in *Azotobacter vinelandii. Journal of Bacteriology,* 164, 866-871.

Goodwin, C., Armstrong, J., Vers, T., Peters, M., Collins, M., Sly, L., Mcconnell, W. & Harper, W. (1989). Transfer of *Campylobacter pylori* and *Campylobacter mustelae* to *Helicobacter* gen. nov. as *Helicobacter pylori* comb. nov. and *Helicobacter mustelae* comb. nov. Respectively. *International of journal of Systematic bacteriology,* 39, 397-405.

Gorkiewicz, G., Feierl, G., Schober, C., Dieber, F., Kofer, J., Zechner, R. & Zechner, E. L. (2003). Species-Specific Identification of *Campylobacters* by Partial 16S rRNA Gene Sequencing. *Journal of Clinical Microbiology,* 41, 2537-2546.

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L., Daly, J. M., Mumford, J. A. & Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science,* 303, 327-32.

Griffiths, P. L. (1993). Morphological changes of *Campylobacter jejuni* growing in liquid culture. *Letters in Applied Microbiology,* 17, 152-5.

Gripp, E., Hlahla, D., Didelot, X., Kops, F., Maurischat, S., Tedin, K., Alter, T., Ellerbroek, L., Schreiber, K., Schomburg, D., Janssen, T., Bartholomaus, P., Hofreuter, D., Woltemate, S., Uhr, M., Brenneke, B., Gruning, P., Gerlach, G., Wieler, L., Suerbaum, S. & Josenhans, C. (2011). Closely related *Campylobacter jejuni* strains from different sources reveal a generalist rather than a specialist lifestyle. *BMC Genomics,* 12, 584.

Guccione, E., Leon-Kempis Mdel, R., Pearson, B. M., Hitchin, E., Mulholland, F., Van Diemen, P. M., Stevens, M. P. & Kelly, D. J. (2008). Amino acid-dependent growth of *Campylobacter jejuni*: key roles for aspartase (AspA) under microaerobic and oxygen-limited conditions and identification of AspB (Cj0762), essential for growth on glutamate. *Molecular Microbiology,* 69**,** 77-93.

Guccione, E. J., Kendall, J. J., Hitchcock, A., Garg, N., White, M. A., Mulholland, F., Poole, R. K. & Kelly, D. J. (2017). Transcriptome and proteome dynamics in chemostat culture reveal how *Campylobacter jejuni* modulates metabolism, stress responses and virulence factors upon changes in oxygen availability. *Environmental Microbiology,* 19**,** 4326-4348.

Guerry, P. 2007. *Campylobacter* flagella: not just for motility. *Trends in Microbiology,* 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. *Journal of Bacteriology,* 173**,** 4757-64.

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.

Guerry, P., Poly, F., Riddle, M., Maue, A. C., Chen, Y. H. & Monteiro, M. A. (2012). *Campylobacter* polysaccharide capsules: virulence and vaccines. *Front Cell Infect Microbiol,* 2**,** 7.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology,* 59**,** 307-21.

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.

Gunther, N. W. T. & Chen, C. Y. (2009). The biofilm forming potential of bacterial species in the genus *Campylobacter*. *Food Microbiology,* 26**,** 44-51.

Haddock, G., Mullin, M., Maccallum, A., Sherry, A., Tetley, L., Watson, E., Dagleish, M., Smith, D. G. & Everest, P. (2010). *Campylobacter jejuni* 81-176 forms distinct microcolonies on in vitro-infected human small intestinal tissue prior to biofilm formation. *Microbiology,* 156**,** 3079-84.

Hanning, I., Jarquin, R. & Slavik, M. (2008. *Campylobacter jejuni* as a secondary colonizer of poultry biofilms. *Journal of Applied Microbiology,* 105**,** 1199-208.

Hanning, I. & Slavik, M. (2009). *Campylobacter jejuni* as a Primary Colonizing Biofilm Former. *International Iournal of Poultry Science,* 8**,** 1-6.

Hansson, I., Persson, M., Svensson, L., Engvall, E. O. & Johansson, K. E. (2008). Identification of nine sequence types of the 16S rRNA genes of *Campylobacter jejuni* subsp. *jejuni* isolated from broilers. *Acta Veterinaria Scandinavica* 50**,** 10.

Harvala, H., Rosendal, T., Lahti, E., Engvall, E. O., Brytting, M., Wallensten, A. & Lindberg, A. (2016). Epidemiology of *Campylobacter jejuni* infections in Sweden, November 2011-October 2012: is the severity of infection associated with *C. jejuni* sequence type? *Infection Ecology & Epidemiology,* 6**,** 31079.

Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. & Shinagawa, H. (2001). Complete Genome Sequence of Enterohemorrhagic *Eschelichia coli* O157:H7 and Genomic Comparison with a Laboratory Strain K-12. *DNA Research,* 8**,** 11-22.

He, Y. & Chen, C. Y. (2010. Quantitative analysis of viable, stressed and dead cells of *Campylobacter jejuni* strain 81-176. *Food Microbiology,* 27**,** 439-46.

Heather, J. M. & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics,* 107**,** 1-8.

Hendrixson, D. R. & Dirita, V. J. (2004). Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Molecular Microbiology* 52**,** 471-84.

Hermans, D., Pasmans, F., Messens, W., Martel, A., Van Immerseel, F., Rasschaert, G., Heyndrickx, M., Van Deun, K. & Haesebrouck, F. (2012). Poultry as a host for the zoonotic pathogen *Campylobacter jejuni*. *Vector-Borne and Zoonotic Diseases* 12**,** 89-98.

Hermans, D., Van Deun, K., Martel, A., Van Immerseel, F., Messens, W., Heyndrickx, M., Haesebrouck, F. & Pasmans, F. (2011). Colonization factors of *Campylobacter jejuni* in the chicken gut. *Veterinary Research,* 42**,** 82.

Hill, L., Veli, N. & Coote, P. J. (2014). Evaluation of *Galleria mellonella* larvae for measuring the efficacy and pharmacokinetics of antibiotic therapies against *Pseudomonas* aeruginosa infection. *International Journal of Antimicrobial Agents,* 43**,** 254-61.

Hofreuter, D. (2014). Defining the metabolic requirements for the growth and colonization capacity of *Campylobacter jejuni*. *Frontiers in Cellular and Infection Microbiology,* 4**,** 137.

Hofreuter, D., Mohr, J., Wensel, O., Rademacher, S., Schreiber, K., Schomburg, D., Gao, B. & Galan, J. E. (2012). Contribution of amino acid catabolism to the tissue specific persistence of *Campylobacter jejuni* in a murine colonization model. *PLOS ONE: accelerating the publication of peer-reviewed science,* 7**,** e50699.

Hofreuter, D., Novik, V. & Galan, J. E. (2008). Metabolic diversity in *Campylobacter jejuni* enhances specific tissue colonization. *Cell Host & Microbe,* 4**,** 425-33.

Holst Sorensen, M. C., Van Alphen, L. B., Fodor, C., Crowley, S. M., Christensen, B. B., Szymanski, C. M. & Brondsted, L. (2012). Phase variable expression of capsular polysaccharide modifications allows *Campylobacter jejuni* to avoid bacteriophage infection in chickens. *Frontiers in Cellular and Infection Microbiology,* 2**,** 11.

Holt, K. E., Parkhill, J., Mazzoni, C. J., Roumagnac, P., Weill, F. X., Goodhead, I., Rance, R., Baker, S., Maskell, D. J., Wain, J., Dolecek, C., Achtman, M. & Dougan, G. (2008). High-throughput sequencing provides insights into genome variation and evolution in *Salmonella Typhi. Nat Genet,* 40**,** 987-93.

Humphrey, S., Lacharme-Lora, L., Chaloner, G., Gibbs, K., Humphrey, T., Williams, N. & Wigley, P. (2015). Heterogeneity in the Infection Biology of *Campylobacter jejuni* Isolates in Three Infection Models Reveals an Invasive and Virulent Phenotype in a ST21 Isolate from Poultry. *PLOS ONE: accelerating the publication of peer-reviewed science,* 10**,** e0141182.

Huson, D. H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution,* 23**,** 254-67.

Huson, D. H. & Scornavacca, C. (2011). A survey of combinatorial methods for phylogenetic networks. *Genome Biology and Evolution,* 3**,** 23-35.

Ica, T., Caner, V., Istanbullu, O., Nguyen, H. D., Ahmed, B., Call, D. R. & Beyenal, H. (2012). Characterization of mono- and mixed-culture *Campylobacter jejuni* biofilms. *Applied and Environmental Microbiology,* 78**,** 1033-8.

Ikeda, N. & Karlyshev, A. V. (2012). Putative mechanisms and biological role of coccoid form formation in *Campylobacter jejuni*. *European Journal of Microbiology & Immunology (Bp),* 2**,** 41-9.

Imlay, J. A. (2003). Pathways of oxidative damage. *Annual Review of Microbiology,* 57**,** 395-418.

Indikova, I., Humphrey, T. J. & Hilbert, F. (2015). Survival with a Helping Hand: *Campylobacter* and Microbiota. *Frontiers in Microbiology,* 6**,** 1266.

Itoh, T., Saito, K., Maruyama, T., Sakai, S., Ohashi, M. & Oka, A. (1980). An outbreak of acute enteritis due to *Campylobacter fetus* subspecies *jejuni* at a nursery school of Tokyo. *Microbiol Immunol,* 24**,** 371-9.

Jagannathan, A., Constantinidou, C. & Penn, C. W. (2001). Roles of *rpoN*, *fliA*, and *flgR* in expression of flagella in *Campylobacter jejuni*. *Journal of Bacteriology,* 183**,** 2937-42.

Jang, K. I., Kim, M. G., Ha, S. D., Kim, K. S., Lee, K. H., Chung, D. H., Kim, C. H. & Kim, K. Y. (2007. Morphology and adhesion of *Campylobacter jejuni* to chicken skin under varying conditions. *Journal of Microbiology and Biotechnology,* 17**,** 202-6.

Janssen, R., Krogfelt, K. A., Cawthraw, S. A., Van Pelt, W., Wagenaar, J. A. & Owen, R. J. (2008). Host-pathogen interactions in *Campylobacter* infections: the host perspective. *Clinical Microbiology Reviews,* 21**,** 505-18.

Jay-Russell, M. T., Mandrell, R. E., Yuan, J., Bates, A., Manalac, R., Mohle-Boetani, J., Kimura, A., Lidgard, J. & Miller, W. G. (2013). Using major outer membrane protein typing as an epidemiological tool to investigate outbreaks caused by milk-borne *Campylobacter jejuni* isolates in California. *Journal of Clinical Microbiology,* 51**,** 195-201.

Jeon, B., Muraoka, W., Sahin, O. & Zhang, Q. (2008). Role of Cj1211 in natural transformation and transfer of antibiotic resistance determinants in *Campylobacter jejuni*. *Antimicrobial Agents and Chemotherapy,* 52**,** 2699-708.

Johnson, T. J., Shank, J. M. & Johnson, J. G. (2017). Current and Potential Treatments for Reducing *Campylobacter* Colonization in Animal Hosts and Disease in Humans. *Frontiers in Microbiology,* 8**,** 487.

Jolley, K. A. & Maiden, M. C. (2010). BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics,* 11**,** 595.

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. *Infection and Immunity,* 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.

Kalmokoff, M., Lanthier, P., Tremblay, T. L., Foss, M., Lau, P. C., Sanders, G., Austin, J., Kelly, J. & Szymanski, C. M. (2006). Proteomic analysis of *Campylobacter jejuni* 11168 biofilms reveals a role for the motility complex in biofilm formation. *Journal of Bacteriology,* 188**,** 4312-20.

Kanji, A., Jones, M. A., Maskell, D. J. & Grant, A. J. (2015). *Campylobacter jejuni* PflB is required for motility and colonisation of the chicken gastrointestinal tract. *Microbial Pathogenesis,* 89**,** 93-9.

Karlsson, E., Larkeryd, A., Sjodin, A., Forsman, M. & Stenberg, P. (2015). Scaffolding of a bacterial genome using MinION nanopore sequencing. *Sci Rep,* 5**,** 11996.

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. (2005). Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Molecular Microbiology,* 55**,** 90-103.

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. *Molecular Microbiology* 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.

Kaspars K., Mati R., Stella S., Aivars B. & Ari H. (2014). *Campylobacter* species and their antimicrobial resistance in Latvian broiler chicken production. *Food Control,* 46**,** 86-90.

Kavanagh, K. & Reeves, E. P. (2004). Exploiting the potential of insects for in vivo pathogenicity testing of microbial pathogens. *FEMS Microbiology Reviews,* 28**,** 101-12.

Kelly, D. J. (2001). The physiology and metabolism of *Campylobacter jejuni* and *Helicobacter pylori*. *Symp Ser Soc Appl Microbiol***,** 16s-24s.

Kennedy, N. A., Walker, A. W., Berry, S. H., Duncan, S. H., Farquarson, F. M., Louis, P., Thomson, J. M., Consortium, U. I. G., Satsangi, J., Flint, H. J., Parkhill, J., Lees, C. W. & Hold, G. L. (2014). The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. *PLOS ONE: accelerating the publication of peer-reviewed science,* 9**,** e88982.

Ketley, J. M. (1997). Pathogenesis of enteric infection by *Campylobacter*. *Microbiology,* 143 (Pt 1)**,** 5-21.

Kim, J., Oh, E., Banting, G. S., Braithwaite, S., Chui, L., Ashbolt, N. J., Neumann, N. F. & Jeon, B. (2016). An Improved Culture Method for Selective Isolation of *Campylobacter jejuni* from Wastewater. *Front Microbiol,* 7**,** 1345.

Kim, J. C., Oh, E., Kim, J. & Jeon, B. (2015a). Regulation of oxidative stress resistance in *Campylobacter jejuni*, a microaerophilic foodborne pathogen. *Frontiers in Microbiology,* 6**,** 751.

Kim, J. S., Park, C. & Kim, Y. J. (2015b). Role of flgA for Flagellar Biosynthesis and Biofilm Formation of *Campylobacter jejuni* NCTC11168. *Journal of Microbiology and Biotechnology,* 25**,** 1871-9.

Kim, N. W. & Chan, V. L. (1989). Isolation and Characterization of the Ribosomal RNA Genes of *Campylobacter jejuni*. *Current Microbiology,* 19**,** 247-252.

Kim, N. W., Lombardi, R., Bingham, H., Hani, E., Louie, H., Ng, D. & Chan, V. L. (1993). Fine mapping of the three rRNA operons on the updated genomic map of *Campylobacter jejuni* TGH9011 (ATCC 43431). *Journal of Bacteriology,* 175**,** 7468-7470.

Kist, M. (1986). [Who discovered *Campylobacter jejuni/coli*? A review of hitherto disregarded literature]. *Zentralbl Bakteriol Mikrobiol Hyg A,* 261**,** 177-86.

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. *Journal of Bacteriology,* 186**,** 3296-303.

Korkmaz, G., Holm, M., Wiens, T. & Sanyal, S. (2014). Comprehensive analysis of stop codon usage in bacteria and its correlation with release factor abundance. *The Journal of Biological Chemistry,* 289**,** 30334-42.

Korolik, V., Alderton, M. R., Smith, S. C., Chang, J. & Coloe, P. J. (1998). Isolation and molecular analysis of colonising and non-colonising strains of *Campylobacter jejuni* and *Campylobacter coli* following experimental infection of young chickens. *Veterinary Microbiology,* 60**,** 239-49.

Kwan, P. S., Barrigas, M., Bolton, F. J., French, N. P., Gowland, P., Kemp, R., Leatherbarrow, H., Upton, M. & Fox, A. J. (2008). Molecular epidemiology of *Campylobacter jejuni* populations in dairy cattle, wildlife, and the environment in a farmland area. *Applied and Environmental Microbiology,* 74**,** 5130-8.

Lang, P., Lefebure, T., Wang, W., Pavinski Bitar, P., Meinersmann, R. J., Kaya, K. & Stanhope, M. J. (2010). Expanded multilocus sequence typing and comparative genomic hybridization of *Campylobacter coli* isolates from multiple hosts. *Applied and Environmental Microbiology,* 76**,** 1913-25.

Lawsonl, A. J., Shafi', M. S., Pathak, K. & Stanley, J. (1998). Detection of *Campylobacter* in gastroenteritis : comparison of direct PCR assay of faecal samples with selective culture. *Epidemiology & Infection* 121**,** 547-553.

Leach, S., Harvey, P. & Wali, R. (1997). Changes with growth rate in the membrane lipid composition of and amino acid utilization by continuous cultures of *Campylobacter jejuni*. *Journal of Applied Microbiology,* 82**,** 631-40.

Lee, R. B., Hassane, D. C., Cottle, D. L. & Pickett, C. L. (2003). Interactions of *Campylobacter jejuni* Cytolethal Distending Toxin Subunits CdtA and CdtC with HeLa Cells. *Infection and Immunity,* 71**,** 4883-4890.

Leon-Kempis, M., R., Guccione, E., Mulholland, F., Williamson, M. P. & Kelly, D. J. (2006). The *Campylobacter jejuni* PEB1a adhesin is an aspartate/glutamate-binding protein of an ABC transporter essential for microaerobic growth on dicarboxylic amino acids. *Molecular Microbiology,* 60**,** 1262-75.

Leuko, S. & Raivio, T. L. (2012). Mutations that impact the enteropathogenic *Escherichia coli* Cpx envelope stress response attenuate virulence in *Galleria mellonella*. *Infection and Immunity,* 80**,** 3077-85.

Levengood, S. K., Beyer, W. F., Jr. & Webster, R. E. (1991). TolA: a membrane protein involved in colicin uptake contains an extended helical region. *Proceedings of the National Academy of Sciences U S A,* 88**,** 5939-43.

Li, L., Mendis, N., Trigui, H., Oliver, J. D. & Faucher, S. P. (2014). The importance of the viable but non-culturable state in human bacterial pathogens. *Frontiers in Microbiology,* 5**,** 258.

Lien, K. A., Sauer, W. C. & Fenton, M. (1997). Mucin output in ileal digesta of pigs fed a protein-free diet. *Z Ernahrungswiss,* 36**,** 182-90.

Loc Carrillo, C., Atterbury, R. J., El-Shibiny, A., Connerton, P. L., Dillon, E., Scott, A. & Connerton, I. F. (2005). Bacteriophage therapy to reduce *Campylobacter jejuni* colonization of broiler chickens. *Applied and Environmental Microbiology,* 71**,** 6554-63.

Loman, N. J. & Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nat Rev Microbiol,* 13**,** 787-94.

Loman, N. J. & Watson, M. (2015). Successful test launch for nanopore sequencing. *Nat Methods,* 12**,** 303-4.

Loomis, W. F. & Magasanik, B. (1967). Glucose-Lactose Diauxie in *Escherichia coli*. *Journal of Bacteriology,* 93**,** 1397-1401.

Loughney, K., Lund, E. & Dahlberg, J. E. (1983). Deletion of an rRNA gene set in *Bacillus subtilis*. *Journal of Bacteriology,* 154**,** 529-32.

Louwen, R., Heikema, A., Van Belkum, A., Ott, A., Gilbert, M., Ang, W., Endtz, H. P., Bergman, M. P. & Nieuwenhuis, E. E. (2008). The sialylated lipooligosaccharide outer core in *Campylobacter jejuni* is an important determinant for epithelial cell invasion. *Infection and Immunity,* 76**,** 4431-8.

Lynch, W. H. & Franklin, M. (1978). Effect of temperature on diauxic growth with glucose and organic acids in *Pseudomonas fluorescens*. *Archives of Microbiology* 118**,** 133-40.

Maćkiw, E., Korsak, D., Rzewuska, K., Tomczuk, K. & Rożynek, E. (2012). Antibiotic resistance in *Campylobacter jejuni* and *Campylobacter coli* isolated from food in Poland. *Food Control,* 23**,** 297-301.

Mahdavi, J., Pirinccioglu, N., Oldfield, N. J., Carlsohn, E., Stoof, J., Aslam, A., Self, T., Cawthraw, S. A., Petrovska, L., Colborne, N., Sihlbom, C., Boren, T., Wooldridge, K. G. & Ala'aldeen, D. A. (2014). A novel O-linked glycan modulates *Campylobacter jejuni* major outer membrane protein-mediated adhesion to human histo-blood group antigens and chicken colonization. *Open Biology,* 4**,** 130202.

Mansfield, L. S., Bell, J. A., Wilson, D. L., Murphy, A. J., Elsheikha, H. M., Rathinam, V. A., Fierro, B. R., Linz, J. E. & Young, V. B. (2007). C57BL/6 and congenic interleukin-10-deficient mice can serve as models of *Campylobacter jejuni* colonization and enteritis. *Infection and Immunity,* 75**,** 1099-115.

Marchant, J., Wren, B. & Ketley, J. (2002). Exploiting genome sequence: predictions for mechanisms of *Campylobacter* chemotaxis. *Trends Microbiol,* 10**,** 155-9.

Martiny, D., Dediste, A., Debruyne, L., Vlaes, L., Haddou, N. B., Vandamme, P. & Vandenberg, O. (2011). Accuracy of the API Campy system, the Vitek 2 Neisseria-Haemophilus card and matrix-assisted laser desorption ionization time-of-flight mass spectrometry for the identification of *Campylobacter* and related organisms. *Clinical Microbiology and Infection,* 17**,** 1001-6.

Mäser, P., Thomine, S., Schroeder, J. I., Ward, J. M., Hirschi, K., Sze, H., Talke, I. N., Amtmann, A., Maathuis, F. J. M., Sanders, D., Harper, J. F., Tchieu, J., Gribskov, M., Persans, M. W., Salt, D. E., Kim, S. A. & Guerinot, M. L. (2001). Phylogenetic Relationships within Cation Transporter Families of Arabidopsis. *Plant Physiology,* 126**,** 1646-1667.

Mccarthy, N. D., Colles, F. M., Dingle, K. E., Bagnall, M. C., Manning, G., Maiden, M. C. & Falush, D. (2007). Host-associated genetic import in *Campylobacter jejuni*. *Emerging Infectious Diseases,* 13**,** 267-72.

Meerburg, B. & Kijlstra, A. (2007a). *Role of rodents in transmission of Salmonella and Campylobacter*.

Meerburg, B. G., Jacobs-Reitsma, W. F., Wagenaar, J. A. & Kijlstra, A. (2006). Presence of *Salmonella* and *Campylobacter* spp. in wild small mammals on organic farms. *Applied and Environmental Microbiology,* 72**,** 960-2.

Meerburg, B. G. & Kijlstra, A. (2007b). Role of rodents in transmission of *Salmonella* and *Campylobacter*. *Journal of the Science of Food and Agriculture,* 87**,** 2774-2781.

Megraud, F., Boudraa, G., Bessaoud, K., Bensid, S., Dabis, F., Soltana, R. & Touhami, M. (1990). Incidence of *Campylobacter* infection in infants in western Algeria and the possible protective role of breast feeding. *Epidemiology & Infection,* 105**,** 73-8.

Mendz, G. L., Hazell, S. L. & Burns, B. P. (1994). The Entner-Doudoroff pathway in *Helicobacter pylori*. *Archives of Biochemistry and Biophysics,* 312**,** 349-56.

251

Meric, G., Yahara, K., Mageiros, L., Pascoe, B., Maiden, M. C., Jolley, K. A. & Sheppard, S. K. (2014). A reference pan-genome approach to comparative bacterial genomics: identification of novel epidemiological markers in pathogenic Campylobacter. *PLoS One,* 9**,** e92798.

Miles, A. A., Misra, S. S. & Irwin, J. O. (1938). The estimation of the bactericidal power of the blood. *The Journal of Hygiene,* 38**,** 732-749.

Miller, C. E., Williams, P. H. & Ketley, J. M. (2009). Pumping iron: mechanisms for iron uptake by *Campylobacter. Microbiology,* 155**,** 3157-65.

Miller, W. G. (2008). Comparative Genomics of *Campylobacter* Species Other than *Campylobacter jejuni.* 73-95.

Miller, W. G., Parker, C. T., Heath, S. & Lastovica, A. J. (2007). Identification of genomic differences between *Campylobacter jejuni* subsp. *jejuni* and *C. jejuni* subsp. *doylei* at the nap locus leads to the development of a *C. jejuni* subspeciation multiplex PCR method. *BMC Microbiology,* 7**,** 11.

Mohan, V., Stevenson, M. A., Marshall, J. C. & French, N. P. (2017). Characterisation by multilocus sequence and *porA* and *flaA* typing of *Campylobacter jejuni* isolated from samples of dog faeces collected in one city in New Zealand. *The New Zealand Veterinary Journal* 65**,** 209-213.

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.

Mortensen, N. P., Schiellerup, P., Boisen, N., Klein, B. M., Locht, H., Abuoun, M., Newell, D. & Krogfelt, K. A. (2011). The role of *Campylobacter jejuni* cytolethal distending toxin in gastroenteritis: toxin detection, antibody production, and clinical outcome. *APMIS,* 119**,** 626-34.

Mughini-Gras, L., Penny, C., Ragimbeau, C., Schets, F. M., Blaak, H., Duim, B., Wagenaar, J. A., De Boer, A., Cauchie, H. M., Mossong, J. & Van Pelt, W. (2016). Quantifying potential sources of surface water contamination with *Campylobacter jejuni* and *Campylobacter coli*. *Water Research,* 101**,** 36-45.

Muldoon, J., Shashkov, A. S., Moran, A. P., Ferris, J. A., Senchenkova, S. N. & Savage, A. V. (2002). Structures of two polysaccharides of *Campylobacter jejuni* 81116. *Carbohydrate Research,* 337**,** 2223-9.

Mullner, P., Spencer, S. E., Wilson, D. J., Jones, G., Noble, A. D., Midwinter, A. C., Collins-Emerson, J. M., Carter, P., Hathaway, S. & French, N. P. (2009). Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach. *Infection, Genetics and Evolution,* 9**,** 1311-9.

Muralidharan, M., Ghosh, A., Singhvi, N., Dhanaraj, P. S., Lal, R., Patel, D. D., Kaicker, A. & Verma, M. (2017). Identification of genus *Campylobacter* up to species level using internal features of 16S rRNA gene sequences. *Molecular Genetics, Microbiology and Virology,* 31**,** 187-196.

Muraoka, W. T. & Zhang, Q. (2011). Phenotypic and genotypic evidence for L-fucose utilization by *Campylobacter jejuni*. *Journal of Bacteriology,* 193**,** 1065-75.

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. *Applied and Environmental Microbiology,* 59**,** 1269-1273.

Naikare, H., Palyada, K., Panciera, R., Marlow, D. & Stintzi, A. (2006). Major role for FeoB in *Campylobacter jejuni* ferrous iron acquisition, gut colonization, and intracellular survival. *Infection and Immunity,* 74**,** 5433-44.

Nappi, A. J. & Christensen, B. M. (2005). Melanogenesis and associated cytotoxic reactions: applications to insect innate immunity. *Insect Biochem Mol Biol,* 35**,** 443-59.

Nemelka, K. W., Brown, A. W., Wallace, S. M., Jones, E., Asher, L. V., Pattarini, D., Applebee, L., Gilliland, T. C., Jr., Guerry, P. & Baqar, S. (2009). Immune response to and histopathology of *Campylobacter jejuni* infection in ferrets (Mustela putorius furo). *Comp Med,* 59**,** 363-71.

Newell, D. G. (2001). Animal models of *Campylobacter jejuni* colonization and disease and the lessons to be learned from similar Helicobacter pylori models. *Symp Ser Soc Appl Microbiol***,** 57S-67S.

Newell, D. G. & Fearnley, C. (2003). Sources of *Campylobacter* colonization in broiler chickens. *Applied and Environmental Microbiology,* 69**,** 4343-51.

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

Nikel, P. I., Chavarria, M., Fuhrer, T., Sauer, U. & De Lorenzo, V. (2015). *Pseudomonas putida* KT2440 Strain Metabolizes Glucose through a Cycle Formed by Enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and Pentose Phosphate Pathways. *The Journal of Biological Chemistry,* 290**,** 25920-32.

Nkogwe, C., Raletobana, J., Stewart-Johnson, A., Suepaul, S. & Adesiyun, A. (2011). Frequency of Detection of *Escherichia coli*, *Salmonella* spp., and *Campylobacter* spp. in the Faeces of Wild Rats (*Rattus* spp.) in Trinidad and Tobago. *Veterinary Medicine International,* 2011**,** 686923.

Ochman, H., Lawrence, J. G. & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature,* 405**,** 299-304.

Oh, E. & Jeon, B. (2014). Role of alkyl hydroperoxide reductase (AhpC) in the biofilm formation of *Campylobacter jejuni. PLOS ONE: accelerating the publication of peer-reviewed science,* 9**,** e87312.

Okuda, J., Kurazono, H. & Takeda, Y. (1995. Distribution of the cytolethal distending toxin A gene (cdtA) among species of *Shigella* and *Vibrio*, and cloning and sequencing of the cdt gene from *Shigella* dysenteriae. *Microbial Pathogenesis,* 18**,** 167-72.

Olson, C. K., Ethelberg, S., Pelt, W. V. & R.V.Vauxe (2008). Epidemiological of *Campylobacter jejuni* Infections in Industialized Nations. *In:* Nachamkin, I. S., Szymanski, C. M. & M. J. Blaser (eds.) *Campylobacter* Washington.USA: ASM Press.

On, S. L. (2001). Taxonomy of *Campylobacter, Arcobacter, Helicobacter* and related bacteria: current status, future prospects and immediate concerns. *Symp Ser Soc Appl Microbiol***,** 1s-15s.

Oyarzabal, O. A. & Fernández, H. (2016). Isolation and Identification of *Campylobacter* spp. in Poultry. 19-35.

Palyada, K., Threadgill, D. & Stintzi, A. (2004). Iron acquisition and regulation in *Campylobacter jejuni. Journal of Bacteriology,* 186**,** 4714-29.

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.

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. *Epidemiology & Infection,* 106**,** 77-82.

Pead, P. (1979). Electron microscopy of *Campylobacter jejuni*. *Journal of Medical Microbiology,* 12**,** 383-385.

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). *Journal of Bacteriology,* 189**,** 8402-3.

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

Pendleton, S., Hanning, I., Biswas, D. & Ricke, S. C. (2013). Evaluation of whole-genome sequencing as a genotyping tool for *Campylobacter jejuni* in comparison with pulsed-field gel electrophoresis and flaA typing. *Poultry Science,* 92**,** 573-80.

Penner, J. L. (1988). The genus *Campylobacter*: a decade of progress. *Clinical Microbiology Reviews,* 1**,** 157-72.

Petkau, A., Stuart-Edwards, M., Stothard, P. & Van Domselaar, G. (2010). Interactive microbial genome visualization with GView. *Bioinformatics,* 26**,** 3125-6.

Phe. (2014). *Laboratory reports of Campylobacter sp by age, England and Wales, 2000-2011* [Online]. Available: http://webarchive.nationalarchives.gov.uk/20140714111045/http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/*Campylobacter*/EpidemiologicalData/campyDataEwAge/ [Accessed November 2017 November 2017].

Phe. (2017a). *Campylobacter data 2006 to 2015, National laboratory data for residents of England and Wales* [Online]. Available: https://www.gov.uk/government/organisations/public-health-england [Accessed November 2017 November 2017].

Phe. (2017b). *The symptoms, diagnosis, management and epidemiology of Campylobacter* [Online]. Available: https://www.gov.uk/government/collections/campylobacter-guidance-data-and-analysis [Accessed November 2017 November 2017].

Phillips, P. L. & Schultz, G. S. (2012). Molecular Mechanisms of Biofilm Infection: Biofilm Virulence Factors. *Advances in Wound Care (New Rochelle),* 1**,** 109-114.

Pickett, C. L., Auffenberg, T., Pesci, E. C., Sheen, V. L. & Jusuf, S. S. (1992). Iron acquisition and hemolysin production by *Campylobacter jejuni*. *Infection and Immunity,* 60**,** 3872-7.

Pike, B. L., Guerry, P. & Poly, F. (2013). Global Distribution of *Campylobacter jejuni* Penner Serotypes: A Systematic Review. *PLOS ONE: accelerating the publication of peer-reviewed science,* 8**,** e67375.

Pittman, M. S. & Kelly, D. J. (2005). Electron transport through nitrate and nitrite reductases in *Campylobacter jejuni*. *Biochemical Society Transactions,* 33**,** 190-2.

Povolotskaya, I. S., Kondrashov, F. A., Ledda, A. & Vlasov, P. K. (2012). Stop codons in bacteria are not selectively equivalent. *Biology Direct,* 7**,** 30.

Preston, M. A. & Penner, J. L. (1989). Characterization of cross-reacting serotypes of *Campylobacter jejuni*. *Canadian Journal of Microbiology,* 35**,** 265-73.

Quick, J., Quinlan, A. R. & Loman, N. J. (2014). A reference bacterial genome dataset generated on the MinION™ portable single-molecule nanopore sequencer. *GigaScience,* 3**,** 22-22.

Rahman, H., King, R. M., Shewell, L. K., Semchenko, E. A., Hartley-Tassell, L. E., Wilson, J. C., Day, C. J. & Korolik, V. (2014). Characterisation of a multi-ligand binding

chemoreceptor CcmL (Tlp3) of *Campylobacter jejuni*. *PLOS Pathogens: A Peer-Reviewed Open-Access Journal,* 10**,** e1003822.

Rainey, F. A., Ward-Rainey, N. L., Janssen, P. H., Hippe, H. & Stackebrandt, E. (1996). *Clostridium paradoxum* DSM 7308T contains multiple 16S rRNA genes with heterogeneous intervening sequences. *Microbiology,* 142 (Pt 8)**,** 2087-95.

Ramonaite, S., Tamuleviciene, E., Alter, T., Kasnauskyte, N. & Malakauskas, M. (2017). MLST genotypes of *Campylobacter jejuni* isolated from broiler products, dairy cattle and human campylobacteriosis cases in Lithuania. *BMC Infectious Diseases,* 17**,** 430.

Ramos, H. C., Rumbo, M. & Sirard, J. C. (2004). Bacterial flagellins: mediators of pathogenicity and host immune responses in mucosa. *Trends in Microbiology,* 12**,** 509-17.

Ratledge, C. & Dover, L. G. (2000). Iron metabolism in pathogenic bacteria. *Annual Review of Microbiology,* 54**,** 881-941.

Reeser, R. J., Medler, R. T., Billington, S. J., Jost, B. H. & Joens, L. A. (2007). Characterization of *Campylobacter jejuni* biofilms under defined growth conditions. *Applied and Environmental Microbiology,* 73**,** 1908-13.

Reuter, M., Mallett, A., Pearson, B. M. & Van Vliet, A. H. (2010). Biofilm formation by *Campylobacter jejuni* is increased under aerobic conditions. *Applied and Environmental Microbiology,* 76**,** 2122-8.

Rhoads, A. & Au, K. F. (2015). PacBio Sequencing and Its Applications. *Genomics Proteomics Bioinformatics,* 13**,** 278-89.

Richardson, P. T. & Park, S. F. (1995). Enterochelin acquisition in *Campylobacter coli:* characterization of components of a binding-protein-dependent transport system. *Microbiology,* 141 (Pt 12)**,** 3181-91.

Riechmann, L. & Holliger, P. (1997). The C-terminal domain of TolA is the coreceptor for filamentous phage infection of *E. coli*. *Cell,* 90**,** 351-60.

Rollins, D. M. & Colwell, R. R. (1986). Viable but nonculturable stage of *Campylobacter jejuni* and its role in survival in the natural aquatic environment. *Applied and Environmental Microbiology,* 52**,** 531-8.

Romaniuk, P. J., Zoltowska, B., Trust, T. J., Lane, D. J., Olsen, G. J., Pace, N. R. & Stahl, D. A. (1987). *Campylobacter pylori,* the spiral bacterium associated with human gastritis, is not a true *Campylobacter* sp. *Journal of Bacteriology,* 169**,** 2137-2141.

Rosenquist, H., Sommer, H. M., Nielsen, N. L. & Christensen, B. B. (2006). The effect of slaughter operations on the contamination of chicken carcasses with thermotolerant *Campylobacter*. *International Journal of Food Microbiology,* 108**,** 226-32.

Rueden, C., Schindelin, J., Hiner, M., Dezonia, B., E. Walter, A. & Eliceiri, K. (2017). *ImageJ2: ImageJ for the next generation of scientific image data*.

Sails, A. D., Swaminathan, B. & Fields, P. I. (2003). Utility of Multilocus Sequence Typing as an Epidemiological Tool for Investigation of Outbreaks of Gastroenteritis Caused by *Campylobacter jejuni*. *Journal of Clinical Microbiology,* 41**,** 4733-4739.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A,* 74**,** 5463-7.

Schröder, I., Johnson, E. & De Vries, S. (2003). Microbial ferric iron reductases. *FEMS Microbiology Reviews,* 27**,** 427-447.

Schurch, A. C. & Van Schaik, W. (2017). Challenges and opportunities for whole-genome sequencing-based surveillance of antibiotic resistance. *Ann N Y Acad Sci,* 1388**,** 108-120.

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. *Journal of Bacteriology,* 184**,** 4187-4196.

Senior, N. J., Bagnall, M. C., Champion, O. L., Reynolds, S. E., La Ragione, R. M., Woodward, M. J., Salguero, F. J. & Titball, R. W. (2011). *Galleria mellonella* as an infection model for *Campylobacter jejuni* virulence. *Journal of Medical Microbiology,* 60**,** 661-9.

Shang, Y., Ren, F., Song, Z., Li, Q., Zhou, X., Wang, X., Xu, Z., Bao, G., Wan, T., Lei, T., Wang, N., Jiao, X. A. & Huang, J. (2016). Insights into *Campylobacter jejuni* colonization and enteritis using a novel infant rabbit model. *Scientific Reports* 6**,** 28737.

Sheppard, S. K., Dallas, J. F., Strachan, N. J., Macrae, M., Mccarthy, N. D., Wilson, D. J., Gormley, F. J., Falush, D., Ogden, I. D., Maiden, M. C. & Forbes, K. J. (2009). *Campylobacter* genotyping to determine the source of human infection. *Clinical Infectious Diseases,* 48**,** 1072-8.

Sheppard, S. K., Didelot, X., Meric, G., Torralbo, A., Jolley, K. A., Kelly, D. J., Bentley, S. D., Maiden, M. C., Parkhill, J. & Falush, D. (2013). Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter. Proceedings of the National Academy of Sciences U S A,* 110**,** 11923-7.

Shimizu, T. & Nakamura, A. (2014). Characterization of LgnR, an IclR family transcriptional regulator involved in the regulation of L-gluconate catabolic genes in *Paracoccus* sp. 43P. *Microbiology,* 160**,** 623-34.

Shou, S. P., Dular, R. & S., K. (1983). Effect of ferrous sulfate, sodium metabisulfite, and sodium pyruvate on survival of *Campylobacter jejuni. Journal of Clinical Microbiology,* 18**,** 986-987.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., Mcwilliam, H., Remmert, M., Soding, J., Thompson, J. D. & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology,* 7**,** 539.

Skarp, C. P., Akinrinade, O., Nilsson, A. J., Ellstrom, P., Myllykangas, S. & Rautelin, H. (2015). Comparative genomics and genome biology of invasive *Campylobacter jejuni*. *Scientific Reports,* 5**,** 17300.

Skirrow, M. B. (2006). John McFadyean and the centenary of the first isolation of *Campylobacter* species. *Clinical Infectious Diseases,* 43**,** 1213-7.

Smalla, K., Wachtendorf, U., Heuer, H., Liu, W. T. & Forney, L. (1998). Analysis of Biolog GN Substrate Utilization Patterns by Microbial Communities. *Applied and Environmental Microbiology,* 64**,** 1220-5.

Sopwith, W., Birtles, A., Matthews, M., Fox, A., Gee, S., Painter, M., Regan, M., Syed, Q. & Bolton, E. (2006). *Campylobacter jejuni* multilocus sequence types in humans, northwest England, 2003-2004. *Emerging Infectious Diseases journal* 12**,** 1500-7.

Spees, A. M., Lopez, C. A., Kingsbury, D. D., Winter, S. E. & Baumler, A. J. (2013). Colonization resistance: battle of the bugs or Menage a Trois with the host? *PLOS Pathogens: A Peer-Reviewed Open-Access Journal,* 9**,** e1003730.

Stahl, M., Butcher, J. & Stintzi, A. (2012). Nutrient acquisition and metabolism by *Campylobacter jejuni. Frontiers in Cellular and Infection Microbiology,* 2**,** 5.

Stahl, M., Friis, L. M., Nothaft, H., Liu, X., Li, J., Szymanski, C. M. & Stintzi, A. (2011). L-fucose utilization provides *Campylobacter jejuni* with a competitive advantage. *Proceedings of the National Academy of Sciences U S A,* 108**,** 7194-9.

Stahl, M., Ries, J., Vermeulen, J., Yang, H., Sham, H. P., Crowley, S. M., Badayeva, Y., Turvey, S. E., Gaynor, E. C., Li, X. & Vallance, B. A. (2014). A novel mouse model of *Campylobacter jejuni* gastroenteritis reveals key pro-inflammatory and tissue protective roles for Toll-like receptor signaling during infection. *PLoS Pathog,* 10, e1004264.

Stahl, M. & Vallance, B. A. (2015). Insights into *Campylobacter jejuni* colonization of the mammalian intestinal tract using a novel mouse model of infection. *Gut Microbes,* 6, 143-8.

Stanton, R. C. (2012). Glucose-6-phosphate dehydrogenase, NADPH, and cell survival. *IUBMB Life,* 64, 362-9.

Steele, E. & Owen, R. (1988). *Campylobacter jejuni* subsp. *doylei* subsp. nov. a Subspecies of Nitrate-Negative *Campylobacters* Isolated from Human Clinical Specimens. *International journal of systematic bacteriology,* 38, 316-318.

Stuart, A., Prescott, C., Macintyre, S., Sethar, A., Neuman, B., Mccarthy, N. D., Wimalarathna, H. & Maiden, M. (2011). *The role of rodents as carriers of disease on UK farms: a preliminary investigation.*

Szymanski, C. M. (2015). Are Campylobacters now capable of carbo-loading? *Mol Microbiol,* 98, 805-8.

Szymanski, C. M. & Wren, B. W. (2005). Protein glycosylation in bacterial mucosal pathogens. *Nature Reviews Microbiology,* 3, 225-37.

Taieb, F., Svab, D., Watrin, C., Oswald, E. & Toth, I. (2015). Cytolethal distending toxin A, B and C subunit proteins are necessary for the genotoxic effect of *Escherichia coli* CDT-V. *Acta Veterinaria Hungarica,* 63, 1-10.

Tam, C. C., Rodrigues, L. C., Viviani, L., Dodds, J. P., Evans, M. R., Hunter, P. R., Gray, J. J., Letley, L. H., Rait, G., Tompkins, D. S. & O'brien, S. J. (2012). Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut,* 61, 69-77.

Taylor, D. E., Eaton, M., Yan, W. & Chang, N. (1992). Genome maps of *Campylobacter jejuni* and *Campylobacter coli. Journal of Bacteriology,* 174, 2332-2337.

Teh, A. H. T., Lee, S. M. & Dykes, G. A. (2017). Identification of potential *Campylobacter jejuni* genes involved in biofilm formation by EZ-Tn5 Transposome mutagenesis. *BMC Research Notes,* 10, 182.

Teh, K. H., Flint, S. & French, N. (2010). Biofilm formation by *Campylobacter jejuni* in controlled mixed-microbial populations. *International Journal of Food Microbiology* 143, 118-24.

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., Angiuoli, S. V., Crabtree, J., Jones, A. L., Durkin, A. S., Deboy, R. T., Davidsen, T. M., Mora, M., Scarselli, M., Margarit Y Ros, I., Peterson, J. D., Hauser, C. R., Sundaram, J. P., Nelson, W. C., Madupu, R., Brinkac, L. M., Dodson, R. J., Rosovitz, M. J., Sullivan, S. A., Daugherty, S. C., Haft, D. H., Selengut, J., Gwinn, M. L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'connor, K. J., Smith, S., Utterback, T. R., White, O., Rubens, C. E., Grandi, G., Madoff, L. C., Kasper, D. L., Telford, J. L., Wessels, M. R., Rappuoli, R. & Fraser, C. M. (2005). Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". *Proc Natl Acad Sci U S A,* 102, 13950-5.

Thomas, C., Hill, D. & Mabey, M. (2002). Culturability, injury and morphological dynamics of thermophilic *Campylobacter* spp. within a laboratory-based aquatic model system. *Journal of Applied Microbiology,* 92, 433-42.

Toth, I., Herault, F., Beutin, L. & Oswald, E. (2003). Production of Cytolethal Distending Toxins by Pathogenic *Escherichia coli* Strains Isolated from Human and Animal Sources: Establishment of the Existence of a New cdt Variant (Type IV). *Journal of Clinical Microbiology,* 41**,** 4285-4291.

Tsai, C. J., Loh, J. M. & Proft, T. (2016). *Galleria mellonella* infection models for the study of bacterial diseases and for antimicrobial drug testing. *Virulence,* 7**,** 214-29.

Tu, Q. V., Mcguckin, M. A. & Mendz, G. L. (2008). *Campylobacter jejuni* response to human mucin MUC2: modulation of colonization and pathogenicity determinants. *Journal of Medical Microbiology,* 57**,** 795-802.

Van Iersel, L., Kelk, S., Rupp, R. & Huson, D. (2010). Phylogenetic networks do not need to be complex: using fewer reticulations to represent conflicting clusters. *Bioinformatics,* 26**,** i124-31.

Van Mourik, A., Bleumink-Pluym, N. M., Van Dijk, L., Van Putten, J. P. & Wosten, M. M. (2008). Functional analysis of a *Campylobacter jejuni* alkaline phosphatase secreted via the Tat export machinery. *Microbiology,* 154**,** 584-92.

Van Putten, J. P., Van Alphen, L. B., Wosten, M. M. & De Zoete, M. R. (2009). Molecular mechanisms of *Campylobacter* infection. *Current Topics in Microbiology and Immunology* 337**,** 197-229.

Van Vliet, A. H. & Ketley, J. M. (2001). Pathogenesis of enteric *Campylobacter* infection. *Symp Ser Soc Appl Microbiol***,** 45s-56s.

Vandamme, P., Falsen, E., Rossau, R., Hoste, B., Segers, P., Tytgat, R. & De Ley, J. (1991). Revision of *Campylobacter, Helicobacter*, and *Wolinella* taxonomy: emendation of generic descriptions and proposal of Arcobacter gen. nov. *International Journal of Systematic and Evolutionary Microbiology,* 41**,** 88-103.

Vegge, C. S., Jansen Van Rensburg, M. J., Rasmussen, J. J., Maiden, M. C., Johnsen, L. G., Danielsen, M., Macintyre, S., Ingmer, H. & Kelly, D. J. (2016). Glucose Metabolism via the Entner-Doudoroff Pathway in *Campylobacter*: A Rare Trait that Enhances Survival and Promotes Biofilm Formation in Some Isolates. *Frontiers in Microbiology,* 7**,** 1877.

Velayudhan, J., Jones, M. A., Barrow, P. A. & Kelly, D. J. (2004). L-serine catabolism via an oxygen-labile L-serine dehydratase is essential for colonization of the avian gut by *Campylobacter jejuni. Infection and Immunity,* 72**,** 260-8.

Velayudhan, J. & Kelly, D. J. (2002). Analysis of gluconeogenic and anaplerotic enzymes in *Campylobacter jejuni*: an essential role for phosphoenolpyruvate carboxykinase. *Microbiology,* 148**,** 685-94.

Veron, M. & Chatelain, R. (1973). Taxonomic Study of the Genus *Campylobacter* Sebald and Vkron and Designation of the Neotype Strain for the Type Species, *Campylobacter fetus* (Smith and Taylor) Sebald and Vkon. *International Journal of Systematic Bacteriology,* 23 122-134.

Vorwerk, H., Huber, C., Mohr, J., Bunk, B., Bhuju, S., Wensel, O., Sproer, C., Fruth, A., Flieger, A., Schmidt-Hohagen, K., Schomburg, D., Eisenreich, W. & Hofreuter, D. (2015). A transferable plasticity region in *Campylobacter coli* allows isolates of an otherwise non-glycolytic food-borne pathogen to catabolize glucose. *Molecular Microbiology,* 98**,** 809-30.

Wagley, S., Newcombe, J., Laing, E., Yusuf, E., Sambles, C. M., Studholme, D. J., La Ragione, R. M., Titball, R. W. & Champion, O. L. (2014). Differences in carbon source utilisation distinguish *Campylobacter jejuni* from *Campylobacter coli*. *BMC Microbiology,* 14**,** 262.

Walker, R. I., Caldwell, M. B., Lee, E. C., Guerry, P., Trust, T. J. & Ruiz-Palacios, G. M. (1986). Pathophysiology of *Campylobacter* enteritis. *Microbiological Reviews,* 50**,** 81-94.

Wang, G., Clark, C. G., Taylor, T. M., Pucknell, C., Barton, C., Price, L., Woodward, D. L. & Rodgers, F. G. (2002). Colony Multiplex PCR Assay for Identification and Differentiation of *Campylobacter jejuni, C. coli, C. lari, C. upsaliensis,* and *C. fetus subsp. fetus*. *Journal of Clinical Microbiology,* 40**,** 4744-4747.

Wang, P. Y. (1990). [Cereal grain preference of rats]. *Gaoxiong Yi Xue Ke Xue Za Zhi,* 6**,** 402-7.

Wanken, A. E., Conway, T. & Eaton, K. A. (2003). The Entner-Doudoroff Pathway Has Little Effect on *Helicobacter pylori* Colonization of Mice. *Infection and Immunity,* 71**,** 2920-2923.

Watson, R. O., Novik, V., Hofreuter, D., Lara-Tejero, M. & Galan, J. E. (2007). A MyD88-deficient mouse model reveals a role for Nramp1 in Campylobacter jejuni infection. *Infect Immun,* 75**,** 1994-2003.

Weingarten, R. A., Grimes, J. L. & Olson, J. W. (2008). Role of *Campylobacter jejuni* respiratory oxidases and reductases in host colonization. *Applied and Environmental Microbiology,* 74**,** 1367-75.

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. *Infection and Immunity,* 66**,** 1934-40.

Wilson, D. J., Gabriel, E., Leatherbarrow, A. J., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Fearnhead, P., Hart, C. A. & Diggle, P. J. (2008). Tracing the source of campylobacteriosis. *PLOS Genetics: A Peer-Reviewed Open-Access Journal,* 4**,** e1000203.

Wingstrand, A., Neimann, J., Engberg, J., Nielsen, E. M., Gerner-Smidt, P., Wegener, H. C. & Molbak, K. (2006). Fresh chicken as main risk factor for campylobacteriosis, Denmark. *Emerging Infectious Diseases,* 12**,** 280-5.

Wood, M. D., Beresford, N. A. & Copplestone, D. (2012). Limit of detection values in data analysis: Do they matter? *Radioprotection,* 46**,** S85-S90.

Wright, J. A., Grant, A. J., Hurd, D., Harrison, M., Guccione, E. J., Kelly, D. J. & Maskell, D. J. (2009). Metabolite and transcriptome analysis of *Campylobacter jejuni* in vitro growth reveals a stationary-phase physiological switch. *Microbiology,* 155**,** 80-94.

Yahara, K., Meric, G., Taylor, A. J., De Vries, S. P., Murray, S., Pascoe, B., Mageiros, L., Torralbo, A., Vidal, A., Ridley, A., Komukai, S., Wimalarathna, H., Cody, A. J., Colles, F. M., Mccarthy, N., Harris, D., Bray, J. E., Jolley, K. A., Maiden, M. C., Bentley, S. D., Parkhill, J., Bayliss, C. D., Grant, A., Maskell, D., Didelot, X., Kelly, D. J. & Sheppard, S. K. (2017). Genome-wide association of functional traits linked with *Campylobacter jejuni* survival from farm to fork. *Environ Microbiol,* 19**,** 361-380.

Yang, Z. & Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nat Rev Genet,* 13**,** 303-14.

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. *Molecular Microbiology,* 14**,** 883-93.

Yao, R., Burr, D. H. & Guerry, P. (1997). CheY-mediated modulation of *Campylobacter jejuni* virulence. *Molecular Microbiology,* 23**,** 1021-31.

Young, K. T., Davis, L. M. & Dirita, V. J. (2007). *Campylobacter jejuni*: molecular biology and pathogenesis. *Nature Reviews Microbiology,* 5**,** 665-79.

Yu, R. K., Usuki, S. & Ariga, T. (2006). Ganglioside molecular mimicry and its pathological roles in Guillain-Barre syndrome and related diseases. *Infect Immun*, 74, 6517-27.

Zautner, A. E., Herrmann, S., Corso, J., Tareen, A. M., Alter, T. & Gross, U. (2011). Epidemiological association of different *Campylobacter jejuni* groups with metabolism-associated genetic markers. *Applied and Environmental Microbiology*, 77, 2359-65.

Zautner, A. E., Tareen, A. M., Gross, U. & Lugert, R. (2012). Chemotaxis in *Campylobacter jejuni. European Journal of Microbiology & Immunology (Bp)*, 2, 24-31.

Zeng, X., Xu, F. & Lin, J. (2009). Molecular, antigenic, and functional characteristics of ferric enterobactin receptor CfrA in *Campylobacter jejuni. Infection and Immunity*, 77, 5437-48.

Zhang, Q., Meitzler, J. C., Huang, S. & Morishita, T. (2000). Sequence polymorphism, predicted secondary structures, and surface-exposed conformational epitopes of Campylobacter major outer membrane protein. *Infect Immun*, 68, 5679-89.

Zilbauer, M., Dorrell, N., Wren, B. W. & Bajaj-Elliott, M. (2008). *Campylobacter jejuni*-mediated disease pathogenesis: an update. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 102, 123-129.

Ziprin, R. L., Young, C. R., Stanker, L. H., Hume, M. E. & Konkel, M. E. (1999). The absence of cecal colonization of chicks by a mutant of *Campylobacter jejuni* not expressing bacterial fibronectin-binding protein. *Avian Dis*, 43, 586-9.